

A blue and black humanoid robot with a camera on its head and a worker in the background. The robot is in the foreground, and a worker in a white hard hat and yellow safety vest is in the background, holding a laptop. The robot has a blue body with black joints and a black head with a camera. The worker is standing in a construction site with concrete walls and a doorway in the background.

Proceedings of the 39th International Symposium on Automation and Robotics in Construction

Bogotá, Colombia, July 13-15, 2022

NYU / ABU DHABI

Proceedings of the 39th International Symposium on Automation and Robotics in Construction

ISSN (for the proceedings series): 2413-5844

ISBN (for this issue of the proceeding series): 978-952-69524-2-0

The proceeding series is Scopus indexed.

Scopus®



Copyrights reserved.

© 2022 International Association on Automation and Robotics in Construction

This work including all its parts is protected by copyright. Any use outside the narrow limits of copyright law without the consent of the individual authors is inadmissible and punishable. This applies in particular to reproductions, translations, microfilming and saving and processing in electronic systems.

The reproduction of common names, trade names, trade names etc. in this work does not justify the assumption that such names are to be regarded as free within the meaning of the trademark and trademark protection legislation and can therefore be used by everyone, even without special identification.

Cover design: Borja García de Soto, Samuel Prieto, Xinghui Xu, Semih Sonkor; Image: Robotic system for the autonomous collection of laser scans in construction sites developed by the S.M.A.R.T. Construction Research Group at NYUAD

Editorial Board

Editors in Chief

Linner, Thomas | García de Soto, Borja | Hu, Rongbo | Brilakis, Ioannis

Editors (Area Chairs)

Bock, Thomas & Pan, Wen

Carbonari, Alessandro

Castro, Daniel & Mesa, Harrison

Feng, Chen

Fischer, Martin & Brosque, Cynthia

Gonzalez, Vicente

Hall, Daniel & Ng, Ming Shan

Kamat, Vineet & Liang, Ci-Jyun

Lafhaj, Zoubeir

Pan, Wei & Pan, Mi

Zhu, Zhenhua

Local Organizing Committee

Guevara Maldonado, José Alberto (Committee Head)

Castro Lacouture, Daniel

García Rodríguez, Salvador

Mesa Hernández, Harrison

Kwon Cho, Yong

Sponsor:

Autodesk



MCad



Construsoft



CONSTRUSOFT

Cámara Colombiana de la
Construcción (CAMACOL)



Foreword

The International Association for Automation and Robotics in Construction (IAARC) and the 39th ISARC organizing committee are pleased to present the Proceedings of the 39th International Symposium on Automation and Robotics in Construction held, in a hybrid mode, on July 13-15, 2022, at Universidad de los Andes, Bogotá, Colombia. The 39th ISARC was proudly hosted by the Department of Civil and Environmental Engineering at Universidad de los Andes. The 2022 ISARC has been the first-ever ISARC in Latin America and it was organized in collaboration with Pontificia Universidad Católica de Chile and Tecnológico de Monterrey. It also received the support of the Purdue Polytechnic Institute and Georgia Institute of Technology. A total of 89 papers from 291 authors/co-authors representing 111 universities, labs, and companies in 26 countries were selected after a rigorous peer-review process that was possible thanks to the great support from the Area Chairs.

In the last 4-5 years, the growing need and interest in construction robotics have become highly evident worldwide. Start-ups, spin-offs, and investors have introduced more than 200 robot systems into the market. This is backed up by an enormous number of activities and projects carried out in the academic area pushing the boundaries of what is technologically possible.

Competency in digital construction, automation and robotics has become a key element for all stakeholders in the construction sector, and many universities worldwide have launched dedicated interdisciplinary programs. Governments and major funding programs such as Horizon Europe massively request and fund the development of robotic solutions for construction, such as drones, mobile robots, 3D-printing solutions, cable-driven robots, and exoskeletons. Regulators and standardization organizations started to develop the first certification and standardization schemes for construction robots, and large software companies attempt to simulate and program robotic construction processes efficiently based on digital building and construction data.

ISARC continues to be the premier global conference in the domain of automation and robotics in construction. To prepare ISARC for the future, further attempts were made this year to restructure IAARC's framework, and the topic of "Applications in Developing Countries" was added as a new submission domain. The technical areas summarize the paper topic areas of interest, representing all the research themes relevant to ISARC/IAARC. This is an important mechanism for the technical committee to consolidate the knowledge accumulated from each year's conference while allowing for the smooth incorporation of new research topics and trends in the community.

We hope you find the papers interesting and inspirational. Enjoy the readings!

Thomas Linner
Borja García de Soto
Rongbo Hu
Ioannis Brilakis

Table of Contents

K - Keynote talks

A framework for a comprehensive mobile data acquisition setting for the assessment of Urban Heat Island phenomenon.....	1
Monic Pena Acosta, Faridaddin Vahdatikhaki, Joao Santos and André G. Dorée	
Using BIM and LCA to evaluate material circularity: Contributions to building design improvement.....	9
Haibo Feng, Qian Chen, Borja García de Soto and Mehrdad Arashpour	
Building Information Model Pre-Processing for Automated Geometric Quality Control.....	17
Martin Bueno and Frédéric Bosché	

P - Plenary talks

Insights into automation of construction process using parallel-kinematic manipulators.....	25
Maike Klöckner, Mathias Haage, Helena Eriksson, Henrik Malm, Klas Nilsson, Anders Robertsson and Ronny Andersson	
Predictive Maintenance Optimization Framework for Pavement Management.....	33
Amit Kumar and Omidreza Shoghli	
Weakly supervised pseudo label generation for construction vehicle segmentation.....	41
Wei-Chih Chern, Vijayan Asari, Tam Nguyen and Hongjo Kim	
Teaching Construction Robotics for Higher Education Students: “Imagine and Make”.....	47
Zoubeir Lafhaj, Wassim Al Balkhy and Thomas Linner	
Integrating VR and Simulation for Enhanced Planning of Asphalt Compaction.....	55
Andre Renato Revollo Dalence, Faridaddin Vahdatikhaki, Seirgei Miller and André G. Dorée	
Direction of Arrival of Equipment Sound in the Construction Industry.....	63
Kehinde Elelu, Tuyen Le and Chau Le	

A - Sensing systems & data infrastructures

Towards BIM-based robot localization: a real-world case study.....	71
Huan Yin, Jia Min Liew, Wai Leong Lee, Marcelo H Ang Jr and Justin Ker-Wei Yeoh	
Current State and Future Opportunities of Data Mining for Construction 4.0.....	78
Keyi Wu and Borja García de Soto	
The Use of Boston Dynamics SPOT in Support of LiDAR Scanning on Active Construction Sites.....	86
Eric Wetzel, Junshan Liu, Tom Leathem and Anoop Sattineni	
Autonomous operation of a robot dog for point cloud data acquisition of scaffolds.....	93
Duhoo Chung, Sunwoong Paik, Juhyeon Kim and Hyungkwan Kim	
Cluster-based Deterioration Prediction of Composite Pavements with Incorporation of Flooding.....	99
Ishan Neema, Fatemeh Banani Ardecani and Omidreza Shoghli	
Predicting construction productivity with machine learning approaches.....	107
Laura Florez-Perez, Zhiyuan Song and Jean C. Cortissoz	
Automated Checking of Scaffold Safety Regulations using Multi-Class 3D Segmentation.....	115
Jeehoon Kim, Juhyeon Kim, Nahye Koo and Hyoungkwan Kim	
A Systematic Classification and Evaluation of Automated Progress Monitoring Technologies in Construction.....	120
Varun Kumar Reja, Megha S Pradeep and Koshy Varghese	
Reducing spatial error in mobile laser scanning by real-time uncertainty visualization and human-machine interaction.....	128
Maciej Trzeciak, Chris Burgoyne and Ioannis Brilakis	
Toward Automation in Crack Detection and Measurements: Benchmarking of CNN-based Algorithms.....	136
Donghoon Ji, Yelda Turkan and Paolo Calvi	

B - Sensing systems & data infrastructures

Innovative Model for Forecasting Trailer Usage for Prefabricated Exterior Wall Panels...	144
Andrew Rener, Aslihan Karatas and Matthew Cole	
Automated Wall Detection in 2D CAD Drawings to Create Digital 3D Models.....	152
Chialing Wei, Mohit Gupta and Thomas Czerniawski	
IFC-based Information Extraction and Analysis of HVAC Objects to Support Building Energy Modeling.....	159
Hang Li and Jiansong Zhang	
Systematic Literature Review of Building Information Modelling and Green Building Certification Systems.....	167
Oludolapo Olanrewaju, Wallace Enegbuma and Michael Donn	
Information Extraction from Text Documents for the Semantic Enrichment of Building Information Models of Bridges.....	175
Phillip Schönfelder, Tariq Al-Wesabi, Andreas Bach and Markus König	
BIM-Integration of Light Construction Equipment.....	183
Maximilian Schöberl, Sebastian Offinger, Theo Goldfuß, Stephan Kessler and Johannes Fottner	

Visualisation and graph-based storage of customised changes in early design phases...	191
Daniel Napps, Ata Zahedi, Markus König and Frank Petzold	
Conceptual Modeling of Lifecycle Digital Twin Architecture for Bridges: A Data Structure Approach.....	199
Inga Maria Giorgadze, Faridaddin Vahdatikhaki and Hans Voordijk	
Towards the development of a digital twin for subsoil monitoring and stability against overturning of a mobile drilling rig.....	207
Francisco Williams Riquer, Duy Anh Dao and Jürgen Grabe	
Closing the Gap Between Concrete Maturity Monitoring and Nonlinear Time-dependent FEM Analysis through a Digital Twin. Case Study: Post-tensioned Concrete Slab of an Office Building, Barcelona, Spain.....	215
Héctor Posada, Rolando Chacón, Lucian-Constantin Ungureanu, David García	
ColombiaClass: proposal for a BIM Classification System for public buildings in Colombia.....	223
María Del Pilar Revuelta Mendoza, Nelly Paola García Lopez and Luis Arturo Salazar Fica	
Towards Automated Pipelines for Processing Load Test Data on a HS Railway Bridge in Spain using a Digital Twin.....	231
Carlos Ramonell and Rolando Chacón	
Potentials of 5D BIM with Blockchain-enabled Smart Contracts for Expediting Cash Flow in Construction Projects.....	238
Jong Han Yoon and Pardis Pishdad-Bozorgi	
Integrating Digital Twins in Construction Education Through Hands-on Experiential Learning.....	246
Steven Kangisser, Javier Irrazary, Kelly Watt, Richard Borger and Amadeus Burger	
Benefits of Open Infra BIM – Finland Experience.....	253
Rauno Heikkilä, Tanja Kolli and Teppo Rauhala	

C - Human factors & human-system collaboration

The Neural Basis of Risk Attitude in Decision-Making Under Risk: fNIRS investigation of the Simulated Electrical Construction Task.....	261
Shiva Pooladvand, Deha Ay and Sogand Hasanzadeh	
Toward Personalized Safety Training: Automating the Classification of Construction Workers' Cognitive Failures.....	268
Kyeongsuk Lee, Yugandhar Shinde, Sogand Hasanzadeh and Behzad Esmaeili	
iVisit-Collaborate: Online Multiuser Virtual Site Visits Using 360-Degree Panoramas and Virtual Humans.....	276
Ricardo Eiris and Masoud Gheisari	
Vulnerability Assessment of Construction Equipment: An Example for an Autonomous Site Monitoring System.....	283
Muammer Semih Sonkor, Xinghui Xu, Samuel A. Prieto and Borja García de Soto	
Fire Evacuation and Management Model Based on Building Information Modeling and Virtual Reality.....	291
Ren-Jie Gao, Kun-Chi Wang, Xiang-Hao Lai and Wei-Hsiang Hung	
Evaluation of Construction Site Layout Using Virtual Reality Linked with 3D CAD and Body Tracking.....	297
Hidetoshi Tsuda and Kosei Ishida	

Automated Recognition of Hand Gestures for Crane Rigging using Data Gloves in Virtual Reality.....	304
Aparna Harichandran and Jochen Teizer	
A Knowledge Graph for Automated Construction Workers' Safety Violation Identification.....	312
Yifan Zhu and Xiaowei Luo	
Wearable Technology for Highway Maintenance and Operation Safety: A Survey of Workers' Perception and Preferences.....	320
Sepehr Sabeti, Omidreza Shoghli, Nichole Morris and Hamed Tabkhi	
Integrated Sensor-Based Interface for Human-Robot Collaboration in Construction.....	328
Xin Wang, Dharmaraj Veeramani and Zhenhua Zhu	

D - Robotic machines, devices, and end-effectors

Requirements analysis of additive manufacturing for concrete printing – A systematic review.....	336
Patricia Peralta Abadia and Kay Smarsly	
Safety Risk Assessment of Drones on Construction Sites using 4D Simulation.....	344
Zixian Zhu, Idris Jeelani and Masoud Gheisari	
Progress Estimation of an Excavation Pit.....	352
Axel Vierling, Tobias Groll, Dennis Meckel, Kristina Heim, Daniel Walter, Lukas Scheidhauer, Karsten Körkemeyer and Karsten Berns	
Impact of Passive Back-Support Exoskeletons on Manual Material Handling Postures in Construction.....	359
Ali Golabchi, Linda Miller, Hossein Rouhani and Mahdi Tavakoli	
Stag hunt game-based approach for cooperative UAVs.....	367
Lanh Van Nguyen, Ignacio Torres Herrera, Trung Hoang Le, Duong Manh Phung, Ricardo Patricio Aguilera and Quang Phuc Ha	
Reinforcement learning with construction robots: A review of research areas, challenges and opportunities.....	375
Xinghui Xu and Borja García de Soto	
Constrained Control Scheme for the Manipulation of Heavy Pre-fabricated Elements with Lightweight Robotic Arm.....	383
Michele Ambrosino, Fabian Boucher, Pierre Mengeot and Emanuele Garone	
Importance of a 5G Network for Construction Sites: Limitation of WLAN in 3D Sensing Applications.....	391
Hyung Joo Lee, Ajith Krishnan, Sigrid Brell-Cokcan, Janina Knußmann, Maximilian Brochhaus, Robert H. Schmitt, Johannes J. Emontsbotz and Johannes Sieger	
Path Generation for Foam Additive Manufacturing of Large Parts with a Cable-Driven Parallel Robot.....	399
Elodie Paquet, Marceau Metillon, Kevin Subrin, Benoit Furet and Stéphane Caro	
Automated material selection based on detected construction progress.....	406
Tyler Stephans, Austin McClymonds, Robert Leicht and Alan Wagner	
Solution Kits for automated and robotic façade upgrading.....	414
Kepa Iturralde, Thomas Bock, Philip Zimmermann, Renzo G. Bazan Santos, Samanti Das, Wenlan Shen, Mahruk Malik, Steven Maio, Agi Hidri and Juncheng Shen	

Concept of a Robot Assisted On-Site Deconstruction Approach for Reusing Concrete Walls.....	422
Hyung Joo Lee, Christoph Heuer and Sigrid Brell-Cokcan	
Subjective Evaluation of Passive Back-Support Wearable Robot for Simulated Rebar Work.....	430
Nihar Gonsalves, Mohamad Khalid, Abiola Akinniyi, Omobolanle Ogunseiju and Abiola Akanmu	
Industry Perspectives of the Potential of Wearable Robot for Pipe Installation Work.....	437
Nihar Gonsalves, Mohamad Khalid, Abiola Akinniyi and Abiola Akanmu	
Development of Framework for Highway Lawn Condition Monitoring using UAV Images	444
Yeseul Kim, Seongyong Kim, Yosuke Yajima, Javier Irizarry and Yong K. Cho	

E - Construction management techniques

YAKE-Guided LDA approach for automatic classification of construction safety reports...	451
Hrishikesh Gadekar and Nikhil Bugalia	
Automated Construction Contract Summarization Using Natural Language Processing and Deep Learning.....	459
Xiaorui Xue, Yiru Hou and Jiansong Zhang	
Construction Progress Monitoring and Reporting using Digital Images and Computer Vision Techniques – A Review.....	467
Dena Shamsollahi, Osama Moselhi and Khashayar Khorasani	
BIM, Twin and Between: a Conceptual Engineering Approach to Formalize Digital Twins in Construction.....	475
Fabiano Correa	
Inspection Data Exchange and Visualization for Building Maintenance using AR-enabled BIM.....	483
Jisoo Park, Soowon Chang, Hyungi Lee and Yong Cho	

F - Services and business applications / Industry papers

Comparison of TLS and Photogrammetric 3D Data Acquisition Techniques: Considerations for Developing Countries.....	491
Eyob Mengiste, Samuel A. Prieto and Borja García de Soto	
Mud Dauber: Prototype of the Mobile Gantry Architecture.....	495
Peter Staritz, Josiah McClurg, Caleb Miller, Moriah Schlenker and Shane Wozniak	
A model driven method for crack detection in robotic inspection.....	500
Erika Pellegrino and Tania Stathaki	
A Case Study: Conception of digitalizing prefabrication processes in the construction industry.....	504
Zhen Cai, Mohamed Ben Tarfaoui, Stephan Kessler and Johannes Fottner	
On the Bar Installation Order for the Automated Fabrication of Rebar Cages.....	508
Mahdi Momeni, Johan Reliefs, Lars Pettersson, Alessandro Vittorio Papadopoulos and Thomas Nolte	
A Taxonomy for Connected Autonomous Plant.....	512
Cormac Browne, Ross Walker, Tim Embley, Muneer Akhtar, Amer Essa, Anette Pass, Simon Smith and Alex Wright	

Digital Commissioning Processes for the Oil and Gas Sector.....	520
Daniel Luiz de Mattos Nascimento, Alessandra Brancher Roeder, Diego Calvetti, Alexander Chavez Lopez, Fernando Rodrigues Gonzalez and Flavio Araujo	
Overview of the State-of-Practice of BIM in the AEC Industry in the United States.....	524
Hala Nassereddine, Makram Bou Hatoum and Awad Hanna	
Machine Learning for Construction Process Control: Challenges and Opportunities.....	532
Bharath Sankaran	
5G Wireless Communication for Autonomous Excavation.....	536
Rauno Heikkilä, Matti Immonen, Heikki Keränen, Olli Liinamaa, Esa Piri and Tanja Kolli	

G - Technology management and innovation

The Use of Drones in the Construction Industry: Applications and Implementation.....	542
Makram Bou Hatoum and Hala Nassereddine	
Intelligent BIM-based spatial conflict simulators: A comparison with commercial 4D tools.....	550
Leonardo Messi, Borja García de Soto, Alessandro Carbonari and Berardo Naticchia	
BIM-assisted, automated processes for commissioning in building services engineering	558
Ralf Becker, Christoph Blut, Christoph Emunds, Jérôme Frisch, Dirk Heidemann, Tristan Kinnen, Alexander Lenz, Michael Möller, Nicolas Pauen, Tobias Rettig, Dominik Schlütter, Matthias Wenthe, Jörg Blankenbach, Günter Bleimann-Gather, Johannes Fütterer, Jörg Jungedeitering and Christoph van Treeck	
Blockchain-supported design tool on building element scale.....	566
Goran Sibenik, Marijana Sreckovic, Thomas Preindl, Martin Kjær, and Wolfgang Kastner	
Digital Transformation in Asset Management – A Case of BIM Adoption in New Zealand Local government.....	574
John Jiang, Theuns Henning and Yang Zou	
Project-based curriculum for teaching construction robotics.....	582
Cynthia Brosque and Martin Fischer	
A case-based reasoning technique for evaluating performance improvement in automated construction projects.....	590
Krishnamoorthi Sundararaman and Benny Raphael	
The Optimized Process for Effective Decision Makings to Minimize Fall From Height (FFH) Accidents on Construction Site.....	597
Qinghao Zeng and Jong Han Yoon	
Construction Robotics Excellence Model: A framework to overcome existing barriers for the implementation of robotics in the construction industry.....	605
Jan-Iwo Jäkel, Shervin Rahnama and Katharina Klemm-Albert	
State of Advances in Reality Capture for Construction Progress Monitoring and Documentation.....	614
Jordan Grogan, Anoop Sattineni and Jeff Kim	

H - Applications in developing countries

An Approach for Estimation of Swing Angle and Digging Depth During Excavation Operation.....	622
Amirmasoud Molaei, Marcus Geimer and Antti Kolu	
Automated Valve Detection in Piping and Instrumentation (P&ID) Diagrams.....	630
Mohit Gupta, Chialing Wei and Thomas Czerniawski	
Suitability and Effectiveness of Visualization Platform-based Construction Safety Training modules.....	638
Kishor Bhagwat and Venkata Santosh Kumar Delhi	
Semiarid Terrain Alteration for Converting Dryland into Arable Land – Construction and Earthmoving Perspectives.....	646
Moshe Alamaro, Joseph Louis and Jochen Teizer	
Action Sequencing in Construction Accident Reports using Probabilistic Language Model.....	653
Quan Do, Tuyen Le and Chau Le	

A framework for a comprehensive mobile data acquisition setting for the assessment of Urban Heat Island phenomenon

M. Pena Acosta, F. Vahdatikhaki, J. Santos, and A. G. Dorée

Department of Construction Management and Engineering, University of Twente, Enschede, the Netherlands
E-mail {m.penaacosta@utwente.nl, f.vahdatikhaki@utwente.nl, j.m.oliveiradossantos@utwente.nl, a.g.doree@utwente.nl}

Abstract –

The debates around the Urban Heat Island phenomenon (UHI) have gained momentum in the context of smart cities and sustainable development. It is crucial to understand the complex interaction between urban features and temperature variation in the city based on reliable and detailed data. Yet, the complex interaction between the UHI of the canopy layer, paved surfaces and urban geometries (e.g., buildings, vegetation, and urban elements) has not been intensively explored to accurately capture their interplay. This is mainly caused by the palpable absence of comprehensive data that can support this type of correlational analysis. This paper proposes a comprehensive data acquisition framework to guide the collection of the required data for the development of a data-driven UHI assessment model, with a specific focus on the contributions of paved roads to UHI.

The framework was tested with a case study in Apeldoorn, the Netherlands, during a period of six months. The data collected, highlights the useability of the proposed framework for collecting high-resolution urban data required to assist local governments and urban planners to make informed decisions. To the best of authors' knowledge, this is the first time the interplay between urban feature, surface and air temperatures has been measured via mobile transects.

Keywords

Data-driven methods; data collection; smart and sustainable cities; mobile sensing systems; urban heat island

1 Introduction

The extreme weather conditions caused by climate change are reshaping the world. There were 38 heat waves in Europe in the last century, 17 of them in the last decade. Only the heat wave of 2003 caused 70,000 excess deaths over 4 months in Central and Western Europe [1].

In urban areas, the negative effects of climate change are greater. This is because changes in the natural environment, render urban areas more prone to store solar radiation [2]. 75% of the world's population is living in fast growing urban areas, and this number is projected to continue to increase in the years to come [3]. While global efforts are focused on climate adaptation and smart, sustainable urban development, climate policies and actions are often based on subjective knowledge due to the absence of sound and comprehensive data [4].

UHI phenomenon is defined as the temperature difference between the suburbs and the inner city. To measure the UHI effects, two main approaches have been widely adopted: (1) air temperatures, which refers to the UHI of the canopy layer (CUHI), and (2) the surface UHI (SUHI), which refers to the thermal emissivity of land surfaces [5]. CUHIs are usually studied by measuring air temperatures, typically at about 2 m above the ground via fixed or mobile weather stations. SUHI, on the other hand, are monitored via remote sensing data. Although techniques to measure and explain both CUHI and SUHI have been successful in explaining these phenomena, they possess a few limitations. First, commonly fixed weather stations are used to continuously measure air temperatures. While weather stations collect the data with a high frequency, the ability of the data coming from limited fixed locations to represent the temperature variability in the city (from the city center to the outskirts) is questionable. Moreover, weather stations are often located in open areas to avoid interference from shading or urban factors [6]. As such, they do not capture the effects of drivers that intensify the UHI phenomena. Remote sensing technologies, on the other hand, do provide sufficient data resolution to characterize inner urban centers and rural surroundings. However, they do not account for temperature differential above the ground.

To overcome these limitations, researchers worldwide have modeled the spatial variability of urban temperatures by applying sensing technologies and spatial and numerical models that integrate data from multiple sources [7]. In addition, the technological

advances of recent years have made it possible to deploy mobile weather stations to map the temperature variation in the city at a higher resolution. The emphasis, however, has been placed on the relationship between urban geometries (e.g., Sky View Factor (SVF), height to width ratio (H/W)) and air temperature [8, 9]. Other studies have looked at the relationship between geographic characteristics, such as proximity to the coast, rivers, and land cover [10, 11]. Nevertheless, the relationship between paved roads and temperature variation in the city has not received much attention. This is a major oversight because paved surface has been commonly identified as one of the main drivers of UHI.

Between 2016 and 2021, a considerable amount of literature was published on the different field campaigns for the collection temperature data via mobile measurements [6, 8, 12-16]. While the majority targeted CUHI temperatures, a few studies focused on air temperatures and thermal comfort, the relation between air temperatures and evapotranspiration, and the interaction between air and surface temperatures.

However, the complex interaction between CUHI, paved surfaces and urban geometries (e.g., buildings, vegetation, and urban elements) has not been intensively explored to accurately capture their interplay. This is mainly caused by the palpable absence of comprehensive data that can support this type of correlational analysis. Therefore, a more comprehensive approach towards data collection is required to enable the collection of the data needed to build a more comprehensive temperature profile of urban areas.

Given the above limitation, this paper aims to develop a comprehensive data collection and processing framework to guide the collection, pre-processing and visualization of the data required for the development of a data-driven UHI assessment model, with a specific focus on the contributions of paved roads to UHI.

The remainder of the paper is structured as follows. First, the proposed framework is presented. This is followed by a brief explanation of a case study that demonstrates the applicability of the framework. The paper ends with conclusions and future work.

2 The framework

To address the research gap presented in Section 1, a comprehensive data collection and processing framework is developed. It encompasses three main steps (Figure 1). The first stage involves the development of a bicycle-based mobile urban data collection station. That is then followed by a data collection campaign. In the last stage, the collected data is pre-processed and visualized.

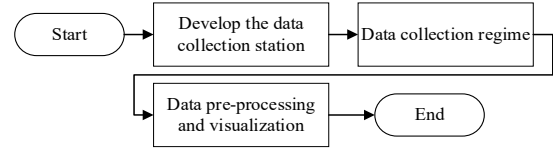


Figure 1: Schematic representation of the proposed framework

2.1 Development of data collection station

The development of the data collection station starts with the analysis of the required data, followed by the analysis of the type of sensor that can meet the data needs, and ends with the assembling of the mobile station.

2.1.1 Data required

The dark materials used to replace natural landscaping in cities store solar radiation during the day and release it when temperatures begin to drop at night. This intrinsic property of the materials used in the built environment is largely responsible for the UHI phenomenon. That has prompted researcher worldwide to devote their efforts to understand the behavior of these materials under different environmental conditions. As highlighted in Section 1, there is an urgent need to expand the body of knowledge on UHI-material interplay based on data that capture as much of the urban environment as possible.

As shown in Figure 2, the urban fabric (i.e., the urban materials) can be categorized in three main groups: building materials, roads (paved or unpaved), and vegetation. Regarding the temperature of buildings, the facade materials are the main concern. Hence, capturing the energy interaction between the building façade and the adjacent urban elements is a key task in the data collection campaign. For instance, the temperature of a building façade next to a green area can be greater or lower than the temperature of a façade in the business district where paved roads are predominant. As such, a thermal camera capable of screening temperatures at a larger distance is best suited for this task. Likewise, there are different types of paved roads, i.e., in terms of function and material type. Hence, a granular measurement of the surface temperature is required.

Finally, the air temperatures at the canopy layer must be taken into account. Air temperature readings should be recorded at the same location that the façade and road surface measurements are taken. This is crucial because the variation of air temperature in relation to the variation of urban geometries and their temperatures can signal the contributions of paved roads to UHI.

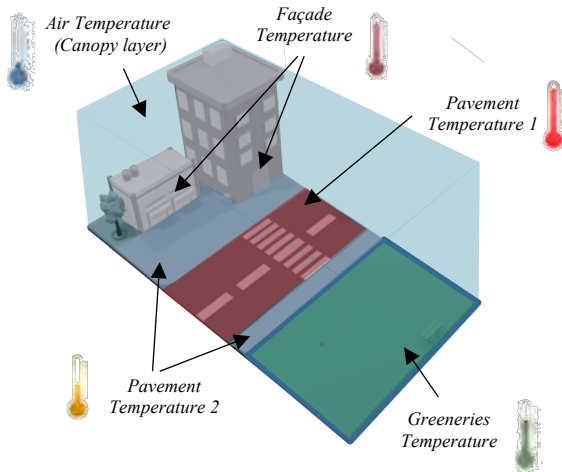


Figure 2: Schematic representation of data required for comprehensive UHI studies

2.1.2 The sensors

A bicycle-based mobile urban data-gathering station is used to roam around the city and collect geo-referenced and time-stamped temperature data at the level of road surfaces and above the ground level. As shown in Figure 3, the sensor kit includes a processing centre, a GPS rover, a mobile weather station, and an infrared camera.

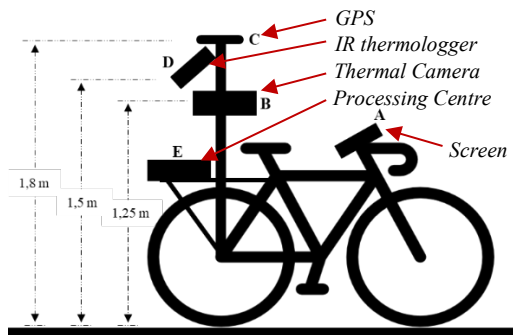


Figure 3: Schematic representation of the bicycle-based mobile urban data-gathering station

To record accurate locations, a GPS antenna is used. A dedicated GPS antenna is chosen instead of a smartphone because the latter do not provide the accurate and frequent enough data needed for the high-resolution data collection campaigns.

A Thermal imaging temperature sensor is used to read thermal images of the surrounding building façades. Thermal imaging cameras are widely used for building inspections, body temperature screening, etc. They are suited to read measurements at a distance without compromising the capabilities of capturing temperature

variations with high accuracy. Figure 4 presents a sample of an image taken from the thermal imaging sensor. Each frame contains temperatures for every pixel in the image. In this research, the average temperature along the horizontal line (as shown in Figure 4) of the middle of the frame was stored. This is because the middle of the image has a higher chance of providing unobstructed view of the façade (no leaves or cars).

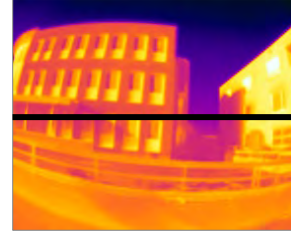


Figure 4: Sample image from the thermal camera

To capture the surface temperature, an infrared sensor is best suited. As shown in Figure 3, an InfraRed Thermometer (D), coupled with an environmental sensor, is placed 1.5 m above ground to read air temperature, relative humidity, and wet bulb.

All the sensors are connected to a dedicated processing centre. This processing centre runs a windows-based application developed to record, synchronize, and store the readings from the sensors in real-time. The data is stored in a comma-separated values (CSV) file format as follows: *Row ID, GPS time, Latitude, Longitude, Altitude, Thermal camera readings, Canopy air temperature, Wet bulb, Relative Humidity, Surface temperature, and time*. In addition, for each row of recorded data, a digital image, similar to Figure 4, is stored in a JPG format. The architecture of the processing centre is shown in Figure 5.

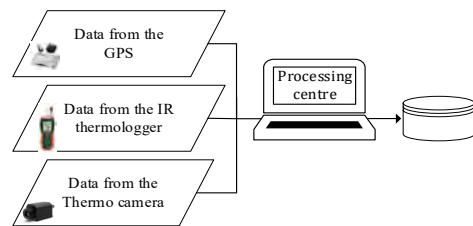


Figure 5: Illustration of the data storage process

2.2 Data collection campaign

Since the objective of the field campaign is to reveal the interplay between urban morphologies and the UHI, the time of day when field campaigns are conducted is critical. In a nutshell, the proposed campaign seeks to capture the variation of temperatures during the day and

across the targeted city as much as possible. Therefore, it is recommended to create three different temperature profiles per day (i.e. morning, afternoon, and evening UHI profiles). The reason for this lies in the process by which solar energy is absorbed and released by the urban fabric, which follows the solar pattern of dawn and dusk. A morning shift (a few hours after sunrise) is required to capture the temperatures of the urban fabric before solar radiation is absorbed. Another shift must be carried out during the hours in the day when the intensity of the sunlight is highest. Finally, another shift during the sunset is required.

As for the speed of the cycling, given that data are stored every second, a constant cycling speed of 8 km/h is recommended to collect data with a spatial resolution of 2 m.

2.3 Data pre-processing and visualization

The goal of this steps is to organize the collected datasets in such a way that they can later be used for data-driven modeling. As illustrated in Figure 6, data pre-processing and visualization involves five main steps. First, all collected field measurements must be assembled and arranged in a data frame.

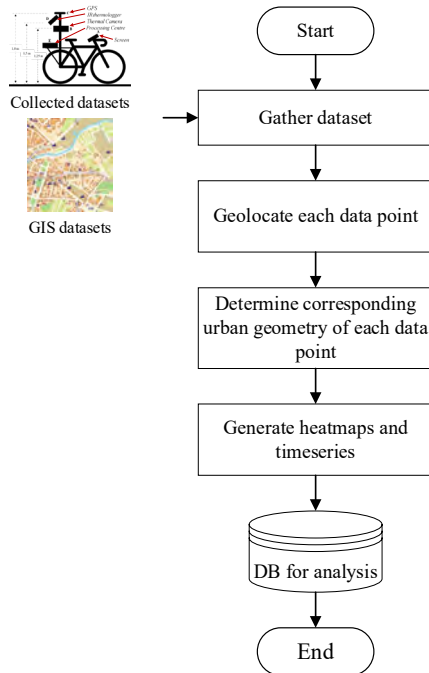


Figure 6: Flow chart of data processing and visualization

Since each measurement point is stored with a corresponding latitude and longitude, it can be geolocated using a GIS software, from where the routes can be examined and compared spatially. Once the data

points are georeferenced, it is possible to map each measurement with its corresponding geometry. This is a two-step process. First, the measurements corresponding to the surface temperature are assigned to its corresponding road. In the second step, the readings from the thermal camera are assigned to each corresponding façade. The proposed method is illustrated in Figure 7. In this method, a line perpendicular to the movement direction of the bike is drawn at each point. If the point intersects with the footprint of the building, the reading is considered to be of the façade. For each façade, the final temperature is taken as the average of all the readings associated with that façade.



Figure 7: Example of the temperature readings for a given façade

Next, the geolocated datasets for each day and temperature profile (i.e., morning, afternoon, and evening) are consolidated into a single data frame where each road has the following attributes: *Road ID*, *Temperature profile*, *Surface temperature*, *Canopy air temperature*, *Façade temperature*, *Wet bulb*, *Relative Humidity*.

3 Case study

To evaluate the applicability of the proposed framework, a case study was conducted in the city of Apeldoorn, the Netherlands. Apeldoorn (52.2112° N, 5.9699° E) is a medium-sized city, located in the middle of the countryside and the De Hoge Veluwe nature reserve. The average elevation of the city is 39 m above sea level, and it has a moderate oceanic climate. It is the 11th largest municipality in the Netherlands, with 165,525 inhabitants (2021), which make it a good example of a midsize city in the Netherlands. However, due to its geographic location, it presents a unique

combination of vegetation and built environment, which offers the basis for an interesting analysis of the UHI. In terms of infrastructure, the heights of buildings are between 10 and 15 meters, and the city center is densely built up. The construction materials of the building facades are homogeneous, varying from dark to light traditional Dutch bricks.

Given the importance of having a comprehensive sample of urban features, an 8 km-long urban circuit was selected to sample the urban area from the center of the city, passing through commercial, recreational, and residential areas (Figure 8).



Figure 8: 8 km-long urban circuit selected for the case study

3.1 The prototype

A bicycle was equipped as described in section 2.1.3 and shown in Figure 9. It features (1) an ANN-MS high performance active GPS antenna, (2) a thermologger (Extech HD500) equipped with air temperature, relative humidity and wet bulb sensors, (3) an infrared thermometer with an accuracy of 30:1 distance to target ratio, and (4) a thermal camera (FLIR A45 FOV 69). As for the processing center, a windows tablet was installed to facilitate on-the-fly access to the data being recorded. Table 2 summarizes the installed equipment.





3.2 Preliminary results

Field campaigns were conducted between 25th of March and 30th of August, 2021, two times per week. For each time the field measurements were undertaken, it started and ended at the same point to make the measurements as consistent as possible. Under heavy weather conditions the field campaigns were not conducted.



Figure 9: Developed mobile data collection unit

Table 2: Summary of measurement equipment installed on the bicycle

Sensor	Reference in Fig. 3	Function	Model	Measured parameter
 GPS	A	Localization	ANN-MS, GPS antenna	Longitude and Latitude
 Thermologger	B	Surface temperature monitoring and weather station	Extech HD500	Surface Air Temperature, Relative Humidity, Wet Bulb
 IR Camera	C	Façade temperature monitoring	FLIR A45 FOV 69	Urban morphology temperatures
 Processing Centre	D	Data Processing	Microsoft surface pro-2	N/A

The processing of the data started with the detection of outliers to identify measurements that deviated substantially from others. This was done by the interquartile range (IQR) method per street. After removing the outliers, the dataset was assembled, as described in section 2.3. As specified in section 2.2, three temperature profiles were collected. Figure 10 presents the summary of the temperatures for the complete dataset in the three profiles measured by the mobile station. The mean morning air temperature was 14.8 °C, while the surface temperature reading was 2.8 degrees lower (12.0 °C). The average temperature reading from the thermal camera in the morning was 13.8 °C with a maximum temperature reading of 29.3 °C, while the maximum morning surface temperature and air temperature were 26.6 °C and 27.2 °C, respectively. In the afternoon the maximum temperatures oscillated between 29.2°C and 30.6 °C. In turn, in the evening the

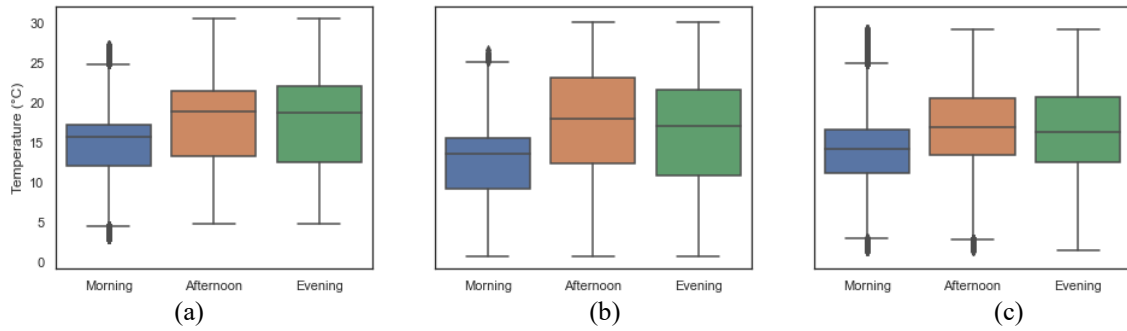


Figure 10: Distribution of the temperature measurements recorded during the field campaigns: a) air temperature; b) surface temperature; and c) façade temperature.

maximum air temperature and surface temperature were 30.6°C and 30.1 °C, respectively. The thermal camera readings were approximately one degree lower (29.2 °C). Figures 11 compares the three different temperature profiles recorded by the mobile station during the six months of data campaign with the air temperature recorded by an on-site weather station located on the surroundings of the city. The air measurements remain constant, while the surface measurements are low in the morning, increased in the afternoon and remain high in the evening. Further, the surface temperatures were highest in the afternoon and decreased during the evening. While this paper concerns mainly with a detailed explanation of the framework for a comprehensive mobile data acquisition regime, the preliminary analysis of the data indicates a pattern of increasing temperatures in the inner city throughout the day. In particular, the surface temperature, which starts with a relatively low temperature in the morning, almost doubles by noon, and stays high until sunset. It is this thermal energy that is released during the night and thus increasing the temperatures in the city.

4 Discussion and future work

The results of this study highlight the variation of temperatures during the day across the city. The granularity of the data collected by the mobile station allows to perform different type of analysis at different scales (i.e., city, urban canyon, street, material level). For example, Figure 12a presents an example of an urban canyon and Figure 12b illustrates the data collected for that part of the city. Figure 12c shows an example of a thermal image at this canyon. Figure 12d summarizes the temperature variation for that urban canyon on the morning of Jun 14th. This level of detail in the data makes it possible to evaluate the different characteristics of the urban materials in conjunction with the different geometric configurations of the city at a very detailed

level. Moreover, as shown in Figure 13, the specific behavior of the paved surfaces themselves can be studied. In the specific example of Figure 13, the surface temperature went from 13 °C in the morning to 48.5 °C in the afternoon, finding a more or less constant temperature in the evening at 31 °C. This is a difference of 18 °C between the morning and evening UHI temperature profiles at that given location. Figure 13 also shows the behavior of different road materials; brick shows a lower temperature in the afternoon profile than asphalt concrete.

With these data at hand, different analyses can be performed, at different scales, and at different times of day (e.g., cooling/warming effects of different types of materials at different locations over different time periods), as well as different urban elements (e.g., vegetation, buildings, cold sinks, roads) on the UHI. Furthermore, the intricate interplay between these decision variables and how the combination of these urban elements could result in optimum context specific design for the built environment. In addition, the authors see great potential for this data collection regime because, once the bicycle is up and running, data collection and analysis do not require a large capital investment. Moreover, since the configuration of the bicycle is quite straightforward, it can be envisaged that larger municipalities could implement this green solution of city bikes. A similar initiative has already been launched in the province of Utrecht with a less sensorized bike that aims to map air quality [17].

Future studies based on the data collected will involve the development of a data-driven approach to study the contribution of paved surfaces to temperature variation in the city. To this end, the authors are already busy analyzing the data collected for the city of Apeldoorn to investigate the impact of pavement material and road design on UHI. Finally, other data collection campaigns will be carried out in other urban areas with different climatic locations.

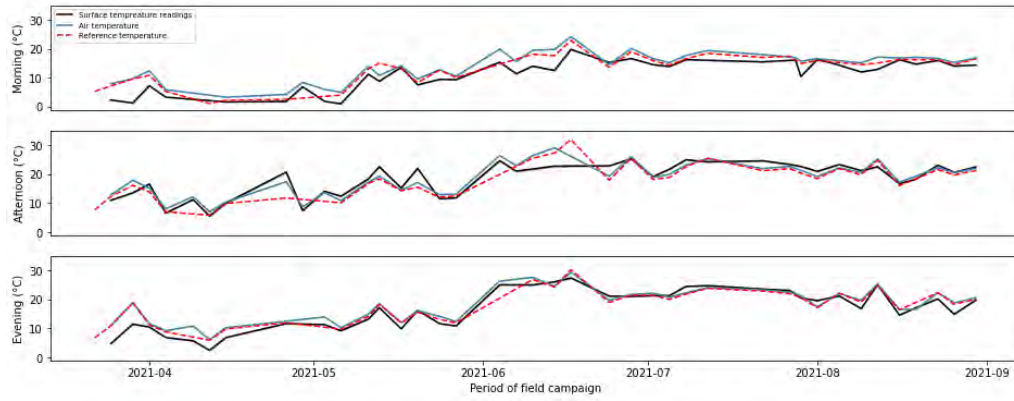


Figure 11: Plots of the three different air temperature profiles obtained with the bicycle-based mobile urban data-gathering station



Figure 12a: Urban canyon



Figure 12b: Data collected in that specific urban canyon



Figure 12c: Example of façade temperature

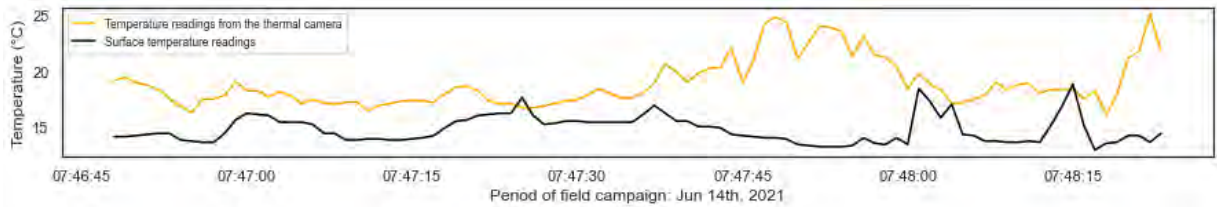


Figure 12d: Surface temperatures and thermal readings in that urban canyon on the morning of Jun 14th, 2021

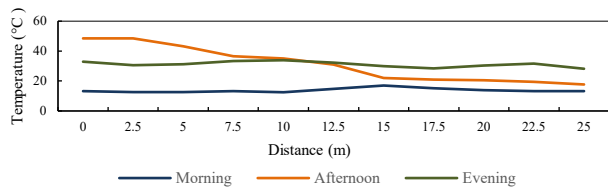


Figure 13: Example of surface temperature measurement taken on the 14th of June 2021. The temperature variation ranges from 13 °C in the morning to 48.5 °C in the afternoon across the different road materials.

5 Conclusions

The aim of this research work was to present a framework for a comprehensive mobile data acquisition

regime to assist the understanding of the UHI with particular interest in paved surfaces. The framework was tested in a case study in Apeldoorn, the Netherlands, during a period of six months. It is shown that the proposed mobile surveying unit provides high-resolution

urban data. Further, the case study demonstrated that the mobile unit enables the acquisition of the data required to build a more comprehensive temperature profile of a given urban area. This level of detail in the data makes it possible to evaluate the different characteristics of urban materials along with the different geometric configurations of the city at a different levels of detail.

The granularity of the data collected by the mobile unit enables the quantitative analysis of the interaction between the UHI, paved surfaces and urban geometries. These quantitative measurements can be useful tools to assist local governments and urban planners to make informed decisions. Moreover, the framework presented in this research offers a clear pathway to feed mobile transect data streams into the UHI discussion with emphasis in urban road infrastructure. This could allow accurate correlations, develop regression models, identify key features by location, and ultimately the development of a data-driven models for comprehensive assessment of Urban Heat Island phenomenon.

References

- [1] Climate Adapt. <https://climate-adapt.eea.europa.eu/knowledge/tools/urban-adaptation/climatic-threats/heat-waves> (accessed 01 June 2021).
- [2] L. Howard, "The Climate of London," 1833. [Online]. Available: <http://www.jstor.org/stable/1793062>.
- [3] United Nations, "World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)," New York: United Nations, 2019. [Online]. Available: <https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf>
- [4] P. Gallinelli, R. Camponovo, and V. Guillot, "CityFeel-micro climate monitoring for climate mitigation and urban design," *Energy Procedia*, vol. 122, pp. 391-396, 2017.
- [5] I. D. Stewart, "Redefining the urban heat island," University of British Columbia, 2011.
- [6] S. W. Kim and R. D. Brown, "Urban heat island (UHI) variations within a city boundary: A systematic literature review," *Renewable and Sustainable Energy Reviews*, vol. 148, p. 111256, 2021.
- [7] T. R. Oke, G. Mills, A. Christen, and J. A. Voogt, *Urban climates*. Cambridge University Press, 2017.
- [8] A. F. Speak and F. Salbitano, "Summer thermal comfort of pedestrians in diverse urban settings: A mobile study," *Building and Environment*, p. 108600, 2021.
- [9] L. Liu *et al.*, "Analysis of local-scale urban heat island characteristics using an integrated method of mobile measurement and GIS-based spatial interpolation," *Building and Environment*, vol. 117, pp. 191-207, 2017.
- [10] Y. Shi, K. K.-L. Lau, C. Ren, and E. Ng, "Evaluating the local climate zone classification in high-density heterogeneous urban environment using mobile measurement," *Urban Climate*, vol. 25, pp. 167-186, 2018.
- [11] T. E. Parece, J. Li, J. B. Campbell, and D. Carroll, "Assessing urban landscape variables' contributions to microclimates," *Advances in Meteorology*, vol. 2016, 2016.
- [12] L. Romero Rodríguez, J. Sánchez Ramos, F. J. Sánchez de la Flor, and S. Álvarez Domínguez, "Analyzing the urban heat Island: Comprehensive methodology for data gathering and optimal design of mobile transects," *Sustainable Cities and Society*, vol. 55, p. 102027, 2020.
- [13] C. Cao *et al.*, "Performance Evaluation of a Smart Mobile Air Temperature and Humidity Sensor for Characterizing Intracity Thermal Environment," *Journal of Atmospheric and Oceanic Technology*, vol. 37, no. 10, pp. 1891-1905, 2020.
- [14] C. D. Ziter, E. J. Pedersen, C. J. Kucharik, and M. G. Turner, "Scale-dependent interactions between tree canopy cover and impervious surfaces reduce daytime urban heat during summer," *Proceedings of the National Academy of Sciences*, vol. 116, no. 15, pp. 7575-7580, 2019.
- [15] L. Klok, N. Rood, J. Kluck, and L. Kleerekoper, "Assessment of thermally comfortable urban spaces in Amsterdam during hot summer days," *International journal of biometeorology*, vol. 63, no. 2, pp. 129-141, 2019.
- [16] L. P. Dorigon and M. C. d. C. T. Amorim, "Spatial modeling of an urban Brazilian heat island in a tropical continental climate," *Urban Climate*, vol. 28, p. 100461, 2019.
- [17] Data Europe. "Sniffer Bike - a project to track air quality in Utrecht." <https://data.europa.eu/en/news/sniffer-bike-project-track-air-quality-utrecht> (accessed.

Using BIM and LCA to evaluate material circularity: Contributions to building design improvements

H. Feng ^a, Q. Chen ^b, B. García de Soto ^b, M. Arashpour ^c

^a Department of Mechanical and Construction Engineering, Northumbria University, Newcastle, UK

^b S.M.A.R.T. Construction Research Group, Division of Engineering, New York University Abu Dhabi (NYUAD), Experimental Research Building, Saadiyat Island, P.O. Box 129188, Abu Dhabi, United Arab Emirates

^c Department of Civil Engineering, Monash University, Melbourne, Australia

E-mail: haibo.feng@northumbria.ac.uk, qc737@nyu.edu, garcia.de.soto@nyu.edu, mehrdad.arashpour@monash.edu

Abstract –

The construction sector has suffered from low productivity and considerable waste due to the fragmentation of its value chain and inefficient design and material usage processes. The circular economy (CE) principles have gained significant attention among researchers and practitioners to help overcome these challenges. Construction materials such as timber and steel elements will lend themselves more easily to reuse and recycle for new construction to reduce carbon emissions. The BIM-based LCA method could be explored and expanded to evaluate design options for circular construction to advance the knowledge about implementing circular economy principles in construction projects. However, this is relatively new and requires proof of concepts to demonstrate precisely how BIM-based LCA could be implemented to help stakeholders decide on optimal design options and material choices. To address this need, this study proposes a BIM-based LCA process for comparing the carbon impact of two design model options considering different material choices, including the possibilities of using virgin materials and recycled and reused materials. Findings show that the timber structure was favored over the precast concrete structure because timber materials entailed less carbon emission; however, the precast concrete structure has great potential of being reused for future new construction projects. Findings also show that Module A (with timber and steel materials) has a slightly higher circularity (39%) than Module B (with concrete materials) with 37% circularity.

Keywords –

BIM; LCA; Circular economy; Design options; Secondary materials; Embodied carbon

1 Introduction

Construction projects and buildings typically use many concrete and steel materials with high embodied carbon values. According to the Intergovernmental Panel on Climate Change (IPCC, 2014), buildings are expected to account for 52% of worldwide energy-related carbon emissions by 2050. A solution to reducing embodied and total carbon emissions is using by-products and waste materials for producing building materials, as advocated by the Ellen McArthur Foundation (2021). Circular economy (CE) in the built environment was defined as the strategic programming of a building to quickly change its configuration for longevity and potentially be susceptible to the loop of reducing, reusing, and recycling for resource efficiency (Pomponi & Moncaster, 2017). Associated is the concept of circular construction, which was understood as an approach to achieving CE targets (slowing, closing, and narrowing the construction resource loops) considering the local or regional capacity to supply and transport (re-)used materials through prioritizing the extension of building service life and recycling, reusing, recovering materials when building functionalities are lost (Stahel, 2016; Chen et al., 2021). By adopting the circular economy principles in the built environment, the resources can be used more efficiently, and therefore material waste and carbon impacts in construction projects could be reduced. Projects guided by CE principles can potentially alter traditional building design processes. For example, many researchers investigated design with reused materials (Brütting et al., 2019), design with recycled materials (Borg et al., 2021), and design for disassembly (Sanchez et al., 2020) methods to ensure optimal design options could be generated to minimize carbon emissions and reduce the exploitation of raw virgin materials.

Building information modeling (BIM) and lifecycle

assessment (LCA) methods have been widely studied and accepted as “business-as-usual” in current design practices (Hollberg et al., 2020). The potential of integrating BIM and LCA is sufficient information about building geometry and functionality for building sustainability assessment purposes (e.g., to achieve a high rating in sustainable building certification systems such as BREEAM or LEED). For assessing the environmental impact of building design options, LCA covers the entire life cycle of buildings from raw materials extraction and processing, manufacturing of building components and transportation to use and end-of-life (Safari & AzariJafari, 2021). Researchers have developed various tools, plug-in functions and software to enable BIM-based LCA by linking the quantity take-off and standard material libraries from a BIM authoring software with the local LCA database to measure the environmental impact during different phases of the project life cycle (Röck et al., 2018; Hollberg et al., 2020). With the development of CE principles, BIM-based LCA could be further explored to evaluate circular design options. However, the research in this direction is relatively new and requires proof of concepts to demonstrate exactly how BIM-based LCA is implemented to help stakeholders make decisions on optimal design options and material choices.

This study proposes a BIM-based LCA process for comparing the carbon impact of two design model options considering different material choices, including the possibilities of using virgin materials and recycled and reused materials. Tekla Structures software and the One-Click LCA platform are used to implement the BIM-based LCA process to perform the carbon emission calculations. The two design model options include 1) Model A: a timber building structure, and 2) Model B: a precast concrete building structure, both of which represent the prevalent building types in the view of sustainable construction. The BIM-based LCA process outputs the carbon emissions for different building life cycle stages when certain building materials are chosen for the building design and when they are again circulated in the next project life cycle. The focus is on understanding how the CE principles support the carbon reductions in buildings through optimal material considerations.

The rest of this paper is structured as follows. Section 2 provides a literature review on circular construction and material circularity as well as the BIM-based sustainable design. Section 3 describes the BIM-based LCA process for comparing the material circularity of two design model options. Section 4 illustrates the findings of the carbon emissions from the two design model options considering different materials choices. Section 5 discusses the potential and limitations of this study, followed by conclusions and future work in Section 6.

2 Literature review

2.1 Circular construction and material circularity

The construction sector is characterized as extremely resource-intensive due to the significant energy consumption, greenhouse gas emissions, and waste generation. The emissions in the construction industry can be reduced by increasing practices of reusing, recycling and recovering materials, in particular the CE model. By adopting the CE principle, the construction sector could play a strategic role in achieving Net Zero by 2050 (Pomponi & Moncaster, 2017). Stahel (2016) suggested that the CE should emphasize reducing product environmental impact, extending the useful life of the products used and employing sustainable resources, all of which are critical for developing climate change mitigation strategies. Circular construction models have been developed around utilizing the embedded economic and environmental value in products and materials as long as possible, such as substituting primary materials with secondary materials (Safari & AzariJafari, 2021). To overcome the resource depletion challenges, there has been a drastic shift towards a CE paradigm in the built environment to reduce the pressure on non-renewable resources (Chen et al., 2021). CE principles seek to maintain building components and resources at their highest intrinsic value for as long as possible. Building components are kept in a continuous loop of use, reuse, repair and then recycled, thereby reducing waste and preventing negative externalities of CO₂ emissions.

Various studies have shown the advantages of adaptive reuse of building materials and the possibility of using the existing built environment as a source of reused components (Brütting et al., 2019; Sanchez et al., 2020). Building components can be reused and circulated in three ways: 1) reusing the existing components on-site through improving them or extending them, 2) relocating the majority or even all of the existing components to a new location, and 3) individual components obtained from the destruction of a building being reused directly in another building (Nußholz et al., 2020). Stahel (2016) states that the prefabrication of components and their modularization could create building products designed for reuse. Strategies for re-entering construction and demolition waste into the production chain concentrate primarily on the recycling process, whereas studies focusing on reuse are less frequent. It has been a challenging task for construction practitioners to understand what and how building materials could be reused as well as how much carbon impact could be reduced through these new design and construction strategies such as design for future reuse or design with reused materials.

2.2 BIM-based sustainable design

BIM has been widely adopted in current construction practices that facilitate the coordination of design and construction information. By adding a sustainability dimension, BIM-based sustainable design centered around design considering the LCA when using the material information and parametric building design information from BIM authoring software.

The BIM-based LCA method is a well-established technique for sustainability studies in the built environment which could be extended to the CE-driven research (Pomponi & Moncaster, 2017). For example, Röck et al. (2018) and Hollberg et al. (2020) have designed BIM-LCA applications to support the design process of real buildings, which allowed the designers to track design decisions on the continuously evolving BIM model. However, studies in this direction have not supported circular concepts in building design processes.

One way to promote CE is to design through material circularity assessment using BIM-based LCA to embed sustainable building regulations and environmental product declaration (EPD) in the BIM authoring software. However, the related research has been missing from the current body of knowledge on CE in the built environment. Considering the benefits of CE, as claimed by many scholars (e.g., Joensuu et al., 2021), there is a need to prove the concept of using BIM and LCA tools to realize new design processes guided by CE. This study investigates a case study of two design model options by using BIM and LCA to evaluate material circularity. The comparative results would help stakeholders understand material choices in optimal design to reduce carbon by maximizing the reusing and recycling of building materials.

3 Methodology

The study proposes a BIM-based LCA process to calculate carbon emissions for buildings with different design options and material choices. Tekla Structural Designer was used for developing the building model. The structural elements such as beams, columns, footings, etc., were designed in this software and applied the required materials and loads for the building. It was also used to calculate the bill of quantities of the materials required based on the building design. The One-Click LCA software was used for performing the life cycle assessment concerning carbon emissions for the building. One-Click LCA provides embedded algorithms for measuring building circularity. The detailed information on building circularity calculations is provided through the One-Click LCA online help center.

Two elemental building models, Model A and Model B (Figure 1) were designed at LOD 300 using the same dimensions according to the British Standards (BS 1192,

2018). Different materials were applied to the same geometrical components of the building. Model B used more concrete than Model A. The internal walls of Model B were applied with ready mix concrete, whereas Model A has timber wall panels. Also, precast concrete elements were used in Model B for slabs and external walls. The total area of the building is 324 m² (18 m x 18 m). The 3-story buildings have a total height of 9 m with a 3 m floor-to-floor spacing. The two design model options were used to perform life cycle assessments to investigate the amount of carbon these buildings emit from manufacturing to demolition. The total material quantities and the material compositions with the design options for Model A and Model B are summarized in Table 1 and Table 2, respectively. It is noted that parts of the ready-mix concrete are used in different locations and for different purposes in Model A and B. For example, Model A uses ready-mix concrete for flooring, while Model B uses pre-casted holly-core concrete slabs. Another slight difference is that Model B uses ready-mix concrete for interior walls but instead, Model A uses timber walls. Despite the differences, the total usage quantities of ready-mix concrete are almost equal in Model A and Model B, as shown in Table 1.

The electricity consumption for each building is set to 25 kWh/m² and the heating consumption to 68 kWh/m². The water consumption and wastewater are set to 25 m³ per annum. It is noted that these values are not chosen according to specific real-world cases but are considered for comparison purposes only based on the average consumption of the household in the UK.

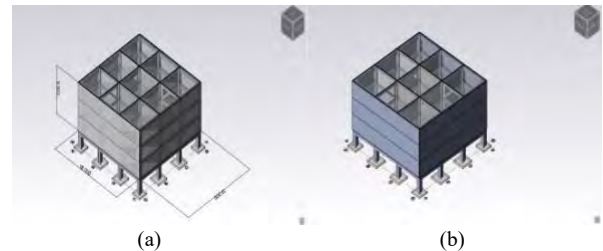


Figure 1. Structural models for Model A (a) and Model B (b) using Tekla Structural Designer

The standard EN 15978:2011 is chosen in this study to specify the LCA framework in the built environment. It divides the life cycle of buildings and infrastructures into the product stage (A1-A3), construction stage (A4-A5), use stage (B1-B7), end-of-life stage (C1-C4), and benefits or loads beyond the system boundary (D) (British Standards Institution, 2011). The product stage (A1-A3) is represented as materials that include the impacts caused due to the extraction of raw materials, transportation and emissions that are caused due to manufacturing. The construction stage is divided into transportation to site (A4) and installation process (A5). A4 covers the effects of transporting materials from the

production site to the construction site. A5 covers the impacts of energy and water consumption during the construction stage, material waste and other environmental implications. Use stage (B1-B7) includes the use phase (B1-B3), material replacement & refurbishment (B4-B5), energy use (B6) and water use (B7). B1-B3 covers the emissions from the use of materials. B4-B5 covers the replacements of entire construction elements bringing them back to their pre-existing performance levels. B6 includes the usage of both electricity and district heating. B7 includes the environmental consequences of water throughout the life cycle, including manufacturing, transportation, and wastewater treatment. End-of-life (C1-C4) covers the process of recyclable construction depending on the kind of material, impacts of pre-processing and landfilling for

waste streams that cannot be recycled. Benefits or loads beyond system boundary (D) are about the external impacts, including the advantages and loads beyond the building's life cycle. It provides various environmental advantages from reusable, recycled and secondary materials.

During the use phase (B1-B3), carbon will be reduced. When exposed to air, materials such as concrete, cement, and mortar absorb carbon dioxide and reverse the calcination step during the cement manufacturing process. The amount of carbon absorbed is determined by the material exposure duration and the original amount of cement. During stage D, carbon emissions have been reduced again because of the environmental benefits of using reused and recycled materials.

Table 1. Total material quantities for the two design model options

Material choices	Model A (units)	Model B (units)	Material Wastage (%)	Transport (km)
Ready-mix Concrete (m ³)	455.7	456.0	4.0	60.0
Precast Concrete (m ³)	0.0	139.9	0	60.0
Timber (m ³)	194.4	0.0	17.9	130.0
Steel (ton)	35.5	18.2	4.9	110.0
Hollow-core concrete for slabs (m ²)	0	1296.0	0.0	60.0

Table 2. The material compositions of the two design model options

Component	Dimensions	Materials used (Model A)	Materials Used (Model B)
Slab	18 m x 18 m D = 0.15 m	Ready Mix Concrete C28/35, Cut & bent steel rebar (104.5 m ³)	Ready Mix Concrete C28/35, Cut & bent steel rebar (104.5 m ³)
Floorings	18 m x 18 m D = 0.15 m	Ready Mix Concrete C28/35, Cut & bent steel rebar (194.3 m ³)	Hollow-core concrete slabs C30/37 (1296 m ²)
Columns	0.8 m x 0.8 m H = 12 m	Ready Mix Concrete C40/50, Reinforcement steel for concrete (60 m ³)	Ready Mix Concrete C40/50, Reinforcement steel for concrete (60 m ³)
Beams	0.25 m x 0.5 m (internal) 0.45 m x 0.7 m (external)	Ready Mix Concrete C28/35, Carbon steel reinforcing bar (62.9 m ³)	Ready Mix Concrete C28/35, Carbon steel reinforcing bar (62.9 m ³)
Footings/Pad Bases	2.1 m x 2.1 m 2.44 m x 2.45m	Ready Mix Concrete C25/30, Reinforcement steel for concrete (34 m ³)	Ready Mix Concrete C25/30, Reinforcement steel for concrete (34 m ³)
Exterior Wall	9 m x 18 m T = 0.23 m	Insulated masonry wall with brick slips and aircrete block (298.1 m ²)	Precast concrete wall elements (139.9 m ³)
Interior Wall	9 m x 18 m T = 0.23 m	Timber wall – structural sawn timber panels (194.4 m ³)	Ready Mix concrete – low strength C12/15 (194.4 m ³)
Roof	18m x 18 m D = 100 mm	Roof panels with QuaCore hybrid insulation core (12.6 kg/m ²)	Roof panels with QuaCore hybrid insulation core (12.6 kg/m ²)

4 Findings

4.1 Comparing carbon emissions from two design model options

After performing the calculations using One-Click LCA, Model A showed a total carbon emission of 1461 Tons CO₂e with an average of 71.15 kg CO₂e/m² per year. Model B showed a total carbon emission of 1533 Tons of CO₂e with an average of 78.87 kg CO₂e /m² per year. The results represent the carbon emitted from the raw material stage to the building's end-of-life stage with an assumed lifespan of 60 years, covering carbon emission from the construction stage along with heat and electricity distribution systems.

Comparing all the life cycle stages in Model A, energy use has the highest carbon emission, followed by the raw material stage. However, in Model B, the raw material stage has the highest carbon emission when compared to the other life cycle stages. Table 3 shows the detailed comparison of embodied carbon emissions (Tons CO₂e) from the two design model options when sourcing virgin materials (Stages A1-A3). During stages B4-B5 and C1-C4, Model B has emitted less carbon than Model A, attributed to the use of precast concrete panels. Model B used ready mix concrete and reinforcement steel for slabs, beams and columns and brick masonry for external walls, which requires less refurbishment during the use stage compared with timber materials. The precast elements can be easily reused, which is why Model B has less carbon emission during the C1-C4 stage. Whereas ready mix concrete needs to be crushed, and they can be recycled as additives or aggregates or for landfilling but cannot be reused as a direct material.

To compare the impact of design with reused materials for the two model options, reused materials were assigned to the models in One-Click LCA to reveal how much carbon can be reduced. It is noted that the CE concept of reuse in this study means the reuse of the product or component as is or through direct remanufacturing of materials without being recycled. The results from One-Click LCA show that approximately 31% of carbon was reduced for Model A, and nearly 35% of carbon was reduced for Model B in the total carbon emissions (operational plus embodied carbon emission). Confirming that using reused materials helps reduce carbon emissions due to various factors such as manufacturing, transportation, etc. As a result, it helps reduce total carbon emissions. The remaining carbon emissions are due to the building's operational use and end-of-life stages.

The recycling factors were applied for material settings in One-Click LCA to compare the impact of design with recycled materials for the two models.

Materials such as concrete and steel were used as virgin materials mixed with recycled binders. For example, concrete was added with 50% of fly ash content or ground granulated blast-furnace slag (GGBS), and steel was produced from secondary scrap. Table 4 shows the details of recycled binders added to the materials along with the carbon emissions those materials produced.

By adding recycled binders to the virgin materials, around 6% of carbon was reduced for Model A, and 6.5% of carbon was reduced for Model B from the total carbon emissions. Another assessment was performed using the same but reused materials to reveal how much additional carbon can be reduced using the reused products along with recycled binders. Results show that a total of 33% of carbon was reduced for Model A, and 36% was reduced for Model B. The total carbon emission of Model A dropped to 991 Tons of CO₂e with an average of 50.98 kg CO₂e/m² per year. For Model B, the emission was reduced to 985 Tons of CO₂e with an average of 50.65 kg CO₂e/m² per year.

Table 3. Comparison of embodied carbon emissions (Tons CO₂e) from the two design model options when sourcing virgin materials (Stages A1-A3)

Material choices (sourcing virgin materials)	Model A	Model B
<i>Ready-Mix concrete</i>		
C28/35	75	26
C40/50	48	48
C25/30	8.4	56
<i>Steel Reinforcement</i>		
Cut & Bent steel rebar	12	N.A
Reinforcement (Rebar)	10.1	10.1
Carbon steel reinforcement bar	4.9	4.9
<i>Precast concrete</i>		
Hollow-core concrete slabs	N.A	64
Concrete wall elements	N.A	47
<i>Other construction materials</i>		
Structural sawn timber	21	N.A
Emulsion for exterior masonry	0.23	N.A
Emulsion matt paint for outdoor	N.A	0.49
Anti-corrosive paints	0.31	0.31

Note: N.A indicates that the field is not applicable.

Table 4. Comparison of embodied carbon emissions (Tons CO₂e) from the two design model options when sourcing recycled materials (Stages A1-A3)

Material choices (sourcing recycled materials)	Model A	Model B
Ready-Mix concrete (Adding 50% GGBS content)		
C28/35	40	14
C40/50	29	29
C25/30	4.7	32
Steel Reinforcement		
Cut & Bent steel rebar	12	N.A
Reinforcement (Rebar, 90% recycled content)	8.1	8.1
Carbon steel reinforcement bar (using secondary production, 97.07% recycled content)	5.2	5.2
Precast concrete (Added 40% recycled binders in cement)		
Hollow-core concrete slabs	N.A	53
Concrete wall elements	N.A	34

Note: N.A indicates that the field is not applicable.

4.2 Evaluation of material circularity

Building circularity is calculated through One-Click LCA based on the end-of-life process for each material. Throughout all the phases of the life cycle, the entire material flow will be treated utilizing the specified processing chain, which defines the implications for life cycle stages C & D. The reason for performing circularity assessment is to reveal the amount of material recovery and reuse capability after the building's end-of-life stage. The end-of-life process for each material was set before performing the circularity assessment for the building. For concrete, the end-of-life process was crushing concrete for aggregate usage in cement, or it can be used for landfilling for concrete blocks. Steel was recycled or reused as a direct material. Materials like emulsion and paints can be used as landfilling inert materials. Timber panels will be reused as material, or they can also be used for wood incineration or landfilling.

According to the embedded definition in One-Click LCA, material circularity refers to the percentages of materials quantities that could be “recovered” and “returned” at the building's end-of-life phase. The term “materials recovered” means the utilization of circular materials. It is indicated by the percentage of total materials used, made up of recycled, reused, and renewable materials. The term “materials returned” is represented by the end-of-life circular treatment of used materials. It is the total sum of recycled or utilized materials, together with 50% of materials downcycled or used as energy. Both materials returned and materials recovered are calculated based on the embedded default factors used for the project in One-Click LCA, such as material wastage on the construction site and other factors, including material replacement and refurbishment. The circularity score is the average of the sum of the materials recovered and the materials returned.

To illustrate the circularity score for both models, including the percentage for each CE path, the scenario with recycled materials sourced to replace virgin materials (Table 4) was used. The results are shown in Figure 2. Results show that 7.5% (5% + 2.5%) of the total quantities of materials could be recovered as is in Model A, and 69.5% (26.9% + 16.6% + 52%/2) of the total quantities could be returned either through reusing, recycling, or downcycling for other projects. The results are similar in Model B; however, its recovery rate is relatively lower than that in Model A, very likely due to the difficulties of recovering precast concrete than timber materials. Model A showed a building circularity of 39% ($\approx (7.5\% + 69.5\%)/2$), and Model B showed a circularity of 37% ($\approx (1.8\% + 71.4\%)/2$). The results indicate that timber and metals (mainly timber and steel used in Model A) showed the highest circularity score, as they can be easily recycled or reused in contrast to concrete which usually needs to be crushed for recycling. This could also be the primary reason why concrete and other materials such as bricks and gypsum have the least percent of material recovery. Also, these materials have a high downcycling rate, which only indicates a single recycling process due to low quality.

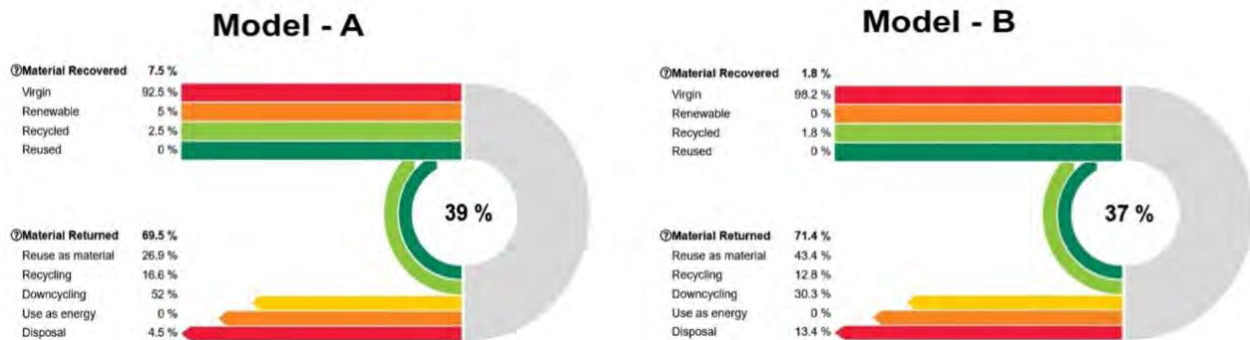


Figure 2. Circularity scores output from One-Click LCA

Behind these numbers, the implications of material choices are critical. Reused and recycled materials can reduce transportation, manufacturing, installation, other construction works, and waste disposals. When reused, precast concrete produced less carbon than ready mix concrete because the precast panels could be easily transported and reused directly without crushing them. As previously shown, adding recycled binders to virgin materials could reduce carbon emissions. Concrete added with fly ash or GGBS could be used as the recycled binder, and steel obtained from secondary scrap can be used for reinforcement.

5 Discussions

The circular economy concept has received significant attention to be restorative and regenerative and aims to keep products, components, and materials at their highest utility and value at all times. Using reused materials to replace virgin materials showed an apparent reduction of nearly 35% for both studied models. Model B showed higher carbon emissions when compared to Model A, whereas when using the reused materials, Model B showed an apparent reduction in carbon emissions compared to Model A. In addition, LCA has been performed for materials containing recycled binders or other recycled products. Recycling involves using a considerable amount of water and energy and the formation of carbon emissions, which may have a more significant environmental impact than reusing materials. The reuse of building components is a potential alternative for reducing construction and demolition waste. Many studies suggest that reused materials are considered environmentally and economically beneficial compared with recycled materials. This study also showed that reused materials entailed fewer carbon impacts than recycled materials because of less reprocessing efforts.

The use of precast elements opens up the possibility of designing the materials to be readily installed, deconstructed, and reused in the future. Materials such as ready-mix concrete can be crushed as aggregate by downcycling when demolished at their end-of-life stage. From the findings, the precast panels and timber walls have shown the highest reuse factor to reduce carbon emissions.

In addition, there are a few limitations of this study:

- Real-world examples and case studies could have been used; however, there is a lack of evidence of CE-oriented projects in the UK construction industry.
- The developed models targeted the concept designs. Similar studies should be carried out on more detailed model designs in the future.
- Evaluating the difference between the expected

and actual service life of the entire building may impact the findings from the One-Click LCA analysis. Although a certain lifetime has been provided for each material in the study, however, in reality, this can be completely different.

- The factors of Design for Disassembly (DfD) and Design for Adaptability (DfA) in One-Click LCA may play a critical role in computing the circularity score, which should be studied further as part of the CE implementation.

Despite the limitations and potential barriers to implementing design for CE solutions in construction (such as difficulties of implementing new business models), the findings from this study indicate that design for material circularity and design with material circularity could help reduce carbon emissions in construction projects and therefore would have an impact to achieve the United Nations' Net Zero Carbon program targets.

6 Conclusions and future work

Circular economy in the built environment is an important solution to overcoming resource depletion and reducing environmental impact. Using a BIM-based LCA process, this study investigated two building design options by applying the recycled and reused materials to reveal how much carbon could be reduced using virgin materials.

Results showed that adaptive reuse of precast concrete and timber elements could reduce carbon emissions through reduced needs for transportation, energy usage and material manufacturing. Comparing the two model design options suggested that the secondary materials should be used to replace the virgin materials to reduce carbon emissions and increase the circularity of materials. Reusing materials could reduce more carbon emissions when compared to recycled materials. As expected, materials such as precast elements have a high reuse capability compared to cast-in-situ and ready-mix concrete.

The circularity score was calculated to reveal the percentage of materials recovered and materials returned. The results indicate that timber and metals (timber and steel in Model A) yield a high circularity score since they can be easily recycled or reused in contrast to concrete, which usually needs to be crushed for recycling. Maximizing the reuse of materials could be considered a major strategy for the designers to choose between material resources.

The carbon emissions factors used for the resources of the two models in this study were based on the average values of water usage, fuel consumption, district heating and cooling consumption, and exported energy for UK households. This study will be extended into a real-life

project using specific real-world project values. Other environmental impacts such as acidification, ozone depletion, eutrophication, and other factors will be investigated in future research work.

Acknowledgments

We thank Mr. Nikhil Lingamaneni for his assistance with case design modeling and lifecycle assessment work.

7 References

- BS 1192. (2018). *Specification for collaborative sharing and use of structured Health and Safety information using BIM*. <https://shop.bsigroup.com/products/specification-for-collaborative-sharing-and-use-of-structured-health-and-safety-information-using-bim/standard/details> (Accessed 21 January 2022).
- Borg, R. P., Cuenca, E., Garofalo, R., Schillani, F., Nasner, M. L., & Ferrara, L. (2021). Performance Assessment of Ultra-High Durability Concrete Produced From Recycled Ultra-High Durability Concrete. *Frontiers in Built Environment*, 7. <https://doi.org/10.3389/fbuil.2021.648220>
- Brütting, J., Desruelle, J., Senatore, G., & Fivet, C. (2019). Design of Truss Structures Through Reuse. *Structures*, 18, 128–137. <https://doi.org/10.1016/j.istruc.2018.11.006>
- Chen, Q., Feng, H., & García de Soto, B. (2021). Revamping construction supply chain processes with circular economy strategies: A systematic literature review. *Journal of Cleaner Production*, 130240. <https://doi.org/10.1016/j.jclepro.2021.130240>
- Ellen MacArthur Foundation. (2021). Delivering the circular economy a toolkit for policymakers - selection of key exhibits. <https://ellenmacarthurfoundation.org/articles/building-a-world-free-from-waste-and-pollution> (Accessed 25 January 2022).
- Hollberg, A., Genova, G., & Habert, G. (2020). Evaluation of BIM-based LCA results for building design. *Automation in Construction*, 109, 102972. <https://doi.org/10.1016/j.autcon.2019.102972>
- Intergovernmental Panel on Climate Change (IPCC) (2014). *Buildings*. In: *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the IPCC*. Cambridge University Press, Cambridge, United Kingdom. https://www.ipcc.ch/site/assets/uploads/2018/02/ipcc_wg3_ar5_chapter9.pdf (Accessed 20 January 2022).
- Joensuu, T., Edelman, H., & Saari, A. (2020). Circular economy practices in the built environment. *Journal of Cleaner Production*, 276, 124215. <https://doi.org/10.1016/j.jclepro.2020.124215>
- Nußholz, J. L. K., Rasmussen, F. N., Whalen, K., & Plepys, A. (2020). Material reuse in buildings: Implications of a circular business model for sustainable value creation. *Journal of Cleaner Production*, 245, 118546. <https://doi.org/10.1016/j.jclepro.2019.118546>
- Pomponi, F., & Moncaster, A. (2017). Circular economy for the built environment: A research framework. *Journal of Cleaner Production*, 143, 710–718. <https://doi.org/10.1016/j.jclepro.2016.12.055>
- Röck, M., Hollberg, A., Habert, G., & Passer, A. (2018). LCA and BIM: Visualization of environmental potentials in building construction at early design stages. *Building and Environment*, 140, 153–161. <https://doi.org/10.1016/j.buildenv.2018.05.006>
- Safari, K., & AzariJafari, H. (2021). Challenges and opportunities for integrating BIM and LCA: methodological choices and framework development. *Sustainable Cities and Society*, 102728. <https://doi.org/10.1016/j.scs.2021.102728>
- Sanchez, B., Rausch, C., Haas, C., & Saari, R. (2020). A selective disassembly multi-objective optimization approach for adaptive reuse of building components. *Resources, Conservation and Recycling*, 154, 104605. <https://doi.org/10.1016/j.resconrec.2019.104605>
- Stahel, W. R. (2016). The circular economy. *Nature*, 531(7595), 435–438. <https://doi.org/10.1038/531435a>

Building Information Model Pre-Processing for Automated Geometric Quality Control

M. Bueno¹ and F. Bosché²

^{1,2}School of Engineering, University of Edinburgh, Scotland

martin.bueno@ed.ac.uk, f.bosche@ed.ac.uk

Abstract -

Geometric quality control (QC) in a construction project is an important but time-consuming and not value-adding task. While significant progress is being made in construction digitalisation, geometric QC processes remain highly manual and inefficient. This manuscript proposes a new methodology to pre-process initial information contained in the as-designed Building Information Model ('BIM model' hereafter) and obtain a full list of the geometric QC tasks that need to be conducted over the duration of the corresponding construction project. The proposed methodology employ a network graph constructed automatically from the BIM model information; a dictionary of types of geometric QC (e.g. dimensions to be checked with tolerances) that apply to the given project; and a QC digital manager that ties both elements together and identifies the list of unique geometric QC tasks that need to be conducted throughout the given project. The workings of the proposed methodology are illustrated with a case study, for some QC specifications extracted from the EN 13670 standard. These demonstrate its usefulness to exhaustively establish and record all geometric QC tasks required over a project.

Keywords -

BIM; IFC; Quality Control; Geometric; Point Clouds; Digital Twin

1 Introduction

During construction, guaranteeing that the built elements meet quality specifications (e.g. geometric tolerances) is critical to ensure they "achieve the intended level of safety and serviceability during their service life" [1]. From a construction delivery viewpoint, this is also critical to ensure that subsequent construction works are not delayed (due to rework), which would also result in additional costs.

Yet, current practice in geometric Quality Control (QC) is human-intensive, prone to error. Besides, quality specifications are recorded across a multitude of documents, including standards, regulations, and bespoke project-specific specifications, which makes it difficult to track all quality specifications and conduct all necessary corre-

sponding QC assessments and measurements.

Digitalisation in the construction industry promises numerous benefits in terms of efficiency and quality improvements [2]. But, despite the current challenges discussed above, Geometric QC processes have seen some, but limited evolution. Automated compliance checking is an active area of R&D for the analysis and validation of design (BIM) models [3, 4]. But this area has not explored compliance checking of actual works, only design (BIM) models. Regarding QC of physical construction, developments have occurred around surveying technologies, with laser scanning (terrestrial and mobile) and photogrammetry offering means to rapidly and effectively collect 3D survey data. Among those, terrestrial laser scanning (TLS) offers high accuracy that is particularly suitable for use in geometric QC, but the analysis of the large point clouds TLS acquires has traditionally remained a manual and tedious process. Digitalisation of this data analysis stage would provide significant benefits to improve the quality and speed of geometric QC activities. To fill this need, works have been exploring the 'scan-vs-BIM' principle to match laser scanned points to components in 3D BIM model in order to (1) recognise those components and subsequently (2) assess their geometric correction [5, 6, 7].

Although it remains an active area of research, the value of the 'scan-vs-BIM' principle to analyse point cloud is now generally accepted. But, the remaining gap in knowledge is in automatically identifying what type of geometric QC needs to be conducted where in the model, and robustly conducting all these geometric QC checks [8].

In this paper, we propose a methodology to automatically analyse 3D BIM models to identify where geometric specifications apply and thus need to be quality-controlled. The methodology includes the development of (1) a simplified dictionary to store the different geometric specifications, or rules; (2) an algorithm to extract from the design 3D BIM models all components and component relationships relevant to the given specifications, stored in a graph structure; (3) an algorithm to process that graph to systematically establish where each specification is applied in the given model. The output of that last step is a list of QC checks that need to be conducted during construction.

The rest of the manuscript is organised as follows. In

Section 2, the GeometricQC tool is briefly introduced and the proposed method for establishing geometric QC tasks given the design BIM model is detailed. Preliminary results obtained with a case study is reported in Section 3. Section 4 concludes the manuscript.

2 Proposed methodology

2.1 GeometricQC tool overview

The work presented in this manuscript is part of a highly automated geometric QC tool, called GeometricQC tool, developed as part of the COGITO project [9].

The GeometricQC tool is designed to conduct automatic geometric QC by comparing the as-designed geometry, contained in the BIM model, and the as-built data, captured in the form of point clouds acquired using TLS. The tool essentially follows an Automated Rule Checking (ARC) approach [10, 11]. The GeometricQC tool first defines what geometric QC needs to be conducted, where and when during the project geometric QC. After components requiring geometric QC are constructed, the surveyor acquires the relevant site-referenced (or geo-referenced) laser scans. At this point, the GeometricQC tool automatically matches the as-built TLS points to the 3D geometry of the components of interest stored in the as-designed BIM model, following a ‘scan-vs-BIM’ process, and performs the scheduled geometrical control. This GeometricQC tool workflow is summarised in Figure 1.

This paper focuses on the first part of the above process that occurs during the planning phase of a construction project. This part aims to define ‘what’ geometric QC needs to be conducted, ‘where’ and ‘when’ during the project geometric QC’.

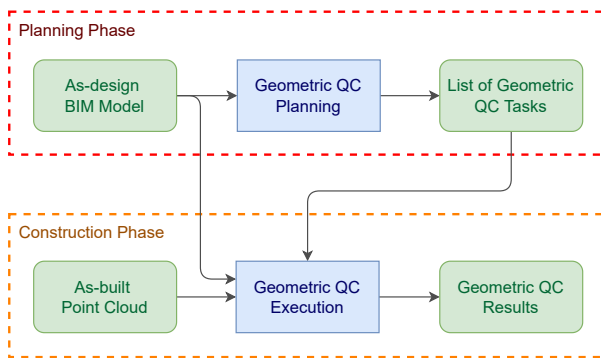


Figure 1. GeometricQC tool workflow. This paper focuses on the planning phase of the process.

For the ‘what’, digital rules first need to be defined. In construction, as discussed earlier, geometric QC specifications and tolerances can be defined in various ways, but standard specifications also exist. For example, EN

13670:2009 [1] and EN 1090-2:2018 [12] provide construction and erection geometric specifications for the execution of concrete and steel structures, respectively. These include all dimensions (or geometry) that need to be checked and tolerances for each of them. For example, EN 13670:2009 [1] details in section 10.4 and annex G the different geometrical tolerances that need to be verified for each of the structural components and the type of deviations. Figure 2 shows two representations of the specifications in EN 13670:2009 [1], for the inclination of a structural column or wall, and for the alignment of stacked structural columns or walls.

	a	b
Type of deviation		
Description	Inclination of a column/wall at any level in a single- or a multi-storey building	Deviation between centres
Permitted deviation Δ	The larger of $h \leq 10 \text{ m} \Rightarrow 15 \text{ mm or } h/400$ $h > 10 \text{ m} \Rightarrow 25 \text{ mm or } h/600$	The larger of $\#30 \text{ or } 15 \text{ mm}$ but not more than 30 mm

Figure 2. EN 13670 [1] Walls and Columns geometrical tolerances example. (a) Inclination/verticality of a single wall/column, (b) Deviation between centres of stacked walls/columns.

In this paper, we present a solution to digitise such geometric specifications as digital rules, thereby creating a dictionary of all geometric specifications (i.e. all rules) that apply to a given project. We focus on 15 rules defined in EN 13670:2009 [1] for the execution of concrete structures.

For the ‘where’, the solution must identify where the above rules apply in the given project. For this, we developed an algorithm that processes the as-design BIM model, in IFC format [13], and detects where the geometric QC rules contained in the pre-defined dictionary apply. The output of this process is a network graph where nodes are the components in the as-design BIM model requiring geometric QC and edges are the component relationships that are relevant to the geometric specifications to be applied to the project.

Finally, for the ‘when’, we take into account the information contained in the project 4D model that links the as-design BIM model components to construction schedule activities, so that the defined list of geometric QC tasks can be laid over time, in line with the construction schedule.

The above three steps are detained in the following subsections.

2.2 What: Establishing the Geometric QC Dictionary

The geometric QC dictionary encompasses all the geometric QC ‘rules’ that need to be checked during the project. Each rule defines the context within which it must be applied. In the case of structures, that context is defined by: the types of components involved in the rule (e.g. wall, column, slab), their material types (e.g. concrete or steel), and their geometric relationships (e.g. stacked, connected). Note that some rules refer to single components (e.g. column inclination), in which case no geometric relationship applies to define the rule’s context. At each location in the as-design BIM model where the defined context is encountered, the corresponding rule must be checked. Since we consider the as-design BIM model to be encoded in IFC format, it is important that the dictionary rules be encoded by employing standard (or pre-agreed) IFC classes and properties, such as *IfcWall* or *IfcColumn* for component types, and material category contained in the *IfcMaterial* field of a component description.

The dictionary rules then contain additional description fields, capturing where the rule comes from (e.g. title of original document and section or specification number with it), and a description of the rule. Finally, the rule is associated to software code that encodes the actual geometric specification rule is to be applied to each instance where the rule context is detected in the BIM model.

In summary, the list of fields for each of the dictionary entries is as follows:

- **Rule Description:**
 - **Source document:** norm or regulation number, or specific ID to identify the original document
 - **Source section:** an ID value to identify the geometrical tolerance within the original document
 - **Description:** geometrical tolerance brief description
- **Rule Context:**
 - **Component Type:** the type of components the entry needs to be applied to (i.e. walls, slabs, etc)
 - **Material Type:** the material type of the structural components that the entry needs to be applied to (i.e. concrete, steel)
 - **Relationship Type:** the geometric relationship the different components need to be connected with that the entry needs to be applied to (i.e. above, below, same level adjacency, etc). Important remark here that in case the tolerance only involves a single element, this field can be left empty.

In our implementation, the dictionary is stored using the JSON open format, which is easy to read by users and most development tools.

Table 1 illustrates an example for a pair of the entries from EN 13670:2009 [1], where several cases of the keywords are represented for clarity.

Table 1. Quality control dictionary entries examples from EN 13670:2009

QC_1	
SourceDocument	EN 13670-2009
SourceSection	ColumnsAndWalls_a
Description	Inclination of element
ComponentType	[COLUMN, WALL]
MaterialType	[CONCRETE]
RelationshipType	[]
QC_2	
SourceDocument	EN 13670-2009
SourceSection	ColumnsAndWalls_b
Description	Deviation between centres
ComponentType	[COLUMN, WALL]
MaterialType	[CONCRETE]
RelationshipType	[ABOVE, BELOW]

2.3 Where: defining all instances of geometric QC to be conducted in a given project

Given the geometric QC dictionary, the BIM model is now analysed to identify all instances where each geometric QC rule applies, i.e. where geometric QC must be conducted. This is done by finding in the BIM model all instances where the ‘context’ defined in each geometric QC rule is found. It is performed in two steps. First, this requires the input as-design BIM model be interpreted to identify the component types and relationships of interest for the given geometric QC rules; this step is detailed in section 2.3.1. The output of the first step can then be further interpreted to identify all individual instances where the geometric specifications need to be controlled; this step is detailed in section 2.3.2.

2.3.1 BIM Model Interpretation

Here, we assume that component types and material types are explicitly encoded in the IFC file. However, component relationships are rarely provided, especially all the ones necessary for the geometric specifications we consider in the paper. As a result, specific algorithms had to be developed to detect such relationships in the model. It is proposed here to represent the outcome of this BIM model interpretation process using a network graph where each node represents a BIM model component (3D component in our case) and each edge represents a relationship

between pairs of components. The nodes and edges have properties that are defined for the intended purpose.

In the context of geometric QC of concrete or steel structure works with the above-defined geometric QC dictionary, the following properties are defined. The nodes contain the following properties: structural component **unique ID**, the **type of component** (e.g. wall, beam, column, slabs), the **material type** (e.g. concrete, steel), and a **label** (if provided in the BIM model) to facilitate the component identification for the users.

The edges contain the following properties: **source UID**, **target UID** (indicating the source and target structural components in the BIM model the edge is connected to), and the **geometric relationship type**. For the latter, we currently consider the following relationship types: *above*, *below*, *adjacent storey level*, *same storey adjacency* or *physical connection*. These relationships are those employed in the different geometric QC specifications defined in EN 13670:2009 [1].

These geometric relationships are computed by analysing each structural component (nodes), requiring various levels of computation effort. However, the details of these computations are beyond the scope of this paper (due to space constraints).

The use of the network graph has multiple advantages: First, it is easy to read. Beside, meaningful statistical information can be obtained from it overall (e.g. graph density) or for each of its nodes (e.g. centrality analysis), which can subsequently help analyse the ‘tightness’ of the geometric specifications and the ‘criticality’ of certain components.

The network graph only needs to be computed once, since the structural components are not going to change once construction is initiated (if they are the network could naturally be recomputed). Figure 3 illustrates a very basic graph example with a couple of nodes and their relationships.

Naturally, the graph structure with its nodes and edges properties can be extended to include other component types and/or properties that can be useful to satisfy other types of geometric QC (and more broadly, other use cases). The different properties that are depicted in this manuscript are the ones that were identified as minimal requirements to obtain the geometric QC of concrete structures according to the EN 13670-2009 along the duration of the project.

2.3.2 QC instances

The second step produces all individual instances of geometric QC tasks that need to be conducted for the given construction project.

For each dictionary rule, the network graph is queried to provide all unique instances of the rule ‘context’, i.e. the structural components with the same *Component Types*

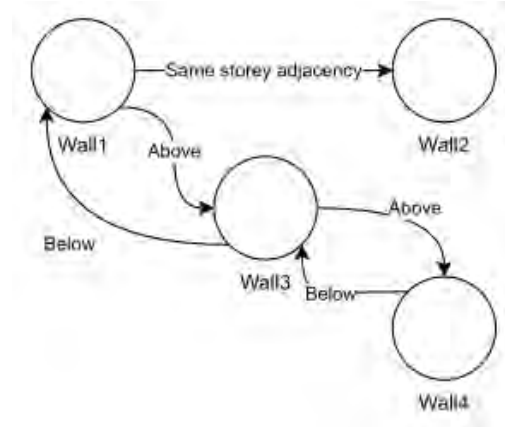


Figure 3. Example network graph of structural components and their relationships.

and *Material Types*, and, in the case the entry requires multiple components, the related components with the *Relationship Type*.

Table 2 shows a small example of the result of the search for the dictionary rules defined in Table 1 (rules QC_1 and QC_2) in the network graph of Figure 3. Table 2 shows the list of QC task associated to ‘Wall1’ only.

Table 2. Geometrical tolerances list example

QC0		
QC_UID	Project_example_QC0	
Dictionary_ID	QC_1	
SourceDocument	EN 13670-2009	
SourceSection	ColumnsAndWalls_a	
Description	Inclination of element	
Result		
Involved_Components	[Wall1]	
TimestampSchedule	01/01/2022	
TimestampPerformed		
Unit	[,]	
ScalarResult	[]	
ToleranceReference	[]	
AuxiliaryOutputFile		
QC1		
QC_UID	Project_example_QC1	
Dictionary_ID	QC_2	
SourceDocument	EN 13670-2009	
SourceSection	ColumnsAndWalls_b	
Description	Deviation between centres	
Result		
Involved_Components	[Wall1, Wall3]	
TimestampSchedule	01/01/2022	
TimestampPerformed		
Unit	[,]	
ScalarResult	[]	
ToleranceReference	[]	
AuxiliaryOutputFile		

2.3.3 When: scheduling of the QC instances

Beyond the as-designed 3D BIM model, projects increasingly develop 4D model that link the design 3D model with the construction schedule. In a 4D model, the different physical components making up the design are linked to the construction activities that construct them. These activities are defined by start date and duration (and therefore finish date). The QC Manager also queries this information from the input (4D) BIM model, so that the overall list of unique geometric QC tasks that is automatically extracted from the 3D BIM model can be organised along the project timeline. This way, once a construction task is completed (e.g. building columns ground floor), then the project manager and QC manager know the exact list of all geometric QC tasks that need to be performed at that specific point, i.e. and thus where surveying needs to be conducted. Here, we refer the reader back to the "Construction Phase" box in Figure 1.

3 Case study: Revit Sample Project Technical School

The proposed methodology has been tested in the Revit Sample Project Technical School [14]. This model is a sample model provided by Autodesk. We focus on the structural model (Figure 4), since the rules considered in these tests focus on structural works.

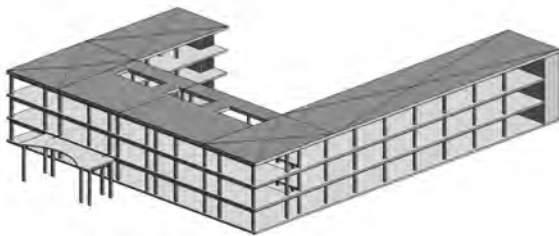


Figure 4. Technical School sample model

The (preliminary) version of the dictionary used in this test contains a total of 15 rules, all of them selected from the EN 13670-2009 document [1]. The selected rules involve four types of structural components: (*slabs, columns, walls, beams*) and five types of geometrical relationships: *above, below, adjacent storey level, same storey adjacency, and physical connection*. The rules include (reference to document section between brackets):

- Inclination of a column/wall (10.4 Columns and Walls No a)
- Deviation between centres (10.4 Columns and Walls No b)
- Curvature of a column/wall between adjacent storey levels (10.4 Columns and Walls No c)

- Location of a column/wall at any storey level w.r.t. base level (10.4 Columns and Walls No d)
- Location of a beam-to-column connection measured relative to the column (10.5 Beams and Slabs No a)
- Position of bearing axis of support (10.5 Beams and Slabs No b)
- Cross-sectional dimensions (10.6 Sections No a)
- Lap-joints (10.6 Sections No c)
- Free space between adjacent columns/walls (Annex G - G.10.4 Columns and Walls No c)
- Horizontal straightness of beams (Annex G - G.10.5 Beams and Slabs No a)
- Distance between adjacent beams (Annex G - G.10.5 Beams and Slabs No b)
- Inclination of a beam/slab (Annex G - G.10.5 Beams and Slabs No c)
- Level of adjacent beams (Annex G - G.10.5 Beams and Slabs No d)
- Level of adjacent floors at supports (Annex G - G.10.5 Beams and Slabs No e)
- Orthogonality of a cross-section (Annex G - G.10.6 Sections No a)

The test consisted of loading the different structural components from the IFC file, compute their geometric relationships and construct the network graph. The graph could then be analysed to generate the list of geometric QC tasks that would need to be conducted over the duration of the project.

The model contains a total of 589 structural components. Our algorithm reads the IFC file (exported from Revit) and generated the network graph containing those 589 components as nodes. Beside, the algorithms found 6,677 relevant geometrical relationships converted into 6,677 corresponding edges in the graph. Those nodes and relationships are broken down as follows:

• Components:

- 6 Walls
- 5 Slabs
- 203 Columns
- 375 Beams

• Relationships:

- 126 Above
- 126 Below
- 5 Adjacent storey level
- 6,013 Same storey adjacency
- 407 Physical connection

Figure 5 shows the graph structure distribution and its colour reference legend, where we can visualise all the structural components contained in the BIM model represented by the coloured nodes, and the geometric relationships connecting them represented by the coloured arrows

between the nodes. We can also identify the high amount of columns (orange nodes) and beams (blue nodes) that is contained in the BIM model, representing more than the 95% of the structural components, and the different relationships between them. Each node has a size proportional to its *degree centrality* (i.e. the number of edges connected to it).

Approximately 90% of the edges belong to the same storey adjacency category, hence the graph is full of those edges making it difficult to visualise the other relationships. Using the graph visualisation and analysis tool Gephi [15], the graph is explored and analysed, extracting interesting statistics and exploring different arrangements. For example, Figure 5 shows a *Force Atlas 2* layout, where we can easily depict the different storeys by the way the columns form three distinct clusters. Degree centrality can also be further investigated. Nodes with high degree centrality can be considered as potentially more critical from a geometric QC viewpoint, so the building team may be interested to know which ones they are and the reasons for their criticality. For example, Figure 6 shows a highlighted component (on the top) with high degree centrality, indicating it can be a critical structural component from a geometrical tolerance compliance viewpoint, while on below, we can see the same component identified in the as-design BIM model, where it may be easier for the contractor to understand its criticality.

After generating the network graph, it can be analysed according to the quality control dictionary to output a full list of geometric QC checks that need to be conducted. In this example, the list contains 9,750 unique geometric QC checks. These unique entries in the list demonstrate the total number of quality control actions that should be carried out during the entire project's execution and that should all output a positive result to certify the quality of execution. Table 3 shows a snapshot of the generated JSON file containing all the project's geometrical tolerances.

It is important to remark that despite the limited number of geometric QC rules in the dictionary, the total number of unique geometric QC checks is still significant. This numbers suggests that all checks are not systematically conducted in practice, with many results potentially implied from some others. The rest of the GeometricQC tool aims to automate these control and will ensure they are systematically conducted.

4 Conclusion and future work

The tool proposed in this paper is meant to be employed during project planning stage and has two steps: First, it automatically extracts all the necessary information from the design BIM model (IFC file) and stores it in a convenient data structure, a network graph. Then, the tool automatically detects within the network graph all

Table 3. Technical School geometrical tolerances list example

QC0		
	QC_UID	Rst_advanced_sample_project_QC0
	Dictionary_ID	QC_1
	SourceDocument	EN 13670-2009
	SourceSection	ColumnsAndWalls_a
	Description	Inclination of element
	Result	
	Involved_Components	[01SfNHv5nEReC9M9Bzo7D_]
	TimestampSchedule	17/01/2022
	TimestampPerformed	
	Unit	[,]
	ScalarResult	[]
	ToleranceReference	[]
	AuxiliaryOutputFile	
QC1		
	QC_UID	Rst_advanced_sample_project_QC1
	Dictionary_ID	QC_15
	SourceDocument	EN 13670-2009
	SourceSection	Sections_annex_a
	Description	Orthogonality of a cross-section
	Result	
	Involved_Components	[01SfNHv5nEReC9M9Bzo7D_]
	TimestampSchedule	17/01/2022
	TimestampPerformed	
	Unit	[,]
	ScalarResult	[]
	ToleranceReference	[]
	AuxiliaryOutputFile	
QC100		
	QC_UID	Rst_advanced_sample_project_QC100
	Dictionary_ID	QC_9
	SourceDocument	EN 13670-2009
	SourceSection	ColumnsAndWalls_annex_c
	Description	Free space between adjacent elements
	Result	
	Involved_Components	[15kM\$yjmD0FxFxNxtEUi7rG8, 15kM\$yjmD0FxFxNxtEUi7rIs]
	TimestampSchedule	17/01/2022
	TimestampPerformed	
	Unit	[,]
	ScalarResult	[]
	ToleranceReference	[]
	AuxiliaryOutputFile	
QC8872		
	QC_UID	Rst_advanced_sample_project_QC8872
	Dictionary_ID	QC_7
	SourceDocument	EN 13670-2009
	SourceSection	Sections_a
	Description	Cross-sectional dimensions
	Result	
	Involved_Components	[3KQkoT3ZD758EdANtsCisU]
	TimestampSchedule	17/01/2022
	TimestampPerformed	
	Unit	[,]
	ScalarResult	[]
	ToleranceReference	[]
	AuxiliaryOutputFile	

unique instances of the contexts within which each geometric specification applies and QC must be conducted. These geometric specifications are stored as rules in a QC

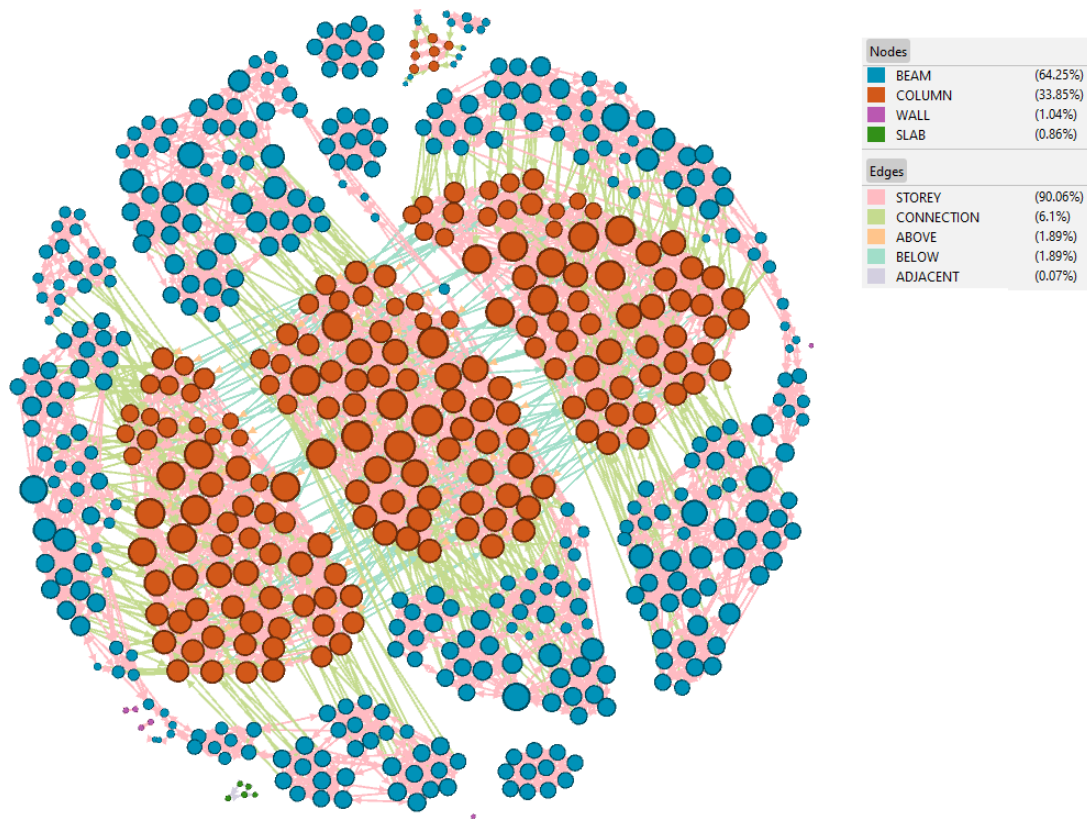


Figure 5. Graph representation with a Force Atlas 2 layout

dictionary, with the overall geometric QC methodology employed by the proposed GeometricQC tool akin to Automated Rule Checking (ARC). The value of the proposed tool to identify all necessary geometric QC tasks within a given project is demonstrated with a realistic mid-size project (school building).

In future works, the planning stage methodology will be extended to include the rest of the structural deviations presented in EN 13670-2009 [1] and EN 1090-2:2018 [12], in addition to the rest of the geometrical relationships, and a fully integrated schedule from the 4D BIM model. Then, the full pipeline of the GeometricQC tool will be delivered that will conduct the automatic geometric QC with point clouds acquired on site. Finally, as part of the large COGITO project [9], the tool will be integrated within a Construction Project Digital Twinning ecosystem enabling construction activities, including QC, to be effectively orchestrated and their output recorded in a structured way with the project Digital Twin.

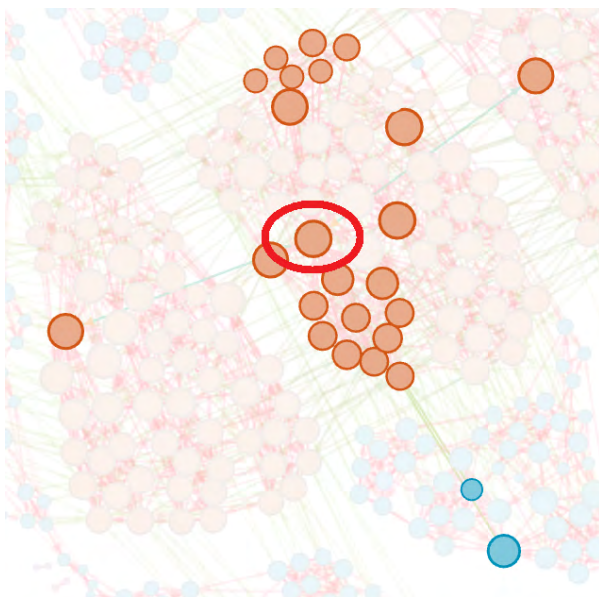
Acknowledgements

The research leading to these results has received funding from the European Union's Horizon 2020 research and

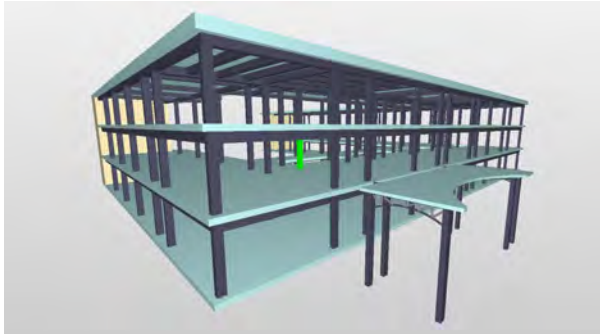
innovation programme under grant agreement No 958310. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

References

- [1] European Committee for Standardization. EN 13670:2009 - Execution of concrete structures. Standard, CEN, 2010.
- [2] Agarwal R., S. Chandrasekaran, and S. Sridhar. Imagining construction's digital future. Technical report, KcKinsey & Company, 2016.
- [3] Aimi Sara Ismail, Kherun Nita Ali, and Noorminshah A. Iahad. A review on bim-based automated code compliance checking system. In *2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*, pages 1–6, 2017. doi:10.1109/ICRIIS.2017.8002486.
- [4] Robert Amor and Johannes Dimyadi. The promise of automated compliance checking. *Developments*



(a) Column node selection with all its geometric relationships



(b) Column selection in BIM model

Figure 6. Column component inspection in (a) graph and (b) model.

in the Built Environment, 5:100039, 2021. ISSN 2666-1659. doi:10.1016/j.dibe.2020.100039.

- [5] Frédéric Bosché. Automated recognition of 3d cad model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction. *Advanced Engineering Informatics*, 24(1):107–118, 2010. ISSN 1474-0346. doi:10.1016/j.aei.2009.08.006. Informatics for cognitive robots.
- [6] M.-K. Kim, Q. Wang, J.-W. Park, J.C.P. Cheng, H. Sohn, and C.-C. Chang. Automated dimensional quality assurance of full-scale precast concrete elements using laser scanning and bim. *Automation in Construction*, 72:102–114, 2016. doi:10.1016/j.autcon.2016.08.035.
- [7] Changmin Kim, Hyojoo Son, and Changwan Kim. Automated construction progress measurement using a 4d building information model and 3d data. *Automation in Construction*, 31:75–82, 2013. ISSN 0926-5805. doi:10.1016/j.autcon.2012.11.041.
- [8] Maarten Bassier, Stan Vincke, Heinder De Winter, and Maarten Vergauwen. Drift invariant metric quality control of construction sites using bim and point cloud data. *ISPRS International Journal of Geo-Information*, 9(9), 2020. ISSN 2220-9964. doi:10.3390/ijgi9090545. URL <https://www.mdpi.com/2220-9964/9/9/545>.
- [9] Construction-phase digital twin model: Cogito. On-line: <https://cogito-project.eu/>, Accessed: 09/02/2022.
- [10] J. Zhang and N. M. El-Gohary. *Automated Extraction of Information from Building Information Models into a Semantic Logic-Based Representation*, pages 173–180. ASCE, 2015. doi:10.1061/9780784479247.022.
- [11] Jiansong Zhang and Nora M. El-Gohary. Semantic-based logic representation and reasoning for automated regulatory compliance checking. *Journal of Computing in Civil Engineering*, 31(1):04016037, 2017. doi:10.1061/(ASCE)CP.1943-5487.0000583.
- [12] European Committee for Standardization. EN 1090-2 - Execution of steel structures and aluminium structures. Technical requirements for steel structures. Standard, CEN, 2018.
- [13] buildingSMART. Industry foundation classes (ifc). On-line: <https://www.buildingsmart.org/standards/bsi-standards/industry-foundation-classes/>, Accessed: 17/02/2022.
- [14] Autodesk. Revit sample project files. On-line: <https://knowledge.autodesk.com/support/revit/getting-started/caas/CloudHelp/cloudhelp/2022/ENU/Revit-GetStarted/files/GUID-61EF2F22-3A1F-4317-B925-1E85F138BE88-htm.html>, Accessed: 12/02/2022.
- [15] Mathieu Bastian, Sebastien Heymann, and Mathieu Jacomy. Gephi: An open source software for exploring and manipulating networks. In *International AAAI Conference on Weblogs and Social Media*, 2009. URL <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>.

Insights into automation of construction process using parallel-kinematic manipulators

M. Klöckner^a, M. Haage^a, H. Eriksson^b, H. Malm^c, K. Nilsson^d, A. Robertsson^e, R. Andersson^f

^aDepartment of Computer Science, Lund University, Sweden

^bWinsome Consulting AB, Lomma, Sweden

^cFojab arkitekter AB, Malmö, Sweden

^dCognibotics AB, Lund, Sweden

^eDepartment of Automatic Control, Lund University, Sweden

^fAChoice AB, Malmö, Sweden

E-mail: maike.klockner@cs.lth.se, mathias.haage@cs.lth.se, helena@winsomeconsulting.se, Henrik.Malm@fojab.se, klas@cognibotics.com, anders.robertsson@control.lth.se, ronny.andersson@achoice.se

Abstract – This paper discusses challenges, experiences and lessons learned so far while transforming a masonry build system based mostly on manual labour into a robot automated build system. Our motivation for selection of this masonry process is to try out how robot automation could impact the architects in their design work by providing a tool to directly manipulate wall expression down to individual brick level. Such manipulation is often much too costly for manual labour today. Moreover, masonry is a challenging application to automate. Understanding the manual processes involved and transforming them into automation equivalents faces several challenges; among them handling and distribution of the different materials involved, selection of tooling, sensing for handling of variation and digital tooling for the programming of the process. A novel parallel-kinematic manipulator (PKM) with computerized numerical control (CNC) is used as target for experiments, because the performance properties in stiffness, workspace and accuracy will allow us to extend work into further construction processes involving heavy and dirty manual labour.

Keywords –

Construction robotics; Parallel-kinematic manipulator; Masonry; Concrete build system

1 Introduction

Attempts at machines to perform automatic masonry have been tried from time to time. Even patents for bricklaying machines have already been announced in 1875 [1]. Despite this, bricklaying machines are not in common use today. There is one commercial machine

available, the SAM100, offering automation of bricklaying for large straight building facades. Another commercial machine, Hadrian X, is usually also mentioned but it uses a build system with much larger bricks. A recent online article called "Where are the robotic bricklayers" [2] suggests several reasons why automation for masonry is not widely spread today. Among others, mortar is highlighted as a difficult material to handle whereby it is difficult to produce clean mortar joints, which also mirrors our experiences.

In general, autonomous machines and robots are not common in the construction industry. Several articles investigate reasons why: [3] lists lack of interoperability, design for human installation procedures, lack of tolerance management, power and communications as hampering factors. [4] and [5] list high initial investment and risk for subcontractors, immature technology, unproven effectiveness, lack of experts, low R&D budgets, among others. But there are indications that automation is needed in construction for continued growth [6]. Our interpretation is that digital, technical and regulatory infrastructure is lacking to lessen the effort of introducing autonomous construction machinery in the construction value chain. In the project, which this article is part of, we therefore work towards a model to bring business, technology and infrastructure together for bringing commercial application of autonomous robots and machinery closer to reality [7]. We also started up the Center for Construction Robotics [8] in 2019 as a forum between actors to meet and a test site for experiments on robot automation of construction processes.

The paper focuses on mapping the explained masonry process (2) into robot equivalents, including adapting a new developed PKM for construction processes (3) as well as performing and discussing experiments (4).

2 Brick masonry fundamentals

The fundament on which our experimental setup is built on, from a masonry point of view, are bricks, mortar, tools and process performance.

Bricks which are used in Sweden have different properties regarding dimensions, structure, surface, weight and colour. Most utilized in the Swedish construction business are standard sized bricks of the following dimensions:

- 250 x 120 x 62 [mm] (Swedish brick) [10]
- 228 x 108 x 54 [mm] (Danish brick) [10]
- 240 x 115 x 71 [mm] (German brick) [11]

Bricks, independent of their size are available as vertical coring brick, horizontal coring brick, solid brick, pre-wall solid brick and pre-wall vertical coring brick [12]. Most available colours are yellowish, brownish, reddish and blackish. The surface itself is rough, sometimes sandy, corny or dusty and the weight of solid bricks is about 2000 kg/m³ [10][11], whereby single solid bricks of the sizes mentioned in the list above weigh between 1.3 kg and 2.0 kg.

Though there exist many different types of bricks with all their dissimilarities used in the construction industry, they have one thing in common: Deviations in their dimensions. We experienced up to +/- 2.5 mm in width and height along their surfaces (see Figure 1).



Figure 1: Real brick with deviations (left) and ideal brick (right)

The brick type we focus on in our experimental setup is a red Danish brick with 228 x 108 x 54 [mm]. Next to the described deviations the brick contains the following characteristics: Red Danish bricks having an upper and a lower side and a front and a back. To achieve the best possible wall impression, it is important to arrange the bricks in the same orientation along a wall to build. Moreover, they have a solid body and have fine dust on their surfaces. These parameters are very important for the automatization since they are influencing the tool design for handling the bricks as well as the choice of sensory. A solid body brick can for instance be handled with vacuum technology whereas a hollow brick needs clamping technology to be handled. The deployed vision technology (Realsense D435) features point cloud acquisition. In particular the point cloud measurements are matched towards brick geometry for brick detection and localization. The camera is fairly robust in the utilized pick situations. The algorithm uses thresholding

in depth to differentiate between brick surface and background.

2.1 Manual brick masonry process

The manual masonry process for building a straight wall with bricks includes different action steps, logistics and use of special tools. To identify parts of the process appropriate for automatization we analyzed this process in detail in a workshop with an expert masoner. The process is divided into three main steps including preparation, performance and post-processing. Sub-steps for the preparation including blending the mortar, building a frame for building a leveled wall, putting a horizontal chord to define a specific height for each layer of mortar and bricks, brick- and tool supply. Sub-steps for post-processing includes grading vertical and horizontal joints, possible wall plastering and disassembly of framework.

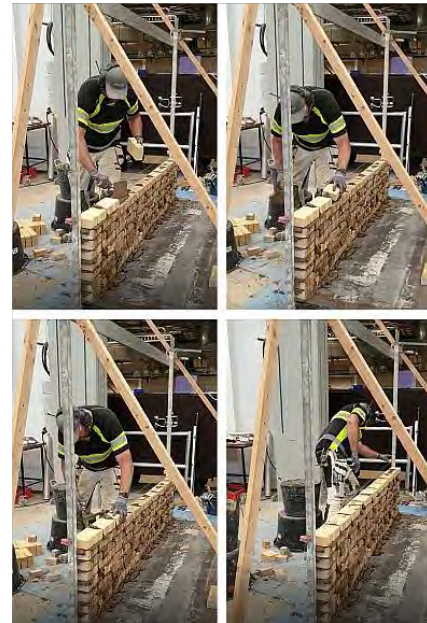


Figure 2: Bricklaying: Apply mortar (upper left); disperse mortar with brick (upper right); push brick in position (down left); set stressed chord for next layer (down right)

For building a layer of bricks (process performance see Figure 2) for a straight wall, we have set the horizontal stressed chord to a specific height. Thereby we defined a continuous height through all wall layers and achieved a regular build wall. For the application of bricks, we have put mortar with a defined dispersion on the prebuild brick layer. The mortar was applied with a tool especially designed for the needed mortar dispersion (Figure 3 (A)) for the used sizes of bricks. Thereafter the brick was applied with a specific application strategy.

This strategy contains tilting the brick behind the wall along its longitudinal and lateral axis (Figure 3 (B & C)), pushing it in this configuration on top of the mortar from the back of the wall, while tilting both axes back to “zero”. Thereby the current placed brick gets aligned along the former placed bricks (Figure 3 (C & D)).

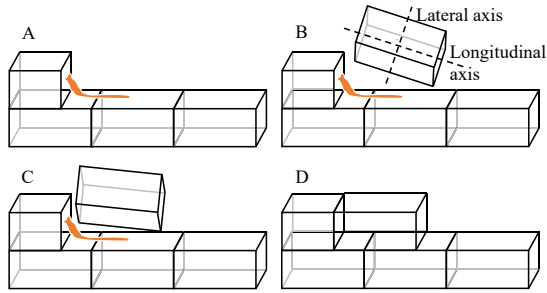


Figure 3: Application strategy mortar and brick – front view

Through this application technology a vertical joint between two side by side bricks as well as horizontal joints between bricks of two sequenced layers are generated. Furthermore, this application strategy offers us to control the flow of excess mortar whereby the joints become as clean as possible. To increase the walls stability vertical joints between bricks of two sequenced layers needs to be displaced. For the displacement we used normal, half sized bricks, which we generated by breaking them with a special hammer (see Figure 4).



Figure 4: Generate normal, half sized bricks

In general bricks are used both as facing bricks and bearing brick structure, i.e., a complete wall. The different bricks are in each of these uses laid in different pattern. Both related to visual and physical reasons. In practice also the length of a wall decides how different individual bricks must be cut and placed. Different thickness of vertical and horizontal joints is used and several different recipes for mortar are available.

2.2 Wall to build

For the brick experiments presented in this paper, we designed a curved masonry wall (Figure 5). The base curve of each layer of bricks is a planar sinusoidal curve, where the amplitude decreases with the height of the layer above ground. This shape offers the possibility to demonstrate the ability of changing brick placement.

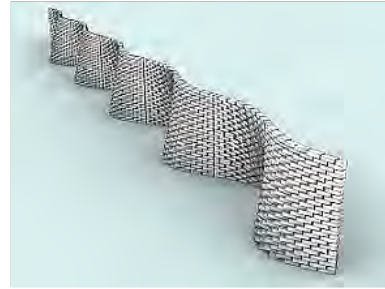


Figure 5: Curved masonry wall

While having a robot being able to directly interpret the wall's design, we can precisely control the exact position and orientation of individual bricks with our robotic masonry system, which is, manually performed, very challenging. The bricks to set are individually rotated out of the tangent direction of the base curves. This individual rotation is proportional to the curvature of the base curve at the location of the current brick, see Figure 6. Thereby we create a wall with differently sized angles between side-by-side placed bricks, which causes differently sized vertical joints.



Figure 6: One wall module used as reference for automated masonry process: Top view (left); Tilted front view (right)

For creating the 3D model of this wall and for calculating the exact position and orientation of each brick, the algorithmic modelling extension Grasshopper to the 3D modelling software Rhinoceros 3D was used, cf. [13]. Algorithmic modelling tools are essential for robotic masonry since manual specification of the orientation and position of each brick would be impossible in larger projects, e.g., in a complete brick facade of a building.

The 3D model can be used for visualising the finished appearance of the wall, but more important, the

algorithmic model also exports the orientation and position of each brick to the robotic system to construct the steering code for the brick-laying robot, which is further described in 4.2.

3 Parallel-kinematic machines

In contrast to a standard arm manipulator where the robot links are arranged in a serial chain from the base of the robot to the tool, parallel-kinematic manipulators typically consist of a robot structure where several parallel links are attached to a common tool plate. The so-called closed kinematic loops together lock degrees of freedom for the position and orientation of the tool. The PKM's fundamentally different design allows for important properties like e.g., less moving mass and significant higher stiffness which may offer important benefits compared to standard industrial manipulators with respect to acceleration, positioning accuracy, structural rigidity with respect to process forces and e.g., footprint/workspace and complementary broadens the applicability and use of robots [14][15]. The PKM configuration used within the project is based on a gantry frame and a novel wrist construction. It will be explained and its benefits within construction robotics for masonry operations and efficient use in Architecture, Engineering and Construction (AEC) applications, will be highlighted and further discussed in section 3.1.

3.1 PKM for brick masonry

In previous work "Parallel-kinematic construction robot for AEC industry" [16] our work-in-progress to adapt a PKM structure to automate a selected masonry process was presented. In the meanwhile, we have set up the PKM in the laboratory and equipped it with necessary hardware and software items to perform experiments.

The PKM is an eight-link parallel-kinematic manipulator that provides 5-axes continuous motion. There are six links that in three pairs connect the three carriages on the 4 m linear guides with the so-called *support platform* that positions the base for the robot wrist mechanism. Each of these six links have a fixed length of 2 m. By controlling carts on the three linear guides, the robot can perform translatory movement with the support platforms keeping a very stiff orientation. To provide stiff and precise tool orientation in two directions (tilting the tool, while keeping the third orientation stiff), two telescopic links are mounted between the upper and lower carts respectively. Together with a cardanic joint between the support platform and the tool platform, this results in controlled rotational motions around x- and y-axis. This type of machine provides a large singularity free workspace, high rigidity and precision, as also described in [17]. For PKM including its workspace see Figure 7.

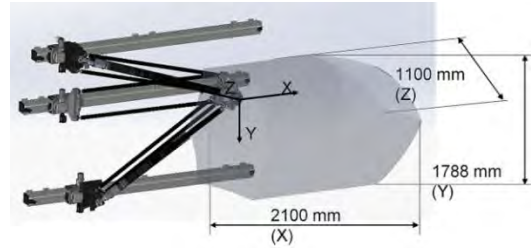


Figure 7: Workspace PKM

To fit with laboratory conditions the PKM is presently mounted on a horizontal support structure. Though it is also possible to mount it in a setup for working from top for example. Figure 8 shows the PKM connected to a support structure in a brick masonry process.

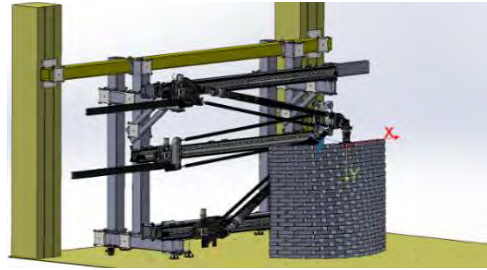


Figure 8: PKM mounted on support structure in brick masonry process configuration

To adapt the PKM for the described masonry process, we designed a tool (Figure 9) containing two motors allowing rotation around z- and y-axis.



Figure 9: Tool side view (left); Front view (right)

Moreover, the brick masonry tool consists of adapter plates and motors that transmit the movement generated by the motors into rotary tool motions via cross-roller bearings. The use of an L-shaped part allows the tool to point down in zero configuration. Furthermore, we decided to use a vacuum gripper consisting of three vacuum cups with 55 mm diameter each, equipped with filters inside to handle the dusty bricks and provided with foam to be able to create vacuum for gripping the rough brick surface. Flow regulation and control of vacuum gripper is realized with Avac injector MFE-300H-AS-1.

With a connected 10 m long and 8 mm diameter air hose and 6 bar pressure, we achieved 60 % flow. For easy and fast tool changing option RSP TC60-8 tool changer and RSP TA60-8 tool changing adapter are included in rotational y-axis.

For the mortar application process, we equipped the tool with an extruder mounted on the L-shaped part mentioned before and connected the extruder to a circular nozzle at its end (Figure 10 left). On the other end a connection for a hose is prepared which connects the mortar pump with the extruder.

4 Experiments

For performing experiments with the developed parallel-kinematic manipulator, we built the experimental setup described in 4.1., generated the brick data needed for the parallel-kinematic manipulators programming and focused on the challenges to solve (section 4.3). Furthermore, we performed experiments described in 4.4 to evaluate our proposed solution. Moreover, we list and discuss our results (section 4.5), including successful realization as well as obstacles which had been occurred and possible solution options, which offer the base for conclusion and future work (section 5) including Technology Readiness Level enhancement of the parallel-kinematic manipulator.

4.1 Experimental setup

Our experimental setup (see Figure 10) located in the Swedish National Center for Construction Robotics [8] contains the PKM equipped with the following peripherals: Described vacuum gripper tool to handle bricks, a palette with bricks placed on stacks for picking application, a double sized palette equipped with a flake board for placing application, a controller, a workstation, acrylic glass walls for safety during robot execution, process peripherals connected to tool for mortar application, computer and Intel RealSense Depth camera for vision integration.



Figure 10: Experimental setup: Mortar glue configuration (left); Pick and place configuration (right)

4.2 Wall data generation & transformation

For creating the steering code for the bricklaying PKM, data on each stone is exported directly from the algorithmic 3D model of the wall (2.2). In this project, this data is exported in the form of an automatically created Microsoft Excel spreadsheet file. Apart from the orientation and position of each brick, this file contains a column specifying the layer, and consecutive number on this layer from left to right, of the current brick. It also specifies if the current brick is a full-sized or half-sized brick. Half-sized bricks are used in the current design for having straight vertical ends of each module of the wall.

To feed the parallel-kinematic machine, which is programmed in G-code based on the data exported as spreadsheet file, a semi-automatic tool chain is used to transform from wall description to executable G-code.

4.3 Challenges

Challenges we are focusing on within our experimental performance contain implementation of a stable process for brick handling and mortar application by building the wall described in 2.2. The processes sub-challenges can be divided into: Pick bricks, move bricks, place bricks and apply mortar.

Though we focus in our application on one specific brick type, bricks are having, as experienced, deviations up to +/-2.5 mm in width and height, while having a dusty and rough surface. Our gripper decision is based on the fact, that it would be most convenient to pick directly from a palette. On the palette the bricks are placed with no space in between each other and their upper or lower side points outwards the palette. For this reason, we decided to design a tool containing a vacuum gripper to be able to easily separate the bricks from the palette. For performing pick experiments with the chosen gripper in combination with the bricks to handle, we started to place bricks in stacks on a palette. To pick a brick with the vacuum gripper we need to be close enough to the brick to get it connected to the gripper. In case we drive too close the very sensible foam (grippability in this solution depends on the foam being able to create enough suction force) placed at the end of the vacuum cups releases its connection to the cup which impairs the flow we need to create the vacuum, by what we are not able to create enough vacuum to grip the brick. In addition, to offer a continuous good flow we have a filter implemented in the vacuum cups, which we clean by blowing of several times before we pick a brick and after we have placed a brick. A blocked filter also decreases air flow, which decreases the needed vacuum.

Moreover, the deviations of the bricks are causing bending in the stacks and displacement of bricks (see Figure 11 (upper middle)). A calibration procedure matches the pick position in camera space with the

corresponding position in robot space. Measurements during operation calculate the relative change in camera space to the calibrated position. The calculated relative change is then applied to the pick position in robot space.

Regarding cycle times and safety, we define speed and acceleration as fast as possible to not disconnect the brick from the gripper during acceleration and braking / emergency stop. We accelerate the bricks while connected to the robots' end-effector with 3.87 m/s^2 and moved them with a constant speed of 1 m/s .

In terms of placing, we investigate placing strategies for dry stacking and for stacking with mortar between each layer as well as general design limits of bricklaying. Since all bricks have deviations, the deviations get bigger with an increasing number of layers during dry stacking performance. For this reason, we decided to place the bricks a few millimeters over the last applied layer and let it drop on top. For placing bricks on a mortar layer, we had to figure out, if the dead load of the bricks is enough to get a good connection to the mortar or if we need to apply a defined pressure for placing the bricks on the mortar.

Due to process requirements and experiences from former experiments we decided to mix mortar glue according to the provided formula. By this we got the needed viscosity to get the mortar glue pumped through the hose as well as applied on the bricks. Furthermore, it is mandatory for a good process flow to apply the right amount of material with the right speed and the right consistency to achieve proper results. The aim is to apply pumpable mortar, which we are still investigating.

Finally, design limits, including wall instability, will be caused by external factors like 10 mm joint height between brick layers, dry behavior of mortar and wall design itself as well as by internal process concerning masonry robotics hard- and software.

4.4 Experiment performance

Our experimental performance is divided into two main parts. First, we investigate dry stacking of the wall described in 2.2. Second, we focus on stacking the wall with different mortar application strategies. The process for dry stacking bricks with the adapted PKM for masonry processes includes driving to a defined position over the stacked bricks, taking a picture of the next brick to pick and calculate the displacement regarding the reference brick defined in the process preparations. Afterwards, the PKM drives to the pick position and picks a brick. Thereafter the brick is moved to the placing position and placed with the defined tilt around the y-axis (Figure 11). The wall we build with this handling technology is shown in Figure 12.

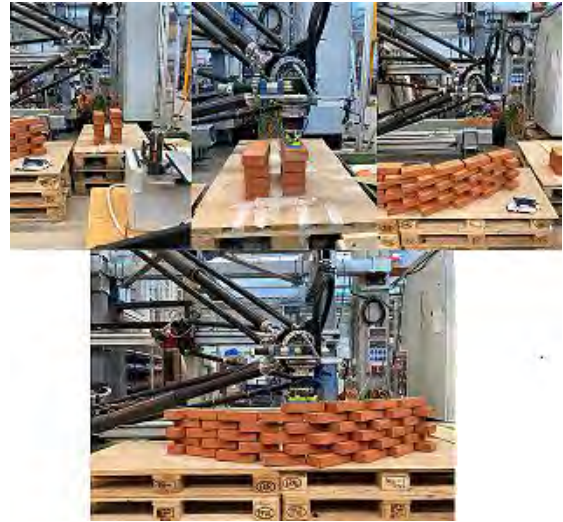


Figure 11: Dry stacking process: Taking picture for vision (upper left); Pick brick (upper middle); Move brick (upper right); Place brick (down)



Figure 12: Dry stacked wall

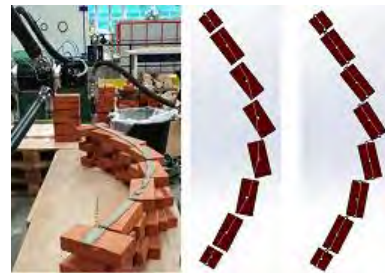


Figure 13: Path A applied on layer of bricks (left); Path A (middle); Path B (right)

For stacking with mortar, we investigate different mortar application strategies with the objective to identify a strategy e.g. without spill and without exposed mortar glue. Therefor we first focus on mortar application on top of a built layer by generating two paths, see Figure 13. Path A goes through the start point of the first stone, the midpoints of all stones and the endpoint of the last stone with a constant speed. Path B goes through all start-, mid-, and endpoints of every stone contained in the layer and is speeded up over the joints.

Since we investigate Path A as the best, we use this as the base to apply another layer of bricks on top, which is shown in Figure 14.



Figure 14: Stones applied on mortar path A

4.5 Results

Results we achieved by performing experiments with regards to the identified challenges are also divided into pick bricks, move bricks, place bricks and apply mortar.

The tool decision including two more rotational axes offer us to handle the bricks in the orientations and positions we need. Furthermore, with the chosen vacuum technology turned it out that the dustiness of the stones caused trouble, because it gets sucked into the air hoses and could, in long term, destroy the injector used for generating the vacuum. A proper solution could be to use another filter unit in front of the injector to offer a long-term use of this part. Furthermore, the vacuum cups turn out very sensible in case we drive too close to the brick while at the same time a good vacuum could just be created when we were close enough to the bricks. Implementing a force-torque sensor into the tool to identify the ideal distance between bricks and the vacuum gripper by control of pressure between these two objects would produce relief. Furthermore, the foam attached to the vacuum cups is very sensible. Investigating a Schmalz FQE-xb-120x60 with foam will be of interest.

Like mentioned the bricks to handle have individualized sides. The identification of the different sides requires the use of cognitive sensory integration into the process. Considering that the bricks need to be placed in a defined orientation and position to offer the best wall impression, a stationary vision system in combination with a further manipulator could be integrated into the setup.

The accuracy for picking the red Danish bricks could be enhanced demonstrably by use of the implemented vision system. The high variety of existing brick types opens the need to enhance the existing vision system for a more flexible future application.

Moving the bricks with the needed acceleration and speed worked very well with our configuration. In dependency on how close human and robot will work together in a demonstrated cooperation an additional gripping solution could be used when the bricks are moved by the robot to avoid disconnection of the brick during movement, which means to avoid possible injuries

by a flying stone to a human for example. Cycle times which we achieved within our setup for continuous brick handling are 145 bricks / hour, with time divided between pick (7 s / brick), place (4 s / brick) and travel between these positions (14 s / brick). Cycle times for manual brick handling are 300 bricks / hour, excluding breaks.

A possible force-torque sensor would also be very useful to enhance placing of the bricks. For dry stacking we would be able to handle the deviations in height by an appropriate control. For stacking bricks on a layer of mortar the force-torque sensor could help to push the brick with a predefined force into the mortar up to a predefined position and orientation, which adapts the stressed chord from the manual equivalent.

The application of mortar has different constraints. For applying mortar on a layer of bricks with a continuous flow and a constant dispersion we need, besides a specific material consistency, also a defined distance of 2 – 4 [mm] between the nozzle exit and the bricks. For path adjustment in height visual depth cognition could be used.

For keeping the needed mortar consistency during experiments, we blended it with a hand blender, in a barrel connected to the pump, frequently. With regards to automation, a solution with continuous blended mortar in a mortar blender and an additional mortar supply mechanism at the end effector in combination with an on purpose-built nozzle could enhance the process quality. Since a lesson learned is that the mortar material needs to be adapted to the robotic process, we are now starting up collaboration with concrete suppliers partly based on our experiences from this work.

Our decision to apply mortar along path A (Figure 13) is based on the fact, that this path overlaps almost with the layer of bricks it is applied on as well as with the layer of bricks which is applied on the mortar (see Figure 14). Apply mortar along Path B ensures less overlapping with the layer of bricks it is applied on, as well as with the layer of bricks which is applied on the mortar path. For enhanced overlapping of mortar and bricks a mathematic model could be used to calculate the best performance mortar path.

Even if the curved wall design used in this project is more stable than a totally straight counterpart, improvements on the stability can, of course, be made. Stability would, for example, be greatly improved by making the wall two bricks deep and placing bricks in the wall which connects the two layers. This is a traditional way of creating a load-bearing masonry wall. The individual rotation of each brick is also introducing instabilities in the wall, both during construction and in the completed wall and while designing the wall we must be convinced that the forces applied on each brick will not push it out of place. It must also be checked that the down-facing side of each brick rests to at least 50 % of

its area on bricks in the layer below. This is essential for the adhesion of the current brick to the bricks below to guarantee stability.

Wall assembly and transportation of prefabricated walls are issues, which are currently performed manually. Improvements remain, at this stage, open issues.

5 Conclusion & future work

Within this paper we have shown our recent work-in-progress in terms of testing, for the brick masonry process adapted parallel-kinematic machine, to Technology Readiness Level 3 – 4. Since we have validated our predefined assumptions through dry stacking of bricks as well as through stacking bricks with mortar our investigations are highlighting the potential of parallel-kinematic manipulator's use in construction robotics. This lays the foundation for former explorations including enhancement of future process performance and parallel-kinematic behaviour.

Future work contains implementation of improvement opportunities into the current setup. This includes content, discussed in 4.5 and will mainly focus on further sensory integration for brick handling (pick, place, move) as well as enhancements of mortar application tool-design and mortar application strategies. Furthermore, we will focus on digital chain improvement to generate executable G-code directly out of the used CAD environment. Moreover, we will implement safety and interaction issues for enabling the PKM to be used for prefabrication processes close to construction site with the aim to bring the PKM up to a higher Technology Readiness Level. Our idea to use the PKM on the construction site itself, is to mount the arm system on site onto specific, application oriented support structures.

6 Acknowledgement

The authors gratefully acknowledge the support received from VINNOVA/”Innovativ Agile Construction for globally improved sustainability” (ACon 4.0 #2019-04750) as well as the support received from Robert Larsson (Cementa AB), Petra Jenning (Fojab), Cognibotics AB, Miklos Molnar (LTH) and Fasadgruppen Group AB. These projects provided the opportunity to investigate the use of parallel-kinematic manipulators in construction applications and to develop a parallel-kinematic manipulator for specific construction tasks.

References

- [1] Franke, C. Brick-Laying Machines: “Improvement in brick-laying machines”, Patent US158838A, 19/01/1875.
- [2] Potter, B. “Where are the robotic brick layers”. On-line: <https://constructionphysics.substack.com/p/where-are-the-robotic-bricklayers>, 29/7/2021, Accessed: 29/1/2022.
- [3] Saidi, K. S. et al (2016). Robotics in construction. In *Springer handbook of robotics*, pp. 1493-1520, Springer, Cham.
- [4] Buchli, J. et al (2018). Digital in situ fabrication- Challenges and opportunities for robotic in situ fabrication in architecture, construction, and beyond. *Cement and Concrete Research*, 112, 66-75.
- [5] Delgado, J. M. D. et al (2019). Robotics and automated systems in construction: Understanding industry-specific challenges for adoption, *Journal of Building Engineering*, 26, 100868.
- [6] Bock, T. (2015). The future of construction automation: Technological disruption and the upcoming ubiquity of robotics. *Automation in Construction*, 59, 113-121.
- [7] ACon 4.0 (#2019-04750), funded by VINNOVA/”Innovativ Agile Construction for globally improved sustainability”.
- [8] Swedish National Center for Construction Robotics, Online: <http://www.lth.se/digitalth/byggrobotik/>, Accessed: 29/1/2022.
- [9] Bruckmann, T. and Boumann, R. “Simulation and optimization of automated masonry construction using cable robots”. *Advanced Engineering Informatics*, Volume 50, 2021, 101388, ISSN 1474-0346, <https://doi.org/10.1016/j.aei.2021.101388>.
- [10] Specification for masonry units - Part 1: Clay masonry units. DIN EN 771-1, 2015.
- [11] Mauerziegel – Beschaffenheit von Ziegeln. DIN 105. 2021.
- [12] Jens Kallfelz. Lebensraum Ziegel. On-line: <https://www.lebensraum-ziegel.de/ziegellexikon/mauerziegel/allgemeine-definitionen-und-begriffe.html>, Accessed: 24/01/2022.
- [13] Robert McNeel & Associates. Rhinoceros 3D. On-line: <https://www.rhino3d.com/>, Accessed: 24/01/2022
- [14] Boer, C.R. et al. *Parallel Kinematic Machines: Theoretical Aspects and Industrial Requirements*, Springer, 1999.
- [15] Neugebauer R. A. *Parallelkinematische Maschinen: Entwurf, Konstruktion, Anwendung*, Springer, 2005.
- [16] Klöckner, M. et al. „Parallel kinematic construction robot for AEC industry“. *Proceedings of the 37th International Symposium on Automation and Robotics in Construction (ISARC)*, p. 1488 - 1495, Kitakyushu, Japan.
- [17] Cognibotics AB “A parallel-kinematic machine with versatile tool orientation”, Patent EP19192225A, US202063051221P, and further based on WO2021032680A1, 25/02/2021.

Predictive Maintenance Optimization Framework for Pavement Management

A. Kumar^a and O. Shoghli^a

^a Department of Engineering Technology and Construction Management, University of North Carolina at Charlotte, USA.

E-mail: akumar28@uncc.edu, oshoghli@uncc.edu

Abstract –

The preservation of highway infrastructure is essential for maintaining its capacity, safety, and efficiency for commerce and defense. Pavements are among the most important elements of highway systems that deteriorate over time. Hence, the goal of pavement asset management is to seek efficient investments where the methods applied will aid in identifying the most appropriate allocation of the resources available to the highway agencies. In the absence of unlimited resources, such decisions will always result in trade-offs in which funding certain assets will be needed at the expense of the other. Decision-makers need data-driven information regarding trade-offs to avoid the reactive solutions that are far from optimum and may be counter-productive over the long run. This paper proposes using a multi-objective predictive maintenance optimization framework using a non-dominated sorting multi-objective evolutionary algorithm (MOEA), for the optimum upkeep of pavements. The algorithm aims to find a spread of Pareto-optimal solutions by concurrently minimizing the life cycle cost and maximizing the level of service (LOS). A case study was developed to compare the model's effectiveness based on the maintenance data from the asset management plan of the California department of transportation. The results from the study will help develop promising techniques for the application of various multi-objective optimization systems and thus pave the way for efficient decision-making tools for the maintenance of highway infrastructure projects.

Keywords –

Asset Management; Life Cycle Cost; Level of Service; Multi-objective Optimization

1 Introduction

Pavement systems constitute one of the most valuable assets in all transportation agencies worldwide. Huge investments are made annually to preserve, expand, and

operate these facilities, which are invaluable for the movement of people, services, and goods. Driving on roads that need repair costs U.S. motorists \$120.5 billion in extra motor vehicle repairs and operating costs. The Federal Highway Administration (FHWA) estimates that each dollar spent on the road, bridge, and highway upgrade returns \$5.20 in the form of lower vehicular upkeep and maintenance costs, time savings, lower fuel consumption, safety, minimized bridge upkeep costs, and lower emissions as a consequence of enhanced traffic flow [1].

Pavement maintenance is essential for extending the service life of deteriorating highway assets. The deterioration of the pavement surface due to aging and extensive use is the main threat to the level of service provided by the highway system networks. Thus, transportation agencies endeavor to renew, repair, and maintain the transportation systems already in place [2]. With the advancement of technology, highway officials and maintenance managers have the opportunity to analyze both the short- and long-term consequences of the maintenance strategies. Utilizing the maintenance strategies to extend the service life of highway systems reduces the frequency of infrastructure replacement and life cycle costs [3].

The funding allocated for maintenance, repair, and rehabilitation is always limited. Therefore, it is necessary to prioritize and select the options that are best aligned with the asset managing organization's objectives, which, in the case of infrastructure, should also reflect the needs of society. The criteria used in this process are often unclear, conflicting, and sometimes subjective, including the type of maintenance intervention, risk and reliability, overall network performance, life cycle costs, desired level of service, budgetary concerns, and construction and social costs [4]. To achieve the best results at both individual and overall system levels, an optimal scheme for fund allocation to individual assets needs to be identified. This necessitates the simultaneous optimization of more than one objective while satisfying all of the necessary constraints [5].

Asset management encourages considering the trade-

offs between deferred maintenance and sustaining current pavement conditions and between short-term fixes and long-term solutions [6]. A well-developed model will enable maintenance managers to consider the impact of selecting one maintenance policy over another. Interventions applied too soon may add little incremental benefit. On the other hand, interventions applied too late may likely be ineffective [7]. It is hypothesized that there is a certain optimal level of performance in between these two extremes of profligacy and parsimony at which the intervention would yield maximum cost-effectiveness [8].

The main objectives of the study are to (1) develop a maintenance optimization framework using a non-dominated multi-objective evolutionary algorithm, non-dominated sorting genetic algorithm II (NSGA-II) and (2) develop an optimization framework that not only minimizes the cost over the lifespan of the highway assets but also maximizes the performance.

2 Literature Review

Asset Management. Highway asset management aims to secure and expand physical highway asset items belonging to transport facilities and sustain a certain level of service for its users. Researchers have proposed several optimization techniques to achieve optimal fund allocation for highway asset maintenance over the last two decades. [9] used an empirical index to combine seven project-level objectives using priority weights. [10] solved the single-objective budget allocation problem using priority weights of several assets based on the prevailing conditions. [11] developed a multi-asset optimization of roadway asset maintenance.

Multi-objective Optimization. Multi-objective optimization (MOO) algorithms are used for handling trade-offs among various objectives. Every multi-objective problem is unique and has the ability to incorporate various user preferences. It can thus prove to be an effective and versatile tool for decision making and can be used by highway agencies while allowing quick evaluation of several competing alternatives and performing a trade-off analysis [12]. Advanced methodologies based on a MOO formulation treat all the performance measures as additional merit objective functions that are not restrictive. As a result, the performance-based maintenance management methodologies lead to a group of non-dominated solutions, each of which represents a unique, optimized trade-off between a large set of alternative solutions.

In the past decade, a number of earnest attempts have been made to carry out the MOO for the purpose of asset management. [13] developed a cross-asset trade-off analysis based on multiple criteria for long-term highway investment. [14] developed a MOO method for bridge-

management investment decision analysis using utility theory.

Life Cycle Cost Analysis (LCCA). LCCA arranges for a framework to specify the projected total incremental cost of constructing, using, developing, and retiring a specific infrastructure project. The six phases in a product's life are: need recognition, design development, production, distribution, use, and disposal [15]. LCCA enables the pavement engineers to conduct a comprehensive assessment of long-term costs, and ideally, agency highway funding can be allocated more optimally. LCCA is applied in road construction to explore the possibility of more efficient investment. LCCA evaluates not only the initial construction cost of the pavement but also all the associated maintenance costs during its service life. Therefore, pavement engineers can choose the pavement type and design with the lowest cost in the long run [16].

Previously, [17] proposed best value award algorithms over low-bid initial costs to choose the pavements. [18] evaluated LCCA practices in the Michigan DOT.

Level of service (LOS). The LOS concept was first proposed in the Highway Capacity Manual (HCM) of version 1965 [19] and then defined by the six levels in relation to a number of traffic conditions in the HCM of version 1985 [20]. The current concept of LOS is applied in a six-level scale (levels of service A-F) that are distinguished in the current HCM by traffic density-the sole criterion used to differentiate between LOS A, LOS B, and LOS C and so on [21]. These measures of LOS used there such as traffic density and traffic flow rate are not the LOS itself but merely characteristics of traffic conditions.

[22] provided a framework to develop tools to reflect road user perception on service quality by defining the quality of service as a function of five performance measures- mobility, perception, of the lack of safety, environment, comfort and convenience, and road user direct cost. [23] determined motorists' views on what aspects of freeway travel are important to them and identified: travel time, density/maneuverability, road safety, and travel information were the most important factors.

3 Method

The primary objective of the proposed model is to minimize the maintenance cost throughout the life cycle of the pavement (LCC) and maximize the level of service (LOS) of the infrastructure. The decision variables are the type of maintenance strategies as shown in Table 2, and the optimization problem has two objectives of LCC and LOS. With this, we established a multi-objective optimization model that will be further discussed in the

next section. Since the study targets optimizing two objectives simultaneously, the research adopted a quantitative approach to investigate the benefits using a multi-objective optimization technique (MOO) for predictive maintenance. It must be acknowledged that the intent of the two objectives are incompatible with each other. Therefore, they are incorporated into the MOO algorithm to generate a set of Pareto optimal solutions that are consistent with the performance goals and resource constraints in a best-suited way while focusing on delivering the best possible results.

A heuristic non-dominated sorting-based multi-objective EA (MOEA), called non-dominated sorting genetic algorithm II (NSGA-II) is deployed. It considers a sustainable assessment of predictive maintenance alternatives and aims to improve the allocation of maintenance resources. NSGA-II has a fast-non-dominated sorting approach with $O(MN^2)$ computational complexity (where M is the number of objectives and N is the population size). Additionally, NSGA-II has a non-elitism approach and does not require specifying a sharing parameter that helps it find a diverse set of solutions and converge near the true Pareto-optimal set.

3.1 Genetic Algorithms for MOO

The Genetic Algorithms (GA's) are a type of heuristic algorithms that follow the survival of the fittest principle and are formulated loosely based on the Darwinian evolution. The search procedure of GAs involves generating an initial pool of feasible solutions that is generated randomly to form a parent solution pool, this is followed by obtaining new solutions and creating new parent pools through the iterative process. The entire iteration process consists of copying, modifying, and exchanging parts of the genetic representation in a pattern similar to the natural genetic evolution.

The solutions generated in the parent pool are evaluated by means of the objective function. The fitness value of each solution is used to determine its potential contribution to the generation of new solutions known as offspring. The next parent pool is formed by selection of the fittest offspring based on their fitness. The process is allowed to continue and repeat itself until the pre-determined stopping criterion is met based on the number of iterations or the magnitude of improvement of the generated solution [24].

3.2 Concept of Pareto Solutions

In the evaluation of a pool of solutions in the multi-objective genetic algorithm (MOGA), is a 2-D curve (for two-objective optimization) or a 3-D surface (for three or more multi-objective problems) which is composed of all the non-dominated solutions. This curve or surface is

known as the Pareto frontier. Each set of Pareto-optimal solutions represents a trade-off among different objectives. The Genetic Algorithm optimization process looks to produce new solutions that can give an improved frontier that dominates the existing frontier. A solution, in which a value of at least one objective is better than the rest of the solutions is known as a non-dominated solution. This process of producing new solutions repeatedly continues until a set of globally non-dominated solutions is found. This globally non-dominated set of solutions is called Pareto optimal set and defines the Pareto optimal front [25].

3.3 Optimization Objectives

The desired MOO model comprises two objectives which form the optimization algorithm in the model development. The two objectives consist of minimizing (1) Life-Cycle Cost (LCC), and maximizing (2) Level of Service (LOS).

3.3.1 Life-Cycle Cost

The total expected costs during the lifetime of a highway, as adopted from [26] is given as:

$$Ct = Cet + Cpm + Cins + Cf \quad (1)$$

Where Cet is the initial cost of construction, Cpm is the expected cost of maintenance, $Cins$ is the expected cost of inspections, and Cf is the expected failure cost—assuming the occurrence of the hazard (e.g., flood, earthquake). In the formulation of this research study, Cet is taken as 0 as the research concentrates on maintenance, not on construction. Additionally, $Cins$ and Cf are also excluded to simplify the problem as they are constants.

Cost of predictive maintenance is calculated as:

$$C_{pm} = \sum_{i=1}^{40} \sum_{j=0}^3 M_{ij} \quad (2)$$

C_{pm} = cost of predictive maintenance

i = number of years (from 1 to 40), j = maintenance types 0, 1, 2, and 3.

M = cost of maintenance associated with maintenance activity for each year.

The total cost is calculated in terms of Present Value (PV), the formula for which as adopted from [27] is:

$$PV = FV \left[\frac{1}{(1+r)^n} \right] \quad (3)$$

Where, PV = present value, FV = future value, r = discount rate = 3%, $n = n^{th}$ year.

3.3.2 Level of Service

The level of service concept in the Highway Capacity Manual (HCM) is used as a qualitative measure representing freeway operational conditions. Many different quantifiable performance measures are currently included in the FHWA Highway Planning and Monitoring Systems (HPMS) database to determine the pavement level of service. In this study, Pavement Condition Rating (PCR) is used to quantify performance of the pavement. The PCR is a composite index (marked on a scale of 0 to 100) derived from monitoring data-pavement roughness and distress rating. Several studies have approached the performance prediction of highways assets [28]. In this study the performance prediction equation for flexible pavement as adapted from [29] is presented by the equation:

$$PCR(t) = 90 - a[\exp \exp (Age^b) - 1] \log \left[\frac{ESAL}{SNC^c} \right] \quad (4)$$

Where,

$$a = 0.6349; b = 0.4203; c = 2.7062$$

$PCR(t)$ = pavement condition rating at time t (in years).

$ESAL$ = traffic volume and weight, which are expressed in terms of yearly equivalent single-axle loads.

SNC = Strength and condition of pavement structure represented by modified structural number.

$$SNC = \sum aihi + SNg \quad (5)$$

Where,

ai = material layer coefficients,

hi = layer thickness (in.),

SNg = subgrade contribution, and $= 3.51 \log CBR - 0.85 (\log CBR)^2 - 1.43 R^2 = 0.75$

CBR = California bearing ratio of subgrade (percent)

3.4 Model Implementation

A graphical representation of the step-by-step procedure for the algorithm's functionality is shown in Figure 1.

3.4.1 Mathematical Formulation of the Model

In this formulation, the pavement condition is considered as a representative of LOS for the highway, which is depicted in terms of PCR which is an ASTM standard for the pavement condition assessment. PCR values are allotted to a scale that ranges from 0 to 100. The PCR for a pavement section at any given time t is

computed as follows:

$$PCR(t) = 90 - a[\exp \exp (Age^b) - 1] \log \left[\frac{ESAL}{SNC^c} \right] \quad (6)$$

The PCR varies from 100 for a perfect pavement condition to 0 for a near failing condition. The optimization model can be represented mathematically as the following equations:

Objective functions:

1. Maximize the average PCR over the design life of the pavement:

$$\text{Maximize } \sum_{t=1}^N PCR_t / N$$

2. Minimize the maintenance cost

$$\text{Minimize } \sum_{j=1}^N C_j$$

Subjected to: $PCR_t \geq \alpha_1 \quad j = 1, 2, N$

Where, N = total number of years = 40; C_j = maintenance cost for pavement section j ; PCR_t signifies the PCR of the pavement section at time t ; and α_1 is the minimum pavement condition threshold for the pavement section (set at 40).

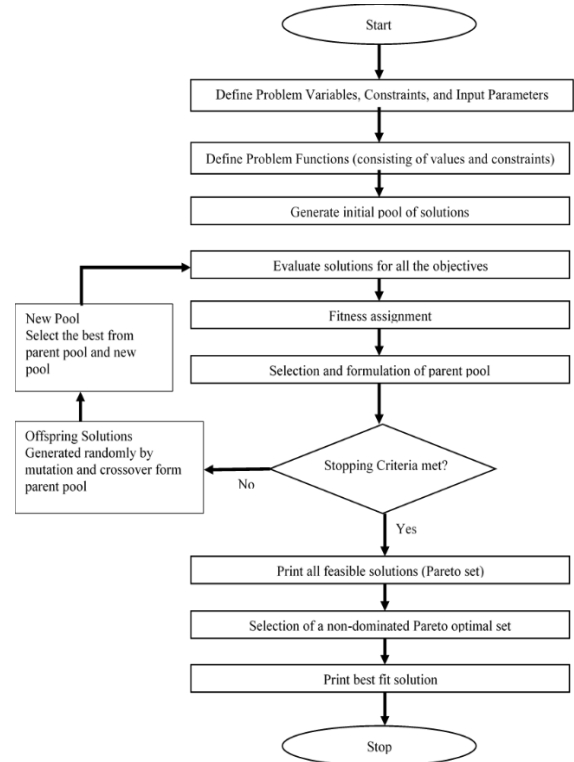


Figure 1. Genetic Algorithm Analysis for Multi-Objective Optimization.

Solution to the objective functions mentioned above will provide a family of Pareto optimal solutions. Each

solution gives the optimal maintenance program and the resultant amount of maintenance cost for corresponding values of average PCR.

4 Result and Discussion

4.1 Application of the model on a Case Study

The developed framework for determining optimal PM technique for the maximum LOS and minimum LCC is illustrated using a case study involving a section of the California Department of Transportation Asset Management Plan (TAMP). The data presented below is adapted from table 4-2 of the California Transportation Asset Management Plan- Fiscal Years 2017/18-2026/27, published in January 2018. The table below depicts a typical example of life cycle treatments (for a 40-year design) for pavements of class I in average climate conditions.

Table 1. Data depicting types of maintenance, schedule and cost in \$/lane mile adopted from [30]

Treatment	Schedule (Years)	Cost (\$/lane mile)	Present Value (\$/lane mile)
Seal surface	4	6000	5129
Thin mill & overlay	8	152000	111065
Seal surface	12	6000	3748
Thin mill & overlay	16	152000	81154
Seal surface	20	6000	2738
Thin mill & overlay	24	152000	59298
Seal surface	28	6000	2001
Thin mill & overlay	32	170000	48460
Dig-out, crack seal, & seal surface	36	76000	18519
Medium overlay	40	325000	67690
Net Present Value \$399,806			

Here, the three major types of maintenance activities are: 1) seal surface, 2) thin mill & overlay, and 3) dig-out, crack seal, and seal surface. The schedule of the type of maintenance activity to be performed is pre-determined on certain interval of years. Furthermore, the cost involved with each kind of maintenance activity is defined in terms of cost \$/lane mile. The total net present

value for the entire 40 years of design life of the pavement system sums up to \$399,806 per lane mile.

Table 2. Data for the case study adopted from [31]

Maintenance	Treatment	Gain in PCR	Budget (\$)
M0	None	0	0
M1	Seal surface	10	6000
M2	Thin mill/Overlay	35	152000
M3	Dig-out, crack seal, seal surface	20	76000

The data depicted in table 2 is applied to the GA NSGAI MOO algorithm. The maintenance activities, treatment type, and cost is obtained from California TAMP (2018). The algorithm was programmed to run 10,000 times. The minimum performance threshold of 40 PCR was selected. The results obtained after the accomplishment of the stopping criteria were imported into a matrix format, and all the feasible, non-dominated optimized solutions were printed. Moreover, to better illustrate the solutions, graphs are plotted for each solution, Table 3 and Figure 2 illustrate the various solutions and graphs.

Table 3. Depiction of the various solution sets (S1 to S16) for 40 years of maintenance

Year	SOLUTIONS															
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	35	35	35	10	35	10	20	35	10	10	10	10	0	0	0
7	10	10	10	35	20	10	10	10	20	10	10	10	10	10	10	10
8	0	20	10	10	10	35	10	10	10	10	10	10	10	10	10	10
9	10	20	20	10	10	10	10	0	10	0	20	20	35	0	0	0
10	0	20	10	35	10	35	0	10	10	35	0	0	10	0	0	0
11	0	35	20	20	35	10	10	10	10	10	35	10	10	0	0	0
12	10	35	35	10	35	35	10	0	10	10	10	0	10	10	10	10
13	10	10	10	20	10	10	10	10	10	10	0	10	0	10	10	20
14	10	10	35	20	10	10	10	10	10	10	10	10	10	10	10	10
15	0	35	10	35	10	0	10	0	10	10	10	10	10	0	0	0
16	10	10	35	20	20	10	10	10	10	10	10	10	20	10	10	10
17	0	10	20	10	20	10	10	10	35	10	10	0	0	0	0	0
18	10	35	0	10	20	10	10	20	35	10	0	10	10	0	0	10
19	0	10	35	10	10	20	10	20	20	10	10	0	10	0	10	0
20	0	10	10	35	0	10	10	10	10	10	10	0	35	20	0	0
21	0	35	20	10	10	10	20	0	20	10	10	10	0	10	10	10
22	10	35	10	10	10	10	10	10	10	10	10	0	10	10	10	10
23	10	35	35	10	10	10	0	0	20	10	10	10	10	10	10	10
24	10	10	10	20	10	10	10	10	10	10	10	10	10	10	20	20
25	0	0	35	20	10	20	10	10	35	20	10	10	10	0	0	10
26	10	20	20	10	10	10	0	10	10	10	20	0	10	0	10	10
27	10	10	20	10	20	35	0	10	10	10	0	0	10	0	10	10
28	0	35	0	10	10	35	10	0	20	35	0	0	10	0	0	0
29	0	35	20	10	0	0	20	0	20	10	35	0	35	0	0	10
30	10	10	10	0	10	10	10	0	10	0	10	0	0	10	10	10
31	10	10	35	35	10	10	10	0	20	10	10	10	10	20	0	0
32	0	20	10	10	0	10	0	20	0	10	10	10	0	0	10	0
33	0	20	10	10	20	20	10	10	20	0	0	0	0	0	10	10
34	0	10	20	0	10	0	10	0	10	0	10	10	0	10	10	10
35	0	35	10	0	0	20	35	10	35	20	10	10	10	0	10	10
36	0	10	10	10	10	0	10	0	0	10	20	35	10	0	0	0
37	0	0	10	10	0	10	10	20	0	10	10	0	10	0	0	0
38	20	0	10	10	0	0	10	0	0	10	0	10	0	0	0	0
39	0	20	10	0	35	10	0	20	0	0	0	10	10	10	35	20
40	0	0	0	0	10	0	10	0	0	0	10	10	10	0	0	0
PCR	73	90	89	88	85	87	81	79	86	84	82	77	83	74	75	76
NPV	7	129	110	94	59	80	24	23	71	48	36	17	39	12	14	17

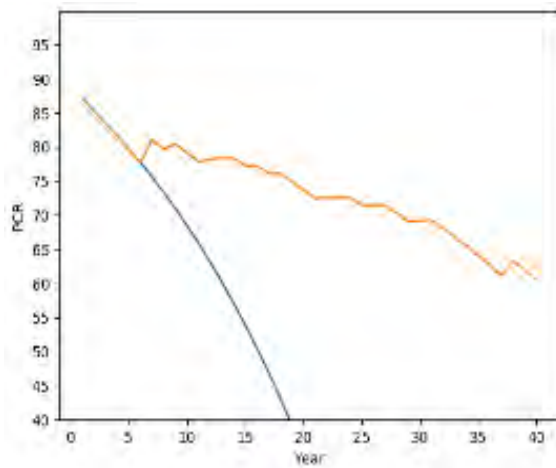


Figure 2. Graph for the S1 solution

Figure 2, gives the graphical representation of the solution 1 (S1) obtained from the algorithm. Similar plots could be made for solution S2 to S16. Each set of the solution gives information about the maintenance activity carried out for a time span of 40 years, under the constraint that at no point of time, the average PCR value dropped below 40. Every set of solution gives detailed data for the gain in value of PCR for the pavement each year, from this data the type of maintenance activity carried out each year can also be found out. The cost is calculated on the basis of net present value (NPV) and depicted (rounded off) to the nearest \$10,000 value at the bottom of Table 3. Depicted in Fig. 3 is a graphical representation of the Pareto set of solutions obtained.

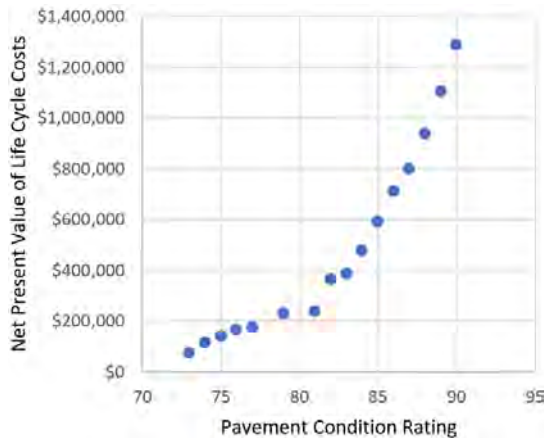


Figure 3. Pareto front obtained via the convergence of the two objectives.

The Pareto set of non-dominated feasible solutions for the two objectives contained 16 set of solutions. The Pareto front obtained covers a vast range of PCR values along with a range of costs associated with them. The range of PCR varies from 73 (lowest) to an excellent

value of 90 (i.e. the highest). Subsequently, the costs associated with the various sets of maintenance solutions range from \$73,196 to \$1,287,603 in terms of present value for a total span of forty years.

A brief comparative summary for the results of the case study has been depicted below in table 4.

Table 4. Summary of the case study result

Metrics	Optimized Solution	Caltrans Report
Total Solutions	16	1
Time-frame	40	40
Constraint	1 (min PCR 40)	None
Objectives	2	1
Test run for algorithm	1000	N/A
Minimum cost (NPV)	\$73196 with PCR of 73	\$399806
Maximum cost (NPV)	\$1287603 with PCR of 90	Unknown
Total 9/16 optimized solutions outperformed the Caltrans estimations in terms of \$ cost		

4.2 Sensitivity Analysis

A sensitivity analysis is conducted by altering the values of benefits in the value of PCR for the various maintenance interventions. The range of the benefits in terms of PCR is associated with the different maintenance types. The types of distress, treatment, and the budget associated with the interventions are explained in table 1 and table 2 of section 4.1 in this research study. The range of the PCR that was taken is depicted in Table 5:

Table 5. Maintenance type (M1, M2, & M3) PCR ranges for sensitivity analysis adopted from [33]

M1 (PCR)	M2 (PCR)	M3 (PCR)
8	36	18
12	38	22
		25

The algorithm was run for 10,000 times at a minimum threshold of a PCR equal to 40 throughout for 40 years. The minimum PCR refers to a threshold defined by a decision maker and represents a level of service that

is considered acceptable to a transportation agency. In future, the decision makers can set a different threshold value for the PCR and the model can provide an alternative set of optimized solutions. The results obtained were put into a .CSV file and a graph was plotted using Microsoft Excel 2016. The various combinations associated with the maintenance types and benefits in PCR have been depicted in the figure 4 below.

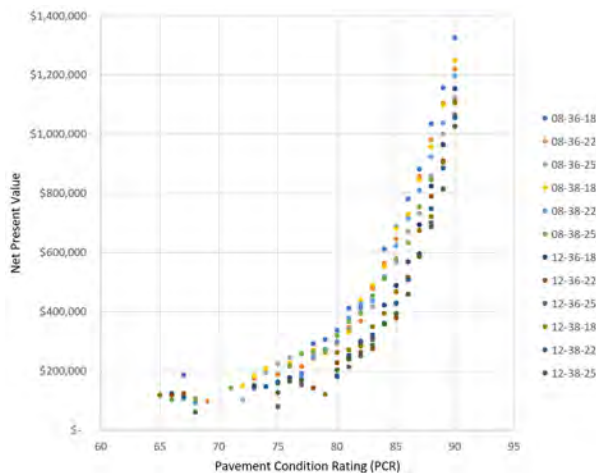


Figure 4. Sensitivity analysis of the Pareto front

It is observable that a wide range of Pareto optimal front is obtained with the various combinations of the benefits in PCR associated with the maintenance types. Through the analysis, it was obvious that a trend followed, that gave better convergence of the Pareto-front with increase in the values of the benefits in the PCR. The minimum value of PCR observed is 63 while the maximum value soared at 90. The minimum cost in terms of NPV was observed to be \$80,447 while the maximum value NPV was documented as \$1325882.

5 Conclusion

The results suggest that the MOEA NSGA-II is capable of tracking the overall pavement performance as well as the maintenance costs much more efficiently as compared to the conventional maintenance techniques.

The algorithm is tested for its effectiveness with the help of data from California Department of Transportation Asset Management Plan. The real data was fit into the model and the program was run for 10000 times to test for solutions that can be compared with the data from the Caltrans report. Based on the constraints and limitations, 16 feasible, non-dominated, optimized solutions were generated on the Pareto-front. The Pareto-front covers a large range of PCR values ranging from 73 to 90 and costs ranging from \$73,196 to \$1,287,603 respectively. Overall, 9 solutions are obtained across the Pareto front that outperformed the Caltrans estimation for LCC on the basis of cost alone. Moreover, the PCR value

in the optimization model has a lower limit constraint for average PCR at 40, while there is no data available on the level of service for the pavement condition in the Caltrans data.

In conclusion, the solutions obtained from the MOO are far superior in comparison with the conventional time-based maintenance warrants used by the California DOT. The Pareto-front gives a wide range of flexibility with the option of trade-offs between the two objectives. The pavement is the only asset item that is being considered in this research study and its formulation whereas in the future, more asset items can be incorporated in the algorithm to acquire a more holistic approach towards the predictive maintenance optimization problem. Moreover, the approach can be applied to more research cases in addition to California DOT data to test the model and its efficiency. Application of the optimization model to a more complex highway asset management scenario with the possibility of multiple intermediate maintenance actions will give a better result that will help in decision-making with greater confidence.

References

- [1] ASCE. (2017). 2017 Report Card for America's Infrastructure. (2017), 111.
- [2] Karimzadeh A, Sabeti S, Shoghli O. Optimal clustering of pavement segments using K-prototype algorithm in a high-dimensional mixed feature space. *Journal of Management in Engineering*. 2021 Jul 1;37(4):04021022.
- [3] Chasey, A. D. (1997). A framework for determining the impact of deferred maintenance and/or obsolescence of a highway system. *Transportation Research Part A*, 1(31), 68.
- [4] Šelih, J., Kne, A., Srdić, A., & Žura, M. (2008). Multiple-criteria decision support system in highway infrastructure management. *Transport*, 23(4), 299-305.
- [5] Fwa, T., & Farhan, J. (2012). Optimal multi-asset maintenance budget allocation in highway asset management. *Journal of Transportation Engineering*, 138(10), 1179-1187.
- [6] Dornan, D. (2001). Accounting Rule Fixes Roads. *ENR*, 246(13), 55-55.
- [7] Labi, S., & Sinha, K. C. (2005). Life-cycle evaluation of flexible pavement preventive maintenance. *Journal of Transportation Engineering*, 131(10), 744-751.
- [8] Khurshid, M. B., Irfan, M., & Labi, S. (2010).

- Optimal performance threshold determination for highway asset interventions: Analytical framework and application. *Journal of Transportation Engineering*, 137(2), 128-139.
- [9] Small, E. P., & Swisher, M. (2000). Integration of bridge and pavement management systems: A proposed strategy for asset management. *Transportation Research Circular*, 498, 8th.
- [10] Sadek, A. W., Kvasnak, A., & Segale, J. (2003). Integrated infrastructure management systems: Small urban area's experience. *Journal of Infrastructure Systems*, 9(3), 98-106.
- [11] Shoghli, O., & De La Garza, J. M. (2017). Multi-Asset Optimization of Roadways Asset Maintenance. In *ASCE International Workshop on Computing in Civil Engineering 2017* American Society of Civil Engineers.
- [12] Jha, M. K., & Maji, A. (2007). *A multi-objective genetic algorithm for optimizing highway alignments*. Paper presented at the Computational Intelligence in Multicriteria Decision Making, IEEE
- [13] Mrawira, D., & Amador, L. (2009). Cross-assets trade-off analysis: Why are we still talking about it? Paper presented at the 88th Annual Meeting of the Transportation Research Board, Washington, DC.
- [14] Patidar, V., Labi, S., Sinha, K. C., & Thompson, P. (2007). NCHRP Report 590: Multi-objective optimization for bridge management systems. *Transportation Research Board of the National Academies*, Washington, DC.
- [15] Alting, L. (1993). Life-cycle design of products: a new opportunity for manufacturing enterprises. *Concurrent engineering*, 1-17.
- [16] Chan, A., Keoleian, G., & Gabler, E. (2008). Evaluation of life-cycle cost analysis practices used by the Michigan Department of Transportation. *Journal of Transportation Engineering*, 134(6).
- [17] Gransberg, D. D., & Molenaar, K. R. (2004). Life-cycle cost award algorithms for design/build highway pavement projects. *Journal of Infrastructure Systems*, 10(4), 167-175.
- [18] Chan, A., Keoleian, G., & Gabler, E. (2008). Evaluation of life-cycle cost analysis practices used by the Michigan Department of Transportation. *Journal of Transportation Engineering*, 134(6).
- [19] Manual, H. C. (1965). *TRB Special Report N 87*. TRB, Washington, DC, 398.
- [20] Manual, H. C. (1985). *Special report 209*. Transportation Research Board, Washington, DC, 1
- [21] Choocharukul, K., Sinha, K. C., & Mannering, F. L. (2004). User perceptions and engineering definitions of highway level of service: an exploratory statistical comparison. *Transportation Research Part A: Policy and Practice*, 38(9-10), 677-689.
- [22] Pfefer, R. C. (1999). Toward reflecting public perception of quality of service in planning, designing, and operating highway facilities. *Transportation Research Record*, 1685(1), 81-89.
- [23] Hall, F., Wakefield, S., & Al-Kaisy, A. (2001). Freeway quality of service: What really matters to drivers and passengers? *Transportation Research Record*, 1776(1), 17-23.
- [24] Fwa, T., Chan, W., & Hoque, K. (2000). Multi-objective optimization for pavement maintenance programming. *Journal of Transportation Engineering*, 126(5), 367-374.
- [25] Yang, N., Kang, M.-W., Schonfeld, P., & Jha, M. K. (2014). Multi-objective highway alignment optimization incorporating preference information. *Transportation Research Part C: Emerging Technologies*, 40, 36-48.
- [26] D. M., & Liu, M. (2004). Life-cycle cost analysis for highways bridges: Accomplishments and challenges *Structures 2004: Building on the Past, Securing the Future* (pp. 1-9).
- [27] Beaves, R. G. (1993). The Case for a Generalized Net Present Value Formula. *The Engineering Economist*, 38(2), 119-133.
- [28] Karimzadeh A, Shoghli O, Sabeti S, Tabkhi H. Multi-Asset Defect Hotspot Prediction for Highway Maintenance Management: A Risk-Based Machine Learning Approach. *Sustainability*. 2022 Apr 21;14(9):4979.
- [29] George, K., Rajagopal, A., & Lim, L. (1989). Models for predicting pavement deterioration. *Transportation Research Record*(1215).
- [30] TAMP, C. (2018). *California Transportation Asset Management Plan*. 126.
- [31] Rajagopal, A., & George, K. (1991). Pavement maintenance effectiveness. *Transportation Research Record*, 1276, 62-68.
- [32] Mayo, G., Shoghli, O., & Morgan, T. (2020). Investigating efficiency utilizing data envelopment analysis: case study of shipyards. *Journal of Infrastructure Systems*, 26(2), 04020013.
- [33] George, K., Rajagopal, A., & Lim, L. (1989). Models for predicting pavement deterioration. *Transportation Research Record*(1215).

Weakly supervised pseudo label generation for construction vehicle segmentation

W.-C. Chern¹, V. Asari¹, T. Nguyen², and H. Kim³

¹Department of Electrical and Computer Engineering, University of Dayton, U.S.A

²Department of Computer Science, University of Dayton, U.S.A

³Department of Civil and Environmental Engineering, Yonsei University, Republic of Korea

chernw1@udayton.edu, vasari1@udayton.edu, tnguyen1@udayton.edu, hongjo@yonsei.ac.kr,

Abstract -

Segmentation tasks in computer vision have been adopted in various studies in the civil engineering domain to provide accurate object locations in images. However, preparing annotation to train segmentation models is a time consuming and costly process, which hinders the use of segmentation models in vision-based applications. To address the problem, this study proposes a fusion model integrating self-supervised equivariant attention mechanism (SEAM) and sub-category exploration (SC-CAM) to generate pseudo labels in the form of polygon annotation from bounding box annotation that is relatively easy to obtain. To test the performance of the fusion model, a public data set—Advanced Infrastructure Management Group (AIM) dataset—for construction object detection was selected to generate pseudo labels; the effectiveness of pseudo labels was measured by the segmentation performance of a feature pyramid network (FPN) trained with the pseudo labels. FPN showed the mean intersection over union (mIoU) score of 86.03%, demonstrating the potential of the proposed fusion model to reduce the manual annotation efforts in preparing training data for segmentation models.

Keywords -

Weakly supervised learning; Semantic segmentation; Pseudo labels; Training data preparation

1 Introduction

Semantic segmentation is an important task in vision-based analysis, providing pixel-level annotation to represent exact object boundaries. There are various applications of segmentation tasks in civil engineering domain such as construction site monitoring and infrastructure damage assessment.

The quantity and quality of training data greatly affects the performance of semantic segmentation models. The time spent for annotation shows a positive correlation to semantic segmentation's performance [1]. This is a general phenomenon in deep learning applications which require a lot of training data to acquire better performance and

generalization capability. However, polygon annotation for semantic segmentation can be overly time consuming in comparison to preparing annotation for object detection and classification tasks. To minimize the time and cost of annotation, weakly supervised learning can be used. In weakly supervised learning, imperfect data are used to train segmentation models. For example, segmentation models can be trained with bounding box annotation by treating the entire box as a target class region. Likewise, segmentation masks of an object of interest can be generated from images with class activation mapping (CAM) to generate pseudo labels for training segmentation models.

CAM often represents a discriminative part of the object only, instead of localizing the entire object, resulting in producing incomplete pseudo labels. Previous studies have utilized consistency regularization on CAMs [2], and a self-supervised task [3] to generate accurate pseudo labels for the Pascal Visual Object Classes Challenges 2012 (VOC2012) [4]. Nevertheless, there has been a lack of detailed investigation on difficult datasets in which the target classes have visual similarity. VOC2012 contains 20 distinctive object classes from animals, transportation vehicles, furniture, and etc. Thus, it presents clear visual differences between target classes for pseudo label generation. On the other hand, a dataset contains 20 different vehicle brands can be more challenging to generate pseudo labels covering entire vehicles. Because models tend to focus on the most distinctive part of vehicles such as the logo of vehicles rather than captures the entire body. This issue can also be found in the AIM dataset having construction machines only. To address the challenge, this study propose a novel architecture integrating SEAM and SC-CAM features to generate accurate pseudo labels.

For experiment, the AIM dataset is used, which contains five object classes—dump trucks, excavators, loaders, mixer trucks, and rollers—with a total of 2,721 images and 2,873 vehicle instances as shown in Table 1. Examples of the AIM dataset is shown in Figure 1. AIM dataset's construction vehicle classes share a considerable degree of visual similarity. For example, wheels and front windows between mixer trucks and dump trucks look similar;



Figure 1. Examples of AIM dataset. The construction vehicles from top to bottom and left to right are: loader, mixertruck, dumptruck, excavator, and roller.

Table 1. AIM Dataset Statistics.	
Vehicle Types	Number of Instances
Dumptruck	762
Excavator	413
Loader	714
Mixertruck	632
Roller	352
Sum	2,873

wheels between loaders and excavators show similar appearances. Construction vehicles from the AIM dataset also possess highly distinct features, such as the boom of excavators, the drum of rollers, the mixing drum of mixer trucks, the dump box of dump trucks, and the bucket of loaders, which can lead to over-attention to the distinctive part in CAM. This can result in low-quality pseudo labels for semantic segmentation as the trained classification models may not learn the entire appearance of objects for classification. Instead, the models only pay attention to the most discriminative features from the target objects, resulting in poor pseudo labels which do not cover entire target objects. It was found that the fusion model of self-supervised equivariant attention mechanism (SEAM) [2] and weakly-supervised semantic segmentation via sub-category exploration (SC-CAM) [3] with random erase data augmentation can mitigate this issue to improve the quality of the pseudo labels. The experimental results showed that the feature pyramid network (FPN) [5] for semantic segmentation model recorded the mIoU score of 86.03%.

This study propose a novel model combining two different architectures (SEAM & SC-CAM) along with the data augmentation of random erase [6] to improve the quality of the pseudo labels as shown in Figure 2.

2 Methodology

The proposed fusion model incorporates methods that encourage models to pay attention to more foreground regions for better pseudo labels from SEAM and SC-CAM as shown in Figure 2. SEAM architecture modifies the self-attention module for weakly supervised pseudo labels generation. SC-CAM approaches the same goal by asking models to distinguish differences within the same class, asking models to look at the entire objects for sub-category classification. A common way to generate pseudo labels for semantic segmentation is to use classification models and their CAM [7]. CAM visualizes the attentions of classification models, and it highlights the visual features contributed to classification. However, CAM itself is hardly used as training data for semantic segmentation as the pseudo labels usually cover the most discriminative features only as shown in Figure 3. mIoU is used as the performance metric to evaluate the performance of segmentation models as it measures the degree of overlap between ground truth and predicted masks across all target classes. It is formulated as follows:

$$IoU = \frac{g_t \cap p_t}{g_t \cup p_t}, \quad (1)$$

$$mIoU = \sum_{n=1}^C \frac{IoU_n}{C}, \quad (2)$$

where g_t represents ground truth masks, p_t represents predicted masks, and mIoU is an average of IoU scores from all target classes.

2.1 Self-supervised Equivariant Attention Mechanism

SEAM architecture proposed a pixel correlation module (PCM) and a smaller-scaled branch siamese network to teach classification models to pay more attention to the entire region of target objects. PCM adopts the concept of self-attention [8] to extract contextual information. It takes features maps from two convolutional blocks and the RGB image to form the self-attention map that is then used to correct original CAM. The formula of PCM can be formulated as:

$$y'_i = ReLU\left(\frac{\sum_{\forall j} e^{\theta(x_i)^T \theta(x_j)} \cdot \phi(y_j)}{\sum_{\forall j} e^{\theta(x_i)^T \theta(x_j)}}\right), \quad (3)$$

where x represents the feature maps input, y' represents the refined CAM, y represents the original CAM, θ , and ϕ are two 1x1 convolutions as embedding functions. The refined CAM further normalized by the sum of the attention map and a ReLU function.

Wang et al. [2] also discovered a unique phenomenon of CAM that classification models' attention on target objects

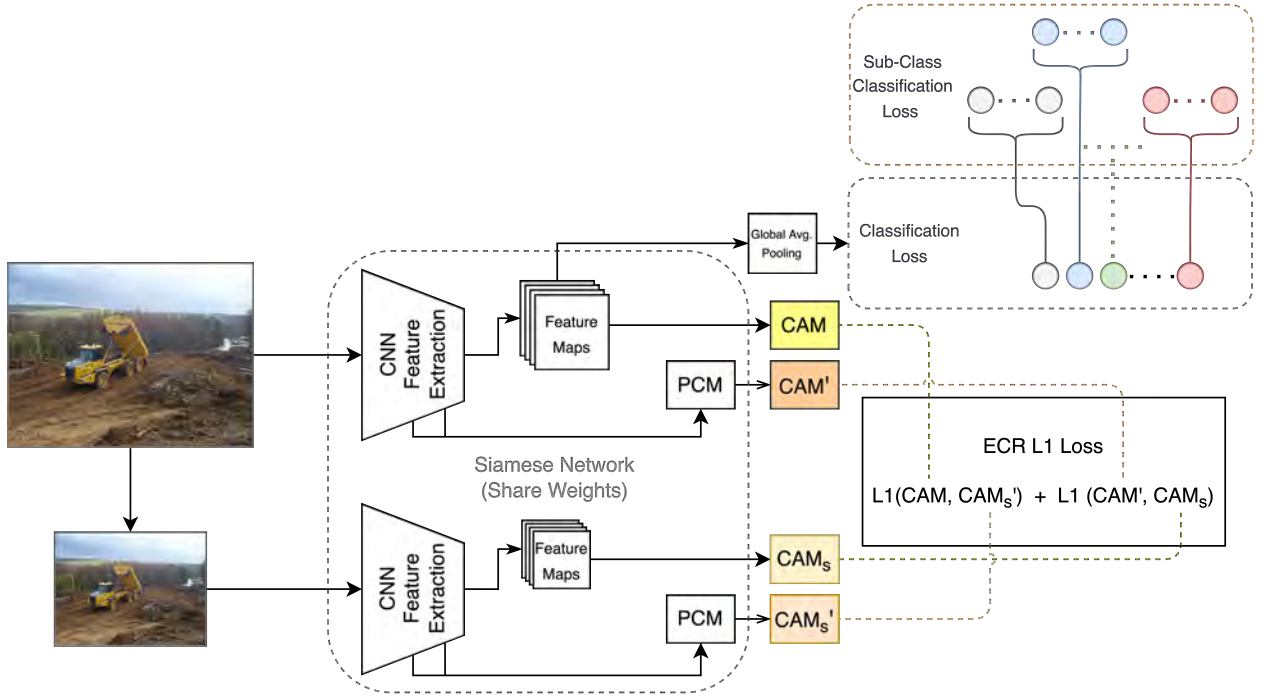


Figure 2. An overview of SEAM + SC-CAM architecture for weakly supervised pseudo labels generation for semantic segmentation.



Figure 3. Examples of models paying attention to the most discriminative features. The construction vehicles from top to bottom and left to right are: roller, mixer truck, excavator, and dump truck

is affected by an input resolution. That is, as input resolution is reduced, the output of its corresponding CAM tends to cover the entire foreground regions more. As the result, SEAM creates a siamese network that takes two input images—an original-scale image, and a down-scaled image. The two images generate two sets of CAM. The two sets including the original CAM and refined CAM

are then regulated by an equivariant cross regularization (ECR) function which is used as a loss function. This function is formed by two L1 norms as follow:

$$ECR = ||CAM - CAM'_s||_1 + ||CAM' - CAM_s||_1, \quad (4)$$

where CAM represents outputs from inputs of the original scale, CAM' represents outputs refined by PCM, CAM_s represents output from inputs of the smaller scale, and CAM'_s represents output refined by PCM from inputs of the smaller scale. ECR uses CAM and CAM_s as ground truth to guide CAM' and CAM'_s (output by PCM), because CAM and CAM_s are guided by the classification labels. In another words, CAM and CAM_s are used to optimize PCM using L1 functions to extract contextual features.

2.2 Sub-Category Exploration

In addition to use CAM from downscaled inputs to guide the original inputs' CAM, this study adopts an idea of SC-CAM called sub-category exploration which also encourage classification models to pay more attention to the entire foreground regions. Instead of training a regular classification model with CAM to visualize and output the features as pseudo labels, SC-CAM also cluster each target class into K sub-clusters for training as follow:

$$\text{Class} = \{C_1, \dots, C_N\}, \quad (5)$$

$$\text{Class}_K = \{C_{11}, C_{12}, \dots, C_{1K}, \dots, C_{N1}, C_{N2}, \dots, C_{NK}\}, \quad (6)$$

where N represent number of classes, K represents the number of sub-cluster to each class. The total number of classes for sub-cluster will be $N \times K$ classes. The sub-class clustering is conducted by K-Mean clustering after images were encoded into feature vectors by a pre-trained ResNet model.

As the result, models are trained with both one-hot labels by Equation 5 and 6. With sub-category during training, models not only learn to recognize differences between original target classes, but also forced to pay more attention to the entire images to distinguish differences between each sub-cluster as shown in Figure 2 at the top-right corner.

2.3 Post-Processing Step

Although the fusion model of SEAM and SC-CAM can improve the quality of CAM, it may still fail on covering some foreground regions. This study follows SEAM and SC-CAM's post-processing procedure of applying dense condition random field (CRF) [9] to the output CAMs.

CRF is able to improve the CAM results from the fusion model as shown in the Figure 4. The output of CRF will then be used as pseudo labels to train FPN models.

3 Experimental Results

In this study VOC2012 was combined with the AIM dataset to increase the visual diversity of the dataset. However, the segmentation performance was only evaluated on the AIM dataset. The complete experiment steps are as follow:

1. Train the fusion model as shown in Figure 2 with classification labels from VOC12 & AIM datasets.

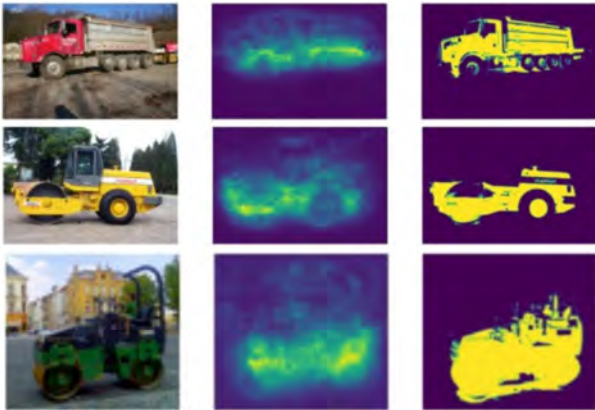


Figure 4. Examples of applying CRF to the fusion model's CAM results. From left to right column: RGB images, CAMs, and CRF results.

2. Generate pseudo labels as shown in the third column in Figure 4 for the AIM dataset only.
3. Train the FPN semantic segmentation model with the generated pseudo labels.
4. Evaluate FPN's mIoU score with human-annotated ground truth segmentation labels of the AIM dataset.

The prediction results of the FPN trained with pseudo labels generated from the fusion model is shown in Figure 5. The results demonstrate that the trained FPN model can capture boundaries of loaders, rollers, and dumptrucks. The mIoU score for the FPN model trained with human-labeled ground truth is also shown as the baseline for performance comparison as shown in Table 2. The proposed method achieves mIoU of 86.03%. This result indicates that segmentation models can be trained without human-annotated segmentation masks.

Table 2. Performance of the trained FPN models. (Second column) The FPN trained with pseudo labels from the fusion model. (Third column) The FPN trained with ground truth labels.

Object Class	P.L. Scores	G.T. Scores
Background	80.78	97.55
Dumptruck	87.19	97.55
Excavator	83.44	97.57
Loader	83.72	97.4
Mixertruck	92.71	99.5
Roller	88.39	99.1
mIoU	86.03	98.11

4 Conclusion

The study aims to increase the quality of semantic segmentation pseudo labels in a weakly-supervised manner with classification labels only. To achieve this, this study proposed a fusion model integrating SEAM and SC-CAM to generate pseudo labels and conducts experiments with the AIM dataset. The proposed method demonstrates the ability to generate effective pseudo labels for semantic segmentation models. The FPN model trained with the pseudo labels achieved the mIoU score of 86.03%.

A potential improvement for better pseudo label quality can be made through the change of loss function to the ECR function as shown in Equation 4. ECR contains two L1 norm functions to regulate PCM for extracting contextual features which can generate better pseudo label quality. However, L1 norm function is a distribution-based function where the error is accumulated from all CAM pixels. Thus, the L1 loss function can be dominated by the classes which contain more instances (data imbalance) such as the dump truck class as shown in Table 1. Region-based loss functions such as Jaccard loss [10], or Dice loss [11] can be used to replace the L1 norm functions in ECR

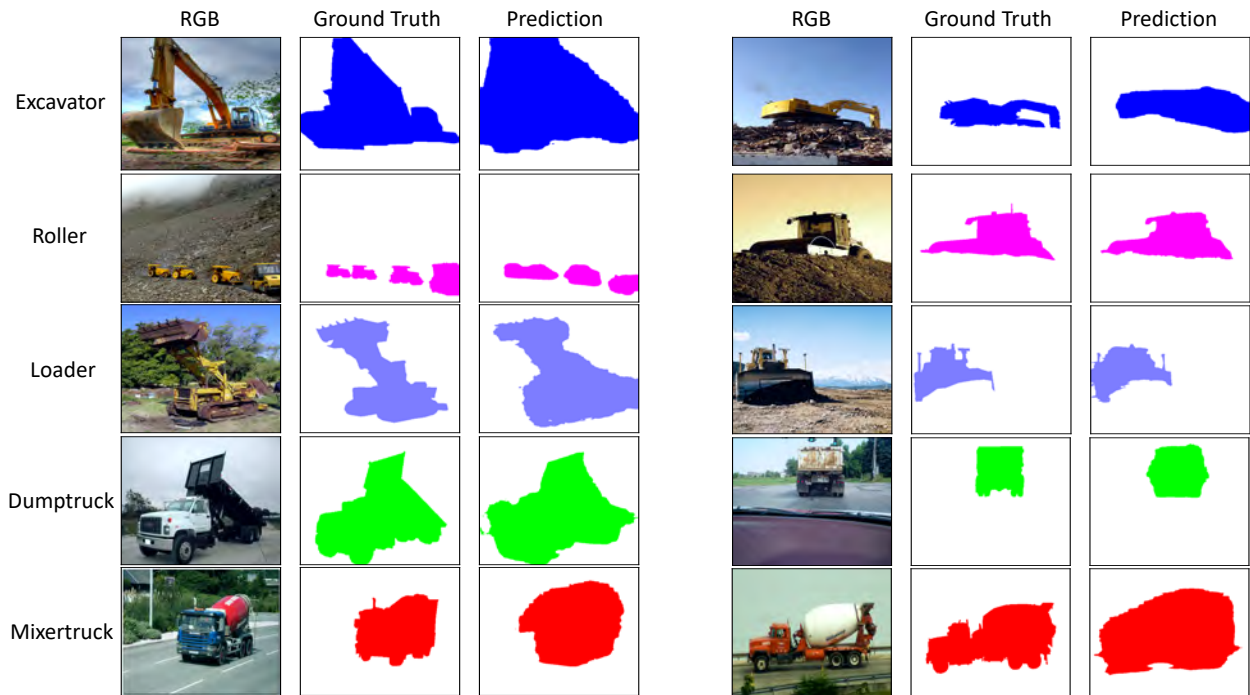


Figure 5. Prediction results of the FPN model trained by pseudo labels.

function as region-based loss functions are commonly used in segmentation tasks.

In addition, Zlateski et al. [1] points out that mixing a small amount of human-labeled fine annotations with a majority of coarse annotations can reach similar performances as using fine annotation. Therefore, it is expected the performance will be better if a small amount of fine annotation is included.

Acknowledgment

This research was conducted with the support of the “2021 Yonsei University Future-Leading Research Initiative (No. 2021-22-0037)” and the “National R&D Project for Smart Construction Technology (No. 22SMIP-A158708-03)” funded by the Korea Agency for Infrastructure Technology Advancement under the Ministry of Land, Infrastructure and Transport, and managed by the Korea Expressway Corporation.

References

- [1] Aleksandar Zlateski, Ronnachai Jaroensri, Prafull Sharma, and Fredo Durand. On the importance of label quality for semantic segmentation. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1479–1487, 2018. doi:10.1109/CVPR.2018.00160.
- [2] Y. Wang, J. Zhang, M. Kan, S. Shan, and X. Chen. Self-supervised equivariant attention mechanism for weakly supervised semantic segmentation. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12272–12281. IEEE Computer Society, 2020. doi:10.1109/CVPR42600.2020.01229.
- [3] Yu-Ting Chang, Qiaosong Wang, Wei-Chih Hung, Robinson Piramuthu, Yi-Hsuan Tsai, and Ming-Hsuan Yang. Weakly-supervised semantic segmentation via sub-category exploration. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8988–8997, 2020. doi:10.1109/CVPR42600.2020.00901.
- [4] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results, 2012.
- [5] Tsung-Yi Lin, Piotr Dollár, Ross B. Girshick, Kaiming He, Bharath Hariharan, and Serge J. Belongie. Feature pyramid networks for object detection. *CoRR*, abs/1612.03144, 2016.
- [6] Zhun Zhong, Liang Zheng, Guoliang Kang, Shaozi Li, and Yi Yang. Random erasing data augmentation. *CoRR*, abs/1708.04896, 2017.

- [7] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. Learning deep features for discriminative localization. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2921–2929, 2016. doi:10.1109/CVPR.2016.319.
- [8] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.
- [9] Philipp Krähenbühl and Vladlen Koltun. Efficient inference in fully connected crfs with gaussian edge potentials. In J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 24. Curran Associates, Inc., 2011.
- [10] Paul Jaccard. Étude comparative de la distribution florale dans une portion des alpes et des jura. *Bulletin del la Société Vaudoise des Sciences Naturelles*, 37: 547–579, 1901.
- [11] Thorvald Julius Sorensen. *A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons*. Biologiske skrifter / Kongelige Danske videnskabernes selskab: bd. 5, nr. 4. I kommission hos E. Munksgaard, 1948. ISBN 0366-3312.

Teaching Construction Robotics for Higher Education Students: “Imagine and Make”

Z. Lafhaj^a, W. Al Balkhy^a, and T.s Linner^b

^a Centrale Lille, CNRS, UMR 9013-LaMcube, Lille, France

^b Professorship in Building Production, OTH Regensburg, Guest Scientist Technical University of Munich, Germany

E-mail: zoubair.lafhaj@centralelille.fr, wassim.al-balkhy@centralelille.fr, thomas.linner@oth-regensburg.de

Abstract –

The use of robotics in construction projects is still in its infancy despite the opportunities that robots can present to the improvement of construction practices. One of the strategies to effectively increase the reliance on robots in construction is increasing the knowledge and improving the educational programs about robotics for university students. This paper contributes to the ongoing efforts around the world to improve the teaching methods about construction robotics through the presentation of a novel method that is called “Imagine and Make”, in which students learn how to integrate robotics in different aspects and practices in construction projects. The method has been applied at Centrale Lille in France since 2018. The results of the application of “Imagine and Make” in the first semester in 2021-2022, evaluation by students, and teaching outcomes are reported in this paper.

Keywords –

Robotics; Construction management techniques; Construction 4.0; Robotics teaching; France

1 Introduction

Poor productivity, resistance to adopting technology, shortage of skilled workers, and poor quality are amongst the most serious challenges that face the construction industry [1–4]. Studies showed that while many sectors (e.g. manufacturing) are witnessing a noticeable increase in productivity rates, productivity has been decreasing in the construction industry for years [1]. Additionally, the industry is threatened by the decreasing numbers of skilled young laborers. Research reported difficulties in replacing aging and retiring laborers in construction projects and resistance by the young workforce to enter the sector due to the great physical efforts while performing repetitive and difficult construction activities, high levels of injuries, and poor safety standards [5], [6].

Robotics can contribute to providing solutions to face

these difficulties. This is because robots can be integrated into countless activities and provide workers with assistance to manage dangerous activities and perform repetitive tasks [3], [7–15]. The use of robots in these activities may not only contribute to productivity and efficiency improvements but also may help to improve safety standards and to attract more workforce to enter the construction industry and face the workforce shortages problems [6]. Moreover, the use of robots in construction operations can help to improve the precision, speed, and quality of construction work and avoid defects due to the use of advanced technologies (e.g. sensors, laser-based methods...etc.) [3], [16–18].

Nevertheless, despite the presence of some studies that expect a remarkable increase in robots’ use in the construction projects in the next few years [6], and despite the presence of some cases in which robots were used on construction sites since the 1980s [4], robotics face several barriers to be normalized in the construction industry and their adoption levels in the industry are still very low. In addition to the impact of the cost and time-related barriers, the lack of sufficient knowledge about new managerial principles and technological advances, lack of skilled team, resistance to change and to adopt new practices and technologies by managers and employees, and poor leadership are amongst the most affecting factors to adopt robotics in the construction industry [19–25].

The role of educators in educational institutes is not only to provide higher levels of knowledge about advanced concepts, technologies, and construction practices but also to prepare future leaders who can create and lead the transformation and reduce the resistance to change. Nevertheless, the literature still has only a limited number of studies that reported construction robotics teaching experiences for students [6], [26–28].

Within the scarcity of the studies about robotics teaching in the field of construction, this paper tries to contribute to the efforts that have been made so far and aims to report an experience about the use of a method that is called “Imagine and Make”. This method was used

to deliver training about construction robotics for students at a school of engineering “Centrale Lille” in France. The following section presents the materials and methods of the study by explaining the bases of “Imagine and Make”, the design of the study, the participants in the study, and the evaluation of the results. Then, the paper reports the prototypes and the results of the study. Finally, the paper reports the conclusion, implications, limitations, and direction for future research.

2 Materials and Methods

This section explains what is “Imagine and Make” and shows how it was developed, the theoretical basis behind it, how it is usually organized, and what are the evaluation criteria that are used to evaluate the provided projects by students. The section also explains the design of the current study, the participants, and the used methods to evaluate the effectiveness of “Imagine and Make”.

As part of previous research, the authors of this paper have done extensive research in developing a systems engineering framework for the holistic development of novel construction robot applications [29–31]. The authors have tested this framework as part of several industry-grade and research-grade robot development projects (for example, [32]). The work presented in this paper, states an attempt to translate the gained experiences in tune with project-based, iterative state-of-the-art teaching methodologies (for example, [33], [34]) into a hands-on teaching methodology. The development of construction robots needs to follow a systematic approach since it is conducted in an extremely interdisciplinary and complex area in which the lack of data and previous experience poses a high risk for developers and investors.

Through “Imagine and Make” students shall learn to understand the underlying value chains of innovation processes. After the participation in the course, they shall be able to handle actively a complex process that interlinks and iterates between analytical thinking, abstract planning skills, project-based integration, and technology transfer strategies. As such the students will be able to initiate and guide digital, robotic innovation in the organizations for which they will work in the future.

2.1 What is “Imagine and Make”

Centrale Lille is a graduate engineering school that is located in the north of France and has roots that go back to 1854. The school has different engineering programs for master’s students and doctoral candidates. The Centralien Engineering Programme is the teaching program that gathers students who finished two/or three years of undergraduate studies and allows them to spend

three years and get a master’s degree in engineering.

In 2018, Centrale Lille started a new teaching method that is called “Imagine and Make” for its students in the Centralien Engineering Programme. This method is based on gathering students for one week (five working days) and allowing them to work on a selected topic during this week.

The targeted topic for the last three years was construction 4.0. In this module, students have opportunities to have lectures about construction 4.0, read and see some practical examples, interact with experts from research and industry, work in teams, make prototypes, and present the results of their work. Accordingly, students have the chance to learn from:

- Lectures
- Reading materials
- Audio-visual tools
- Demonstration
- Discussion within the team and with experts
- Conducting research
- Practicing by doing prototypes
- Presenting their results

The lectures are mainly divided into two types of lectures; the first one is the introductory lecture, which aims to explain the overall structure of “Imagine and Make”, objectives, and evaluation criteria for the prototypes (3 hours). While the second type of lecture is presentations that are provided by experts from industrial partners (9 hours). The lectures provide students with audio-visual tools, some useful references, success stories, and practical examples about some new managerial principles and technological advances in the field of construction. Apart from the time that is devoted to the lectures, students have to use their time freely to develop their ideas (Imagine) in two days and to translate the idea into a physical prototype in the other three days (Make).

2.2 Design of the study

Figure 1 shows the structure of “Imagine and Make”. The discussed topic for the year 2021-2022 was “Construction robotics”. During this year, two rounds of “Imagine and Make” were conducted; the first one was in October and the second was in November. To study the targeted topic, students were asked to use “Spot®” from “Boston Dynamics” (shown in Figure 2) as an example of a type of robot that can be used in the construction industry and to develop a prototype that explains one of the applications of the robot in the industry. Spot® is a new robot that was presented to the market in late 2019 and its launching was faced by the impact of the COVID-19 crisis; however, it has shown promising applications in different fields and operations [35–37]. Therefore, Spot® serves as a clear example of mobile robots that can be understood by students, and its use may bring positive attitudes about the usability of

robotics in the construction industry.

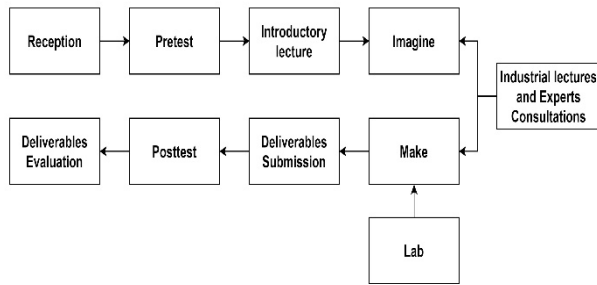


Figure 1. Structure of "Imagine and Make"



Figure 1. Spot® Robot [38]

The work in the two rounds started with the reception of the participants (179 students: 93 in the first round and 86 in the second round). Students were in their second year in Centralien Engineering Programme.

The introductory lecture presented an introduction to the concepts of industry 4.0, construction 4.0, advanced techniques in construction management, and robotics in construction. It also presented some examples of the use of different types of robots in construction and the use of Spot® in other industries. It also summarizes the objectives of the week, the structure of "Imagine and Make", the requested deliverables, and the evaluation criteria for the deliverables. The deliverables of the week and the grade given for each deliverable are as follows:

- Physical Prototype (40%)
- Final report (30%)
- 3-minute video about the prototype (15%)
- 5-minute oral presentation (pitching) (5%)
- Business plan and annexes (10%)

The deliverables have to address the following criteria:

- Clarity of the design and the idea
- Functionality of the prototype
- Collaborative work and roles within the team
- Sustainability and environmental impact
- Benchmarking, originality, and relation to the literature and research
- Clarity of the used methodology
- Marketing and commercial considerations.

During the whole week, students were able to reach eight experts and researchers in the fields of construction 4.0 and project management. The role of experts included explaining the objectives, answering questions about research sources, and supervising the overall progress.

Students were also able to get assistance from the lab of the engineering school (FabLab), which helped students to find needed materials for the prototypes and to cut and handle materials.

2.3 Survey development

In addition to the value of the deliverables provided by the students and the live experience of dealing directly with robots, two surveys were used at the beginning (pretest) and the end of the week (posttest) to evaluate the effectiveness of "Imagine and Make". The two surveys shared two sections, which are "self-efficacy" and "knowledge gain", and an additional section was added to the posttest and which aimed to assess the levels of satisfaction among students with the experience of "Imagine and Make".

Self-efficacy is one of the most important measures in business, educational, and psychological sciences [39], [40]. Self-efficacy refers to the beliefs of the individuals about their capacities in a given situation and their abilities to organize and execute a specific set of tasks and actions [41], [42]. High levels of self-efficacy are associated with better quality and higher efficiency and effectiveness in the work environment and low levels of self-efficacy are among the reasons for having incomplete tasks even with the presence of high levels of knowledge and skills [42]. With students, self-efficacy affects objectives setting, goals achievement, future choices of activities, and learning outcomes [39], [43], [44].

According to Selby et al [45], teaching robotics should aim at increasing self-efficacy for learners to increase their intrinsic interest in dealing and interacting with robots. Therefore, the pretest and post-test surveys in this study used a set of items to assess the improvements in the levels of self-efficacy due to the presentation of "Imagine and Make", and the current study tests the following hypothesis:

H1: "Imagine and Make" helped to improve self-efficacy levels toward robotics use among the students.

To assess the levels of self-efficacy before and after "Imagine and Make", ten items were adopted from the studies of Selby et al [45] and Mallik et al [46]. For this section, students were asked to put their answers using a five-point Likert scale ranging from "Not confident at all to do the following task" to "Completely confident that I can do the task".

The second section of the survey aimed to assess the knowledge gained by using six questions that were developed to study the improvement in familiarity with some terms in the field of robotics. The six questions were about the term "robotics", robots' characteristics, robots' components, degrees of freedom, and design of robots. For each question, the choices carried only one correct answer, one choice of "I do not know", and other wrong choices (four wrong choices for all questions

except one question that was a True/False and “I do not know” question”). The knowledge gain was assessed based on the improvements in the number of correct answers.

The post-test had an additional section to assess the levels of satisfaction among the students with “Imagine and Make”. The section had seven items, in which a five-point Likert scale was used to test the satisfaction with different statements and ranges from “strongly dissatisfied” to “strongly satisfied”. The section also included two open-ended questions to report the most liked things about “Imagine and Make” and the areas for improvement in future experiences.

The survey was developed using Google Forms and distributed to the students at the beginning and at the end of the week to be filled online. Out of 179 students, 166 responded to the survey with a response rate of 92.74%.

3 Results

3.1 Deliverables

By the end of each week in the two rounds of “Imagine and Make”, students were able to deliver all the requirements including prototypes, videos, reports, presentations, and business models for their work. The total number of the groups was 22 delivering different ideas about the use of the dog robot in the construction.

The ideas aimed to solve several problems and challenges faced in construction projects and covered different aspects such as safety and security on the construction site, productivity improvement, logistics management, and sustainability. Furthermore, in addition to making things by hand in the lab (Figure 3), students used several applications to deliver their ideas; including artificial intelligence (AI) and machine learning (ML), 3D printing, sensors, 2D and 3D plans, and drawings, simulation, cameras, thermal analysis, and others.



Figure 3. A picture from the lab during “Imagine and Make”

Examples of the developed prototypes included detection of construction tools and delivery them to workers, detection of physical wastes that can hinder the safe movement of workers on site (Figure 4), inspection of cracks in buildings and elements, identification of

potential hazards onsite (ledges, under cranes areas, misplaced tools) (Figure 5) inspection for intruders on site (Figure 6), inspection for commitments to personal protective equipment (PPE) onsite (Figure 7), presenting cooling system for workers that are working in a hot climate (Figure 8), development of a system to repair gas pipes (Figure 9), guiding people with vision impairment (Figure 10), marking on construction site (Figure 11-12), managing materials storage using Radio Frequency Identification (RFID)-based system (Figure 13), and others.



Figure 4. Detection and collection of wastes onsite

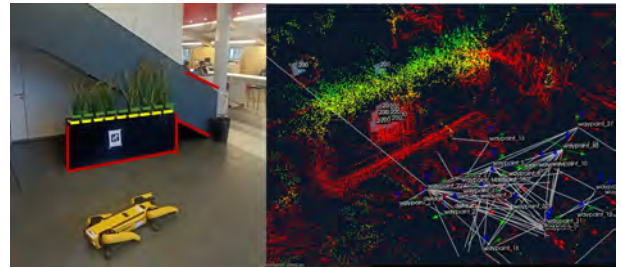


Figure 5. AI-controlled robot to make a Simultaneous localization and mapping (SLAM) of the environment to identify potential hazards on site



Figure 6. Use of MA to inspect for intruders onsite



Figure 7. Inspection for hard hat onsite



Figure 11. Use of Spot® Robot to do marking onsite

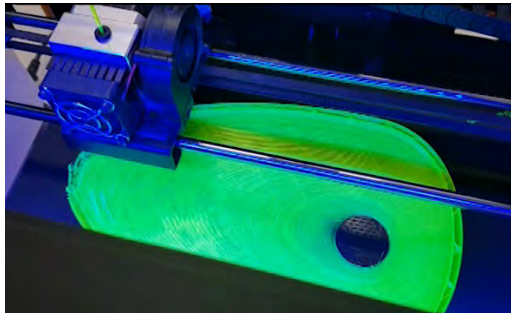


Figure 8. Use of 3D printing to design a cooling system for workers in hot climates



Figure 12. RFID-based system to manage materials



Figure 9. Development of a system to repair leaks areas in gas pipes



Figure 10. Use of robot dog to guide people with vision impairments

3.2 Survey Results

The analysis of the collected data was done using the statistical package for the social science (SPSS) version 25.0, and it covered the follows:

- Testing the reliability of the measurement tool by testing the consistency in the tool across the time and the different items using Cronbach's Alpha, which is supposed to be higher than 0.6 [47].
- Calculating the means and standard deviations to identify the levels of pre and post-training self-efficacy and satisfaction among students at the end of the training.
- Testing the hypothesis H1 (improvements in self-efficacy levels) using "Paired t-test", which is an inferential statistical method that compares the differences in the mean values when the data is collected in pairs (e.g. pretest and post-test) [48].
- Calculating the frequencies of the correct answers to the questions about the familiarity with robotics to assess the knowledge about the topic before and after the training.

Additionally, a qualitative analysis was conducted to analyze the results from the open-ended questions. The analysis was based on inductive thematic coding analysis through analyzing the text to find the frequent, significant, and emerging themes that are inherent in the data [49].

3.2.1 Reliability

The analysis showed that Cronbach's Alpha was (0.872) for the pretest survey and (0.920) for the posttest

survey. As the value of Cronbach's Alpha exceeds 0.6, this indicates that the used scale is reliable and has a high level of internal consistency.

3.2.2 Self-efficacy

Self-efficacy levels are shown in Figure 13. The analysis revealed that the overall self-efficacy had a mean of (2.35), and raised after the training to reach (3.27). The analysis showed increases in the means for all items that were used to assess self-efficacy. The highest increase was in "Know how to program a robot or additive tool", "Perform a design for a robot", and "construct a robotic prototype".

The results from paired t-test showed significant differences between the pretest and post-test means for all items and the overall self-efficacy P-value was less than (0.001) for all comparisons. This indicates that hypothesis H1 could not be rejected and that "Imagine and Make" helped to improve self-efficacy levels toward robotics use among the students.

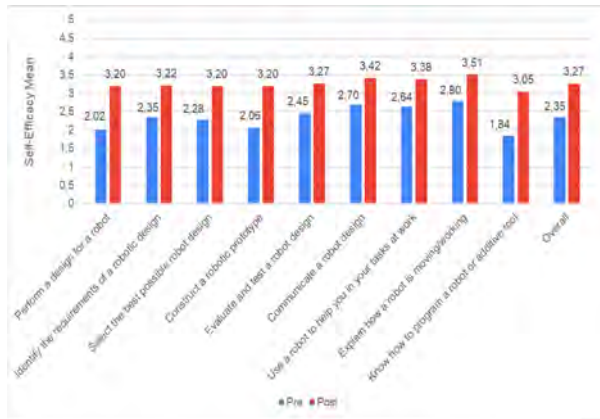


Figure 13. Self-Efficacy levels before and after "Imagine and Make"

3.2.3 Knowledge gain

Figure 14 shows the number of correct answers for each question used in the pre and post-tests. The figure shows an increase in the number of correct answers for all questions; especially question three, which was about the components of the robot.

3.2.4 Satisfaction with "Imagine and Make"

Table 1 shows the level of satisfaction with "Imagine and Make". The analysis showed that, overall, students were satisfied with the experience of "Imagine and Make" as the mean value was (3.97) out of (5.00). The highest rate was for the opportunity the students had to think about the different practices in the construction projects, then helping them improve their teamworking skills.

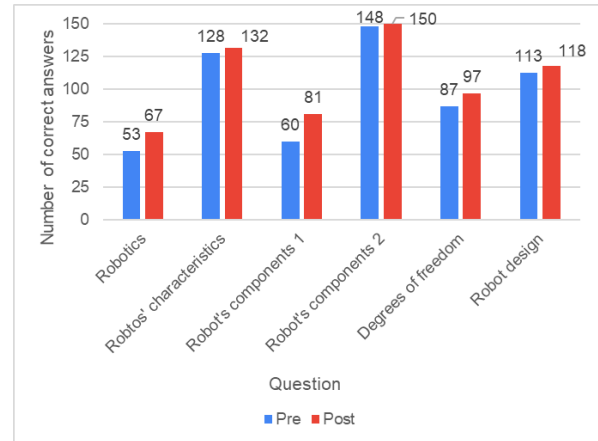


Figure 14. Knowledge gain before and after "Imagine and Make"

Table 1. Satisfaction levels with "Imagine and Make"

Item	Mean	SD
I am satisfied with "Imagine and Make"	4.084	0.81
My interest in robotics increased after "Imagine and Make"	3.663	0.99
"Imagine and Make" helped me to understand the use of robots in construction	3.910	0.89
"Imagine and Make" helped me to think about different practices in the construction industry	4.187	0.87
"Imagine and Make" helped me to enhance my innovative thinking	3.934	0.84
"Imagine and Make" helped me to improve my team working skills	4.072	0.92
"Imagine and Make" met my expectations	3.940	0.96
Overall Satisfaction	3.970	0.69

The qualitative analysis revealed that team working and collaboration, freedom of prototype development and topic selection for the teams, challenging deliverables, ability to work on a practical topic and deal with a real robot, presentations by the experts, and ability to do things by hand in the lab were among the most frequent themes mentioned by the students when they were asked about the things that they like most in "Imagine and Make". In turn, the most frequent suggestion was increasing the time allocated to the training for more than one week.

4 Conclusion

This paper reports a novel method to teach construction robotics by giving the students the opportunity to freely and collaboratively search for the applications that can integrate robotics into the construction field. This method

helped students to be in touch with both academic literature and practical applications. It also allowed them to integrate different types of technologies and tools while making their prototypes. This pedagogic method can be applied in several areas in construction 4.0 and even in other fields and sciences.

References

- [1] T. Bock, "The future of construction automation: Technological disruption and the upcoming ubiquity of robotics," *Autom. Constr.*, vol. 59, pp. 113–121, Nov. 2015.
- [2] W. Albalkhy and R. Sweis, "Assessing lean construction conformance amongst the second-grade Jordanian construction contractors," *Int. J. Constr. Manag.*, pp. 1–13, 2019.
- [3] Z. Dakhli and Z. Lafhaj, "Robotic mechanical design for brick-laying automation," *Cogent Eng.*, vol. 4, no. 1, Jan. 2017.
- [4] E. Forcael, I. Ferrari, A. Opazo-Vega, and J. A. Pulido-Arcas, "Construction 4.0: A Literature Review," *Sustain.* 2020, Vol. 12, Page 9755, vol. 12, no. 22, p. 9755, Nov. 2020.
- [5] A. H. Behzadan, A. Iqbal, and V. R. Kamat, "A collaborative augmented reality based modeling environment for construction engineering and management education," *Proc. - Winter Simul. Conf.*, pp. 3568–3576, 2011.
- [6] C. J. Liang, V. R. Kamat, and C. C. Menassa, "Teaching robots to perform quasi-repetitive construction tasks through human demonstration," *Autom. Constr.*, vol. 120, p. 103370, Dec. 2020.
- [7] S. N. Yu, B. G. Ryu, S. J. Lim, C. J. Kim, M. K. Kang, and C. S. Han, "Feasibility verification of brick-laying robot using manipulation trajectory and the laying pattern optimization," *Autom. Constr.*, vol. 18, no. 5, pp. 644–655, Aug. 2009.
- [8] J. K. Oh, G. Jang, S. Oh, J. H. Lee, B. J. Yi, Y. S. Moon, J. S. Lee, and Y. Choi, "Bridge inspection robot system with machine vision," *Autom. Constr.*, vol. 18, no. 7, pp. 929–941, Nov. 2009.
- [9] S. Hou, B. Dong, H. Wang, and G. Wu, "Inspection of surface defects on stay cables using a robot and transfer learning," *Autom. Constr.*, vol. 119, p. 103382, Nov. 2020.
- [10] D. Lattanzi and G. Miller, "Review of Robotic Infrastructure Inspection Systems," *J. Infrastruct. Syst.*, vol. 23, no. 3, p. 04017004, Feb. 2017.
- [11] O. Kontovourkis and G. Tryfonos, "Robotic 3D clay printing of prefabricated non-conventional wall components based on a parametric-integrated design," *Autom. Constr.*, vol. 110, p. 103005, Feb. 2020.
- [12] T. Linner and T. Bock, "Evolution of large-scale industrialisation and service innovation in Japanese prefabrication industry," *Constr. Innov.*, vol. 12, no. 2, pp. 156–178, 2012.
- [13] T. Linner, W. Pan, C. Georgoulas, B. Georgescu, J. Güttler, and T. Bock, "Co-adaptation of Robot Systems, Processes and In-house Environments for Professional Care Assistance in an Ageing Society," *Procedia Eng.*, vol. 85, pp. 328–338, Jan. 2014.
- [14] C. Georgoulas, T. Linner, and T. Bock, "Towards a vision controlled robotic home environment," *Autom. Constr.*, vol. 39, pp. 106–116, Apr. 2014.
- [15] T. Linner, J. Güttler, T. Bock, and C. Georgoulas, "Assistive robotic micro-rooms for independent living," *Autom. Constr.*, vol. 51, no. C, pp. 8–22, Mar. 2015.
- [16] L. Song and D. Liang, "Lean construction implementation and its implication on sustainability: A contractor's case study," *Can. J. Civ. Eng.*, vol. 38, no. 3, pp. 350–359, Mar. 2011.
- [17] R. Bogue, "What are the prospects for robots in the construction industry?," *Ind. Robot An Int. J.*, vol. 45, no. 1, pp. 1–6, Jan. 2017.
- [18] M. Jeong Kim, H.-L. Chi, X. Wang, M. Jeong, K. H.-L. Chi, and Lieyun Ding, "Automation and Robotics in Construction and Civil Engineering Special Issue for Sensors 'Intelligence Sensors and Sensing Spaces for Smart Home' View project Automation and Robotics in Construction and Civil Engineering," *Artic. J. Intell. Robot. Syst.*, vol. 79, pp. 347–350, 2015.
- [19] W. Al Balkhy, R. Sweis, and Z. Lafhaj, "Barriers to Adopting Lean Construction in the Construction Industry—The Case of Jordan," *Buildings*, vol. 11, no. 6, p. 222, May 2021.
- [20] W. Albalkhy and R. Sweis, "Barriers to adopting lean construction in the construction industry: a literature review," *Int. J. Lean Six Sigma*, vol. 12, no. 2, pp. 210–236, Mar. 2020.
- [21] D. J. Trujillo and E. Holt, "Barriers to Automation and Robotics in Construction," *Epic Ser. Built Environ.*, vol. 1, pp. 257–265, 2020.
- [22] R. Mahbub, "Framework on the Barriers to the Implementation of Automation and Robotics in the Construction industry," *Int. J. Innov. Manag.*, vol. 3, no. 1, pp. 21–36, 2015.
- [23] J. M. Davila Delgado, L. Oyedele, A. Ajayi, L. Akanbi, O. Akinade, M. Bilal, and H. Owolabi, "Robotics and automated systems in construction: Understanding industry-specific challenges for adoption," *J. Build. Eng.*, vol. 26, p. 100868, Nov. 2019.
- [24] J. Buchli, M. Giftthaler, N. Kumar, M. Lussi, T. Sandy, K. Dörfler, and N. Hack, "Digital in situ fabrication - Challenges and opportunities for robotic in situ fabrication in architecture, construction, and beyond," *Cem. Concr. Res.*, vol. 112, pp. 66–75, Oct. 2018.
- [25] G. Carra, A. Argiolas, A. Bellissima, M. Niccolini, and M. Ragaglia, "Robotics in the construction industry: State of the art and future opportunities," *ISARC 2018 - 35th Int. Symp. Autom. Robot. Constr. Int. AEC/FM Hackathon Futur. Build. Things*, 2018.
- [26] B. Bogosian, L. Bobadilla, M. Alonso, A. Elias, G. Perez, H. Alhaffar, and S. Vassigh, "Work in

- Progress: Towards an Immersive Robotics Training for the Future of Architecture, Engineering, and Construction Workforce,” *EDUNINE 2020 - 4th IEEE World Eng. Educ. Conf. Challenges Educ. Eng. Comput. Technol. without Exclusions Innov. Era Ind. Revolut. 4.0, Proc.*, Mar. 2020.
- [27] P. Adami, P. B. Rodrigues, P. J. Woods, B. Becerik-Gerber, L. Soibelman, Y. Copur-Gencturk, and G. Lucas, “Effectiveness of VR-based training on improving construction workers’ knowledge, skills, and safety behavior in robotic teleoperation,” *Adv. Eng. Informatics*, vol. 50, p. 101431, Oct. 2021.
- [28] H. Yi, “Robotics application for the advanced integration of design and technology in architecture,” *Comput. Appl. Eng. Educ.*, vol. 29, no. 5, pp. 1146–1162, Sep. 2021.
- [29] C. C. Cheah, T. Bock, J. Cao, T. Linner, Y. H. Liu, A. Yamashita, and L. Yang, “Guest Editorial Introduction to the Focused Section on Mechatronics and Automation for Constructions,” *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 6, pp. 2819–2825, Dec. 2021.
- [30] T. Linner, W. Pan, R. Hu, C. Zhao, K. Iturralde, M. Taghavi, J. Trummer, M. Schlandt, and T. Bock, “A technology management system for the development of single-task construction robots,” *Constr. Innov.*, vol. 20, no. 1, pp. 96–111, Jan. 2020.
- [31] M. Pan, T. Linner, W. Pan, H. Cheng, and T. Bock, “Structuring the context for construction robot development through integrated scenario approach,” *Autom. Constr.*, vol. 114, p. 103174, Jun. 2020.
- [32] K. Iturralde, M. Feucht, R. Hu, W. Pan, M. Schlandt, T. Linner, T. Bock, J. B. Izard, I. Eskudero, M. Rodriguez, J. Gorrotxategi, J. Astudillo, J. Cavalcanti, M. Gouttefarde, M. Fabritius, C. Martin, T. Henninge, S. M. Normes, Y. Jacobsen, A. Pracucci, J. Canāda, J. D. Jimenez-Vicaria, C. Paulotto, R. Alonso, and L. Elia, “A cable driven parallel robot with a modular end effector for the installation of curtain wall modules,” *Proc. 37th Int. Symp. Autom. Robot. Constr. ISARC 2020 From Demonstr. to Pract. Use - To New Stage Constr. Robot*, pp. 1472–1479, 2020.
- [33] L. O. Wilson, “Anderson and Krathwohl Bloom’s Taxonomy Revised Understanding the New Version of Bloom’s Taxonomy,” 2016. [Online]. Available: https://quincycollege.edu/wp-content/uploads/Anderson-and-Krathwohl_Revised-Blooms-Taxonomy.pdf. [Accessed: 28-Feb-2022].
- [34] R. Moen and C. Norman, “The History of the PDCA Cycle,” in *The 7th Asian Network for Quality*, Tokyo, Japan, 2009.
- [35] BBC, “Robot cheerleaders support Japanese baseball team Fukuoka SoftBank Hawks - BBC Sport,” 2020. [Online]. Available: <https://www.bbc.com/sport/av/football/53352676>. [Accessed: 03-Nov-2021].
- [36] offshore-mag.com, “‘Spot’ deployed on Aker BP’s Skarv FPSO | Offshore,” 2020. [Online]. Available: <https://www.offshore-mag.com/business-briefs/equipment-engineering/article/14187978/spot-deployed-on-aker-bps-skarv-fps-o-offshore-norway>. [Accessed: 03-Nov-2021].
- [37] Boston Dynamics, “Boston Dynamics COVID-19 Response | Boston Dynamics,” 2021. [Online]. Available: <https://www.bostondynamics.com/resources/COVID-19>. [Accessed: 03-Nov-2021].
- [38] Boston Dynamics, “Spot® - The Agile Mobile Robot | Boston Dynamics,” 2021. [Online]. Available: <https://www.bostondynamics.com/spot>. [Accessed: 03-Nov-2021].
- [39] C. Salavera, P. Usán, and L. Jarie, “Emotional intelligence and social skills on self-efficacy in Secondary Education students. Are there gender differences?,” *J. Adolesc.*, vol. 60, pp. 39–46, Oct. 2017.
- [40] J. Cherian and J. Jacob, “Impact of Self Efficacy on Motivation and Performance of Employees,” *Int. J. Bus. Manag.*, vol. 8, no. 14, 2013.
- [41] A. Bandura, W. H. Freeman, and R. Lightsey, “Self-Efficacy: The Exercise of Control,” *J. Cogn. Psychother.*, vol. 13, no. 2, pp. 158–166, Jan. 1999.
- [42] A. Bandura, “Self-efficacy mechanism in human agency,” *Am. Psychol.*, vol. 37, no. 2, pp. 122–147, Feb. 1982.
- [43] J. Jakešová, P. Gavora, J. Kalenda, and S. Vávrová, “Czech Validation of the Self-regulation and Self-efficacy Questionnaires for Learning,” *Procedia - Soc. Behav. Sci.*, vol. 217, pp. 313–321, Feb. 2016.
- [44] K. R. Wentzel, K. Muenks, D. McNeish, and S. Russell, “Peer and teacher supports in relation to motivation and effort: A multi-level study,” *Contemp. Educ. Psychol.*, vol. 49, pp. 32–45, Apr. 2017.
- [45] N. S. Selby, J. Ng, G. S. Stump, G. Westerman, C. Traweek, and H. H. Asada, “TeachBot: Towards teaching robotics fundamentals for human-robot collaboration at work,” *Heliyon*, vol. 7, no. 7, p. e07583, Jul. 2021.
- [46] A. Mallik, S. B. Rajguru, and V. Kapila, “Fundamental: Analyzing the effects of a robotics training workshop on the self-efficacy of high school teachers,” *ASEE Annu. Conf. Expo. Conf. Proc.*, vol. 2018-June, Jun. 2018.
- [47] U. Sekaran and R. Bougie, *Research methods for business: A skill building approach*. 2019.
- [48] H. Hsu and P. A. Lachenbruch, “Paired t Test,” *Wiley StatsRef Stat. Ref. Online*, Sep. 2014.
- [49] D. R. Thomas, “A General Inductive Approach for Analyzing Qualitative Evaluation Data:,” <http://dx.doi.org/10.1177/1098214005283748>, vol. 27, no. 2, pp. 237–246, Jun. 2016.

Integrating VR and Simulation for Enhanced Planning of Asphalt Compaction

A.R. Revollo Dalence^a, F. Vahdatikhaki^a, S. Miller^a, A. G. Dorée^a

^a Department of Civil Engineering and Management, University of Twente, The Netherlands
E-mail: arrevollo@gmail.com, f.vahdatikhaki@utwente.nl, s.r.miller@utwente.nl, a.g.doree@utwente.nl

Abstract

The current decision-making practices in road construction, are largely based on tacit knowledge, craftsmanship, tradition, and custom. This results in considerable variability in the execution of projects and deviation between as-planned and as-executed practices. The current simulation-based planning techniques are limited because they tend to present spatial and temporal characteristics of projects separately. This segregated approach ignores the interdependencies between spatial and temporal aspects of projects specially with respect to safety and process quality assessment. This is more palpable in the asphalt compaction projects because the quality of the compaction depends on a myriad of temporal (e.g., compaction speed) and spatial (e.g., homogenous compaction of the mat) parameters. Therefore, this research aims to develop a novel framework to capture the factors affecting the compaction process in a holistic manner and translate them into relevant decision variables. This framework achieves this objective by integrating simulation and virtual reality technologies. In this framework, simulation is responsible for capturing the affecting factors and generating temporal decision variables, whereas VR virtualizes them and provides high (3D) spatial assessment and awareness. A prototype is developed and tested with ASPARi case studies to demonstrate the feasibility of the framework. It is shown that compared to current planning practices, the integrated model can significantly improve various aspects of planning the construction process, especially by improving awareness among decision-makers concerning the development of more standardized compaction patterns.

Keywords –

Simulation; Virtual Reality; Compaction; Planning

1 Introduction

The ultimate performance of asphalt pavements is predicated on the asphalt mix (i.e., design) and the

compaction process (i.e., construction). While a good design is crucial for high-quality pavement, it is shown that the compaction process plays a more important role in explaining variability in asphalt quality [1,2,3]. This is mainly due to the fact that while the production of asphalt mix benefits from the highly controlled industrial environment of asphalt plants, the compaction process on the site is largely unstructured, experience-driven, uncertainty-laden, and more difficult to control [3,4,5,6]. Besides, compaction is a highly complex process that depends on and is sensitive to a myriad of factors, e.g., material properties, initial density, equipment, traffic, and environment [1,7]. This renders the control of the compaction process very difficult and strongly ties the success of compaction to meticulous and detailed planning.

Miller [5] suggested that proper planning of the compaction process needs to (1) make the operational behavior explicit, (2) visualize the compaction process, and (3) engage all construction crew in the decision making. In recent years, construction simulation methods have been deployed for the planning of different types of operations. These methods are shown to be able to meet the above-mentioned requirements [8]. However, the current simulation-based planning techniques have a major shortcoming for application in compaction operations. It is mainly because the current methods tend to present spatial and temporal characteristics of projects separately. This segregated approach ignores the interdependencies between spatial and temporal aspects of compaction operations. This is especially important because proper planning of compaction requires concurrent consideration of temporal (e.g., compaction speed and delivery schedule) and spatial (e.g., homogenous compaction of the mat) parameters. In other words, while spatial modeling concerns itself mainly with planning at the strategic level of operation, tactical planning requires decision-making at the spatial level (i.e., precise compaction route planning). The fragmented modeling approach keeps strategic and tactical planning separate and forces planners to take a sequential planning approach, where first strategic decisions are made and then the tactical decisions are made. This separation ignores the intricate interplay between strategic and

tactical planning in compaction operations. Therefore, it seems to be more optimal to have an integrated platform that supports comprehensive and concurrent planning. Nevertheless, the existing compaction simulation models primarily focus on the temporal perspective (i.e., strategic planning) of the process and offer limited support for tactical planning.

Virtual Reality (VR) has been successfully employed to increase the spatial awareness of construction processes [9]. Rekapalli [10] used VR to study complex earthmoving operations and argued that simulation-based VR could be an effective method to test and validate complex simulation models (i.e., identify and rectify possible errors in the simulation). Turner [11] and Akpan [12] demonstrated the potential benefits of integrating simulation and VR in different projects. However, to the best of the authors' knowledge, the integration of VR representation of the asphalt process with the temporal simulation models has been seldom investigated.

Therefore, the objective of this research is to develop a framework for the integration of temporal simulation of the compaction process and 3D spatial representation (i.e., through VR) to enhance the efficiency and consistency of the decision-making process and to support concurrent strategic and tactical planning. It is argued that this approach would open dialogue between compaction planners and operators and help them better sensitize themselves to the intricacies involved at different levels of compaction planning. In this integrated approach, the simulation could deal with the temporal decision variables, such as quantity and speed of equipment, roller passes, and translate them into relevant assessment indicators, i.e., time and cost. VR, on the other hand, could deal with the spatial decision variables, such as the areal output of equipment, rollers' trajectory, asphalt cooling rate, and translate them into relevant quality indicators, i.e., compaction efficiency and process consistency.

The remainder of the paper is structured as follows. First, the research methodology applied in this research is presented. Next, the requirements of the new framework are discussed. Subsequently, the proposed framework is explained in detail. This is followed by the presentation of the implementation and case study. Finally, the conclusions are discussed.

2 Research Methodology

This research applied a variation of the design research methodology [13]. Accordingly, the research is divided into four phases, i.e., literature review, requirement analysis, framework development, and synthesis. The first phase focused on acquiring relevant knowledge regarding the asphalt pavement industry, simulation methods, and VR platforms. Particularly, the

interconnections, interdependencies, current decision-making practices, and characteristics of the hot mix asphalt. This phase resulted in the identification of research gaps in the current body of knowledge concerning the integration of simulation and VR techniques for asphalt paving operations. In Phase Two, the essential requirements for the integrated model were obtained through interviews with asphalt experts in The Netherlands at both tactical and strategic levels. In Phase Three, the framework was developed. In this framework, Agent-Based Simulation (ABS) is integrated with a VR representation of the asphalt compaction process. To this end, the existing VR developed by previous research of the ASPARi group [14] was used. In Phase Four, the framework was implemented, verified, and validated. This was done by developing a prototype and then using it in a case study. The model was then validated in two different ways. First, its accuracy was assessed based on data from an actual compaction operation. Second, the usefulness of the model was assessed with the input of asphalt experts.

3 Requirement Analysis for the New Framework

To identify and analyze the requirements of the new framework, a set of four interviews were conducted with experts from BAM and Heijmans, i.e., two of the largest construction companies in The Netherlands. The interviewees from the former company belong to the tactical realm of project planning and their focus lies on asphalt construction projects. Both of them, while being part of academia, conducted research focusing on asphalt pavement. On the other hand, the experts from the latter company have been working on asphalt-related projects for several years. One interviewee had a more tactical background and gave a broader overview of the process, whereas the other was more experienced in the operational realm, providing relevant feedback from an operational point of view.

Based on these interviews, the overall requirements of the integrated model have been identified as follows: (1) The model should provide the user with resource allocation alternatives, (2) The model should allow the user to evaluate different strategies with different fixed parameters, e.g., provide different strategies for two allocated (fixed) rollers, or provide different strategies for (fixed) paver speed, (3) The model should provide the user with relevant feedback about the quality of the compaction. The interviewees also pointed out the importance of considering the temperature issue during the compaction process. Every single phase of the process is affected by the asphalt-mix temperature. That being said, to properly represent the process in a simulation model, the asphalt cooling behavior needs to

be captured in the model.

To be able to translate these high-level requirements to specific parameters that need to be molded in the integrated simulation, the experts were asked to identify the operational factors involved in asphalt compaction processes. Then, these factors are divided into fixed input parameters and decision variables. The fixed parameters refer to the factors that cannot be altered in the planning phase, such as layer thickness, mixture design, road geometry. In contrast, the decision variables, i.e., equipment speed, equipment output, compaction strategy, are the factors that fall into the decision-makers' hands and, consequently, can be modified within the model. To illustrate, Figure 1 and Table 1 present the parameters and decision variables, respectively. In essence, the integrated model needs to capture all these parameters and allow the planners to capture them in their decision-making.

4 Proposed framework

This research proposes a hybrid, i.e., simulation and VR, planning model for the simulation of asphalt compaction operations.

In this platform, simulation is responsible for capturing the parameters described in Table 1, and generating temporal decision variables, i.e., the quantity of trucks, equipment output, and average speed. Then, the simulation feeds all the factors to the VR, in which virtualization of all the planning decisions takes place. Within this environment, the decision-maker is able to evaluate the spatial decision variables, i.e., roller trajectory, length of roller track, and distance Paver-Roller. On top of that, the VR offers operational quality feedback for the work done.

Figure 2 depicts an overview of the proposed framework, which is divided into five phases, i.e., Data Collection, Simulation, Integration, VR, and Strategy Assessment. The first phase begins with the collection of the parameters shown in Table 1. Afterward, the data is organized and initial values for the decision variables are computed. Specifically, the paver and roller speed, length of roller track, and roller trajectory.

Then, the data is fed to the ABS simulation, and the

user can assess, in virtual real-time and 2D, the equipment movement, paved surface, and the asphalt mat cooling. Once the simulation is finished, it is possible to evaluate the paver and roller output (m²/h) along the entire process. Subsequently, the integration phase translates the thermal and equipment behavior into data-driven physics and data-driven agents for the VR, respectively [15].

Table 1. Relevant factors for the asphalt compaction modeling

Factor	Unit	Type of factor	Category
Available time for operations		Parameter	-
Mixture properties	-	Parameter	Mat. properties
Road geometry	m	Parameter	Mat. properties
Pavement dimensions	m	Parameter	Mat. properties
Delivery temperature	°C	Parameter	Mat. Properties
Asphalt cooling rate		Parameter	Environmental
Distance asphalt plant-site	km	Parameter	Equipment
Truck capacity	m ³	Parameter	Equipment
Truck cycles	-	Parameter	Equipment
Waiting periods	-	Parameter	Equipment
Number of trucks	#	Variable	Equipment
Quantity of pavers	#	Parameter	Equipment
Width/screed width	m	Parameter	Equipment
Flow stoppers	-	Parameter	Equipment
Areal output	m ² /h	Variable	Equipment
Average paver speed	m/min	Variable	Equipment
Quantity of rollers	#	Parameter	Equipment
Number of passes	#	Parameter	Equipment
Roller's width	m	Parameter	Equipment
Overlap	m	Parameter	Equipment
Average roller speed	m/min	Variable	Equipment
Length of roller track	m	Variable	Equipment
Distance paver-roller	m	Variable	Equipment
Roller trajectory	-	Variable	Equipment
Areal output (capacity)	m ² /h	Variable	Equipment

Next, the user can evaluate the impact of their choices on the asphalt mat, such as equipment movement, mat cooling, and rolling pattern with high spatial resolution in the VR environment. Finally, the strategy assessment phase takes place. The user receives resource output feedback and operational quality feedback. The resource output feedback concerns the equipment output (m²/h), and their speed (m/min). The operational quality feedback concerns compaction efficiency (%) and process consistency (%), i.e., whether the mat has been

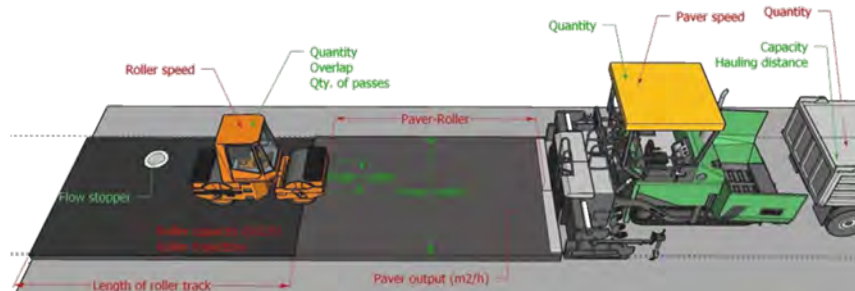


Figure 1. Schematic overview of the HMA-CP parameters (green) and decision variables (red)

effectively compacted and within an appropriate the temperature range. A detailed description of how compaction efficiency and process consistency are calculated is presented in the previous work of the authors [14].

4.1 Data Collection

The available time to deliver the hot asphalt to the site is the starting point for compaction planning. Then, the geometry of the road, i.e., road length, width, and layer thickness must be stated using a parametric model.

The temperature of the asphalt mix is of utmost importance in asphalt compaction. If the mixture is too hot, it can be overstressed, hence the mat would spread laterally rather than being compacted. If the mixture is cold while being compacted, it can be under-stressed, hence the roller cannot create sufficient shear force to increase the density of the mat [5]. Therefore, the compaction of the mat must be achieved within a certain temperature range, which can be obtained from the asphalt cooling curve.

This cooling behavior of asphalt can be determined either through the use of data from actual construction sites or cooling curve prediction tools such as PaveCool [16].

To guarantee uniformity in the paving operations and to have a constant delivery rate of asphalt mix, the planner needs to evaluate the number of trucks that are going to be employed in the operations. The user needs to state the distance from the asphalt plant to the construction site, truck capacity, the estimated time for hauling operations, loading and unloading time. The above information is employed to simulate the hauling operations and to provide the minimum number of trucks required to guarantee a constant delivery of asphalt mix. Moreover, depending on the particular characteristics of the road, there is a transition period between trucks. That is, when a truck is ready to pour the asphalt mix into the paver, there is an interruption in the paver flow between the empty-truck departure and the full-truck steering for

unloading. This transition is captured and represented in the simulation model. The user needs to state an average time for this shift.

The paver features, such as type, number, and desired screed width need to be specified. With that information retrieved, the model offers two ways to choose the paver speed: (1) the initial average speed is computed based on the productivity goal, (2) the user can state the initial average speed based on their expertise. Furthermore, some specific road sections have a certain degree of complexity, which impacts the productivity of the paver, and consequently, the uniformity and continuity of operations are affected [20]. Therefore, these so-called flow stoppers, e.g., roundabouts, intersections, curves, must be considered when representing the behavior of the paver. Runneboom [20] proposed an approach for the consideration of the number and type of flow stoppers in a simulation model. Based on this approach, the impact of different flow stoppers is identified, categorized, and translated to output rate parameters. Thus, the simulation model quantifies the flow stoppers and translates them into the paver behavior, i.e., reduction in the average speed.

The user also needs to provide information about the roller characteristics, e.g., width, number, and type of rollers as well as the number of required passes. Then, the overlap, length of roller track, and distance of paver-roller can be calculated using equations proposed by BOMAG [17]. The roller speed can be specified in two different ways. Either the user specifies an initial average speed, or the model suggests a value based on the premise that the paver output and the roller capacity should be aligned, again using equations proposed by BOMAG. Finally, the compaction strategy must be chosen. Compaction strategy refers to the trajectory that the roller will follow to cover the mat completely and with the desired number of passes. The simulation model offers different compaction strategies based on the number of lanes, the width coverage of rollers, the number of passes, and the number of rollers. The suggested paths are based on the standard compaction strategies available within the body of knowledge [17].

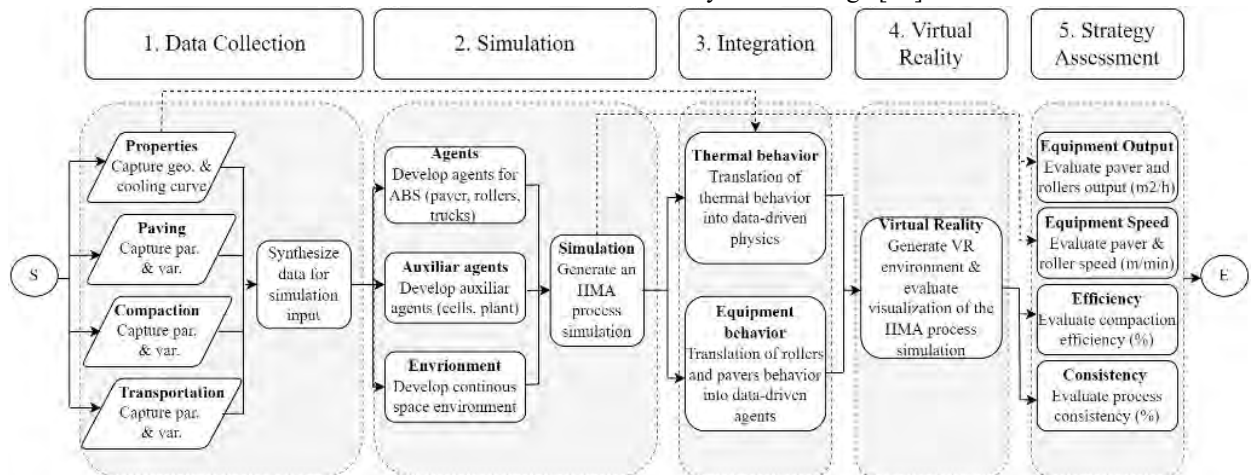


Figure 2. Overview of the proposed framework

4.2 Simulation

Regarding the agents in the ABS model, five agents have been modeled, namely truck, roller, paver, asphalt plant, and cells, which represent the discretized paved surface. Figure 3 shows the overview of the agent's flowcharts.

The truck is dispatched from the plant at a given rate. Then, the truck awaits a signal from the paver to accommodate itself next to the paver-hopper and later dump the material while moving along with the paver. Once the job is completed, the truck leaves the construction site. When the truck is pouring the asphalt mix into the paver-hopper, the paver starts paving at the initial speed. Each paved unit is represented in the model with a cell. The paver evaluates regularly whether the current speed is sufficient to finish the job, if not, the paver modifies its speed, based on the target productivity.

When the paver moves, cells appear in form of discretized lanes. The width of the lane corresponds to the width of the road and it is divided into ten cells. As for the length, each lane has a unitary length, in this case, it is one meter. Afterward, the cell cools down according to the cooling curve.

To start compacting, two conditions need to be met. Firstly, there should be a minimum distance between paver and roller before the roller can start the compaction. Secondly, the cell temperature is equal to or less than the upper limit of the compaction temperature window. To assess the latter, the model considers the time passed from freshly paved asphalt to the upper threshold, which is obtained from the asphalt cooling curve. After rolling each section, the roller evaluates whether its capacity is aligned with the paver output, and its distance with the paver is within the limits. Then, the roller modifies its

speed accordingly.

The simulation allows the user to partially evaluate their strategy, i.e., assess it from a temporal point of view. On one hand, the user can visualize in virtual time and 2D environment, the equipment movement, the paved surface, and the asphalt cooling. Besides, they can evaluate whether the resource allocation is capable of successfully completing the job with the allocated time. On the other hand, once the simulation is completed, the user receives graphical feedback in terms of paver and roller output (m²/h) and their speed along the entire process. If the user is happy with the partial results, they can move forward to the next phase.

4.3 Integration

In general, the integration phase is responsible for the conversion of the simulation output data into input data that feeds the VR. That being said, three main conversions are needed, i.e., agent conversion, physics conversion, and logistics conversion. The agent conversion is about translating the movement of the paver and roller to the timestamp location series. Then, the values are translated into local VR environment coordinates. The physics conversion translates the cooling curve obtained in the previous phase to timestamped values. Then, the VR environment uses this data to represent the temperature of each cell. Finally, the logistics features, such as equipment quantity, and available time, are converted into data that the VR environment can use.

4.4 Virtual Reality

The VR environment allows the user to evaluate their

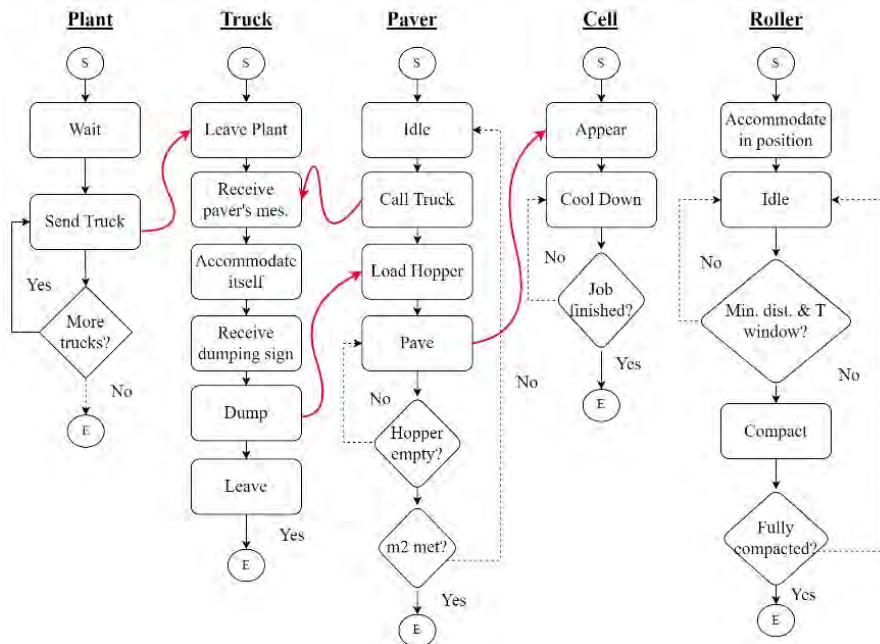


Figure 3. Overview of agents' behavior

strategy in virtual real-time and with high spatial resolution in 3D. VR allows planners to assess the asphalt mat cooling, and the compaction completeness interactively. The quality metrics used for this phase are compaction efficiency and process consistency. The former refers to how many cells have been compacted within the allowable temperature window. To clarify, if the compaction efficiency is 40 %, it means that 40 % of the cells have had all compaction passes within the temperature window [14]. The remaining 60% had at least one pass outside this window. Process consistency show how much time is left for a cell to have successful compaction [15].

4.5 Strategy Assessment

Once the user evaluated the simulation and the VR environment, they can assess their overall strategy. In general, the expert can evaluate the performance of their choices with four metrics, i.e., equipment areal output (m²/h), equipment speed (m/min), compaction efficiency (%), and process consistency (%). Further, for the last metric, the model provides the percentage of cells that have been compacted below and/or above the compaction window.

5 Implementation and Case Study

A prototype is built to test and validate the proposed framework. In this prototype, the data collection, integration, and strategy assessment phases are performed with Excel. Whereas the simulation phase is developed with AnyLogic [18], and the VR environment is built with 3D Unity [19].

In short, the parameters shown in Table 1 are collected and the initial values of the decision variables are computed, both with an excel-based Graphical User Interface (GUI). Then, AnyLogic reads the data from excel and performs the ABS simulation. Once the simulation is finished, AnyLogic provides to Excel the cyclic timestamped location of the equipment. Next,

Excel integrates the data and generates the input for the VR environment. Then, Unity creates the VR environment and allows the user to evaluate their strategy. Finally, excel generates an output PDF file with the strategy assessment values.

To evaluate the performance of this model, the case study was performed for a 250 m surface rehabilitation of the A-15 highway in Rotterdam, The Netherlands. The total allocated time to execute the job is one hour. In this project, the temperature window of 120~80 °C was specified. The other input values are shown in Table 2. It is assumed that at the start of the project the trucks are already on the construction site.

Table 2. Data collection summary

Parameter	Value
Available time	1h
Road length	250m
Road width	8m
Layer thickness	50mm
Available time for compaction	16min
Minimum time for start compaction	9min
Truck capacity	27t
Transition time	3min
Paver quantity	1
Width	8m
Initial speed	4m/min
Roller quantity	1-2
Roller width	2m
Number of passes	2
Number of lanes	4
Initial speed	18m/min
Length of roller track	73m

Four different compaction strategies can be evaluated, as shown in Figure 4. In scenario a, the compaction is performed from one edge of the road to the other. In Scenario b, the roller compacts the inside lanes first and leaves the outer lanes to be compacted last. This would allow the edges to be slightly cool down before the compaction and thus have a smoother edge [17]. In Scenario c, the first roller, or master roller, is the one that leads and the second roller, or slave roller, follows. Finally, in Scenario d, two rollers compact in parallel each from one edge inward. For this case study, given the

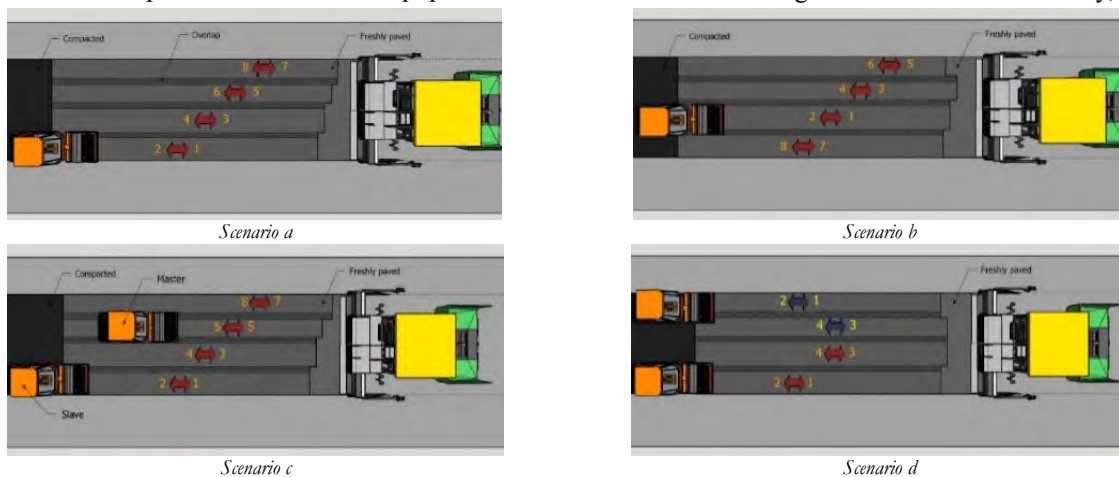
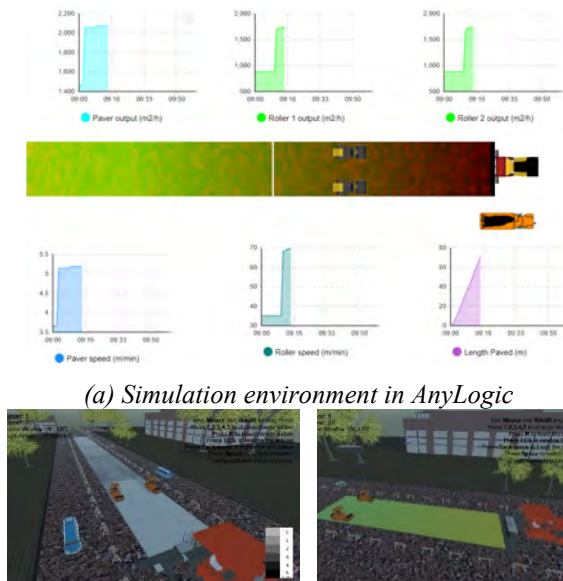


Figure 4. Different compaction strategies

characteristics of the actual project, Scenario d was selected. The simulation model is constructed in AnyLogic and includes all the agents presented in Figure 3. The model reads the input parameters from the excel-based GUI. The simulation stops at the specified target time. Different interactive graphs present the areal output, the equipment speed, and the total surface paved, as shown in Figure 5(a).

Once the simulation is completed, the integration part takes place with the help of the GUI in Excel. During this process, the data generated by Anylogic are imported into Excel, where they are translated to the format readable by Unity. Afterward, the user runs the Unity executable file. Within this environment, it is possible to assess the compaction completeness and the cooling of the asphalt mat. Figure 5(b) presents the snapshots of the VR model.

The user can navigate through the VR model and slow and speed up the process to find the potential bottlenecks. In the end, the analysis of the entire operation, which incorporates the results of both spatial and temporal analysis is presented to the user, as shown in Figure 6.



(a) Simulation environment in AnyLogic



(b) VR representation

Figure 5. System Interface

In terms of accuracy, the simulated results are compared to the actual project statistics. Table 3 presents the result of this comparison. As shown in this table, the average estimation error is about 9%, i.e., considering rollers and pavers together. This would represent high estimation accuracy.

Also, the usability of the proposed method was assessed through a workshop with expert planners. The experts were asked to assess the current and proposed planning method in terms of user-friendliness, usefulness, versatility, teamwork between tactical and strategic planners, and helping planners become more aware of the

compaction process. Figure 7 presents the result of the user assessment. In this figure, orange line represents the current situation and the blue line represents the proposed method.

6 Conclusions and future work

This research offered a framework for the integration of simulation of asphalt-pavement compaction and VR. A review of the asphalt compaction practices and a detailed description of the framework for capturing the relevant factors and providing productivity and operational quality feedback was presented. A prototype was developed, and a previous actual case study was used to demonstrate the feasibility of the proposed framework. The prototype was presented to asphalt experts and it was shown that the integrated model has a great potential to improve the operational quality and planning of compaction operations. The model was also found to have high estimation accuracy.

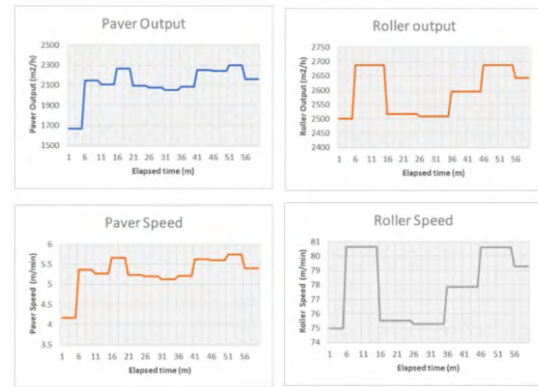


Figure 6. Report of the strategy assessment

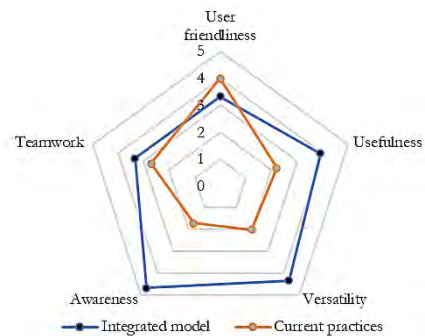


Figure 7. Usability assessment of the framework

It is shown that historical data and collected data from the site can be synthesized in an integrated model. It is demonstrated that this data can be converted into computer agents able to replicate the thermal behavior of the asphalt mat as well as the equipment operator's behavior. However, there are some limitations in the

current research. The prototype should be further tested in more case studies, and its validity needs to be further assessed. Also, the number of compaction strategies supported in the model is limited at the moment. It is highly advised to collect real data from more compaction operations to build a comprehensive library of compaction strategies.

Table 3. Simulation results Vs. the actual data

Index	Actual	Simulated	Error rate
Roller output (m ² /h)	2790	2629	6%
Paver output (m ² /h)	2000	2180	9%
Roller speed (km/h)	5.00	4.74	5%
Paver speed (m/min)	5.00	5.45	9%
Compaction efficiency	20%	23%	15%

Moreover, bearing in mind the quickness of Artificial Intelligence (AI) evolution, the prototype could eventually use AI methods for automatic reasoning and, more interestingly, the development of AI-based compaction strategies.

References

- [1] Hughes C.S. *Compaction of asphalt pavement*, volume 152 of Synthesis of highway practice. National Research Council, Transportation Research Board, Washington DC, 1989.
- [2] Beainy F. Commuri S. Zaman M. Asphalt compaction quality control using Artificial Neural Network. In *Proceedings of 49th IEEE Conference on Decision and Control (CDC)*, pages 4643-4648, Atlanta, United States of America, 2010.
- [3] Vasenev A. Hartmann T. Dorée A. Employing a Virtual Reality Tool to Explicate Tacit Knowledge of Machine Operators. In *Proceedings of the 30th International Symposium on Automation and Robotics in Construction and Mining (ISARC 2013)*, pages 248-256, Montréal, Canada, 2013.
- [4] Bijleveld F. *Professionalising the asphalt construction process: Aligning information technologies, operators' knowledge and laboratory practices*. University of Twente, Enschede, 2017.
- [5] Miller S.R. *Hot Mix Asphalt – Towards a more professional approach*. University of Twente, Enschede, 2010.
- [6] Miller S.R. Dorée A. Improving logistics in the asphalt paving process-what can we learn from the planner's logic? In *Proceedings of the 24th Annual ARCOM Conference*, pages 381-390, Cardiff, United Kingdom, 2008.
- [7] ter Huerne H.L. *Compaction of asphalt road pavements: using finite elements and critical state theory*. University of Twente, Enschede, 2004.
- [8] Nassar K. Thabet W. Beliveau Y. Simulation of asphalt paving operations under lane closure conditions. *Automation in Construction*, 12:527-541, 2003.
- [9] Retik A. Shapira A. VR-based planning of construction activities. *Automation in construction*, 8(6):671-680, 1999.
- [10] Rekapalli P.V. Martinez J. Discrete-event simulation-based virtual reality environments for construction operations: Technology introduction. *Journal of Construction Engineering and Management*, 137(3): 214-224, 2011.
- [11] Turner S.F. Cardinal L.B. Burton R.M. Research Design for Mixed Methods. *Organizational Research Methods*, 20:243-267, 2016.
- [12] Akpan I.J. Shanker M. Razavi R. Improving the success of simulation projects using 3D visualization and virtual reality. *Journal of the Operational Research Society*. 71(12):1900-1926, 2020.
- [13] Blessing L.T. Chakrabarti A. *DRM: A design research methodology*. Springer, London, 2009.
- [14] Makarov D.S. Vahdatikhaki F. Miller S.R. Dorée A. A Methodology for Real-Time 3D Visualization of Asphalt Thermal Behaviour During Road Construction. *Advances in Informatics and Computing in Civil and Construction Engineering*, 797-804, Springer, Cham, 2019.
- [15] Vahdatikhaki F. El Ammari K. Langroodi A.K. Miller S. Hammad A. Dorée A. Beyond data visualization: A context-realistic construction equipment training simulators. *Automation in construction*, 106:102853, 2019.
- [16] Timm D.H. Voller V.R. Lee E.B. Harvey J. Calcool: a multi-layer asphalt pavement cooling tool for temperature prediction during construction. *International Journal of Pavement Engineering*, 2(3):169-185, 2001.
- [17] Kloubert H. *Basic Principles of Asphalt Compaction: Compaction Methods, Compaction Equipment, and Rolling Technique*. BOMAG FYAT Group, 2009.
- [18] Abar S. Theodoropoulos G.K. Lemarinier P. O'Hare G. M. Agent Based Modelling and Simulation tools: A review of the state-of-art software. *Computer Science Review*, 24:13-33, 2017.
- [19] Yang K. Jie J. The designing of training simulation system based on unity 3D. In *Proceedings of the Fourth International Conference on Intelligent Computation Technology and Automation*, 1:976-978, Shenzhen, China, 2011.
- [20] Runneboom, P. Alle assets op een rij, de basis voor parkmanagement : een beter en sneller inzicht in de aanwezige assets voor Klapwijk Parkmanagement op bedrijventerreinen, *Master Thesis, University of Twente*, 2018

Direction of Arrival of Equipment Sound in the Construction Industry

K. Elelu^a, T. Le^a and C. Le^b

^aGlenn Civil Engineering Department, Clemson University, USA.

^bDepartment of Civil, Construction and Environmental Engineering, North Dakota State University, USA

E-mail: kelelu@g.clemson.edu, tuyenl@g.clemson.edu, chau.le@ndsu.edu

Abstract

The construction industry is often affected by unanticipated struck-by accidents, which often cause severe injuries and fatalities to the workers. Therefore, monitoring and tracking struck-by hazards in terms of the spatial relationship between a worker and a heavy vehicle is crucial to prevent such accidents. Current studies focus on using active sensors and implementing computer vision but not on the audibility of their safety signals. To address this issue, this paper utilizes sound, a ubiquitous data source present in every construction site, to track and separate equipment sound into different types and determine the direction of arrival (DOA) using the Open embedded Audition System (ODAS) framework. Each circular array performs DOA estimation independently using commercial software on two equipment sound sources, bulldozer (mobile) sound and hammer (stationary) sound. The DOAs are fed to a relational database, pre-processed, and used to perform the source tracking. This process provides a step towards monitoring the spatial relationship between workers and equipment with few labels of source location for calibration. The results of our study showed that this method was effective in identifying activities of multiple pieces of equipment in real-time in construction sites without the need for separating sound signals in advance. Future studies can focus on triangulating the exact location of the sound source with less computation power and monitoring how this helps improve workers' awareness of surrounding equipment.

Keywords –

Direction of Arrival; Machine Equipment; Struck-by Accident; Local coordinate

1 Introduction

For a long time, researchers have considered struck-by

heavy vehicle accidents as one of the leading causes of death in the labor-intensive construction industry [1]. This is mostly due to the unpredictable nature of construction equipment and limited workspaces resulting in lifelong impairment, and fatalities [2]. Bureau of Labor Statistics in 2020 published that 15% of fatal occupational injuries resulted from hazardous contact with equipment and objects [3]. Also, according to Occupational Safety and Health Administration (OSHA) figures, struck-by-equipment hazards accounted for 58 % of struck-by incidents from 1995 to 2008 [4]. Therefore, contact collisions between construction workers on foot and equipment pose a significant risk to the safety and health of construction workers. However, the capacity to spot these collision accidents in advance on a site is vital to any construction project since preventing unanticipated catastrophes is always the best way to avoid them. Therefore, an automated safety monitoring of struck-by hazards has emerged as a potential option for effective safety management on a construction site [5]. In this paper, the authors propose the use of audio sound to extract important and useful information regarding equipment activities performed in the site. This is possible as audio-based activity identification is very easy to collect regardless of dynamic occlusions and different tools have signature distinguishable sounds which make it suitable for task identification [6].

2 Related Studies

Some past work has focused on the usage of visual sensors for localizing and characterizing the behavior of resources and has been extensively applied in the construction industry and achieved promising results by various methods [7]. Several active technologies have also been developed to provide proximity sensing and alerts for workers and equipment operators, such as image wearable devices [8], tactile-based wearable devices [9], Bluetooth-low energy [7], and global positioning system [11]. Some studies have used GPS

data for long-distance detection of collision hazards. For example, an unsafe-proximity detection model focused on decreasing false alarms has been developed using a GPS-aided inertial navigation system sensor as the state tracking module [12]. Another GPS-related method for recording, identifying, and analyzing interactive hazardous near-miss situations between workers-on-foot and heavy construction equipment was presented using spatiotemporal data [13]. Shorter sensing devices like Bluetooth Low Energy (BLE)-based proximity sensing have been applied to address work-zone safety to allow understanding of dynamic spatial relationships among equipment, operator, workers, and a surrounding work environment [10]. Also, a spatiotemporal network-based model was developed at both entity and network levels to perform dynamic risk analysis on the struck-by-equipment hazard [14]. Sakhakarmi et al. develop a proximity sensor using a wearable tactile-based system for workers to help improve their hazards perception [9]. These smart and automated systems are effective but prone to false alarms. Wang et al. [12] presented two novel four-dimensional models, a time-sphere model and a time-cuboid model that effectively reduces the rate of false alarms; still, the data not wirelessly communicated made it challenging to apply in real-life situations. There is also a growing application of artificial intelligence in safety management such as using computer vision technology in image detection to monitor a comprehensive view of the area surrounding equipment via cameras installed on every side of the equipment body. This allows displaying the surrounding environment on the operator's monitor to protect the workers from potentially dangerous situations involving equipment operations [15]. Researchers also developed a digital twin using the Bayesian network that fed and updated real-time data from sensors that pro-actively forecast dangerous scenarios based on collision probability on affected workers-on-foot [16]. Most of these methods were implemented by installing sensors or electronic devices such as cameras directly on equipment or construction workers. Besides, those devices are relatively expensive; thus, they cannot be used on a wide scale in reality. Also, some base the alert detection on the information obtained from available hazard detection records, as more factors can cause dangers to workers that are not factored into the study. To address the gaps, there is a need to investigate the surveillance approach that is less expensive and adaptable to rigorous construction effect like during the excessive dust in visual equipment monitoring. Auditory surveillance using the sound collected from construction sites could address these issues. However, there is a lack of such a method of processing audio signals for equipment monitoring and safety against collision hazards in the construction field.

A few efforts have been made in other sectors' research in auditory surveillance for collision hazard detection. A real-time framework was created to detect multi-vessel collisions using a spatial clustering process to detect clusters of encounter vessels within each cluster from the vast number of monitored vessels in a surveyed sea area [17]. The framework effectively and efficiently detects encounter vessels and ranks collision risk indexes within each cluster. Another method for detecting collision hazards amongst motorcycles via accelerometer measures was proposed using a machine learning model [18]. The system was designed using data from an instrumented vehicle and validated in simulation. Nonetheless, research on auditory surveillance focusing on detecting collision hazards in construction job sites has not gained much attention in the research community.

Acoustic emissions from construction activities were used to calculate the working hour to allow field managers to know workers' work progress and productivity [19]. Success from this study shows how using auditory surveillance can prove positively in detecting equipment sound type. Another using surveillance technology in the construction industry helps support the construction industry's safety performance since the lack of sufficient visibility is the principal factor leading to fatalities [20]. As stated earlier, most sound-based surveillance technologies were only focused on monitoring construction work activities and equipment operations. For instance, a hybrid system was proposed for recognizing multiple construction equipment activities [18]. A supervised machine learning-based sound identification algorithm was implemented to enhance monitoring and performance of construction site activity [21]. A few studies attempted to develop new approaches for conducting an audio-based event detection system for safety. Experimental trials were designed to deploy sensing technology to provide alerts to proximity detection when heavy construction equipment and workers are in close proximity [22]. Nonetheless, the devices were installed on construction equipment only, not equipped on construction workers. Another approach using a machine learning algorithm can categorize sound events and make construction workers aware of possible safety risks and hazards [23]. Still, the sound data relating to collision hazards were only collected from a particular worksite. Such an approach is restrained because the sounds emitted by the equipment from various construction sites may differ and contribute a different amount of noise. Studies have been done on determining the optimum position for construction noise barrier location, and a comparison was made during each stage of the construction process. All in all, there is a need to develop an audio-based localization of stationary and mobile equipment framework. In doing so, this paper particularly focuses

on extracting the local coordinate of the sound source and process this data to locate the position of construction equipment.

3 Methodology

This study introduces a framework for audio base activity detection to prevent the struck-by construction equipment hazards, which is illustrated in Figure. 1. This framework consists of three phases that include equipment sound extraction, data processing for DOA estimation, and worker's danger notification signal based on distance computed. First, the extraction of the sound record is done in real-time, which is possible with the microphone array mounted on the raspberry pi. The second phase has these audio data separated into different equipment sounds, containing each sound's identification with their local coordinate value uploaded into the database for data cleaning and processing. This is done by separating the mobile and stationary equipment with their unique sound identification and entering each sound type into a separate relational database where the DOA of the sound from the workers is estimated. Lastly, based on the DOA, the location and distance from workers are calculated as shown in equation 7 to notify the workers based on the danger zone specified. Each of these steps is described in further detail below.



Figure 1: The overall framework for audio-based activity identification

3.1 Sound Processing Pipeline

We employ a ReSpeaker Mic Array v2.0 device with four high-performance digital microphones and twelve configurable RGB LED indications (see Figure 2) coupled to a Raspberry Pi 3 processor. This RP3 runs its own instance of the ODAS framework, which outputs direction of arrival (DOA) represented by a 3D unit vector in the array's local coordinate system, as shown in Figure 3. The DOA of the arrays lies on a virtual unit hemisphere with the z-axis facing up and lies on the positive side and is defined in equation 1 [24].

$$e = [e_x \ e_y \ e_z]^T \quad (1)$$

where $e_x^2 + e_y^2 + e_z^2 = 1$, and $e_z > 0$.



Figure 2: Experiment Setup

To maximize DOA localization on the microphone array, MacBook Pro and Android smartphone speakers are utilized to play a bulldozer (mobile) sound and a hammer (stationary) sound, respectively. The loudness of the bulldozer and hammer sound using a decibel-meter is in the range of 75dB to 79dB. The ODAS sound source tracking module is designed to detect both static and moving sources. Although it is capable of tracking up to 4 sources at this time, we only make use of two sound sources for this experiment. A single microphone array is installed in a controlled space (see Figure 4 for calibration locations). This array is positioned in the top-left middle part of the room, while the mic array is oriented with the z-axis facing the ceiling and the x-axis facing the northeast direction. It should be noted that the processing is independent and unaffected by the array's positions and orientations. These relations are useful for interpreting the findings and are not employed to locate the sound source. Microphone arrays are usually omnidirectional, which make their orientation relevant [23], and the multi-channel raw audio is sampled at 44,100 Hz from the ReSpeaker array, resampled by ODAS at 16,000 Hz, which then returns an updated DOA estimation [24]

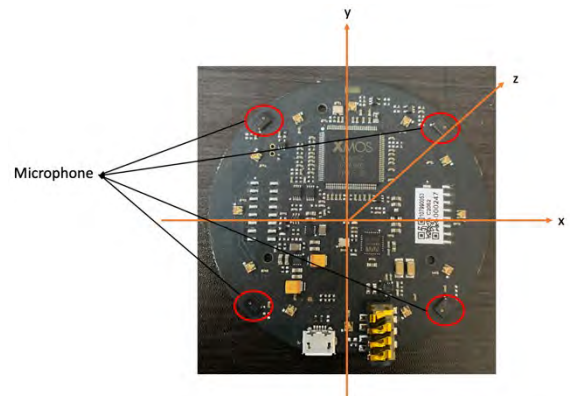


Figure 3: ReSpeaker Microphone Array

We also estimate the angle in degrees to determine the direction in a cartesian plane with the x and y local coordinate collected using equation 2.

$$\theta = \tan\beta - 1 \left(\frac{y}{x} \right) * \left(\frac{180}{\pi} \right) \quad (2)$$

3.2 Experimental Set-up

This section describes how we set up our experiment as well as the technique we used to determine the DOA of construction equipment sound. Firstly, a bulldozer sound device is placed sequentially at four different locations away from the microphone array, as shown in Figure 4; after every 30sec each, the microphone array is moved to these new points to record a 3-component DOA of approximately 130 records estimated every 1s. The device is moved from position 1 to a new position 2, and this process is repeated for positions 3 and 4 while keeping the hammer sound at a single position 4 for the entire duration of the experiment. The procedures are repeated five times, and the measurements were performed in nearly a horizontal plane. After the experiment's conclusion in a controlled environment free from noise and environmental factors, a DOA local coordinate record between the range of 15,100-15,200 for both the stationary and mobile sound was recorded. A timer on an iPhone device is synchronized with the time on RP3. When left running continuously, each calibration point is estimated to record 3700-4000 for the 30s. The timestamp, id, energy, and local coordinate of DOA estimates are saved locally as a text file which is pushed to the cloud and loaded to a relational database for data cleaning and pre-processing. The data is then sorted and joined based on the time stamp. The result is a chronological table that contains the DOAs from all arrays synchronized using the timestamps. This table can be queried to get the full information for any period of time. ODAS tracks the loudest sound source and records the DOAs with energy greater than zero as dynamic. Once the experiment can verify the direction of arrival of construction equipment, which was done in this study, the experiment will be further expanded to detect collision hazard using two microphone array set-ups on the construction site. The first mic-array is a fixed at a point, and the second mic-array is a mobile mounted on the construction workers. To pinpoint the location and distance of the equipment from the worker, we considered several scenarios to present the derivation of the distance to the worker. One of the case scenarios is when a bulldozer and hammer equipment sound are played simultaneously, the bulldozer changes location during the period of operation, if a construction worker is in the danger zone, the framework should be able to notify the

workers of the current status and vice versa. The performance will be tested both in a control environment and a construction site to examine the amount of false alarm generated from the notification which will help to determine wide range application on the construction site.

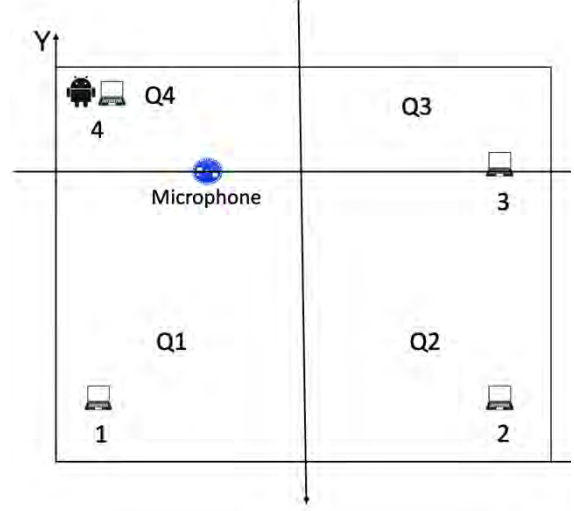


Figure 4: Position 1 to 4 and the location of microphone array.

4 Result and Discussion

As mentioned in Section 3, two sound devices were used for five different scenarios considering equipment like bulldozer and hammer. Figure 5 present the performance analysis result for all the scenarios, and its shows the bulldozer sounds being moved manually from calibration point one to four. The first position of the mobile device is at the positive x and y coordinate in Q1 with a value above the zero line. At position two, the coordinate enters the second quadrant Q2, with the x-axis changing to a negative value slightly below the zero point, as shown in Figure 4. At position three, even though there is a drop close to the zero line, it shows a positive y-axis and still slightly occupying the second quadrant's space. The fourth position shows a positive x-axis and a negative y-axis with values above zero and below zero, respectively. Lastly, the DOAs at experiment four at some point shows some roughness due to slight environmental disturbance and other issues. This is due to a substantial disturbance of the sound source at and near point three from the audio being played, confirmed by re-playing the audio sound along with that duration. A slightly straight line is observed for the stationary hammer sound as this maintained position four without any location changes.

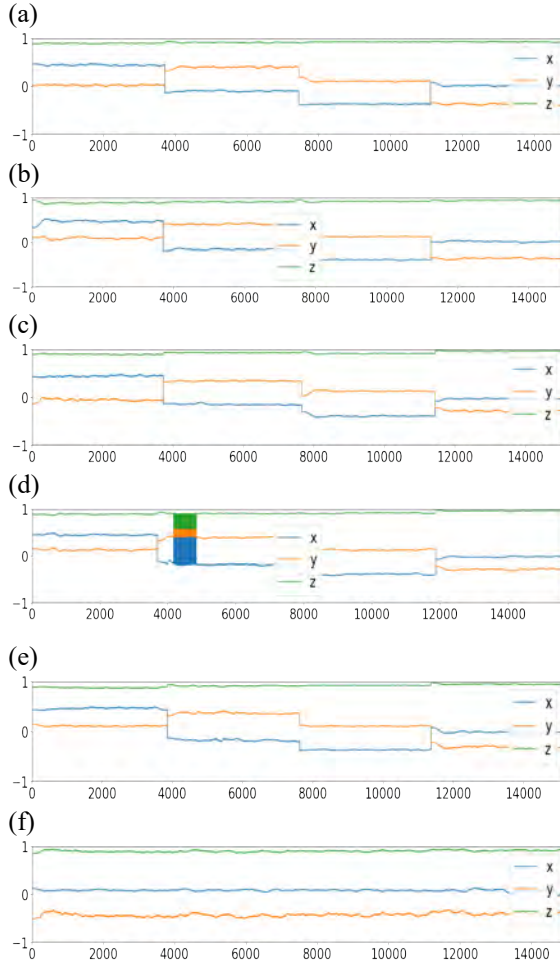


Figure 5: (a-e) Mobile sound DOAs measured from the five arrays and shown measurement from point 1 to 4 and (f) Stationary sound DOAs measured at a single location.

4.1 Angle of Arrival of Sound

This experiment shows we considered four planar wavefronts at direction-of-arrival of 1.47° , 106° , 163° , and 268° for points one, two, three, and four, respectively. The fourth point for mobile and the stationary device were coherent, as they were positioned at the same location in the experiment space.

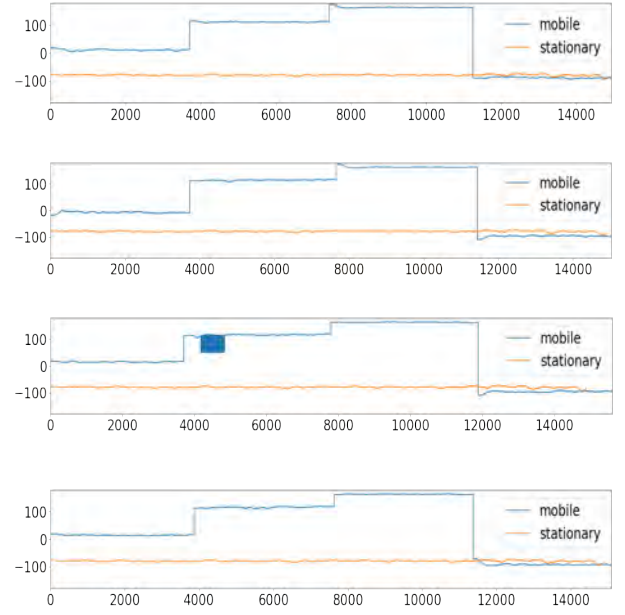
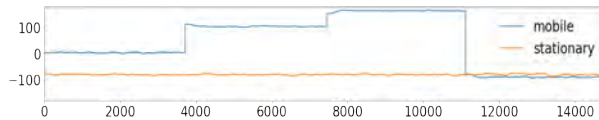
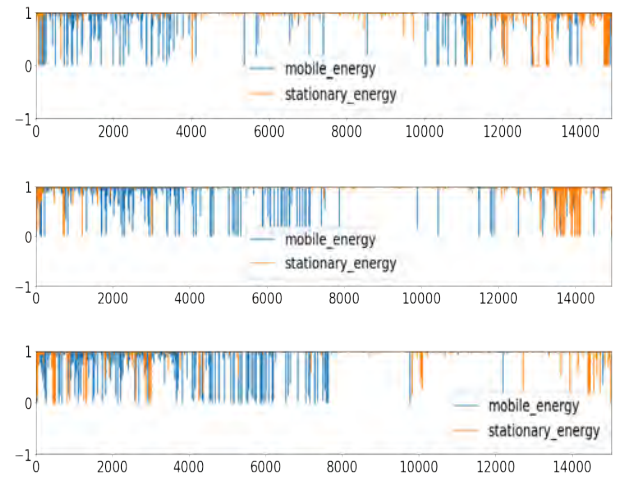


Figure 6: Mobile and stationary sound angle measurement in degree for the five arrays and shown measurement along the point 1 to 4

4.2 Energy Timeseries

The framework gives additional information about the loudness of the sound in a normalized form, with a scale from 0 to 1 for both the mobile and stationary sounds. Figure 7 shows the high stationary energy value, which is attributed to the uniqueness in the sound type, which is a loud intermittent sound that is not affected by sand or concrete noise. The end position of the experiment has a relatively lower energy value for the hammer sound; playing back the audio to confirm the information, we notice a decrease in sound while in operation.



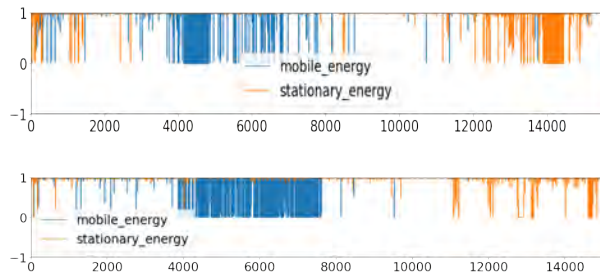


Figure 7: Energy propagation for experiment 1 to 5

4.3 Local coordinate mapping

In each of the DOA local coordinates for the x and y-axis, the projected shape, and the movement of bulldozer sound along the point indicated are recognized, allowing for determination of the movement with few labels on the initial source location. The magnitude of the vector's entries shows a deeper presence at these four calibration points as the device maintains some 30seconds—also, a path of the movement of the equipment sound. Since the experiment has entries for the five sets of measurements, their mapping is different and color-coded, as shown in Figure 8.

4.4 Discussion

The ODAS framework exploits the directivity of microphones to compute generalized cross-correlation phase transform (GCC-PHAT) between pairs of microphones [25] and compute time delay of arrival (TDOA) based on the microphones that are close to each other, which indirectly impacts the performance as an array with less microphone and evidently reduce the accuracy of detection [24]. Equipment sound like a hammer that has strong intermittent sound is easily picked up by the microphone and produces high energy value compared to continues prolong sound like a bulldozer which level of sound can be highly impacted by the materials it is working on and the condition of the machine. The fast, optimized processing strategies make it possible for this framework to perform all processing on low-cost hardware like Raspberry Pi 3, which is cheaper and more economical than the already available sound detection device. Lastly, due to the scope of this study, one particular limitation of this research is that the system does not account for noise filtering on the reSpeaker v2, and further work is needed to be done to eliminate background noise disrupting the sound capture.

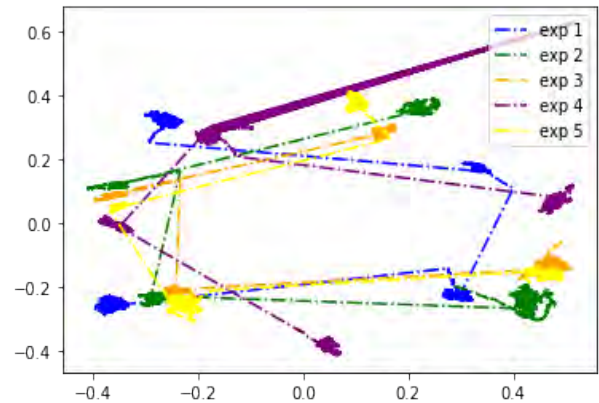


Fig 8: Equipment mapping of positions 1 to 4 showing the movement of equipment sound along the line.

5 Conclusion

This research focused on pro-active struck-by hazard detection in construction and introduced an approach to simulate a sound source localized from a series of local coordinate system arrays. Each computed array performs the DOA estimation independently and feeds its measurements to a data center where the DOA from all arrays are timestamped and preprocessed for data exploration and analysis. The method was demonstrated with a single circular microphone array in a controlled environment. The methods demonstrated here provide a step towards monitoring activities in construction sites with no training effort, as the device has an inbuilt algorithm to separate four sound sources. To make the device more stream-lined, we plan to design a flexible device to allow to be able to mount it on construction workers and will ultimately contribute to promoting a safer working environment for construction workers.

References

- [1] W. Wu, H. Yang, Q. Li, and D. Chew, "An integrated information management model for proactive prevention of struck-by-falling-object accidents on construction sites," *Automation in Construction*, vol. 34, pp. 67–74, 2013, doi: 10.1016/j.autcon.2012.10.010.
- [2] B. Shao, Z. Hu, Q. Liu, S. Chen, and W. He, "Fatal accident patterns of building construction activities in China," *Safety Science*, vol. 111, pp. 253–263, Jan. 2019, doi: 10.1016/j.ssci.2018.07.019.
- [3] Bureau of Labor Statistics, "NATIONAL CENSUS OF FATAL OCCUPATIONAL INJURIES IN 2020," 2020. [Online]. Available:

- www.bls.gov/iif/oshwc/cfoi/cfoi_rates_2020hb.xlsx.
- [4] W. Wu, H. Yang, D. A. S. Chew, S. hua Yang, A. G. F. Gibb, and Q. Li, "Towards an autonomous real-time tracking system of near-miss accidents on construction sites," *Automation in Construction*, vol. 19, no. 2, pp. 134–141, Mar. 2010, doi: 10.1016/j.autcon.2009.11.017.
- [5] J. Park, K. Kim, and Y. K. Cho, "Framework of Automated Construction-Safety Monitoring Using Cloud-Enabled BIM and BLE Mobile Tracking Sensors," *Journal of Construction Engineering and Management*, vol. 143, no. 2, p. 05016019, Feb. 2017, doi: 10.1061/(asce)co.1943-7862.0001223.
- [6] K. M. Rashid and J. Louis, "Activity identification in modular construction using audio signals and machine learning," *Automation in Construction*, vol. 119, Nov. 2020, doi: 10.1016/j.autcon.2020.103361.
- [7] D. Roberts and M. Golparvar-Fard, "End-to-end vision-based detection, tracking and activity analysis of earthmoving equipment filmed at ground level," *Automation in Construction*, vol. 105, Sep. 2019, doi: 10.1016/j.autcon.2019.04.006.
- [8] K. Kim, H. Kim, and H. Kim, "Image-based construction hazard avoidance system using augmented reality in wearable device," *Automation in Construction*, vol. 83, pp. 390–403, Nov. 2017, doi: 10.1016/j.autcon.2017.06.014.
- [9] S. Sakhakarmi, J. W. Park, and A. Singh, "Tactile-based wearable system for improved hazard perception of worker and equipment collision," *Automation in Construction*, vol. 125, May 2021, doi: 10.1016/j.autcon.2021.103613.
- [10] J. Park, X. Yang, Y. K. Cho, and J. Seo, "Improving dynamic proximity sensing and processing for smart work-zone safety," *Automation in Construction*, vol. 84, pp. 111–120, 2017, doi: https://doi.org/10.1016/j.autcon.2017.08.025.
- [11] N. Pradhananga and J. Teizer, "Automatic spatio-temporal analysis of construction site equipment operations using GPS data," *Automation in Construction*, vol. 29, pp. 107–122, 2013, doi: https://doi.org/10.1016/j.autcon.2012.09.004.
- [12] J. Wang and S. N. Razavi, "Low False Alarm Rate Model for Unsafe-Proximity Detection in Construction," *Journal of Computing in Civil Engineering*, vol. 30, no. 2, p. 04015005, Mar. 2016, doi: 10.1061/(asce)cp.1943-5487.0000470.
- [13] O. Golovina, J. Teizer, and N. Pradhananga, "Heat map generation for predictive safety planning: Preventing struck-by and near miss interactions between workers-on-foot and construction equipment," *Automation in Construction*, vol. 71, pp. 99–115, 2016, doi: https://doi.org/10.1016/j.autcon.2016.03.008.
- [14] J. Wang and S. Razavi, "Spatiotemporal Network-Based Model for Dynamic Risk Analysis on Struck-by-Equipment Hazard," *Journal of Computing in Civil Engineering*, vol. 32, no. 2, p. 04017089, Mar. 2018, doi: 10.1061/(asce)cp.1943-5487.0000732.
- [15] H. Son, H. Seong, H. Choi, and C. Kim, "Real-time vision-based warning system for prevention of collisions between workers and heavy equipment," *Journal of Computing in Civil Engineering*, vol. 33, no. 5, p. 4019029, 2019.
- [16] M. Leonardo, N. Berardo, C. Alessandro, R. Luigi, and D. G. Giuseppe Martino, "Development of a Digital Twin Model for Real-Time Assessment of Collision Hazards," 2020, pp. 14–19. doi: 10.3311/ccc2020-003.
- [17] R. Zhen, M. Riveiro, and Y. Jin, "A novel analytic framework of real-time multi-vessel collision risk assessment for maritime traffic surveillance," *Ocean Engineering*, vol. 145, pp. 492–501, 2017, doi: https://doi.org/10.1016/j.oceaneng.2017.09.015.
- [18] D. Selmanaj, M. Corno, and S. M. Savaresi, "Accelerometer-based data-driven hazard detection and classification for motorcycles," in *2014 European Control Conference (ECC)*, 2014, pp. 1687–1692. doi: 10.1109/ECC.2014.6862549.
- [19] C. Cho, Y.-C. Lee, T. Zhang, and P. D. Student, "Sound Recognition Techniques for Multi-Layered Construction Activities and Events."
- [20] J. W. Hinze and J. Teizer, "Visibility-related fatalities related to construction equipment," *Safety Science*, vol. 49, no. 5, pp. 709–718, Jun. 2011, doi: 10.1016/j.ssci.2011.01.007.
- [21] T. Zhang, Y.-C. Lee, M. Scarpiniti, and A. Uncini, "A supervised machine learning-based sound identification for construction activity monitoring and performance evaluation," in *Construction Research Congress 2018*, 2018, pp. 358–366.
- [22] E. D. Marks and J. Teizer, "Method for testing proximity detection and alert technology for safe construction equipment operation," *Construction Management and Economics*, vol. 31, no. 6, pp.

- 636–646, Jun. 2013, doi: 10.1080/01446193.2013.783705.
- [23] Y. C. Lee, M. Shariatfar, A. Rashidi, and H. W. Lee, “Evidence-driven sound detection for prenotification and identification of construction safety hazards and accidents,” *Automation in Construction*, vol. 113, May 2020, doi: 10.1016/j.autcon.2020.103127.
- [24] P. Gerstoft *et al.*, “Audio scene monitoring using redundant ad-hoc microphone array networks,” *IEEE Internet of Things Journal*, pp. 1–1, Aug. 2021, doi: 10.1109/jiot.2021.3103523.
- [25] F. Grondin *et al.*, “ODAS: Open embeddeD Audition System,” Mar. 2021, [Online]. Available: <http://arxiv.org/abs/2103.03954>

Towards BIM-based robot localization: a real-world case study

H. Yin^{1,2}, J.M. Liew¹, W. L. Lee¹, M.H. Ang Jr¹ and J. K.W. Yeoh¹

¹College of Design and Engineering, National University of Singapore, Singapore

²Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, HKSAR

eehyin@ust.hk, mpejml@nus.edu.sg, englwl@nus.edu.sg, mpeangh@nus.edu.sg, justinyeoh@nus.edu.sg

Abstract -

Conventional mobile robots rely on pre-built point cloud maps for online localization. These map points are generally built using specialized mapping techniques, which involve high labor and computational costs. While in the architectural, engineering and construction (AEC) industry, as-planned building information modelings (BIM) are available for management and operation. In this paper, we consider the use of the digital representations of BIM for robot localization in built environments. First, we convert BIM data into localization-oriented point clouds, which is easy to implement and operate compared to relatively complex SLAM systems. Then, we perform iterative closest point (ICP)-based localization on the metric map using a laser scanner. The experiments are tested using collected laser data and BIM in the real world. The results show that ICP-based localization can track the robot pose with low errors ($< [0.20\text{m}, 2.50^\circ]$), thus demonstrating the feasibility of BIM-based robot localization. In addition, we also discuss the reasons for errors, including the deviations between as-planned BIM and as-built status.

Keywords -

BIM; Robot Localization; LiDAR; Deviation

1 Introduction

Precise localization is a fundamental capability for mobile robots. Almost all mobile robots, whether teleoperated or autonomous, require the robot pose to be estimated by the localization module to achieve safe human operation or self navigation in complex environments. In the robotics community, many simultaneous localization and mapping (SLAM) systems [1, 2] have been developed to achieve both mapping and localization when the robot is traveling in the real world.

For some long-term robots that operate under stable conditions such as a quadruped robot working on building inspection, the mapping process of SLAM is redundant because the generated map is almost invariant in each time of SLAM. To solve this problem, researchers in the robotics community proposed to achieve mapping first and then

robot localization in the known map [3, 4]. In this context, map building is required only once and localization in the map could handle the pose estimation for long-term operation, reducing the complexity of repetitive SLAM processes. This two-step workflow has been widely used in various fields of robotics and a typical application is self-driving cars [5].

The mapping step in this workflow is generally based on SLAM or other techniques, which can be considered as a measuring or sensing process of the environment. However, some modeling or representations are directly available in the AEC industry, such as computer-aided design (CAD) or BIM. These map-like representations also contain informative measurements. Thus, we hypothesize SLAM may not always be necessary when these are available. Moreover, BIM has been raised to replace CAD in recent years. We believe that robot localization in a as-designed BIM could be a good choice in built environment.

One might argue that BIM is designed for construction and building management, which is not a localization-oriented map essentially. In this paper, we present a BIM-to-Map process to convert the digital representations of BIM into point cloud maps for robot localization. We also utilize a point-to-plane ICP-based method to localize the robot on the BIM-generated map, thus bridging the gap from design modeling to robotic navigation in the real world. In addition, there are deviations between as-planned BIM and as-built buildings in the real-world, which brings potential difficulties to online robot localization. To address this problem, we present a real-world case study to test BIM-based robot localization using a rotating Light Detection and Ranging (LiDAR) scanner. Overall, the contributions can be summarized as follows:

- A BIM-based robot localization workflow is presented to achieve precise pose estimation in the built environment. The prior maps are built with BIM-to-Map conversion without complex SLAM systems.
- We conduct experiments in the real world. The experimental results show that the proposed workflow can track the robot pose with only a LiDAR scanner

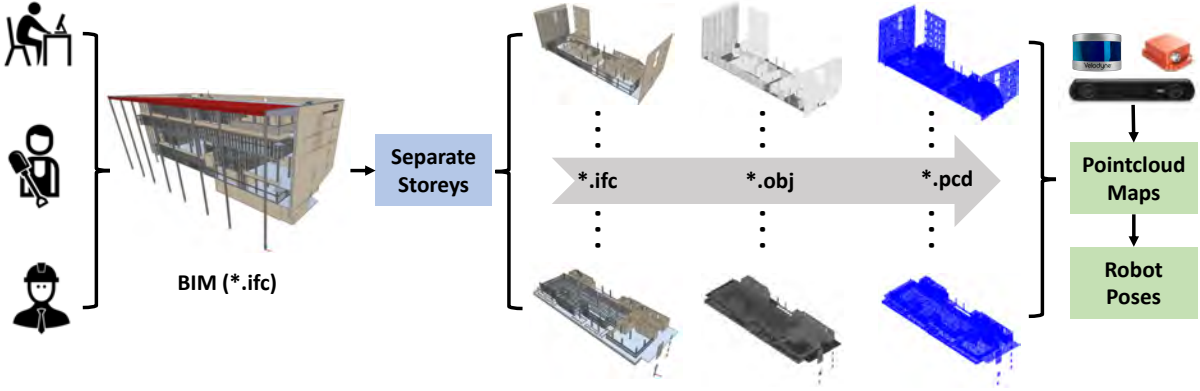


Figure 1. The workflow of BIM-based localization in this paper. We first convert BIM to metric point cloud maps storey by storey. Then based on the localization-oriented maps, mobile robot localization can be achieved using sensor units. In this paper, we perform a case study on 3D LiDAR localization in BIM at NUS campus.

and BIM.

2 Related Work

Many research publications have reviewed related robot navigation topics from different standpoints, including deep learning-based [6], specific sensor-based [7, 8, 9], etc. In this paper, we present some related work on CAD or BIM-based mobile robot navigation.

Intuitively, floor plans or 2D points can be generated from CAD models for lightweight 2D localization. In [10], point clouds were extracted from CAD models to achieve radar localization via multi-modal registration. Researchers in [11] proposed to localize a 2D laser scanner on floor plans and hand-drawn maps using stochastic gradient descent. As for localization in 3D space, ICP-based alignment is considered as an effective method to track the robot pose [12]. Despite the point-based method, meshes were also used for robot global localization without initial guess in [13]. Recently, researchers in [14] proposed a novel interface to connect building construction and map representation, which could also detect deviations between as-designed and as-built models via localization results.

Compared to traditional CAD models, BIM is more interoperable in the construction industry and contains more semantic information for robot navigation. For visual-based pose estimation, photogrammetric point clouds can be aligned to BIM model [15] for camera pose estimation from scratch. With sequential input images, visual-based pose tracking was demonstrated to be effective [16], in which camera poses were estimated by aligning images to BIM models. However, visual-based localization methods are easily affected by illumination changes, while LiDAR-based is more robust in long-term operation. In [17], BIM was combined with LiDAR SLAM system to localize the

robot, but the experiments were conducted in simulated environments. Researchers in [18] extracted semantic features of BIM and achieved robot localization using 2D laser scans in the real world. The results showed that the robot can track its pose in BIM but the localization performance was not evaluated quantitatively.

Inspired by the related works above, we can conclude that it is feasible to use BIM for robot localization tasks. However, in some previous works [16, 17], robot localization modules were built on existing SLAM systems, which makes the localization module complicated for real robot applications. To address the problems, we propose to localize the robot in BIM-based maps using lightweight point cloud registration. Besides that, we evaluate the localization accuracy quantitatively in the real world.

3 Workflow Description

As shown in Figure 1, the proposed workflow consists of two parts: offline point cloud map generation from BIM and online ICP-based localization.

3.1 BIM to Point Cloud Map

As a promising direction in the construction industry, BIM is supported by many tools and used in various construction processes, such as building inspection [19] and quality management [20]. For mobile robot localization, metric maps are required rather than digital representations. In this context, the first challenge is that how to generate localization-oriented maps from BIM files.

In this paper, we propose to convert BIM to localization-oriented maps in three steps. The pipeline is shown in Figure 1. Given a whole BIM of one building, we first split the whole BIM into several separate BIM according

to different storeys. After that, we use the open source tool IfcOpenShell¹ to convert multiple BIM to CAD files. Finally, 3D point clouds are sampled from triangular meshes with a density [21]. There are many other sampling strategies in some software²³, such as Monte-Carlo Sampling. Considering that density value is easy to be understood and tuned, we decide to use this strategy in this paper. The final point cloud maps can be regarded as submaps of each floor in the building.

Note that it is better not to change the sequence of this conversion. If we convert the whole BIM to CAD model first without separation, the storey information of BIM is not used. It is more challenging to split large CAD or point cloud maps into storey-based submaps in the follow-up steps. To simplify the conversion process, a promising research direction is to generate point clouds or features directly from the original BIM, which we conclude as a future work in Section 5.

3.2 ICP-based Localization

With generated point cloud maps, there are many existing methods to localize the robot based on the onboard sensors of mobile platform [4, 22]. Generally, a classical robot localization consists of two parts: odometry as a motion module and data matching as a measurement model. In this paper, to validate the effectiveness of the proposed BIM-to-Map conversion process, the localization system is simplified without odometry, which makes the validation easy and efficient to use.

We use a mobile LiDAR scanner to validate the proposed workflow. With the measured LiDAR scans, an ICP algorithm is performed to register the laser points to generated maps from BIM. ICP is a widely used point cloud registration method in the robotics community [23, 24]. Specifically, we use a point-to-plane ICP to achieve pose estimation since there are many planar structures in the building environment. Overall, the point-to-plane ICP-based pose estimation can be formulated as follows:

$$(\mathbf{R}, \mathbf{t}) = \arg \min_{(\mathbf{R}, \mathbf{t})} \left(\sum_{k=1}^K \|(\mathbf{R}\mathbf{p}_k + \mathbf{t} - \mathbf{q}_k) \cdot \mathbf{n}_k\|_2 \right) \quad (1)$$

where K is the number of matched data associations; (\mathbf{R}, \mathbf{t}) is the rotation and translation of estimated robot pose; \mathbf{p} is the lidar points reading; \mathbf{q} and \mathbf{n} are the map points and normal vectors respectively. At each timestamp, ICP will minimize the error metric in Equation 1 as close to zero as possible within a number of iterations.

As for implementation, open source library libpointmatcher [25]⁴ is utilized. At each timestamp s , we use

¹<https://github.com/IfcOpenShell/IfcOpenShell>

²<https://www.meshlab.net/>

³<https://www.cloudcompare.org/>

⁴<https://github.com/ethz-asl/libpointmatcher>



Figure 2. The devices that we used for data collection and experimental validation.



Figure 3. Interior scenes in NUS SDE4 building

the estimated pose at timestamp $s - 1$ as the initial guess of ICP registration. Random sampling on \mathbf{p} is also used to accelerate the online localization process.

4 Experiments

In this section, we first introduce the devices for data collection and the places where we collected the data. Then localization accuracy is evaluated by comparing it to state-of-the-art LiDAR SLAM methods.

4.1 Set-up

To validate the effectiveness of the proposed workflow, we collect several sequences using a handheld Velodyne VLP-16 sensor in the real world. The data collection devices are shown in Figure 2. All the data Sequences are collected in the School of Design and Environment 4

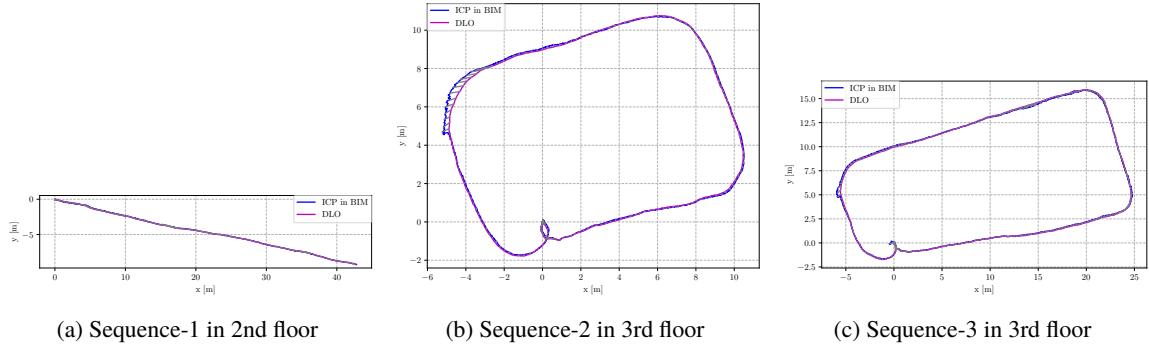


Figure 4. BIM-based localization trajectories compared to DLO

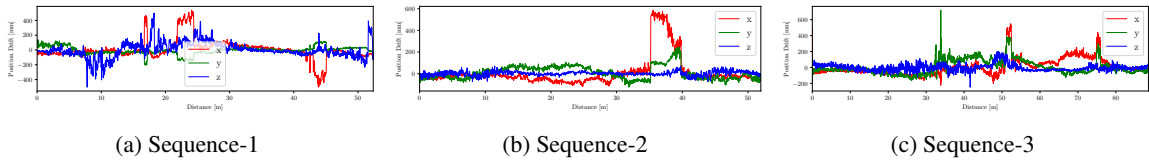


Figure 5. Translation Errors

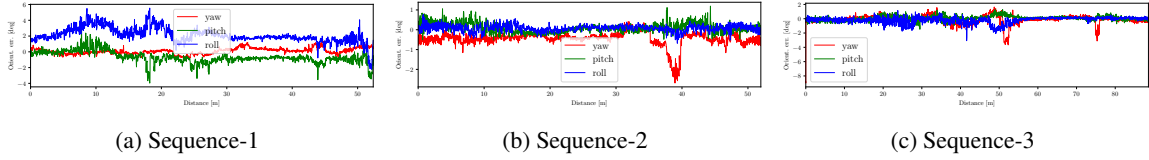


Figure 6. Rotation Errors

Table 1. Localization Errors Compared to DLO

Sequence	Translation Error (m)				Rotation Error (°)			
	Max	Mean	RMSE	Std	Max	Mean	RMSE	Std
1	0.56	0.13	0.17	0.12	6.82	2.37	2.50	0.80
2	0.60	0.09	0.14	0.10	2.70	0.57	0.64	0.30
3	0.76	0.11	0.14	0.08	9.00	0.61	0.84	0.58

(SDE4) building at NUS. The traveled distance is tens of meters in each Sequence. We present some interior scenes in Figure 3. The SDE4 BIM can be viewed in Figure 1. All the online localization experiments are performed using a laptop with Intel I5-8265U and 16G RAM.

Ground truth poses are required to evaluate the ICP-based localization. However, compared to outdoor autonomous vehicles equipped with GPS/INS, it is challenging to collect ground truth poses in indoor scenes, especially for traveling across rooms and corridors in this paper. We notice that a state-of-the-art lidar SLAM system, direct

LiDAR odometry (DLO) [26]⁵, can provide accurate pose estimation in DARPA Subterranean Challenge. According to the error analysis in [26], DLO achieves the best performance compared to other LiDAR SLAM systems. Thus, we set DLO as the “ground truth” for evaluation in this paper.

4.2 Performance Evaluation

An open source toolbox `rpg_trajectory_evaluation` [27]⁶ is used to measure the quantitative results. All poses

⁵https://github.com/vectr-ucla/direct_lidar_odometry

⁶https://github.com/uzh-rpg/rpg_trajectory_evaluation

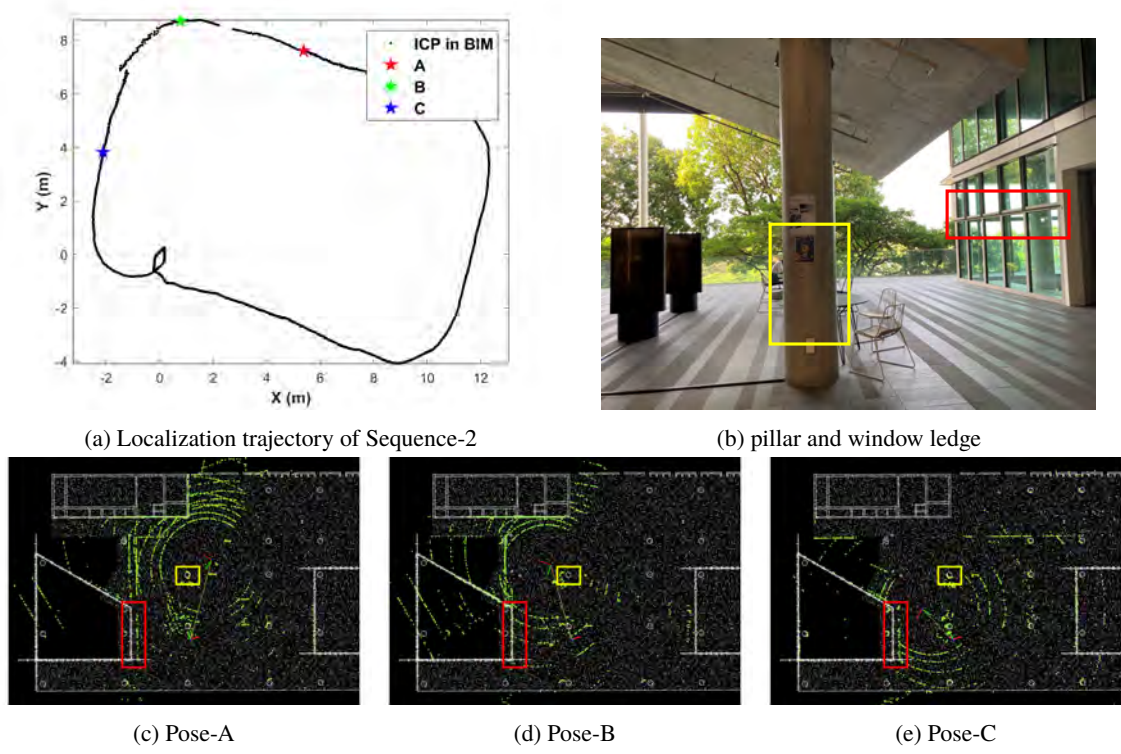


Figure 7. We find that deviations between as-planned and as-built cause a sudden “jump” in the localization trajectory. In 7b, the pillar is shown in the yellow box and the window ledge is shown in the red box for clearance. In 7c, 7d and 7e, the green points represent measured LiDAR points and white points are generated map points from BIM.

(\mathbf{R} , \mathbf{t}) are used to align the trajectories of DLO and BIM-based localization. The initial robot poses are manually fixed in the point cloud maps.

As shown in Figure 4, three data sequences are collected for localization evaluation. Among the three sequences, one is collected on the 2nd floor of SDE4 building and two with loops are collected on the 3rd floor. We also present numerical errors of entire trajectories in Table 1, in which mean error is the mean of the absolute value of each error. The translation and rotation error variations are also presented in Figure 5 and Figure 6 with respect to the traveling distance.

As observed from the errors, the proposed BIM-based localization method can track the lidar scanner successfully with minor errors. In Table 1, the rotation errors of Sequence-1 are larger than errors of Sequence-2 and 3. We consider this is due to several reasons, such as the differences in map point distributions, traveling trajectories, etc. Overall, most of the translation errors are below 0.2m and rotation errors are below 2° , which is acceptable for indoor positioning systems, but still needs to be improved for navigation applications in the future. Compared to previous BIM-based pose estimation meth-

ods [15, 17], the proposed LiDAR localization can track the sensor pose continuously and more accurately, even though only scan matching is involved in our method.

4.3 As-Planned vs As-Built

We also notice that there are two large discrepancies in Sequence-2 during robot localization, resulting in the large errors seen in Figure 5b and Figure 6b. Three robot poses A, B and C are selected in chronological order for investigation, as shown in Figure 7a. Specifically, Pose-B exhibits a large error compared to the ground truth. The bird’s-eye-view of aligned LiDAR points and map points are presented in Figure 7c, 7d and 7e for visualization.

In Figure 7c and 7e, the point cloud of the pillar (yellow box) is aligned correctly but points of the window ledge (red box) are not aligned. While in Figure 7d, the pillar points are unaligned. We measure the distance between the unaligned ledge points using ROS Rviz and the deviation distance is around 0.7m. Based on the analysis above, we conclude that there is a deviation between as-planned BIM and as-built construction status on the 3rd floor which lead to the large localization errors in Sequence-2. On the other

hand, compared to SLAM-based maps, we consider that BIM-to-Map conversion will not involve measurement and estimation errors, which will provide higher-quality local point clouds.

5 Conclusion and Future Work

A workflow of BIM-based robot localization is presented in this paper. We first convert BIM to metric point cloud maps and then perform ICP-based localization to localize a LiDAR sensor. In the experimental section, we conduct a real-world case study at NUS campus. We also find that the deviations between as-planned BIM and as-built buildings bring localization errors in this workflow.

We consider there are several research directions to improve the workflow, categorized as follows:

- More robust point cloud registration or alignment to overcome the deviations from BIM. We consider the semantic information of BIM could help build robust registration. On the other hand, the registration algorithm can be improved using fine-tuned parameters or other outlier filters.
- Multi-sensor fusion for a more accurate localization system. Generally, inertial measurement unit or other odometry modules can help build a more complete system, i.e., providing a high-frequency motion model, which will definitely improve the localization performance.

Acknowledgments

This project is supported by Building Construction Authority (Singapore) and National Robotics Programme under its Built Environment Robotics R&D programme (Award W2122d0154).

Huan Yin is supported by the Hong Kong Center for Construction Robotics (Innohk center supported by HONG KONG ITC).

References

- [1] Ji Zhang and Sanjiv Singh. Loam: Lidar odometry and mapping in real-time. In *Robotics: Science and Systems*, volume 2, 2014. doi:10.15607/RSS.2014.X.007.
- [2] Tong Qin, Peiliang Li, and Shaojie Shen. Vins-mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, 34(4): 1004–1020, 2018. doi:10.1109/TRO.2018.2853729.
- [3] Philipp Krüsi, Bastian Bücheler, François Pomerleau, Ulrich Schwesinger, Roland Siegwart, and Paul Furgale. Lighting-invariant adaptive route following using iterative closest point matching. *Journal of Field Robotics*, 32(4):534–564, 2015. doi:10.1002/rob.21524.
- [4] Xiaqing Ding, Yue Wang, Rong Xiong, Dongxuan Li, Li Tang, Huan Yin, and Liang Zhao. Persistent stereo visual localization on cross-modal invariant map. *IEEE Transactions on Intelligent Transportation Systems*, 21(11):4646–4658, 2019. doi:10.1109/TITS.2019.2942760.
- [5] David Pannen, Martin Liebner, Wolfgang Hempel, and Wolfram Burgard. How to keep hd maps for automated driving up to date. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2288–2294. IEEE, 2020. doi:10.1109/ICRA40945.2020.9197419.
- [6] Changhao Chen, Bing Wang, Chris Xiaoxuan Lu, Niki Trigoni, and Andrew Markham. A survey on deep learning for localization and mapping: Towards the age of spatial machine intelligence. *arXiv preprint arXiv:2006.12567*, 2020. doi:10.48550/arXiv.2006.12567.
- [7] Stephanie Lowry, Niko Sünderhauf, Paul Newman, John J Leonard, David Cox, Peter Corke, and Michael J Milford. Visual place recognition: A survey. *IEEE Transactions on Robotics*, 32(1):1–19, 2015. doi:10.1109/TRO.2015.2496823.
- [8] Mahdi Elhousni and Xinming Huang. A survey on 3d lidar localization for autonomous vehicles. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, pages 1879–1884. IEEE, 2020. doi:10.1109/IV47402.2020.9304812.
- [9] Huan Yin, Runjian Chen, Yue Wang, and Rong Xiong. Rall: end-to-end radar localization on lidar map using differentiable measurement model. *IEEE Transactions on Intelligent Transportation Systems*, 2021. doi:10.1109/TITS.2021.3061165.
- [10] Yeong Sang Park, Joowan Kim, and Ayoung Kim. Radar localization and mapping for indoor disaster environments via multi-modal registration to prior lidar map. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1307–1314. IEEE, 2019. doi:10.1109/IROS40897.2019.8967633.
- [11] Zhikai Li, Marcelo H Ang, and Daniela Rus. Online localization with imprecise floor space maps using stochastic gradient descent. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 8571–8578. IEEE, 2020. doi:10.1109/IROS45743.2020.9340793.

- [12] Hermanncc Blum, Julian Stiefel, Cesarcc Cadena, Roland Siegwart, and Abelcc Gawel. Precise robot localization in architectural 3d plans. In *38th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 755–762. International Association for Automation and Robotics in Construction, 2021. doi:10.22260/ISARC2021/0102.
- [13] Marc Dreher, Hermanncc Blum, Roland Siegwart, and Abelcc Gawel. Global localization in meshes. In *38th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 747–754. International Association for Automation and Robotics in Construction, 2021. doi:10.22260/ISARC2021/0101.
- [14] Selen Ercan, Hermann Blum, Abel Gawel, Roland Siegwart, Fabio Gramazio, and Matthias Kohler. Online synchronization of building model for on-site mobile robotic construction. In *37th International Symposium on Automation and Robotics in Construction (ISARC 2020)*, pages 1508–1514. International Association for Automation and Robotics in Construction, 2020. doi:10.22260/ISARC2020/0209.
- [15] Junjie Chen, Shuai Li, and Weisheng Lu. Align to locate: Registering photogrammetric point clouds to bim for robust indoor localization. *Building and Environment*, 209:108675, 2022. doi:10.1016/j.buildenv.2021.108675.
- [16] Khashayar Asadi, Hariharan Ramshankar, Mojtaba Noghabaei, and Kevin Han. Real-time image localization and registration with bim using perspective alignment for indoor monitoring of construction. *Journal of Computing in civil Engineering*, 33(5):04019031, 2019. doi:10.1061/(ASCE)CP.1943-5487.0000847.
- [17] Mateus Sanches Moura, Carlos Rizzo, and Daniel Serrano. Bim-based localization and mapping for mobile robots in construction. In *2021 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, pages 12–18. IEEE, 2021. doi:10.1109/ICARSC52212.2021.9429779.
- [18] RWM Hendrikx, Pieter Pauwels, Elena Torta, Herman PJ Bruyninckx, and MJG van de Molen-graft. Connecting semantic building information models and robotics: An application to 2d lidar-based localization. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 11654–11660. IEEE, 2021. doi:10.1109/ICRA48506.2021.9561129.
- [19] Yi Tan, Silin Li, Hailong Liu, Penglu Chen, and Zhixiang Zhou. Automatic inspection data collection of building surface based on bim and uav. *Automation in Construction*, 131:103881, 2021. doi:10.1016/j.autcon.2021.103881.
- [20] Zhiliang Ma, Shiyao Cai, Na Mao, Qiliang Yang, Junguo Feng, and Pengyi Wang. Construction quality management based on a collaborative system using bim and indoor positioning. *Automation in Construction*, 92:35–45, 2018. doi:10.1016/j.autcon.2018.03.027.
- [21] Paolo Cignoni, Claudio Rocchini, and Roberto Scopigno. Metro: measuring error on simplified surfaces. In *Computer graphics forum*, volume 17, pages 167–174. Wiley Online Library, 1998. doi:10.1111/1467-8659.00236.
- [22] Martin Magnusson. *The three-dimensional normal-distributions transform: an efficient representation for registration, surface analysis, and loop detection*. PhD thesis, Örebro universitet, 2009.
- [23] Paul J Besl and Neil D McKay. Method for registration of 3-d shapes. In *Sensor fusion IV: control paradigms and data structures*, volume 1611, pages 586–606. International Society for Optics and Photonics, 1992. doi:10.1109/34.121791.
- [24] François Pomerleau, Francis Colas, and Roland Siegwart. A review of point cloud registration algorithms for mobile robotics. *Foundations and Trends in Robotics*, 4(1):1–104, 2015. doi:10.1561/23000000035.
- [25] François Pomerleau, Francis Colas, Roland Siegwart, and Stéphane Magnenat. Comparing icp variants on real-world data sets. *Autonomous Robots*, 34(3):133–148, 2013. doi:10.1007/s10514-013-9327-2.
- [26] Kenny Chen, Brett Lopez, Ali-akbar Aghamohammadi, and Ankur Mehta. Direct lidar odometry: Fast localization with dense point clouds. *IEEE Robotics and Automation Letters*, 2022. doi:10.1109/LRA.2022.3142739.
- [27] Zichao Zhang and Davide Scaramuzza. A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7244–7251. IEEE, 2018. doi:10.1109/IROS.2018.8593941.

Current State and Future Opportunities of Data Mining for Construction 4.0

K. Wu^a and B. García de Soto^a

^aS.M.A.R.T. Construction Research Group, Division of Engineering, New York University Abu Dhabi (NYUAD), Experimental Research Building, Saadiyat Island, P.O. Box 129188, Abu Dhabi, United Arab Emirates
E-mail: keyi.wu@nyu.edu, garcia.de.soto@nyu.edu

Abstract –

In the context of Construction 4.0, the data-intensive nature of the AEC/FM industry puts data-related issues at its core. The powerful ability of data mining to process and utilize data makes it the preferred option for solving data-related problems. Currently, there is no summary of data mining applications in terms of Construction 4.0. To figure out the current state and future opportunities of data mining for Construction 4.0, this study conducts a bibliometric analysis with three steps: (1) determining research context and scope, (2) retrieving literature from Web of Science, and (3) modeling and visualizing the word similarity network using VOSviewer. Three main research areas, namely data mining for intelligence, digitalization, and automation, are identified. The main research topics and objects are summarized for each research area. Furthermore, two promising research fields, namely construction robots and construction cybersecurity, are discussed as future opportunities. This study reveals the current body of knowledge concerning the applications of data mining and points out its future development directions in the context of Construction 4.0.

Keywords –

Knowledge discovery; Digitalization; Automation; Virtualization; Decentralization; Artificial intelligence; BIM; Robot

1 Introduction

With the emergence and evolution of digitalization, automation, virtualization, and decentralization technologies, the world is experiencing a new industrial revolution, commonly labeled as Industry 4.0 [1]. The counterpart of Industry 4.0 in the Architecture, Engineering and Construction, and Facilities Management (AEC/FM) industry is known as

Construction 4.0 [2]. The AEC/FM industry has many opportunities to benefit from Construction 4.0, and traditional industry practices are expected to be highly smart to avoid excessive human intervention for achieving concerned targets [3].

The AEC/FM industry is a typical data-intensive domain. Especially, it undergoes rapid growth in terms of data generation and collection in the information age [4]. Effective data utilization can contribute to the AEC/FM industry's added value. In the context of Construction 4.0, data-related issues are central [5]. Due to the adoption of digitalization, automation, virtualization, and decentralization technologies, large-sized, multi-attributed, and unstructured data from diverse information sources (e.g., text, graph, image, audio, video) significantly increases the difficulty of data utilization [6][7]. Data mining is a process of discovering knowledge such as patterns from large data sets, which incorporates multiple fields, including statistics, pattern recognition, and machine learning [8]. Data mining can powerfully deal with a wide variety of data, and the knowledge it discovers can be used for information management, query optimization, decision support, etc. Therefore, data mining is a promising instrument for Construction 4.0.

In recent years, the interest in data mining application status in the AEC/FM industry has increased, and several studies have explored it [4][6][7][9]. The potential of using data mining over traditional basic statistical methods and pure analytical methods to provide quick and useful insights for the AEC/FM industry has been indicated. However, there is a lack of a summary of data mining applications in terms of Construction 4.0. To fill this gap, this study conducts a literature review to identify the current state and future opportunities of data mining for Construction 4.0. The remainder of the paper is structured below. First, each step of the bibliometric analysis is explained. Then, the current state and future opportunities of data mining for Construction 4.0 are discussed. Finally, conclusions and outlook are provided.

Table 1. Search rules used in Web of Science

Root	AND	Where (OR)	AND	When (OR)	AND	What (OR)
Data mining		Infrastructure		Planning		Construction 4.0
		Building		Design		Digitalization
		Road		Fabrication		Automation
		Railway		Construction		Virtualization
		Bridge		Operation		Decentralization
		Tunnel		Maintenance		Building information modeling
		Water supply		Renovation		Geographic information system
		Sewage		Demolition		Digital twin
		Electrical grid				Artificial intelligence
		Telecommunication				Big data

2 Research Methodology

The bibliometric analysis process consists of three steps, and they are clarified as follows.

The first step is to determine the research context and scope used to identify search keywords and narrow the research focus. From the perspective of technology applications, Construction 4.0 is seen as a means of finding a coherent complementarity between the main emerging technological approaches in the AEC/FM industry [5]. Because digitalization, automation, virtualization, and decentralization technologies are closely related to Construction 4.0, the literature on those technologies falls within the research scope. When studying a technology-related research topic, it is necessary to focus on a recent publication period to highlight the latest developments [10]. Therefore, the publication period of literature is limited to the past ten years (January 2012 to December 2021) to summarize recent technological advances and identify the current state.

The second step is to retrieve literature from Web of Science according to the search rules presented in Table 1. The dimensions data mining could be achieved, namely “Where”, “When”, and “What”, are investigated in parallel using the logical operator “AND”. In the dimension “Where”, ten keywords, such as “Building” and “Railway”, are connected using the logical operator “OR”. In the dimension “When”, eight keywords, such as “Design” and “Construction”, are connected using the logical operator “OR”. In the dimension “What”, 19 keywords, such as “Building information modeling” and “Robot”, are connected using the logical operator “OR”. The initial search returned 3,653 pieces of literature, and

the number was reduced to 274 by filtering to specific categories on Web of Science. The selected categories were Engineering Civil, Construction Building Technology, and Architecture. Fields such as Engineering Mechanical and Engineering Aerospace were excluded.

The third step is to model and visualize the word similarity network using VOSviewer for identifying research areas. The words used for constructing the similarity network are extracted from the titles and abstracts of the retrieved literature [11].

The similarity between two words is calculated using a similarity measure known as the association strength [12], as expressed in Equation (1). s_{ij} refers to the similarity between words i and j ; c_{ij} refers to the number of co-occurrences of words i and j ; and w_i and w_j refer to the total number of occurrences of words i and j , respectively.

$$s_{ij} = \frac{c_{ij}}{w_i w_j} \quad (1)$$

The similarity of words is visualized by their distance from each other in VOSviewer. The higher the similarity between two words, the shorter their distance. VOSviewer is to minimize a weighted sum of the squared Euclidean distances of all words for mapping and to impose that the average distance between two words equals 1 for avoiding all words having the same location [12], as expressed in Equations (2) and (3), respectively. n refers to the number of words; and $\mathbf{x}_i = (x_{i1}, x_{i2})$ refers to the location of word i in a two-dimensional map.

$$V(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{i < j} s_{ij} \|\mathbf{x}_i - \mathbf{x}_j\|^2 \quad (2)$$

$$\frac{2}{n(n-1)} \sum_{i < j} \|\mathbf{x}_i - \mathbf{x}_j\| = 1 \quad (3)$$

The weight of words is visualized by the size of their corresponding nodes. The larger a word and the corresponding node, the higher the weight of the word. The nodes of words are grouped in clusters of different colors. Words are assigned to clusters by maximizing $V(c_1, \dots, c_n)$ [13], as expressed in Equation (4). n refers to the number of words; c_i and c_j refer to the clusters to which words i and j are assigned, respectively; $\delta(c_i, c_j)$ refers to a function equaling 1 if $c_i = c_j$ and 0 otherwise; and γ refers to a resolution parameter determining the detail level of clusters.

$$V(c_1, \dots, c_n) = \sum_{i < j} \delta(c_i, c_j) (s_{ij} - \gamma) \quad (4)$$

3 Findings

The word similarity network based on the retrieved literature is presented in Figure 1. The blue (upper portion), red (right side), and green (left side) nodes indicate the words associated with clusters 1, 2 and 3, respectively. It is worth noting that the words clustered in one specific cluster can share similar concepts with another cluster through strong connections, meaning that they are not exclusively isolated from each other. These clusters could be seen as three main research areas in terms of Construction 4.0: (1) data mining for intelligence, (2) digitalization, and (3) automation, as listed in Table 2. The three research areas and future data mining opportunities for Construction 4.0 are interpreted and analyzed below.

3.1 Data Mining for Intelligence

Intelligence is defined as the ability to acquire and apply knowledge and skills [14]. In the context of artificial intelligence (AI), it allows to make decisions through data processing and get feedback in a manner similar to human thinking when facing complex and random environments. Data mining can nontrivially extract previously unknown and potentially useful information from data, and the application of data mining is one of the mainstream in intelligence. In the research area of data mining for intelligence, there are four main research topics: descriptive intelligence [17][18], diagnostic intelligence [19][20], predictive intelligence [21][22], and prescriptive intelligence [23][24], as listed in Table 2.

Descriptive intelligence refers to describing data, and it answers the question “What happened in the past?”. Descriptive intelligence can be used to understand situations, performances, levels, and so on [17][18]. Occupant presence status is essential information for the simulation of building energy use. For gathering the actual information on occupant presence, a recognition method was designed using C4.5 Decision Tree (C4.5) and Curve Description (CD) from environmental data and the usage information of light and air conditioning [17]. Bad conditions of roads are one of the causes of fatal traffic accidents, and it is required to monitor the road state and detect road damages to enhance the safety of road traffic. Aiming to characterize the road condition, the classifiers applying Support Vector Machine (SVM) were developed to distinguish different types of road surfaces and obstacles [18].

Diagnostic intelligence refers to explaining data, and it answers the question “Why something happened in the past?”. Diagnostic intelligence can be used to determine the causes of trends, correlations between variables, why

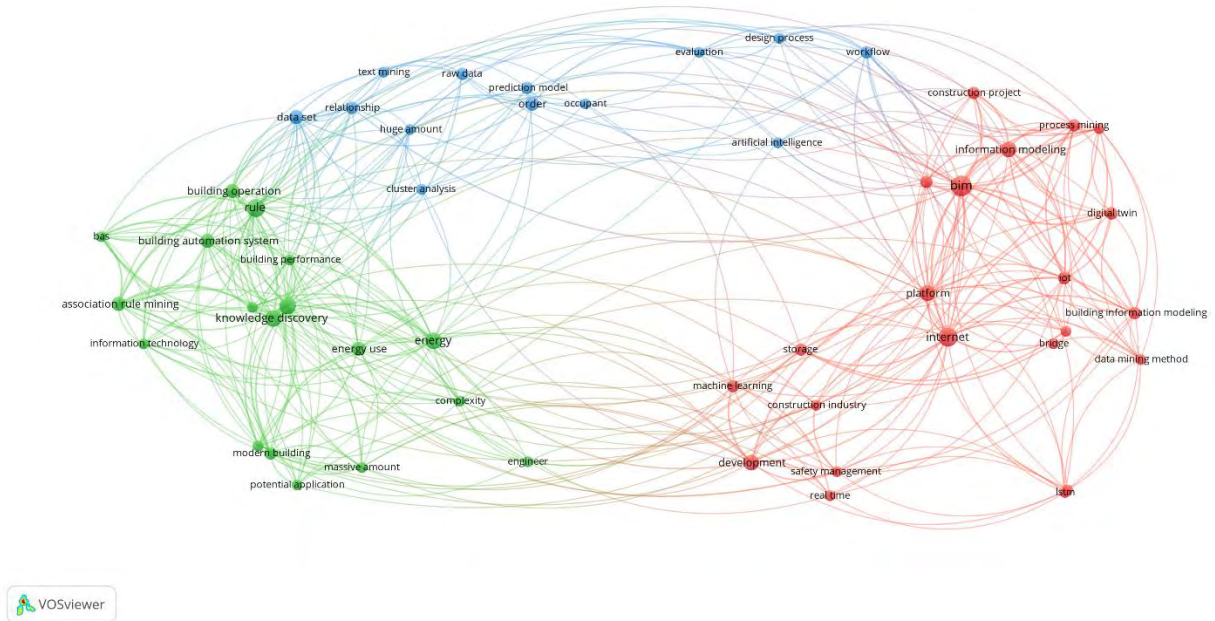


Figure 1. Word similarity network based on the retrieved literature

Table 2. Research areas, topics, and objects identified and summarized from the retrieved literature

Research area	Research topic	Research object
Data mining for intelligence (blue cluster)	Descriptive intelligence	Occupant presence status [17]; Road conditions [18]
	Diagnostic intelligence	Building operation behaviors [19]; Electricity load patterns [20]
	Predictive intelligence	Rebar amount [21]; Rock mass types [22]
	Prescriptive intelligence	Sustainable building design decision-marking [23]; Building maintenance management [24]
Data mining for digitalization (red cluster)	Digitalization demand	BIM innovation directions [25]; BIM user needs [26]; BIM manager role viability [27]; BIM labor costs [28]
	Digitalization modeling	BIM modeling progress [29]; BIM and GIS schemas mapping [30]
	Digitalization management	BIM-based collaborative design [31]; BIM construction data analytics [32]; BIM-based facility management [33]
Data mining for automation (green cluster)	Automation detection	Road cracks [34]; Railway tunnel elements [35]
	Automation equipment	Panelized home prefabrication facility production [36]; Tunnel boring machine construction safety and efficiency [37]
	Building automation system	Building energy management [38]; BAS alarm management [39]

anomalies occur, and so on [19][20]. Discovering underlying building operation data structures and relationships is beneficial to understanding building operation behaviors. In order to realize such abilities, a generic framework with Quantitative Association Rule Mining (QARM) was established, which helps detect and diagnose building operation strategies, non-typical and abnormal building operations, and sensor faults [19]. A considerable amount of real-time electricity consumption data provides a promising way to figure out energy usage patterns and improve building energy management. With this in mind, a general framework integrating Density-Based Spatial Clustering Application with Noise (DBSCAN), K-means, and Classification and Regression Tree (CART) was proposed to extract typical electricity load patterns and discover insightful information hidden in the patterns [20].

Predictive intelligence refers to identifying the likelihood of outcomes based on data, and it answers the question “What is likely to happen in the future?”. Predictive intelligence can be used to forecast unknown future features, activities, trends, and so on [21][22]. The amount estimation of rebar is essential for the cost determination of reinforced concrete structures during the design stage. Taking into account the existing limitations of rebar in 3D modeling, Decision Tree (DT) and Case-Based Reasoning (CBR) were adopted to estimate the amount of rebar in reinforced concrete structures, and the amount can be statistically classified by parameters through generating decision tree nodes [21]. Achieving safe and efficient tunneling needs geological conditions (i.e., rock mass types) ahead of the

tunnel face. To determine rock mass types, Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH), K-means++, and Support Vector Classifier (SVC) were employed based on the operation data of a tunneling boring machine (i.e., cutterhead speed, cutterhead torque, thrust, and advance rate) [22].

Prescriptive intelligence refers to recommending actions to affect likely outcomes based on data, and it answers the question “What is the best course of action?”. Prescriptive intelligence can be used to suggest decision options and show the implication of decision options [23][24]. In sustainable building design, an interplay between multidisciplinary input and the fulfillment of diverse criteria is required for aligning into one high-performing whole. However, design decision-making still relies heavily on rules of thumb. Therefore, the design decision-making based on knowledge discovery in disparate building data was proposed with Multivariate Motif Discovery (MMD) and Temporal Association Rule Mining (TARM) [23]. Building maintenance data such as maintenance requests is a valuable means to assess building performance and gain insights for preventive maintenance actions. To allow facility managers to shrink and limit the area where faults usually occur and determine what building elements and systems are the most problematic, a Text Mining (TM) approach was applied by analyzing textual data contained in unstructured maintenance management systems [24].

3.2 Data Mining for Digitalization

Digitalization is defined as the adaptation of a system,

process, etc., to be operated using computers and the internet [15]. Digitalization greatly facilitates data generation and collection, thus driving an unprecedented growth rate of data. Traditional data analysis cannot handle a large amount of data, which promotes the application of data mining in digitalization. In the research area of data mining for digitalization, there are three main research topics: digitalization demand [25][26][27][28], digitalization modeling [29][30], and digitalization management [31][32][33], as listed in Table 2.

To enable digitalization growth, extra demands, such as innovative directions, user needs, personnel roles, and labor costs, have appeared [25][26][27][28]. The awareness of application hotspots and the forecast of development trends can drive BIM innovations. Based on patent analysis of BIM applications, a framework integrating Social Network Analysis (SNA) and Latent Dirichlet Allocation (LDA) was established to identify the technological development and innovation of BIM [25]. The high threshold of domain knowledge and the information asymmetry between developers and users make it difficult for technology developers of BIM applications to fully understand user needs. Combined with domain knowledge, text mining techniques, including Sentiment Analysis (SA) and Topic Modeling (TM), were utilized to capture user needs from BIM app attributes and user comments [26]. In BIM-enabled projects, the BIM manager has emerged as a necessary adjunct role. To test the likelihood of a long-term market demand for the BIM manager as a distinct role, Singular Value Decomposition (SVD) and Frequent Pattern Growth (FP-growth) were conducted according to BIM-related job advertisements [27]. When adopting BIM in a project, additional labor costs need to be involved. To reduce the prediction risk of BIM labor costs, a hybrid approach integrating Random Forest (RF) and Simple Linear Regression (SLR) was used to improve the prediction accuracy of a project's BIM labor costs [28].

Digitalization changes traditional modeling methods, so new concerns are raised from a modeling perspective, including modeling progress and interoperability [29][30]. In building design practices, the close monitoring of modeling processes and the correct measurement of modelers' performance are required. Having an objective measurement system to quantify modeling progress contributes to performance monitoring. Hence, a sequence mining algorithm based on Generalized Suffix Tree (GST) was implemented to identify implicit 3D modeling patterns from unstructured temporal BIM log data [29]. The functionality between BIM and GIS can be enhanced through their interoperability, and data mapping is critical for seamless information sharing between BIM and GIS models. Given the complexity of BIM and GIS schemas, mapping candidates were generated using text mining techniques

such as Cosine Similarity (CS), Market Basket Analysis (MBA), and Jaccard Coefficient (JC) [30].

Digitalization creates a new work mode and generates new types of information; thus, some management-related thinking emerges [31][32][33]. BIM technically supports multiple designers to model together and exchange opinions. Considering the network-enabled event log mining is beneficial for a deep understanding of the BIM-based collaborative design work, a novel algorithm combining node2vec and Gaussian Mixture Model (GMM) was proposed to discover and analyze potential clusters of designers within a network from BIM log data, which provides support for BIM-based design monitoring and reliable decisions to increase collaboration opportunities [31]. BIM is an effective tool that improves communication and information flow between construction parties. To efficiently retrieve useful information from raw project data within the BIM environment, association, clustering, and trend analyses were performed to identify hidden patterns and detect relationships between different attributes (e.g., the correlation between construction elements or correspondence subjects) [32]. In current building operation and maintenance activities, complex and non-intuitive data records and inaccurate manual inputs raise difficulties in making full use of the information stored in BIM models. For improving facility management, K-means, local density-based outlier detection, and Apriori were conducted to extract meaningful patterns and detect improper records in a data warehouse transformed from the BIM database [33].

3.3 Data Mining for Automation

Automation is defined as the use of largely automatic operations, equipment, or systems [16]. To achieve predetermined goals in accordance with human requirements, automation needs to be driven by data. It is common to collect, process, and utilize various types of complex data in automation, and the application of data mining is widespread. In the research area of data mining for automation, there are three main research topics: automation detection [34][35], automation equipment [36][37], and building automation system [38][39], as listed in Table 2.

Traditional manual detection is time-consuming, error-prone, and in some cases, dangerous. To overcome these shortcomings, data mining is used to automate the detection process [34][35]. Road cracks potentially reduce road performance and threaten traffic safety. For road crack detection automation, there exist several challenges such as intense inhomogeneity along cracks, topology complexity of cracks, and inference of noises with a similar texture to cracks. Taking these issues into consideration, an automatic road crack detection framework was built using structured Random Forest

(RF) and Support Vector Machine (SVM) upon road crack image datasets [34]. Performing regular inspections of railways is essential to avoid extreme events. In order to enhance the automation of railway inspections to reduce the human component, Principal Component Analysis (PCA) and Support Vector Machine (SVM) were applied for the detection and decomposition of railway tunnels according to mobile laser scanning datasets [35].

Automation equipment can facilitate workflow and make work-related tasks perform independently. Data mining can play a positive role in the functional improvement of automation equipment [36][37]. Due to the complex and unique nature of the home building process, existing manufacturing concepts do not apply to the panelized home prefabrication facility. Accordingly, production planning and control were developed for the characteristics of a panelized home prefabrication facility, and Random Sample Consensus (RANSAC) was used to extract models from the data collected using RFID [36]. The tunnel boring machine (TBM) has become a preferred equipment in the construction of long and large tunnels. However, its inability to mine massive information leads to a prevalence of unsafe and uneconomical TBM construction. Aiming to realize real-time safety warnings, deviation corrections, and excavation controls, the rules, such as the interaction between rock mechanics properties and machine characteristics, were mined [37].

Building automation system (BAS) provides a network-based platform for automatically monitoring and controlling various complex building systems, and data mining is applied to enhance the performance of BAS [38][39]. Building operations are typically dynamic, and therefore BAS data is multivariate time-series data in essence. Because the temporal knowledge discovery in BAS data receives little attention, a time-series data mining methodology, including Symbolic Aggregate approXimation (SAX), Motif Discovery (MD), and Temporal Association Rule Analysis (TARA), was presented for building energy management [38]. When building systems behave differently from design values, BAS will raise alarms, usually generating an excessive number of alarms every day. The lack of actionable alarm information makes it difficult for building operators to take action, so a data mining framework was constructed to preprocess raw alarm data, categorize the alarms based on affected objects, and prioritize the alarms with quantitative impacts [39].

3.4 Future Opportunities

There are two promising research fields, i.e., construction robots and construction cybersecurity, to be exploited using data mining for Construction 4.0.

The construction robot is an important part of

Construction 4.0-related technologies. With the continuous development of robotics in recent years, more and more projects have used, or are considering using, construction robots. Construction robots can effectively improve work productivity and reduce safety risks by replacing or assisting workers in performing construction tasks. In a complex and dynamic on-site environment, localization and navigation are the main challenges for construction robots. Localization refers to the robot's ability to identify its location, and navigation refers to its ability to monitor and control its movement from one place to another [40]. To avoid obstacles, especially moving ones, these abilities require construction robots to be equipped with extra sensors to collect data from the on-site environment. Depending on intended applications, the collected sensory data may be visual, thermal, and so forth. A high volume of such unstructured data with noise needs to be processed and utilized in real-time, which provides opportunities for data mining. Although some researchers have begun to apply data mining techniques to deal with these problems, there is still great space for research exploration, considering the random nature of the on-site environment [41].

Construction 4.0 is making the AEC/FM industry more vulnerable to cyber attacks, significantly increasing concerns about construction cybersecurity [42]. Construction cybersecurity is the practice of protecting the critical systems (e.g., digital twin system) and sensitive information (e.g., BIM model) used in projects from cyber threats. These potential threats include denial-of-service, functional modification, reading forgery, and data theft. Intrusion detection is one of the primary means for construction cybersecurity, which has two basic approaches: misuse detection and anomaly detection. Misuse detection refers to matching monitored events with attack signatures in the database, and anomaly detection refers to identifying events that mismatch expected patterns. Undoubtedly, data mining can provide an effective solution for these detections. Although there has been existing research in the cybersecurity domain, the AEC/FM industry still has unique challenges. The application of data mining in construction cybersecurity is worth investigating in this cutting-edge field generally overlooked [43].

4 Conclusions and Outlook

With the increased generation and accumulation of data in the era of Construction 4.0, the computing paradigm is shifting to data-oriented. Large enough data contains valuable information. However, finding this valuable information is not trivial. Data mining provides the ability to discover knowledge from large data.

This study takes a systematic literature review to provide an overall view of data mining for Construction 4.0. A bibliometric analysis was conducted based on the

literature retrieved from Web of Science, and the word similarity network was generated using VOSviewer. The outcome shows three main research areas identified in terms of Construction 4.0: data mining for intelligence, digitalization, and automation. Data mining for intelligence is characterized by four research topics: descriptive intelligence, diagnostic intelligence, predictive intelligence, and prescriptive intelligence. Similarly, data mining for digitalization has three main research topics: digitalization demand, digitalization modeling, and digitalization management. Finally, data mining for automation includes three main research topics: automation detection, automation equipment, and building automation system. Moreover, the fundamental issues of construction robots and construction cybersecurity that can be investigated by data mining techniques are discussed. This study summarizes the current state and future opportunities of data mining for Construction 4.0, reveals the current body of knowledge concerning data mining applications, and proposes development directions in the context of Construction 4.0.

It should be highlighted that this study focuses on the application of data mining for Construction 4.0 from a global perspective. In future research, non-obvious local details (e.g., comparison between data mining techniques applied to a specific field) will be further investigated.

References

- [1] Ghobakhloo, M., Fathi, M., Iranmanesh, M., Maroufkhani, P. and Morales, M.E. Industry 4.0 ten years on: A bibliometric and systematic review of concepts, sustainability value drivers, and success determinants. *Journal of Cleaner Production*, 302:127052, 2021.
- [2] García de Soto, B., Agustí-Juan, I., Joss, S. and Hunhevicz, J. Implications of Construction 4.0 to the workforce and organizational structures. *International Journal of Construction Management*, 1-13, 2019.
- [3] Oesterreich, T.D. and Teuteberg, F. Understanding the implications of digitisation and automation in the context of Industry 4.0: A triangulation approach and elements of a research agenda for the construction industry. *Computers in Industry*, 83:121-139, 2016.
- [4] Yan, H., Yang, N., Peng, Y. and Ren, Y. Data mining in the construction industry: Present status, opportunities, and future trends. *Automation in Construction*, 119:103331, 2020.
- [5] Boton, C. and Forgues, D. Construction 4.0: The next revolution in the construction industry. *CanBIM Innovation Spotlight Publication 2020*, 2020.
- [6] Ahmed, V., Aziz, Z., Tezel, A. and Riaz, Z. Challenges and drivers for data mining in the AEC sector. *Engineering, Construction and Architectural Management*, 25(11):1436-1453, 2018.
- [7] Bilal, M., Oyedele, L.O., Qadir, J., Munir, K., Ajayi, S.O., Akinade, O.O., Owolabi, H.A., Alaka, H.A. and Pasha, M. Big Data in the construction industry: A review of present status, opportunities, and future trends. *Advanced Engineering Informatics*, 30(3):500-521, 2016.
- [8] Dean, J. *Big data, data mining, and machine learning: Value creation for business leaders and practitioners*. John Wiley & Sons, 2014.
- [9] Bilge, E.Ç. and Yaman, H. Research trends analysis using text mining in construction management: 2000-2020. *Engineering, Construction and Architectural Management*, 2021.
- [10] Sonkor, M.S. and García de Soto, B. Operational technology on construction sites: A review from the cybersecurity perspective. *Journal of Construction Engineering and Management*, 147(12):04021172, 2021.
- [11] Chen, Q., García de Soto, B. and Adey, B.T. Construction automation: Research areas, industry concerns and suggestions for advancement. *Automation in Construction*, 94:22-38, 2018.
- [12] Van Eck, N.J. and Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2):523-538, 2010.
- [13] Van Eck, N.J. and Waltman, L. Visualizing bibliometric networks. *Measuring Scholarly Impact*, 285-320, Springer, Cham, 2014.
- [14] Oxford University Press. Lexico. On-line: <https://www.lexico.com/en/definition/intelligence>, Accessed: 20/02/2022.
- [15] Oxford University Press. Lexico. On-line: <https://www.lexico.com/en/definition/digitalization>, Accessed: 20/02/2022.
- [16] Oxford University Press. Lexico. On-line: <https://www.lexico.com/en/definition/automation>, Accessed: 20/02/2022.
- [17] Zhou, H., Yu, J., Zhao, Y., Chang, C., Li, J. and Lin, B. Recognizing occupant presence status in residential buildings from environment sensing data by data mining approach. *Energy and Buildings*, 252:111432, 2021.
- [18] Masino, J., Thumm, J., Levasseur, G., Frey, M., Gauterin, F., Mikut, R. and Reischl, M. Characterization of road condition with data mining based on measured kinematic vehicle parameters. *Journal of Advanced Transportation*, 2018.
- [19] Fan, C., Xiao, F. and Yan, C. A framework for knowledge discovery in massive building automation data and its application in building diagnostics. *Automation in Construction*, 50:81-90, 2015.
- [20] Liu, X., Ding, Y., Tang, H. and Xiao, F. A data

- mining-based framework for the identification of daily electricity usage patterns and anomaly detection in building electricity consumption data. *Energy and Buildings*, 231:110601, 2021.
- [21] Cho, J. and Chun, J. Cost estimating methods for RC structures by quantity takeoff and quantity prediction in the design development stage. *Journal of Asian Architecture and Building Engineering*, 14(1):65-72, 2015.
- [22] Zhang, Q., Liu, Z. and Tan, J. Prediction of geological conditions for a tunnel boring machine using big operational data. *Automation in Construction*, 100:73-83, 2019.
- [23] Petrova, E., Pauwels, P., Svidt, K. and Jensen, R.L. Towards data-driven sustainable design: Decision support based on knowledge discovery in disparate building data. *Architectural Engineering and Design Management*, 15(5):334-356, 2019.
- [24] Marocco, M. and Garofolo, I. Operational text-mining methods for enhancing building maintenance management. *Building Research & Information*, 49(8):893-911, 2021.
- [25] Pan, X., Zhong, B., Wang, X. and Xiang, R. Text mining-based patent analysis of BIM application in construction. *Journal of Civil Engineering and Management*, 27(5):303-315, 2021.
- [26] Zhou, S., Ng, S.T., Lee, S.H., Xu, F.J. and Yang, Y. A domain knowledge incorporated text mining approach for capturing user needs on BIM applications. *Engineering, Construction and Architectural Management*, 27(2):458-482, 2020.
- [27] Hosseini, M.R., Martek, I., Papadonikolaki, E., Sheikhhoshkar, M., Banihashemi, S. and Arashpour, M. Viability of the BIM manager enduring as a distinct role: Association rule mining of job advertisements. *Journal of Construction Engineering and Management*, 144(9):04018085, 2018.
- [28] Huang, C.H. and Hsieh, S.H. Predicting BIM labor cost with random forest and simple linear regression. *Automation in Construction*, 118:103280, 2020.
- [29] Yarmohammadi, S., Pourabolghasem, R. and Castro-Lacouture, D. Mining implicit 3D modeling patterns from unstructured temporal BIM log text data. *Automation in Construction*, 81:17-24, 2017.
- [30] Cheng, J.C., Deng, Y. and Anumba, C. Mapping BIM schema and 3D GIS schema semi-automatically utilizing linguistic and text mining techniques. *Journal of Information Technology in Construction*, 20(14):193-212, 2015.
- [31] Pan, Y., Zhang, L. and Skibniewski, M.J. (2020). Clustering of designers based on building information modeling event logs. *Computer-Aided Civil and Infrastructure Engineering*, 35(7):701-718, 2020.
- [32] Marzouk, M. and Enaba, M. Analyzing project data in BIM with descriptive analytics to improve project performance. *Built Environment Project and Asset Management*, 9(4):476-488, 2019.
- [33] Peng, Y., Lin, J.R., Zhang, J.P. and Hu, Z.Z. A hybrid data mining approach on BIM-based building operation and maintenance. *Building and Environment*, 126:483-495, 2017.
- [34] Shi, Y., Cui, L., Qi, Z., Meng, F. and Chen, Z. Automatic road crack detection using random structured forests. *IEEE Transactions on Intelligent Transportation Systems*, 17(12):3434-3445, 2016.
- [35] Sánchez-Rodríguez, A., Riveiro, B., Soilán, M. and González-deSantos, L.M. Automated detection and decomposition of railway tunnels from mobile laser scanning datasets. *Automation in Construction*, 96:171-179, 2018.
- [36] Altaf, M.S., Bouferguene, A., Liu, H., Al-Hussein, M. and Yu, H. Integrated production planning and control system for a panelized home prefabrication facility using simulation and RFID. *Automation in construction*, 85:369-383, 2018.
- [37] Li, J., Jing, L., Zheng, X., Li, P. and Yang, C. Application and outlook of information and intelligence technology for safe and efficient TBM construction. *Tunnelling and Underground Space Technology*, 93:103097, 2019.
- [38] Fan, C., Xiao, F., Madsen, H. and Wang, D. Temporal knowledge discovery in big BAS data for building energy management. *Energy and Buildings*, 109:75-89, 2015.
- [39] Li, H., Aziz, A., Cochran, E. and Lasternas, B. Building automation system alarm management for operation and maintenance decision making. *ASHRAE 2018 Winter Conference*, 2018.
- [40] Bräunl, T. (2008). *Embedded robotics: Mobile robot design and applications with embedded systems*. Springer, Berlin, Heidelberg, 2008.
- [41] Liang, C.J., Wang, X., Kamat, V.R. and Menassa, C.C. Human-Robot collaboration in construction: Classification and research trends. *Journal of Construction Engineering and Management*, 147(10):03121006, 2021.
- [42] Mantha, B.R.K. and García de Soto, B. Cybersecurity in construction: Where do we stand and how do we get better prepared. *Frontiers in Built Environment*, 7:612668, 2021.
- [43] Yao, D. and García de Soto, B. A preliminary SWOT evaluation for the applications of ML to cyber risk analysis in the construction Industry. *In IOP Conference Series: Materials Science and Engineering*, 1218(1):012017, 2022.

The Use of Boston Dynamics SPOT in Support of LiDAR Scanning on Active Construction Sites

E. M. Wetzel, J. Liu, T. Leathem and A. Sattineni

McWhorter School of Building Science, Auburn University
Auburn, Alabama 36849, USA

E-mail: emw0009@auburn.edu, liujuns@auburn.edu, tml0002@auburn.edu, anoop@auburn.edu

Abstract –

With the recent commercial availability of autonomous mobility platforms, construction researchers have focused their attention to the application of advanced robotic tools on jobsites. One such mobility platform is Boston Dynamic’s robot, “SPOT.” The software development kit (SDK) enabled, quadruped robot has the infrastructure to attach interchangeable payloads including LiDAR (Light Detection And Ranging) scanners. Researchers have conducted a pilot study comparatively analyzing terrestrial LiDAR scans from a human-based tripod scan system and the scans executed by SPOT in both manual and autonomous modes. The research looked at three metrics – quality of scans, productivity savings, and robot accuracy. The result shows that although scan quality is slightly diminished due to the height and shape of the robot, the productivity gains from an autonomous robot could offset the scan quality with additional scans. In addition, in small sample testing, the robot was accurate in returning to a pre-defined location in autonomous mode. Due to the page limit, this paper only presents the results and findings of quality of scans of this this research study.

Keywords –

Autonomous Robots; Construction Robotics; Construction Site Monitoring; Data Capture; LiDAR Scanning

1 Introduction

Light Detection and Ranging, or more commonly known as “LiDAR”, is an important data capture instrument on construction projects. This technology is a remote sensing method that uses light in the form of a pulsed laser to measure ranges (variable distances) from the scanning device to an object. These light pulses generate precise, three-dimensional information about the shape of the items within range of the scanner and its surface characteristics [1]. The resultant data collected by

the scanner is rendered by specialized software creating a 3-dimensional “point cloud.” Such point clouds created from terrestrial LiDAR have an accuracy measured in millimeters and is a reliable mechanism for existing condition assessment and as-built documentation.

Historically, terrestrial LiDAR scanning (TLS) in the construction industry relies on a human setting the scanner on a tripod in predetermined locations in order to capture a comprehensive view of the space requiring the scan (Figure 1). This process works, but does require human intervention and depending on the number of scans for coverage of a designated area, can take a significant amount of time.



Figure 1: Example terrestrial LiDAR scanning (TLS) approach using a tripod.

With the recent availability of autonomous mobility platforms, researchers have focused their attention to the application of advanced robots on construction sites. One such mobility platform is Boston Dynamic’s SPOT. Described by Boston Dynamics [1] as, “an agile mobile robot that navigates terrain with unprecedented mobility,

allowing [users] to automate routine inspection tasks and data capture safely, accurately, and frequently.” SPOT is a SDK (software development kit) enabled, quadruped robot, allowing for the attachment and integration of payloads. One such payload is the FARO S-350-series Laser Scanner (Figure 2). By mounting scanners like the FARO S-350 onto an autonomous mobility platform, interfacing the payload to SPOT, and executing “actions” through the integrated autonomous mode, the possibility exists for scans to be executed without the need for human intervention.



Figure 2: Boston Dynamic’s SPOT with FARO S-350 payload.

Based on hours of testing, under ideal conditions, executing actions with SPOT while in autonomous mode is a fairly reliable process; however, executing autonomous actions, on an active construction site, with industry standard hardware has yet to be evaluated. To date, the published literature on SPOT has been mostly conjecture. This is likely due to the relatively recent availability of SPOT for purchase and the high cost. The goal of this research is to conduct a pilot study comparatively analyzing human-controlled terrestrial LiDAR scans from a tripod and the scans executed by SPOT. This research is significant because there is currently no published experimental research evaluating application of the Boston Dynamics SPOT on an active construction site. For this research, the quality between the tripod based scan and robot based scan are being

evaluated. The research question being evaluated:

- Is there a quality difference between the point clouds produced by the tripod-based scan and the SPOT-based scan?

2 Literature Review

Semi and fully autonomous robotics have made a significant emergence within the industry in the last decade. Current market robots are designed to execute a wide arrange of construction activities in both the administrative and skilled labor spaces.

The use of LiDAR scanning is a growing trend in the built environment. Research has investigated its application in civil construction [2-6] and building construction [6-8], and has shown that it can provide an efficient alternative to traditional surveying options [6-8]. It has been recognized to be suitable in multiple applications such as project progress monitoring, developing as-built documentation, quality control, historical documentation, and existing conditions analysis [6, 8–12]. Since the commercial availability of LiDAR scanners in the early 2000s, implementation has focused mainly on terrestrial laser scanning (TLS) [2, 3, 13]. TLS involves manual setup, deployment, and relocation of a tripod-mounted scanner by an individual to capture the necessary project elements. Although more efficient than alternative methods, the approach requires considerable time investment and has been identified to have project accessibility limitations [14]. Photogrammetry has been an alternative approach to laser scanning for similar project applications – especially when deployed with a UAV [15]. It has advantages over TLS in terms of portability and price; but it also presents a number of limitations in terms of accuracy, data completeness, scaling, robustness to various material textures, etc.

Mobile applications of laser scanning (MLS) have been developed as a means to resolve the limitations of TLS. Gargoum and Karsten [16] investigated using LiDAR scanners mounted to vehicles for scanning large sections of highways for site distancing analyses. LiDAR-mounted unmanned aerial vehicles (UAVs), have shown beneficial for exterior scanning work – especially in hard-to-access areas such as conducting inspections on bridges or exterior building walls [12, 17-19]. However, most of the research using LiDAR-mounted UAVs has focused on industries other than building construction due to identified limitations for interior building applications because of communication issues with the related global positioning systems on which they often operate [17, 20-22].

Alternative strategies are currently being evaluated to address these issues. Xin et al [23] investigated the use of LiDAR as a navigation system for deploying UAVs in

indoor applications where GPS typically falls short. Unmanned Ground Vehicles (UGVs) have been explored as one solution to the limitations associated with UAVs for interior implementation on construction projects (Xin, et al., 2020). The UGV solution has presented challenges related to automation though, especially when ground obstacles are presented. Lee, Park, and Jang [24] utilized a combined drone and wheeled robot mounted with a VR camera and LiDAR to test monitoring capabilities of interior construction progress. Asadi et al. [25] noted numerous studies that combined UGVs and UAVs, but none provided an autonomous solution. They used this previous research as the justification for their investigation of an autonomous combined UAV/UGV approach.

All the previous studies required the use of a UAV and UGV to complete the navigation and monitoring process. One potential solution to this limitation is Boston Dynamic's SPOT; a SDK enabled, quadruped robot, allowing for the attachment and integration of payloads such as LiDAR scanners and cameras [26]. It appears the current research is very limited on a singular autonomous solution like SPOT for construction monitoring [27]. This study is significant as it provides a first look into the application of an autonomous quadruped robot on an active construction site.

3 Methodology

This research project used an experimental setup in order to achieve the goals.

3.1 Experiment Location

The experiment took place on the first floor of a multi-level active construction site with workers present. An active site was preferred in order to identify any limitations to the robot in autonomous mode. At the time of the experiment, the structural elements for the first floor were complete and much of the in-wall and overhead work was taking place. There were numerous stored materials laying around the floor space and no drywall had yet to be installed. This phase of the project is ideal for the "productivity" and "accuracy" portion of the research as the lack of visual markers for the robot to situate itself within the space and the amount of possible route obstructions could put the robots AI (Artificial Intelligence) to the test during autonomous walks.

3.2 Experiment Setup

This research compared the quality and productivity differences between a tripod and SPOT mounted TLS. In addition, the research looked at how accurate SPOT, when placed in autonomous mode, was at positioning itself in relation to the benchmark set in manual mode. In

order to quantify these differences, the researchers divided the data collection into the three constituent parts (scan quality, productivity, and accuracy). All experiments used 4 scan locations on the first floor of the active construction site. Figure 3 presents the 4 scan locations as well as the distance between them. In order to get to Location 4 from Location 3, both the human and robot needed to work around some stored materials obstructing the path.

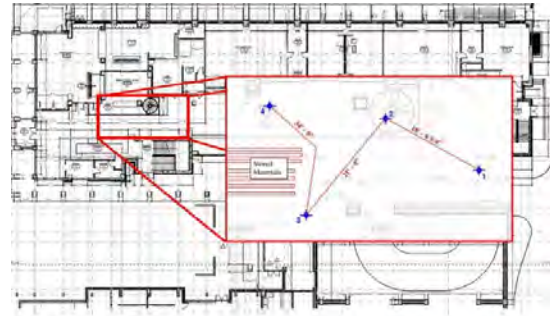


Figure 3: Four scan locations on project floor plan w/ distances between markers.

3.3 Scan Quality

At the 4 predetermined locations, a scan was taken from a FARO S-350 scanner mounted on a tripod (set up by a human) and then again with the same FARO S-350 scanner mounted on SPOT. Each scan used the identical scanning profile as described in Table 1. The series of scans were in immediate succession to avoid possible disruptions that could affect the result, such as difference of sunlight. As an active construction site, the expectation that people would be passing through the scans was assumed. No effort was made to prevent this. Any scan taken during working hours would encounter this issue. The quality analysis in this research was based on the coverage and clarity of the resultant point clouds, not on the specific items that the scan captured. The only variable for this portion of the research was the platform in which the FARO scanner was mounted. The resultant point clouds were processed and registered in FARO SCENE and then exported to Autodesk Recap Pro for evaluation. The quality analysis used two metrics for evaluation, a visual assessment by each researcher and the registration report.

Table 1. Details of Scanning Profile

Setting	Value
Scan Resolution	1/5
Scan Quality	4 x
Scan Duration	> 7:09 minutes
Scan Size	8192 x 3413
Photo Capture	On
HDR Photo	Off
MPts	28.0
Point Distances	0.276 in / 30 ft
Unambiguity Interval	2014.354 ft

3.4 Hardware Setup

The tripod based scanner used a standard hardware setup, with the height of the scanner lens at 5'-5 3/8" from the concrete floor. This elevation was for comfort of the user and was the standard protocol used in TLS.



Figure 4: Hardware setup of tripod based scanner and robot based scanner.

The SPOT based scanner has the scanner lens at 2'-7 1/2". This height was a function of the stand height of the robot and the custom mount used to attach the FARO scanner to the robot. Other hardware - A Velodyne LiDAR scanner and SPOTCore processor - on the robot was used to improve the autonomous vision of the robot during Autowalks. This setup improved the vision of the robot from 2 meters, from the onboard cameras, to 100 meters [28]. A week before the experiment, the research team ran checks on the robot in order to verify proper functionality. This included a load cell check, camera check, and camera recalibration. At the time of the experiment, the robot had no internal errors. Figure 4 presents the hardware setup of the tripod-based scanner and the robot-based scanner.

4 Results and Analysis

Scans were taken at four different pre-determined locations as identified on the building floorplan in the Methodology section. However, during execution of the experiment there was a number of materials, tools, and miscellaneous artifacts located in the work area. All items were left in place as a means to capture the full effects of replicating the scans on an active construction site. Some of these items presented specific obstacles that SPOT would have to maneuver around in order to access the scan locations.

The scan quality was evaluated on two metrics, visual comparative analysis and registry report data developed through the registration in FARO SCENCE. The visual comparison yielded mostly inconsequential differences with two visual issues. Firstly, due to the shape of SPOT, in comparison to a tripod-based scan, the rear shadow was substantially larger as shown in Figure 5. When the scanner rotated to the back side of SPOT and pulsed at an angle slightly below 0 degrees, the scan was interrupted by the internal LiDAR scanner that SPOT was using for extended vision. Another issue noticed by the researchers, arguably more substantial, was that the height of the FARO scanner on SPOT (measured as 2'-7 1/2") was almost three feet lower than that of the scanner mounted on the tripod (measured as 5'-5 3/8"), which affected the effective coverage of the captured point clouds. As shown in Figure 6, the top surfaces of the jobsite storage boxes and other items behind those boxes in the background were captured by the scan on the tripod setup but not by the scan on SPOT.

The second quality metric was the registration report comparison developed by FARO SCENE software while the point cloud was being processed and registered. The results of point cloud registration reports, as shown in Figure 7, revealed that although the robot-based scans had a mean point error that was more than double that of the tripod-based scans, both sets of scans were showing a Mean Point Error under 1.6 millimeters. Two possible causes of the larger Mean Point Error of the SPOT-based scan are 1) the slightly instability of the mount setup on the robot compared to the very steady tripod setup, and 2) fewer points were captured close to the scanner due to the shadow under the robot which could have been used to improve scan registration. The researchers also expected that the error of the scan registration would be fairly minimal due to the following factors:

- A very small amount of scans were captured for each round of the field testing
- The scan stations were set up relatively close with the longest distance between any two scans being less than 38 feet
- There were plenty of unique rough and hard surfaces in the testing area captured by the scanner

which were ideal for scan registration



(a) Scan captured on a tripod



(b) Scan captured on SPOT

Figure 5: Comparison of shadow cast on tripod-based scan vs. SPOT-based scan.

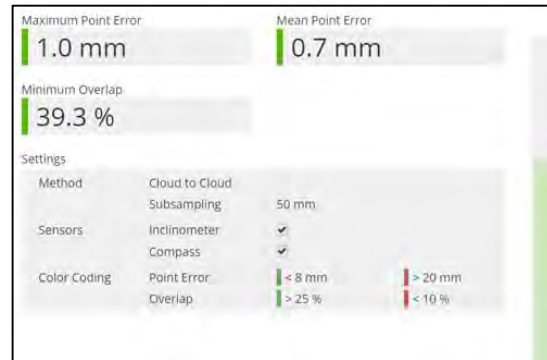


(a) Scan captured on a tripod

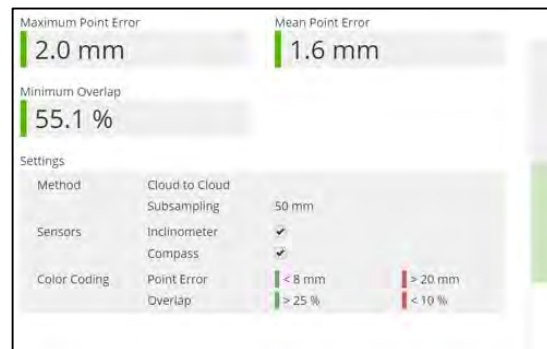


(b) Scan captured on SPOT

Figure 6: Comparison of capture limitations on tripod-based scan vs. SPOT-based scan.



(a) Tripod Based Scans



(b) Robot Based Scans

Figure 7: Registration report comparison.

With other metrics such as overlap, maximum point error, and acceptable color matrix results, the two sets of scans were yielding results that were, quality-wise, indistinguishable.

5 Conclusions and Discussion

The goal of this research was to conduct a pilot study comparatively analyzing terrestrial LiDAR scans (TSL) from a human-based tripod setup and the scans executed using the Boston Dynamics SPOT. Analysis was conducted on the metrics of quality, productivity, and accuracy. However, one the conclusion and discussion regarding the quality of the scan data captured from the study are included in this paper.

5.1 Scan Quality

The scans taken with SPOT identified some challenges with quality of capture related to vantage point and shadow cast. This was due to the physical shape of SPOT, autonomous payloads, and reduced height of the scanner when mounted on SPOT's back. Ideally, LiDAR scans for as-built of the interior of a construction

project are captured by scanners elevated to five feet or higher using a tripod. At such height, the scanners are able to capture the surfaces of common features in the buildings, such as desktops, countertops, lavatories, etc. However, when mounted on SPOT, the highest surfaces that the scanner can capture are roughly 2'-8", resulting in null data for elevated finishes. Certainly there are mechanical ways to elevate the scanner on SPOT, such as an altered mount, however this will create additional issues that could be far worse than the initial problem. First, mechanically raising the scanner will significantly alter the balance the SPOT system and make it more vulnerable to fall. Second, the programming of the SPOT control system needs to be altered to change its sense of clearance. A potential solution to this could be a reticulating mount, but this product currently does not exist and would need to be manufactured specifically for this application.

The other issue of capturing LiDAR scans on SPOT identified through this research is the relatively large shadow in the point clouds that was caused by the rear payloads and shape of the robot. This issue resulted in a less amount of effective points captured by each robot based scan. Mitigation of this issue can be done in two ways. First, removal of the internal LiDAR scanner (rear payload) from the robot would yield a slightly smaller shadow; however, this may actually be counter-productive to the overall effectiveness and productivity gains acquired from the robot. Removing the rear payloads will likely improve the coverage of the point cloud behind the robot but, as the internal LiDAR gives SPOT additional vision, removal of this system may require additional fiducials, resulting in more human intervention. Perhaps the more effective mechanism for improving the scan coverage in regards to the rear shadow, is to decrease the spacing between scan stations to compensate the shadow areas, resulting in additional scans. Additional scans would obviously effect cycle time, but ultimately, an accurate point cloud is the superseding goal.

5.2 Summary

This study focused on providing an introductory look into employing the Boston Dynamics SPOT autonomous robot on an active construction site to analyze practical uses for terrestrial LiDAR scanning. Recognizing the current literature on this new robot is extremely limited, this research is significant in providing a first look at quality, productivity, and accuracy on an active construction site. It also serves as a proof of concept for a methodology to conduct future research. The results of the study suggest some possible usages for SPOT to be implemented on active construction sites. However, future research needs to be conducted to collect more repetitive and larger samples for analysis quality,

productivity, and accuracy. Some specific areas need to focus on how obstacles affect these metrics, consistency and reliability of accuracy, and human intervention required to implement the robot on to achieve consistent and accurate results. Other research should look at options to improve the flexibility of payloads that can be used. Specific to LiDAR scanners, mounting options need to be researched to get the vantage point of the scan higher and to provide a more stable platform for the LiDAR scanner. Future research may want to look at using SPOT for surveying, but that would require greater accuracy and in its current form, does not appear possible without additional programming and/or human intervention. Other robotics companies are focusing their efforts on layout, possibly making SPOT an inappropriate tool for the application.

References

- [1] National Oceanic and Atmosphere Administration. (2021). *What is LiDAR*.
- [2] Panella, F., Roecklinger, N., Vojnovic, L., Loo, Y., & Boehm, J. (2020). Cost-Benefit Analysis of Rail Tunnel Inspection for Photogrammetry and Laser Scanning. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B2-2020*, 1137–1144. <https://doi.org/10.5194/isprs-archives-XLIII-B2-2020-1137-2020>
- [3] Slattery, K. T., & Slattery, D. K. (2013). Modeling Earth Surfaces for Highway Earthwork Computation Using Terrestrial Laser Scanning. *International Journal of Construction Education and Research*, 9(2), 132–146. <https://doi.org/10.1080/15578771.2012.700298>
- [4] Boyu, W., Chao, Y., Han, L., Jack, C. P. C., & Qian, W. (2021). Fully automated generation of parametric BIM for MEP scenes based on terrestrial laser scanning data. *Automation in Construction*, 125, 103615. <https://doi.org/10.1016/j.autcon.2021.103615>
- [5] Liu, J., Jenness, M., & Holley, P. (2016). *Utilizing Light Unmanned Aerial Vehicles for the Inspection of Curtain Walls: A Case Study*. 2651–2659. <https://doi.org/10.1061/9780784479827.264>
- [6] Mani, G.-F., Jeffrey, B., Jochen, T., Silvio, S., & Feniosky, P.-M. (2011). Evaluation of image-based modeling and laser scanning accuracy for emerging automated performance monitoring techniques. *Automation in Construction*, 20(8), 1143–1155. <https://doi.org/10.1016/j.autcon.2011.04.016>
- [7] Afrooz, A., Frédéric, B., & Pingbo, T. (2021). Planning for terrestrial laser scanning in construction: A review. *Automation in Construction*, 125, 103551.

- <https://doi.org/10.1016/j.autcon.2021.103551>
- [8] Burcu, A., Frank, B., Chris, G., Daniel, H., Catherine, L., & Kuhn, P. (2006). A formalism for utilization of sensor systems and integrated project models for active construction quality control. *Automation in Construction*, 15(2), 124–138. <https://doi.org/10.1016/j.autcon.2005.01.008>
- [9] Guo, J., Yuan, L., & Wang, Q. (2020). Time and cost analysis of geometric quality assessment of structural columns based on 3D terrestrial laser scanning. *Automation in Construction*, 110. <https://doi.org/10.1016/j.autcon.2019.103014>
- [10] Hao, Y., Xiangyang, X., & Ingo, N. (2014). The Benefit of 3D Laser Scanning Technology in the Generation and Calibration of FEM Models for Health Assessment of Concrete Structures. *Sensors*, 14(11), 21889–21904. <https://doi.org/10.3390/s141121889>
- [11] Liu, J., & Willkens, D. S. (2021). Reexamining the Old Depot Museum in Selma, Alabama, USA. *4th International Conference on Building Information Modelling (BIM) in Design, Construction and Operations*. 4th International Conference on Building Information Modelling (BIM) in Design, Construction and Operations, Spain.
- [12] Zaychenko, I., Smirnova, A., & Borremans, A. (2018). Digital transformation: The case of the application of drones in construction. *MATEC Web Conf.*, 193, 05066.
- [13] Starek, M. J., Mitasova, H., Hardin, E., Weaver, K., Overton, M., & Harmon, R. S. (2011). Modeling and analysis of landscape evolution using airborne, terrestrial, and laboratory laser scanning. *Geosphere*, 7(6), 1340–1356. <https://doi.org/10.1130/GES00699.1>
- [14] Hugenholtz, C.H., et al.(2015). *Earthwork Volumetrics with an Unmanned Aerial Vehicle and Softcopy Photogrammetry*. *Journal of Surveying Engineering*, 2015. 141(1): p. 06014003
- [15] Duque, L., Seo, J., & Wacker, J. (2018). Bridge Deterioration Quantification Protocol Using UAV. *Journal of Bridge Engineering*, 23(10), 04018080. [https://doi.org/doi:10.1061/\(ASCE\)BE.1943-5592.0001289](https://doi.org/doi:10.1061/(ASCE)BE.1943-5592.0001289)
- [16] Gargoum, S. A., & Karsten, L. (2021). Virtual assessment of sight distance limitations using LiDAR technology: Automated obstruction detection and classification. *Automation in Construction*, 125. <https://doi.org/10.1016/j.autcon.2021.103579>
- [17] Tatum, M. C., & Liu, J. (2017). Unmanned Aircraft System Applications in Construction. *Procedia Engineering*, 196, 167–175. <https://doi.org/10.1016/j.proeng.2017.07.187>
- [18] Van Valkenburgh, P., Cushman, K. C., Butters, L. J. C., Vega, C. R., Roberts, C. B., Kepler, C., & Kellner, J. (2020). Lasers Without Lost Cities: Using Drone Lidar to Capture Architectural Complexity at Kuelap, Amazonas, Peru. *Journal of Field Archaeology*, 45, S75–S88.
- [19] Roca, D., Armesto, J., Lagüela, S., & Díaz-Vilariño, L. (2014). LIDAR-EQUIPPED UAV FOR BUILDING INFORMATION MODELLING. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 45, 523–527.
- [20] Grubinger, S., Coops, N. C., Stoehr, M., El-Kassaby, Y. A., Lucieer, A., & Turner, D. (2020). Modeling realized gains in Douglas-fir (*Pseudotsuga menziesii*) using laser scanning data from unmanned aircraft systems (UAS). *Forest Ecology and Management*, 473. <https://doi.org/10.1016/j.foreco.2020.118284>
- [21] Camarretta, N., A. Harrison, P., Lucieer, A., M. Potts, B., Davidson, N., & Hunt, M. (2020). From Drones to Phenotype: Using UAV-LiDAR to Detect Species and Provenance Variation in Tree Productivity and Structure. *REMOTE SENSING*, 12(19). <https://doi.org/10.3390/rs12193184>
- [22] Leuenberger, D., Haefele, A., Omanovic, N., Fengler, M., Martucci, G., Calpini, B., Fuhrer, O., & Rossa, A. (2020). Improving High-Impact Numerical Weather Prediction with Lidar and Drone Observations. *BULLETIN OF THE AMERICAN METEOROLOGICAL SOCIETY*, 101(7), E1036–E1051. <https://doi.org/10.1175/BAMS-D-19-0119.1>
- [23] Xin, C., Wu, G., Zhang, C., Chen, K., Wang, J., & Wang, X. (2020). *Research on Indoor Navigation System of UAV Based on LIDAR*. IEEE.
- [24] Lee, J. H., Park, J.-H., & Jang, B.-T. (2018). Design of Robot based Work Progress Monitoring System for the Building Construction Site. IEEE.
- [25] Asadi, K., Kalkunte Suresh, A., Ender, A., Gotad, S., Maniyar, S., Anand, S., Noghabaei, M., Han, K., Lobaton, E., & Wu, T. (2020). An integrated UGV-UAV system for construction site data collection. *Automation in Construction*, 112. <https://doi.org/10.1016/j.autcon.2019.103068>
- [26] Boston Dynamics. (2021a). + LIDAR: Spot EAP enhances range and accuracy of Spot's autonomy system.
- [27] Afsari, K., Ensafi, M., Devito, S., and Serdakowski, J. (2021) Fundamentals and Prospects of Four-Legged Robot Applications in Construction Progress Monitoring. in 57th Annual Associated Schools of Construction International Conference. Online: EasyChair.
- [28] Boston Dynamics. (2021b). SPOT: Transformative Mobility.

Autonomous operation of a robot dog for point cloud data acquisition of scaffolds

D. Chung^a, S. Paik^a, J. Kim^a, and H. Kim^a

^aDepartment of Civil and Environmental Engineering, Yonsei University, Korea

E-mail: jungduho1@yonsei.ac.kr, tjsdnd4066@yonsei.ac.kr, kah5125@yonsei.ac.kr, hyoungkwan@yonsei.ac.kr

Abstract –

Scaffolds are essential temporary structures on construction sites. Since scaffolds are frequently installed and dismantled, the inspection needs to be performed in real-time. This paper proposes a framework to automate the acquisition process of scaffold point cloud data using a robot dog. First, a Simultaneous Localization and Mapping (SLAM) algorithm (LIO-SAM) is deployed for real-time map creation based on laser-based 3D data. Scaffolds are automatically detected using the bird's eye view (BEV) projection images of the registered 3D point clouds. A scanning distance is also determined for each detected scaffold to move the robot dog to an optimal location. The robot dog can successfully scan the scaffolds on construction sites by using the proposed framework.

Keywords –

Autonomous Operation; Mobile Laser Scanning; Robot Dog; Real-time Detection; Scaffold

1 Introduction

Scaffolds are an indispensable factor on construction sites, and it is one of the major risk factors for construction safety management. According to the statistics by the Korean Ministry of Employment and Labor, more than half (51.5%) of fatalities in the construction industry are falling accidents, and scaffolds are the contributing factor (19.9%) to falling fatalities [1]. Because scaffolds are frequently installed and dismantled during construction, safety management is difficult. Real-time inspection is ideal for a thorough inspection, but it rarely becomes a reality due to its labor-intensive and costly nature. Automating the inspection process using a mobile robot could be a solution for the effective monitoring of scaffolds. By adding repeatability to the labor-intensive inspection process, fast and efficient monitoring could become a reality.

Kim et al. [2] proposed a framework for automatic scaffold segmentation and 3D reconstruction based on 3D point clouds acquired by Mobile Laser Scanning

(MLS). A robot dog was used in the study for the point cloud data acquisition process, but the robot dog was teleoperated. Teleoperation can reduce human labor for data acquisition but still requires human intervention. There are some studies that use autonomous operation for data acquisition. Kim et al. [3] used an Unmanned Aerial Vehicle (UAV) to make a map of the construction site for calculating optimal scanning points. The map was given to an Unmanned Ground Vehicle (UGV) for autonomous scanning. The UGV relied on the map for its scanning process. Kim et al. [4] provided a fully automatic 3D data acquisition and registration system using a UGV. 2D SLAM was used for localization and navigation, and 3D reconstruction was performed based on the SLAM result. The study was intended to produce a general scan result of the construction site without specific target objects.

This paper proposes a new framework to automate the data acquisition process for scaffold point clouds using a robot dog. The proposed framework aims to bring a focused attention to a specific construction component - scaffolds, without the need for any prior knowledge such as scaffold location or construction site map. This would enable truly dynamic and real-time scan planning, practically applicable in scaffold inspection. To the best of the authors' knowledge, it is difficult to find previous studies in which a mobile robot was tried to automatically acquire point cloud data of specific construction objects. The overview of this framework is shown in Figure 1.

2 Methodology

2.1 System architecture

The scanning platform for this study uses a Unitree A1 robot dog, with Ouster OS1-128 Mobile LiDAR and Microstrain Inertial Measurement Unit (IMU), and an on-board computer (NVIDIA Jetson TX2), as described in Figure 2. Robot dogs can have two major advantages compared to wheeled robot for mobile laser scanning. First, robot dogs can walk stably through rough terrains and small obstacles. Unlike refined workspaces, construction sites generally have uneven surfaces, and a

wheeled robot's bump can easily affect the results of mobile laser scanning. Second, robot dog has more Degree of Freedom (DoF) in joints, which can easily increase the Field of View (FoV) of scanning without additional actuators. In our study, we controlled the roll and pitch simultaneously to increase the robot's FoV as much as possible.

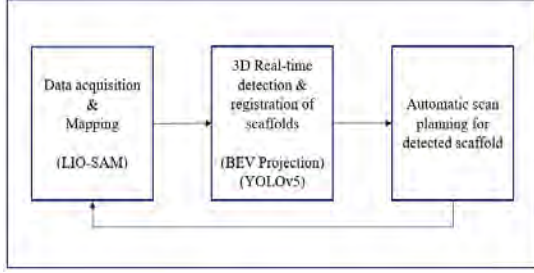


Figure 1. Overview of the proposed framework

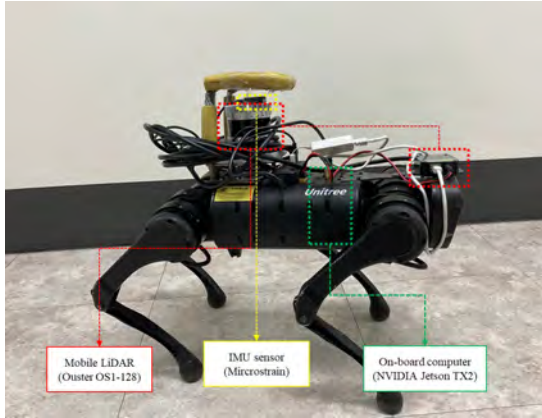


Figure 2. Scanning platform description

3D Simultaneous Localization and Mapping (SLAM) algorithm is used for the localization of the robot dog and for registration of the obtained data into a 3D point cloud map. The 3D point cloud map enables the robot dog to understand the environment, and it also works as the final product of the scanning process. LIO-SAM [5], a 3D SLAM algorithm based on sensor fusion between LiDAR and IMU, was used in this study.

LIO-SAM receives 3D point cloud and IMU sensor data as input. LiDAR odometry is initially predicted based on the motion estimated by IMU data, and the IMU bias is repeatedly corrected based on LiDAR odometry. The LiDAR odometry is sent to a path planning algorithm for the localization of the robot dog. The registered 3D point cloud map is used to perform scaffold detection and allow the robot dog to understand the location of the scaffold. Figures 3(a) and 3(b) show the scaffolds used for the experiment and its registered point clouds, respectively.

2.2 SLAM based BEV 3D scaffold detection

To automate the scaffold data acquisition process, the robot needs to understand the goal of scanning. In this study, deep learning-based object detection is used to detect scaffolds. Understanding the environment can be divided into two categories: image-based and point-cloud-based. Image-based methods are accurate and fast, but it lacks spatial information of the object. Point-cloud has very accurate spatial information but has sparse visual information. 3D point cloud detection also suffers from high computational costs. To reduce the computing cost, some studies attempted bird's eye view (BEV) projection-based object detection [6, 7]. The main idea is to translate 3D point cloud into 2D images by projecting the points vertically, and to apply convolutional neural network (CNN) for object detection. This idea can detect objects from the 3D point cloud in real-time, but a lack of visual information can lower detection performance.

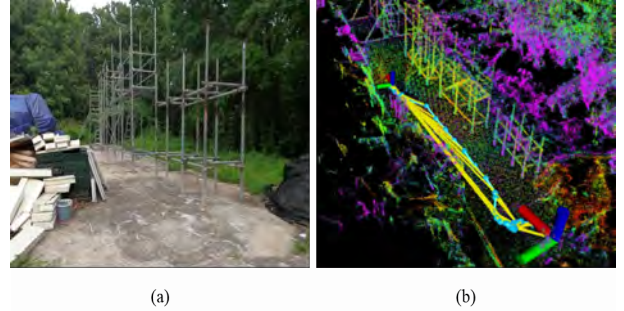


Figure 3. Scaffolds at Yonsei University; (a) photogrammetry, (b) registered point clouds

To overcome this problem, we used SLAM-based registered points instead of raw point cloud data to generate BEV images. By using registered points, the visual information becomes denser and can detect objects more accurately in real-time. In the proposed method, registered points are projected into 2D images, and each pixel value represents the height, density, and intensity features of the registered point. YOLOv5, a real-time object detection algorithm, is then applied to the generated BEV images [8] for detecting scaffolds. The scaffold detection method is shown in Figure 4.

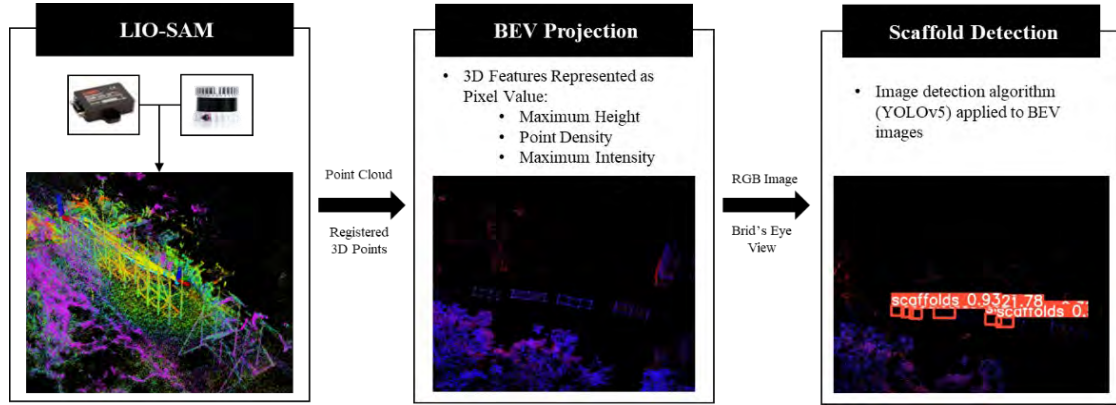


Figure 4. Scaffold detection using the SLAM data

2.3 Implementation

The proposed framework was implemented using a robot operating system (ROS). LIO-SAM subscribes to sensor data from LiDAR and IMU, then publishes registered 3D point cloud to the BEV projection node. The BEV projection node projects registered point cloud data to 2D images and publishes them to the detection node. The detection node detects scaffolds from the subscribed images and calculates the scaffold's Cartesian coordinates and maximum height. To prevent overlapping results from the same scaffolds, detected scaffolds are registered only when it's more than a threshold value away from each other. Maximum height is calculated by detecting the highest pixel inside the detection bounding box.

The path planning node subscribes to the scaffold data, calculates the optimal scanning distance, and publishes the command for scanning. The optimal scanning distance is decided based on the FoV of the LiDAR and the maximum control range of the robot's

pitch by Eq. (1). If there is no recognized scaffold, the robot performs a pre-defined scanning motion to get more information about the environment. After the robot recognizes the scaffold, the robot decides the closest scaffold as a goal, and moves towards the scaffold until it reaches the optimal scanning distance. Once the robot reaches the scanning distance, the robot performs the scanning motion and moves on to the next scaffold. Fig. 6 describes the flowchart of the path planning node. The control node subscribes to the published command to control the robot hardware.

$$\text{Scan distance} = \frac{\text{Scaffold height}}{\tan(\text{LiDAR FoV} + \text{Maximum Pitch})} \quad (1)$$

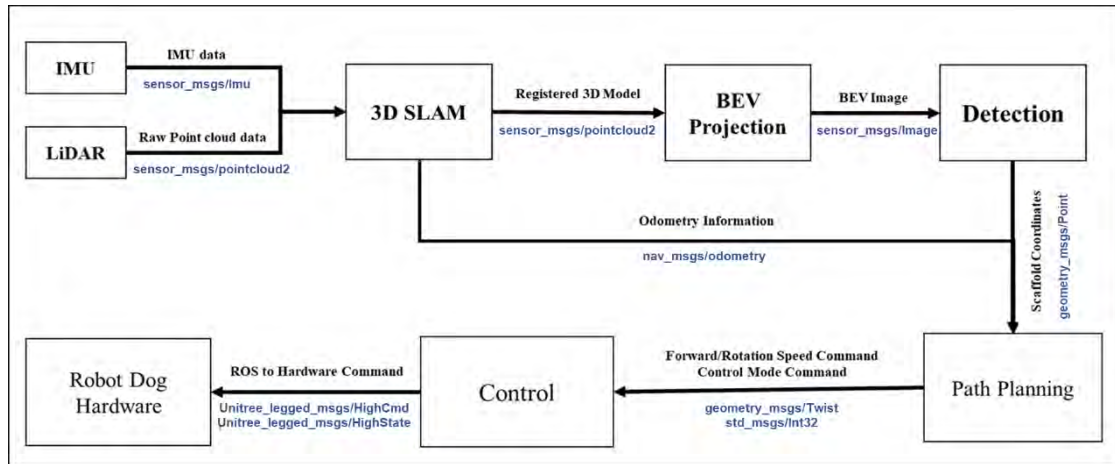


Figure 5. ROS-based implementation for the proposed method

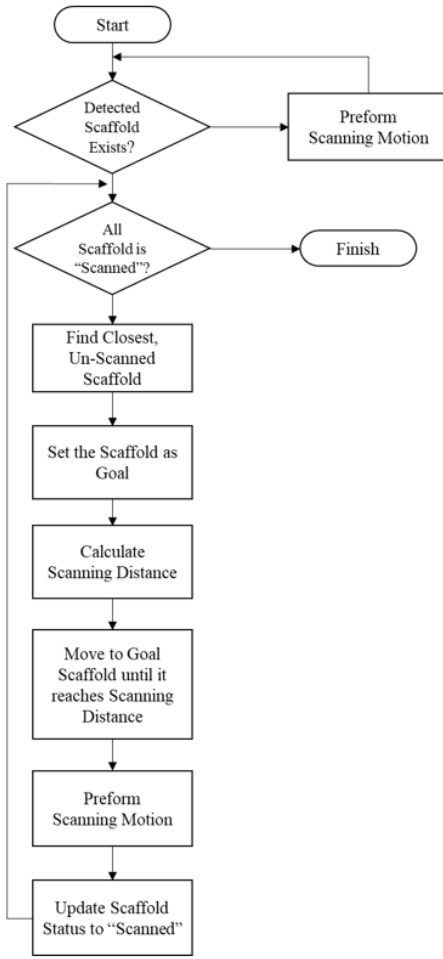


Figure 6. Flowchart for path planning node

3 Experiments and Results

For training the object detection model, 300 BEV images from scaffolds at Chung-Ang University were used, and 58 images from scaffolds at Yonsei University were used to test the model. All datasets were gathered by the robot dog, as conducted in [2]. Figures 7(a) and 7(b) are examples for training and testing sets, respectively.

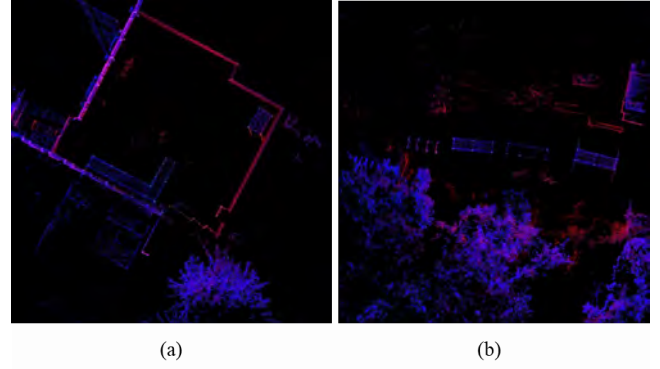


Figure 7. BEV images; (a) an example for training, (b) an example for testing

	Precision	Recall	F1-score
Scaffold	86.9%	73.5%	79.6%

Table 1. The performance of the scaffold detection model

The scaffold detection model is trained for 50 epochs, with pre-trained weights based on the COCO dataset [9]. Table 1 shows the performance of the scaffold detection model. The model achieved 86.9% precision, 73.5% recall, and 79.6% F1-score on scaffolds. Figure 8 shows an example of scaffold detection results, proving that the framework effectively detects scaffolds in real-time.



Figure 8. Scaffold detection results

For this experiment, we used a single scanning motion in which we changed the roll and pitch of the robot joint for 10 seconds. The scanning motion is shown in Figure 9. The robot dog also had a fixed region of interest of 30m x 30m square with the robot's starting point as origin.

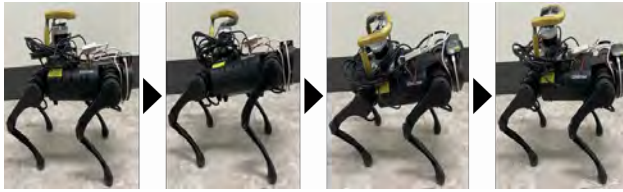


Figure 9. Scanning motion

Figure 10(a) shows an example case of the robot's trajectory and detected scaffolds, Figure 10(b) and 10(c) show the result of automatic data acquisition. As shown in Figure 10, the scaffold points have been successfully obtained. Even though there are still some limitations in the navigation algorithm, the experiment shows that the proposed framework allows the robot dog to automatically move around the site for the successful scanning of scaffolds.

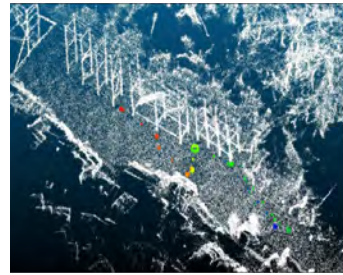
4 Conclusions

This study proposed a new framework for automating the scaffold point cloud data acquisition process using a robot dog. The proposed framework with a real-time 3D scaffold detection algorithm with an SLAM-based BEV image was implemented for a robot dog, and it was tested on a real-world outdoor construction site. The experiments show that the robot dog can automatically perform the end-to-end data acquisition process without any human intervention.

This study currently has some limitations in the navigation system. First, an obstacle avoidance system needs to be developed. Second, the path needs to be optimized. Third, the scan planning algorithm needs to be more generalized for a range of construction sites. With the improvement, the proposed method is expected to enable a fully autonomous operation of smart mobile robots designed to monitor construction sites for safety and productivity management.



(a)



(b)



(c)

Figure 10. Results; (a) robot's trajectory represented by the numbers and detected scaffolds with scanning points connected to its corresponding scaffolds by the same colors, (b) the 3D point cloud of the site, (c) the 3D point cloud of the scaffolds

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education (No. 2018R1A6A1A08025348) and the "National R&D Project for Smart Construction Technology (No.22SMIP-A156488-03) funded by the Korea Agency for Infrastructure Technology Advancement under the Ministry of Land, Infrastructure and Transport, and managed by the Korea Expressway Corporation.

References

- [1] Korea Ministry of Employment and Labor. Industrial Accident Analysis for 2020. Online: https://www.moel.go.kr/news/enews/report/enews_View.do?news_seq=12149, Accessed:02/23/2022
- [2] Kim, J. H., Chung, D. H., Kim, Y. H., and Kim, H.

- K., Deep learning-based 3D reconstruction of scaffolds using a robot dog. *Automation in Construction*, 134: 104092, 2022.
- [3] Kim, P., Park, J., and Cho, Y. K., UAV-assisted autonomous mobile robot navigation for as-is 3D data collection and registration in cluttered environments. *Automation in Construction*, 106: 102918, 2019.
 - [4] Kim, P., Chen, J., and Cho, Y. K., SLAM-driven robotic mapping and registration of 3D point clouds. *Automation in Construction*, 89: 38-48. 2018.
 - [5] Shan, T., Englot, B., Meyers, D., Wang, W., Ratti, C., and Rus, D. LIO-SAM: Tightly-coupled lidar inertial odometry via smoothing and mapping. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5135-5142, Las Vegas, USA, 2020.
 - [6] Beltran, J., Guindel, C., Moreno, F. M., Cruzado, D., Garcia, F., & De La Escalera, A. Birdnet: a 3d object detection framework from lidar information. In *International Conference on Intelligent Transportation Systems (ITSC)*, pages 3517-3523, Maui, USA, 2018
 - [7] Simony, M., Milzy, S., Amendey, K., & Gross, H. M. Complex-yolo: An euler-region-proposal for real-time 3d object detection on point clouds. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops* page 0-0, Munich, Germany, 2018
 - [8] Jocher, Glenn, et al. "yolov5." Code repository <https://github.com/ultralytics/yolov5> (2020).
 - [9] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. Microsoft coco: Common objects in context. In *European conference on computer vision*, page 740-755, Zurich, Switzerland. 2014

Cluster-based Deterioration Prediction of Composite Pavements with Incorporation of Flooding

I. Neema^a, F. Banani Ardecani^a, and O. Shoghli^a

^aDepartment of Engineering Technology and Construction Management,
University of North Carolina at Charlotte, USA

E-mail: ineema1@uncc.edu, fbanania@uncc.edu, oshoghli@uncc.edu

Abstract –

Natural disasters lead to severe deterioration of valuable highway assets, including pavements that should quickly return to service after extreme events such as flooding. Various prediction models were developed to predict pavement performance for several purposes, including maintenance management, budget allocation, and investment strategy. However, limited studies focused on developing a deterioration model for flood-affected composite pavements. This paper proposes a framework for evaluating and predicting the change in composite pavements' roughness due to the flood probability. To this end, a cluster-based pavement deterioration model was developed and applied to a case study of 102 pavement sections from the LTPP database in the United States' eastern region from 2015 to 2019. Then, we used Markov Chain and Monte Carlo simulation on three generated clusters to predict the flood impact on three groups of pavements with different characteristics. The pivotal role of the proposed framework is predicting IRI values due to varying flooding probabilities in different pavement clusters. The results indicate that the pavement tends to deteriorate faster in the initial post-flood years if subjected to heavy or moderate traffic loading and precipitation conditions. This rate will tend to decrease as the age of the pavement increases. For the sections subjected to low traffic loading and low precipitation, the rate of deterioration for the initial post-flood years is less. Still, it will tend to increase as the age of pavement increases.

Keywords –

Pavement deterioration; Markov Chain; Monte Carlo simulation; Composite pavements; LTPP

1 Introduction

Natural disasters and extreme weather events such as flooding, frequent heavy rainfall, and snow contribute to

deterioration in pavement more quickly than normal weather conditions. Some studies have been conducted in the past to understand the impact of flooding on the pavement network [1, 2, 3]. In 2005, two hurricanes, Katrina & Rita, hit New Orleans and the southeastern part of Louisiana in the US, submerging approximately 2,000 miles of road length in flood runoff for five weeks [2]. Highway maintenance management and optimization, especially in the presence of extreme events, are critical [4]. Hence, decision-makers endeavor to develop efficient deterioration models, a key element of maintenance optimization, and establish pre-and-post flood strategies to predict pavement performance under flood conditions to decrease the loss of life and the physical loss of the assets themselves, damage to transport infrastructure, and socio-economic losses. Also, traffic loading, material quality, and surrounding geographical & environmental conditions are among the factors that cause pavement deterioration throughout their lifespan. Due to the fact that these required variables are stochastic, the Markov chain model as a stochastic model can describe the sequence of possible events [5]. Also, various researchers applying the Markov chain theory to construct facilities such as pavements and bridges to predict their deterioration [6, 7, 8, 9] have used the Markov Chain model for predicting pavement deterioration. Although many studies have developed deterioration models, we identified some gaps in the existing frameworks: (1) limited research was conducted on developing probabilistic pavement deterioration models for composite pavement networks. Some of these studies have considered single deterioration models for all the pavement sections; (2) most of the Pavement Management Systems (PMS) used by transportation agencies do not incorporate the effect of flooding in their prediction models.

Considering a single deterioration model for various pavements underestimates or overestimates the pavement condition [10, 11]. Also, when large pavement stretches need to be maintained, prioritizing a particular pavement section's maintenance work becomes complicated. In such a situation, pavement sections' clustering is a

valuable tool for developing a section-wise maintenance strategy [12].

Various clustering algorithms were utilized in the previous studies. Overall, the *K*-means clustering algorithm is easy to apply, accurate, and effective in handling a large amount of data. Also, the *K*-means clustering algorithm was found suitable for unlabeled and non categorized such as LTPP [10, 12, 13].

Furthermore, precipitation and flooding would increase as tropical cyclones' frequency increases [3]. Zhang et al. assessed the effect of the hurricane that occurred in New Orleans. They found a significant difference in the structural strength of pavement between the submerged and non-submerged pavement sections [14]. A recent study was conducted on these flood-affected pavements by Chowdhury et al. to understand pavement's pre-flood and post-flood structural and surface conditions. They developed a deterministic deterioration model and found that pavement tends to lose its strength more rapidly due to flooding [15]. Also, Khan et al. conducted a study on these flood-affected pavements and developed a probabilistic road deterioration model by incorporating flooding effects [6]. To address these challenges, our main research objectives are to create a clustering algorithm to identify pavement sections with similar characteristics and to develop a cluster-based probabilistic deterioration prediction model in composite pavements under different flooding probabilities by leveraging the LTPP data in multiple states.

2 Methodology

We collected the data from the LTPP database, which does not contain the flood-affected pavement sections' information. Then, we grouped pavement sections into three different clusters using the *K*-means clustering algorithm. Then, with the application of Markov chain analysis and Monte Carlo simulation, we developed a pavement deterioration model for each cluster and utilized it to predict the pre-and-post flood IRI (International Roughness Index) values of flood-affected pavement sections. We selected Markov chain, due to the continuous nature of pavement deterioration over time and the fact that the state-space of the deterioration process is finite in number. Furthermore, The Markov chain model focuses on the transition probabilities and the factors responsible for this transition instead of the factors accountable for condition degradation [17]. All these steps are shown in Figure 1. The steps utilized in this research are separately described in the following section.

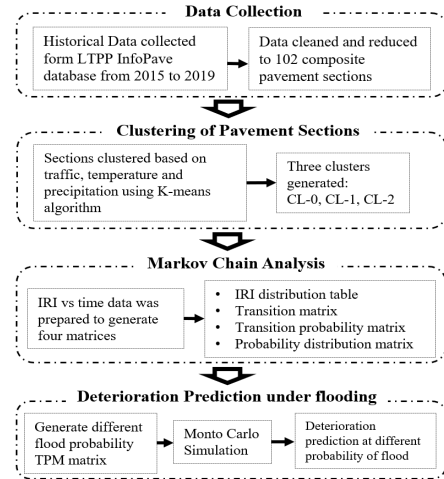


Figure 1. Overall methodology of the analysis

2.1 Data Collection and Clustering

We extracted the IRI, traffic loading (AADTT), temperature, and precipitation data from the LTPP database test sections, belonging to 102 different geographically and spatially composite pavement sections from 2015 to 2019. Figure 2 shows the extent of collected data and their geographical location.

We selected sections data based on three criteria: (1) sections should be composite pavement, (2) sections must be in active monitoring status, and (3) sections should not have maintenance after 2015. The reason behind selecting sections with no maintenance was to remove the improving impact of maintenance (reduction in IRI) in the study. Due to the unavailability of the flood-affected pavement data, we assumed a hypothetical flooding event between 2020 and 2021.



Figure 2. Location of selected pavement sections

It should be noted that all the sections will not follow the same deterioration pattern due to the spatial diversity of collected sections and, in turn, the difference in the contributing factors to their deterioration. To address this issue, we grouped the sections based on the collected historical data (traffic, temperature, and precipitation) using the *K*-means clustering method. The sections within each cluster are homogeneous with each other and heterogeneous between other clusters. Before clustering, we scaled datasets containing traffic, precipitation, temperature, and IRI information between 0 to 1. Then, the *K*-means clustering algorithm was used to cluster the sections. We used the *K*-means method since it is easy to

apply, accurate, and effective in handling a large amount of data [10, 13, 16]. In this study, based on evenly distributing pavement sections, we derived three clusters, CL_0, CL_1 & CL_2, as the optimal number of clusters and as the number of separate deterioration models. Table 1 shows the properties of these clusters.

Table 1. properties of each cluster

Cluster name	Traffic (AADTT)	Temperature (°F)	Precipitation (In)
CL_0	Moderate	Low	High
CL_1	Low	High	Low
CL_2	High	Moderate	Moderate

2.2 Markov Chain Analysis

The Markov chain model is a stochastic model that describes the sequence of possible events. The probability of the next event depends on the current event and not on the event before it [17]. The variables such as traffic loading, environmental aspects, and surface characteristics of the pavements are stochastic. Therefore, we developed the Markov Chain analysis, which comprises four matrices: IRI distribution table, transition matrix, transition probability matrix, and probability distribution matrix. First, we created the IRI distribution table and then developed the other three matrices based on it. The process acquired for developing these matrices is explained below.

2.2.1 IRI Distribution Table

The IRI distribution table was prepared by analysing each section's historical IRI data and then grouping it into each year's respective IRI bucket. A bucket is the IRI range value in m/km. We derived the IRI bucket range based on the following reasons: Pavement sections generally deteriorate with incrementing the IRI value within the range of 0.10 to 0.25 m/km every year [18]. The maximum number of pavement sections in each cluster falls within the IRI range of 0.5 to 1.75 m/km, and due to the low range of the IRI bucket, small changes in the sections were monitored. For uniformly distributing sections into each range bucket, we selected range buckets. The selected IRI bucket range is 0.25 m/km for cluster CL_0 & CL_2, while 0.2 m/km for cluster CL_1; this range is smaller for the cluster CL_1 because more than 90% of the sections have the IRI value less than 1.1 m/km.

2.2.2 Transition Matrix

The transition matrix is an $m \times m$ matrix, where m represents the number of IRI range buckets. This matrix represents the number of sections that will change their IRI value from one range bucket to another in the next year. The IRI data in each cluster were analyzed and

grouped into their respective IRI range bucket for developing the transition matrix. IRI range buckets for cluster CL_0 & CL_2 is 10, while CL_1 is 6. We used the five-year IRI data from 2015 to 2019 of each cluster's sections to develop the transition matrix. We combined the IRI data into four groups, representing the change in IRI between each consecutive year. The pavement sections that show such a decrease in IRI value without any maintenance are not realistic for pavement deterioration, so we did not consider these sections in the analysis.

2.2.3 Transition Probability Matrix

The transition probability is a probability of a pavement section changing its state from the condition i at time t to condition j at time $t + 1$, combined in a matrix called the transition probability matrix (TPM). Several research studies were conducted to derive TPMs using various mathematical methods: the simplest proportion method [19], the expected value method [20], the minimum error method, the percentage transition method, the ordered Probit model, Bayesian technique, and conversion from the deterministic models. In this research, we utilized the percentage transition method. The percentage transition method was feasible for generating the Markov Chain transition probability matrix to derive the change in road condition state with respect to the previous state. This method addresses different explanatory variables used to develop a pavement deterioration model [7]. The transition probability of each pavement section can be calculated using this equation:

$$p_{ij} = \frac{N_{ij}}{N_i} \quad (1)$$

Where p_{ij} is the transition probability from state i to j , N_{ij} is the number of sections transition from state i at time t to state j at time $t + 1$.

For generating the Markov Chain transition probability matrix, we assumed: The condition of the pavement sections cannot be improved without receiving any maintenance treatment, i.e., $p_{ij} = 0$ for $i > j$, the pavement sections which reached their worst condition cannot deteriorate further, i.e., $p_{nn} = 1$, and the pavement section cannot deteriorate by more than one state in a duty cycle.

The TPM is associated with time-independent and time-dependent Markov chain models. This research performed a time-independent Markov chain analysis to develop a pavement deterioration model.

2.2.4 Probability Distribution Matrix

The probability distribution matrix is used to predict the future condition of the pavement at any given year. The pavement network's current condition is termed as

the initial state and described in terms of the initial state vector. The initial state vector of the pavement network is given by [21]:

$$\alpha_0 = [\alpha_1, \alpha_2, \dots, \alpha_n]$$

Initial state vectors assume that all the α_i must be non-negative numbers, and their sum must be equal to one. The TPM is denoted by P and given by [21]:

$$P = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \dots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix}$$

Where p_{ij} indicates the probability that a road is currently in state i and will be in state j next year, as the initial state vector, all the TPM numbers must be non-negative, and the sum of each row must be equal to one. The probability distribution of the states at a future time, say $t = 1$ and at time t , may be calculated from the TPM generated and the initial state vector and is shown in equations 1 and 2 [21].

$$\alpha_1 = \alpha_0 P^1 \quad (2)$$

$$\alpha_t = \alpha_0 P^t \quad (3)$$

Where, α_1 is the probability distribution at time $t = 1$, α_t is the probability distribution at time t , α_0 is the initial state vector at time $t = 0$, and P^t is TPM raised to the power of t . In equations (2) and (3), we assumed that the transition probability matrix (P) of the pavement is constant throughout the time; we assumed the deterioration pavement according to this single transition probability matrix P throughout its lifespan. This equation is used for performing time-independent Markov chain analysis. We assumed that the initial condition of the pavement was perfect. $\alpha(0)$ is the initial state vector.

$$\alpha(0) = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$$

We generated the probability distribution matrix by substituting the variables $\alpha(0)$ and P^t in equation 3.

2.3 Probabilistic Deterioration Model

The main objective of this research is to generate a probabilistic pavement deterioration model. To this end, we used the Monte Carlo simulation to generate a probabilistic deterioration. In the Monte Carlo simulation, an uncertain variable, roughness (IRI), is assigned multiple values by random variables' intervention to achieve multiple results. The implementation of the Monte Carlo simulation was done by transforming the TPM into a cumulative TPM. In each iteration, we generated 20 uniformly distributed random numbers

between 0 and 1 to have the pavement deterioration model for the next 20 years; we generated 1500 deterioration models. Therefore, we generated 30,000 random numbers in the simulation. These random numbers represent the IRI probability, and we used them to predict future IRI values.

2.4 Modeling Flooding events

We designed a framework (a pavement deterioration model showing a change in the IRI of the pavement surface) for incorporating the effect of different flooding probabilities in the deterioration model. Therefore, we used two types of TPMs: non-flooding TPM and flooding TPM. Both TPMs needed to be developed based on the flood-affected pavement sections' historical IRI data.

For making this framework, we considered these assumptions: (1) the initial pavement condition is excellent, and it was developed based on the year 2020, (2) roughness is majorly affected by the accumulation of flooded water on the pavement surface, (3) the annual flooding probability increases, (4) hypothetical flooding event will occur between the year 2020 and 2021 because the LTPP database does not contain the IRI data of flood-affected pavement sections, (5) for the next 3-4 years rehabilitation work will not be done.

To develop a flood-affected predict model: first, we studied the different flood recurrence intervals: 2-years, 5-years, 10-years, and 20-years. Second, we determined the annual flooding probability for developing a deterioration model. The probability of the above-specified flooding events is 1, 0.5, 0.2, 0.1, and 0.05, respectively. Third, we generated the deterioration model of pavement sections at these specified flooding probabilities. Four, we generated the state vectors representing the transition of the pavement sections into various states. Five, we utilized a set of random numbers, are compared with the flooding probability to determine a non-flooding or flooding TPM; the chance of selecting flooding TPM depends upon the chance of flood occurrence. Six, a second random variable is generated to estimate the pavement's future state. We generated the final pavement state by taking the average of all the simulated states. We repeated this process for 10,000 trials and 20 years to generate the deterioration model for different flooding probabilities.

3 Results

The results derived from the suggested methodology for generating the pavement deterioration model in the previous section are illustrated in this section. We depict sections' IRI value, transition, transition probability, and probability distribution matrix for one of the clusters as an example. Then we compare the results of the deterioration model for the composite pavement with or

without flood affecting all clusters.

3.1 Pavement sections clustering

Figure 3 shows the scatterplot of sections through their cluster identity using their geographical location. Table 2 shows a cluster summary based on each of the cluster's traffic, temperature, and precipitation characteristics. Each cluster comprises sections from various states of the United States.

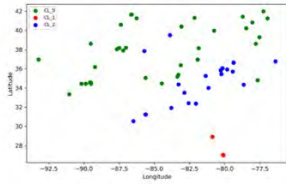


Figure 3. Geographical locations of pavement clusters

Table 2 Section summary of each cluster

cluster name	Number of Sections	Traffic (AADTT)			Temperature (°F)			Precipitation (in)		
		Max	Min	Range	Max	Min	Range	Max	Min	Range
CL_0	44	4748	4	4744	65.8	45.1	20.7	91.3	40.7	50.6
CL_1	14	163	62	101	77.2	71.6	5.6	63.5	31.3	32.2
CL_2	44	5731	17	5714	69.8	52.5	17.3	77.2	34	43.2

3.2 Sections' IRI value

To understand and validate the IRI data quality of each cluster, we prepared the IRI descriptive statistic for each cluster of each year. Table 3 shows The IRI distribution for cluster CL_2. As per the descriptive statistics, sections' IRI value in the specified interquartile range tends to increase every year. It suggests that the pavement sections tend to shift towards the right side of the curve, representing higher IRI values. Higher IRI values represent deterioration in pavement sections. Hence, the IRI data collected showed pavement deterioration and was suitable for developing a deterioration model.

Table 4 shows the IRI distribution table for the CL_0 clusters. This table consists of six columns. The first column represents the IRI bucket range in m/km units. The second column represents the number of sections divided into five sub-sections representing the number of sections in five different years from 2015 to 2019. The third column represents the entire sections, calculated by adding the number of sections each year in a particular range bucket. The fourth column represents the percentage of sections, calculated by taking the summation of total sections in each range bucket and dividing it by the sum of the total sections row. The fifth row represents the cumulative percentage of the sections, calculated by adding the percentage value of the section in each range bucket to the sum of all previous percentage

values. The sixth row represents the lower end range of the bucket. For example, in Table 4, eight sections in the range bucket of 0.5 to 0.75 m/km in 2015, while six sections in 2017 are in the same bucket. The IRI distribution table shows that pavement sections tend to shift in the higher IRI range bucket as time increases. This trend indicates that the IRI of pavement sections is deteriorating with time.

Table 3 IRI descriptive statistics of cluster CL_2

Descriptive statistic	IRI				
	2015	2016	2017	2018	2019
Mean	1.027	1.07	1.141	1.24	1.343
St. Deviation	0.336	0.342	0.344	0.357	0.396
Minimum	0.554	0.55	0.595	0.637	0.665
Maximum	2.126	2.181	2.213	2.561	2.673
25th percentile	0.802	0.837	0.923	1.026	1.085
50th percentile	0.963	1.012	1.063	1.159	1.276
75th percentile	1.206	1.262	1.288	1.438	1.573

Table 4 IRI Distribution table for cluster CL_0

IRI Range (m/km)	Number of Sections					Total Sections	Percent Section	Cumulative Percent	Lower Limit
	2015	2016	2017	2018	2019				
0.500 - 0.750	8	8	6	2	2	26	0.118	0.000	0.500
0.751 - 1.000	15	10	6	2	2	35	0.159	0.118	0.751
1.001 - 1.250	8	11	9	8	7	43	0.195	0.277	1.001
1.251 - 1.500	4	3	9	11	3	30	0.136	0.473	1.251
1.501 - 1.750	3	5	4	8	12	32	0.145	0.609	1.501
1.751 - 2.000	5	4	6	5	7	27	0.123	0.755	1.751
2.001 - 2.250	1	2	3	5	4	15	0.068	0.877	2.001
2.251 - 2.500	0	1	0	1	2	4	0.018	0.945	2.251
2.501 - 2.750	0	0	1	2	3	6	0.027	0.964	2.501
2.751 - 3.000	0	0	0	0	2	2	0.009	0.991	2.751
Total	44	44	44	44	44	220	1.000	1.000	

3.3 Transition Matrix

Table 5 shows the transition matrix for the CL_0 cluster. This table represents the deterioration of pavement sections. Each cell's values represent the number of sections that transitioned its state from one IRI bucket range to another in the next year. This matrix satisfies the requirements of the Markov property, and we further used this matrix in developing the Markov Chain prediction model. In all Transition Matrix tables, the cells showing a zero represent no transition of pavement sections in this IRI bucket range for the next year. For example, in Table 5, 3 pavement sections transitioned from the IRI bucket range of 0.5-0.75 m/km to 0.751-1.0 m/km range, while one sections transitioned from 0.5-0.75 m/km to 1.0-1.250 m/km range.

3.4 Transition Probability Matrix

For developing the transition probability matrix, we used the transition matrix. Table 6 shows the transition probability matrix for the CL_0 cluster. For example, in Table 6, 72.7% of pavement sections remain in the same IRI bucket range of 0.5 – 0.75 m/km for the next year, while 13.6% of pavement sections change their state to

0.751-1.000 m/km IRI range bucket. We used this matrix to develop a probability distribution matrix for a pavement deterioration model in non-flood conditions.

Table 5 Transition Matrix for the Cluster CL_0

IRI Bucket Range (m/km)	Pavement section transition details									
	Number of sections in next year									
Current Year IRI Bucket	0.500 - 0.750	0.751 - 1.000	1.001 - 1.250	1.251 - 1.500	1.501 - 1.750	1.751 - 2.000	2.001 - 2.250	2.251 - 2.500	2.501 - 2.750	2.751 - 3.000
0.500 - 0.750	16	3	1	0	2	0	0	0	0	22
0.751 - 1.000	0	17	9	4	1	2	0	0	0	33
1.001 - 1.250	0	0	24	9	2	0	1	0	0	36
1.251 - 1.500	0	0	1	12	11	1	2	0	0	27
1.501 - 1.750	0	0	0	1	13	6	0	0	0	20
1.751 - 2.000	0	0	0	0	0	13	5	1	1	20
2.001 - 2.250	0	0	0	0	0	0	6	3	1	11
2.251 - 2.500	0	0	0	0	0	0	0	0	3	2
2.501 - 2.750	0	0	0	0	0	0	0	0	2	1
2.751 - 3.000	0	0	0	0	0	0	0	0	0	2

Table 6 Transition Probability Matrix of Cluster CL_0

IRI Bucket Range (m/km)	Transition Probability of Sections									
	Percentage of Sections in Next Year									
Current Year IRI Bucket	0.500 - 0.750	0.751 - 1.000	1.001 - 1.250	1.251 - 1.500	1.501 - 1.750	1.751 - 2.000	2.001 - 2.250	2.251 - 2.500	2.501 - 2.750	2.751 - 3.000
0.500 - 0.750	0.727	0.136	0.045	0	0.091	0	0	0	0	0
0.751 - 1.000	0	0.515	0.273	0.121	0.030	0.061	0	0	0	0
1.001 - 1.250	0	0	0.667	0.250	0.056	0	0.028	0	0	0
1.251 - 1.500	0	0	0.037	0.444	0.407	0.037	0.074	0	0	0
1.501 - 1.750	0	0	0	0.050	0.650	0.300	0	0	0	0
1.751 - 2.000	0	0	0	0	0	0.650	0.250	0.050	0.050	0
2.001 - 2.250	0	0	0	0	0	0	0.545	0.273	0.091	0.091
2.251 - 2.500	0	0	0	0	0	0	0	0	0.600	0.400
2.501 - 2.750	0	0	0	0	0	0	0	0	0.667	0.333
2.751 - 3.000	0	0	0	0	0	0	0	0	0	1.000

3.5 Probability Distribution Matrix

Table 7 shows the probability distribution matrix for the CL_0 cluster and represents the prediction of pavement sections in a particular IRI range bucket. For example, in Table 8, in year 1, 72.7% of sections will remain in the IRI range of 0.5-0.75 m/km, while in year 5, 20.3% of the section will remain this IRI range, and so on forth. After generating these three matrices, the Markovian Chain analysis is completed and used in developing the deterioration model for all the clusters.

3.6 Monte Carlo simulation logic

We compared the random number in each iteration with the cumulative TPM values to find the IRI value of the following year's pavement section. Table 8 shows the cumulative TPM for cluster CL_0.

The IRI range bucket of 0.5 to 0.75 m/km represents the perfectly smooth pavement surface. So, the comparison starts from this IRI range until the cumulative TPM value was greater than the random number; this process continued. The next iteration will start from the same IRI range in which the last iteration

stopped. We continued this procedure for each of the 20 random numbers (number of prediction years) in a trial and repeated for 1500 trials (number of deterioration models). The final deterioration model was generated by taking the average of all the IRI values in their respective year in each iteration.

Table 7 Probability Distribution Matrix of Cluster CL_0

IRI Range (m/km)	Year1	Year2	Year3	Year4	Year5
0.500 - 0.750	0.727	0.529	0.385	0.280	0.203
0.751 - 1.000	0.136	0.169	0.159	0.135	0.107
1.001 - 1.250	0.045	0.101	0.138	0.156	0.157
1.251 - 1.500	0.000	0.032	0.067	0.091	0.105
1.501 - 1.750	0.091	0.132	0.158	0.177	0.191
1.751 - 2.000	0.000	0.036	0.074	0.108	0.135
2.001 - 2.250	0.000	0.001	0.015	0.035	0.057
2.251 - 2.500	0.000	0.000	0.002	0.008	0.015
2.501 - 2.750	0.000	0.000	0.002	0.008	0.018
2.751 - 3.000	0.000	0.000	0.000	0.003	0.012
Total	1.000	1.000	1.000	1.000	1.000

Table 8 Cumulative Transition Probability Matrix for the Cluster CL_0

IRI Bucket Range (m/km)	Cumulative TPM next year									
	0.500 - 0.750	0.751 - 1.000	1.001 - 1.250	1.251 - 1.500	1.501 - 1.750	1.751 - 2.000	2.001 - 2.250	2.251 - 2.500	2.501 - 2.750	2.751 - 3.000
Current Year IRI Bucket										
0.500 - 0.750	0.7273	0.8636	0.9091	0.9091	1	1	1	1	1	1
0.751 - 1.000	0	0.5152	0.7879	0.9091	0.9394	1	1	1	1	1
1.001 - 1.250	0	0	0.6667	0.9167	0.9722	0.9722	1	1	1	1
1.251 - 1.500	0	0	0.037	0.4815	0.8889	0.9259	1	1	1	1
1.501 - 1.750	0	0	0	0.05	0.7	1	1	1	1	1
1.751 - 2.000	0	0	0	0	0	0.65	0.9	0.95	1	1
2.001 - 2.250	0	0	0	0	0	0	0.5455	0.8182	0.9091	1
2.251 - 2.500	0	0	0	0	0	0	0	0	0.6	1
2.501 - 2.750	0	0	0	0	0	0	0	0	0.6667	1
2.751 - 3.000	0	0	0	0	0	0	0	0	0	1

For example, in Table 8, the first random number generated was 0.52. This number was compared with the cumulative TPM value of 0.727, located in the leftmost corner, in the 0.5 to 0.75 m/km IRI range. The random number 0.52 is less than 0.727. Therefore, the IRI transition in the first year did not happen, so we allocated 0.5 to the IRI value in this trial. If the next random number generated was 0.75, it is compared with 0.727, i.e., 0.5 to 0.75 m/km IRI range. The random number 0.75 is greater than 0.727, so the comparison moves to the next IRI range, 0.75 to 1.0 m/km. The cumulative TPM value in this IRI range is 0.864, greater than 0.75; therefore, the comparison stops here, so we allocated 0.751 to the IRI value for the second year. If the third random number again stopped in the same IRI range of 0.75 to 1.0 m/km, the IRI value allocated for the third year would be 0.752. This kind of pattern will continue until a random number stops in a different IRI range.

3.6.1 Results of Deterioration Model with No-flood

Each year's average IRI value is plotted against time

to obtain the deterioration model. Figure 4 shows the deterioration model of each cluster. The trend in Figure 4 for the CL_0 illustrates that the IRI will increase throughout 20 years, representing continuous pavement deterioration. From 2020 to 2025, the IRI of these sections will increase by the average rate of 0.150 m/km each year. After 2026, it will increase by the average rate of 0.135 m/km till 2029; then from 2030, it will increase by the average rate of 0.093 every year till 2034; and then from 2034, it will increase by the average rate of 0.052 m/km every year till 2039. Due to the characteristics of this cluster (heavy precipitation and moderate traffic loading), the increment in the deterioration rate will be highest for this cluster compared to the other two clusters.

The CL_1's roughness value will start at 0.50 m/km in 2020 and reach 1.15 m/km in 2039, representing continuous pavement deterioration over 20 years. From 2020 to 2025, the IRI of these sections will increase by the average rate of 0.02 m/km each year; after 2026, it will increase by the average rate of 0.043 m/km till 2034; and then from 2035, it will increase by 0.030 m/km every year until 2039. Due to the characteristics of this cluster (lower traffic and precipitation), this cluster's deterioration rate is low compared to the other two clusters. However, the deterioration rate will increase as the pavement age increases because of high temperatures.

The CL_2's roughness value will start at 0.585 m/km in 2020 and reach 2.005 m/km in 2039, representing continuous pavement deterioration over 20 years. From 2020 to 2026, the IRI will increase by the average rate of 0.103 m/km each year; after 2026, it will increase by the average rate of 0.067 m/km till 2032; and then from 2033, it will increase by the average rate of 0.057 m/km each year until 2039. Sections in this cluster are subjected to heavy loading conditions; therefore, this trend aligns with expectations. This cluster represents the classic example of a newly constructed pavement. As the age of pavement increases, it tends to deteriorate faster just after its construction, and as the age of pavement increases, the deterioration rate tends to decrease. Hence, this cluster's deterioration rate will be high from 2020 to 2026, and then it will start falling as the year progresses.

3.6.2 Results of Deterioration Model with Flood

Figure 5 shows the predicted roughness by a jump in the IRI value due to hypothetical flooding events of sections in clusters CL_0, CL_1, and CL_2, respectively, at different flooding probabilities between 2020 to 2021. Table 9 shows the hypothetical flooding matrix for the cluster CL_0. For example, in Figure 5, the CL_0 pavement roughness in 2021 will be 1.603 m/km at a 5% probability of flood, while 1.747 m/km at a 50 % probability of flood. The roughness-based deterioration model tends to decrease when post-flood maintenance is applied to the sections; hence, we show roughness

prediction for the first few years at different flooding probabilities. In all clusters of Figure 5, the flood's maximum impact is shown in 2021 because it occurred between 2020 and 2021. This impact tends to reduce as time increases.

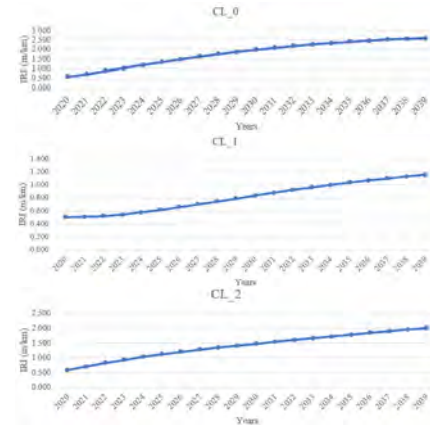


Figure 4. Deterioration Model of the three clusters

Table 9 Hypothetical Flooding TPM for cluster CL_0

Current Year IRI Bucket	Percentage of Sections in Next Year										Total
	0.500 - 0.750	0.751 - 1.000	1.001 - 1.250	1.251 - 1.500	1.501 - 1.750	1.751 - 2.000	2.001 - 2.250	2.251 - 2.500	2.501 - 2.750	2.751 - 3.000	
0.500 - 0.750	0.200	0.8	0	0	0	0	0	0	0	0	1
0.751 - 1.000	0	0.297	0.703	0	0	0	0	0	0	0	1
1.001 - 1.250	0	0	0.48	0.52	0	0	0	0	0	0	1
1.251 - 1.500	0	0	0	0.567	0.433	0	0	0	0	0	1
1.501 - 1.750	0	0	0	0	0.727	0.273	0	0	0	0	1
1.751 - 2.000	0	0	0	0	0	0.611	0.389	0	0	0	1
2.001 - 2.250	0	0	0	0	0	0	0.615	0.385	0	0	1
2.251 - 2.500	0	0	0	0	0	0	0	0.556	0.444	0	1
2.501 - 2.750	0	0	0	0	0	0	0	0	0.444	0.556	1
2.751 - 3.000	0	0	0	0	0	0	0	0	0	1	1

4 Conclusion

The primary objective of this research was to develop a cluster-based probabilistic flood-affected pavement deterioration model for composite pavements. The results indicate that the pavement tends to deteriorate faster in the initial years if subjected to heavy or moderate traffic loading and precipitation condition. This rate will tend to decrease as the age of the pavement increases. Similar trends were shown by the deterioration model of clusters CL_1 and CL_2. Suppose the pavement sections are subjected to low traffic loading and low precipitation. In that case, the rate of deterioration for the initial years is less, but it will tend to increase as the age of pavement increases. Also, the LTPP data illustrate that the impact of flooding on the pavement's roughness is maximum when the probability of flooding is maximum, which confirms the literature.

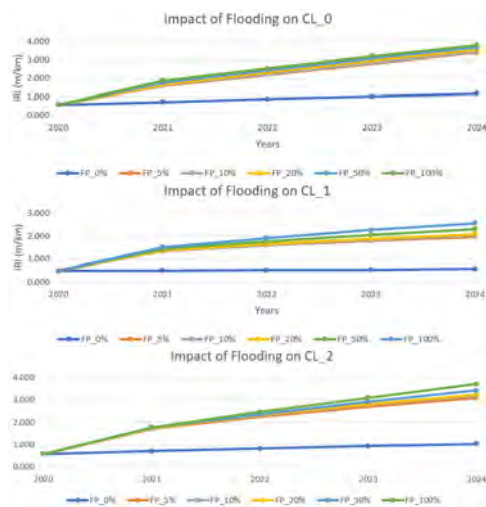


Figure 5. Pavement deterioration model at different probability of flooding for CL_0, CL_1, and CL_2

References

- [1] Lu, D., Tighe, S. L., & Xie, W.-C. International Journal of Pavement Engineering, Impact of flood hazards on pavement performance, 1-7, 2018.
- [2] Sultana, M., Chai, G., Martin, T., & Chowdhury, S. A study on the flood affected flexible pavements in Australia. Paper presented at the 9th international conference on road and airfield pavement technology, 2015.
- [3] Paerl, H. W., Hall, N. S., Hounshell, A. G., Luettich, R. A., Rossignol, K. L., Osburn, C. L., & Bales, J. Recent increase in catastrophic tropical cyclone flooding in coastal North Carolina, USA: Long-term observations suggest a regime shift. Scientific reports, 9(1), 1-9, 2019.
- [4] Shoghli O, De La Garza JM. Multi-Asset Optimization of Roadways Asset Maintenance. In Computing in Civil Engineering (pp. 297-305), 2017.
- [5] Kerali, H., & Snaith, M. NETCOM: the TRL visual condition model for road networks, 1992.
- [6] Khan, M. U., Mesbah, M., Ferreira, L., & Williams, D. J. Developing a new road deterioration model incorporating flooding. Paper presented at the Proceedings of the Institution of Civil Engineers-Transport, 2014a.
- [7] Khan, M. U., Mesbah, M., Ferreira, L., & Williams, D. J. Development of road deterioration models incorporating flooding for optimum maintenance and rehabilitation strategies. Road & Transport Research: A Journal of Australian and New Zealand Research and Practice, 23(1), 3, 2014b.
- [8] Madanat, S., Bulusu, S., & Mahmoud, A. Estimation of infrastructure distress initiation and progression models. Journal of Infrastructure Systems, 1(3), 146-150, 1995.
- [9] Saha, P., Ksaibati, K., & Atadero, R. Developing pavement distress deterioration models for pavement management system using Markovian probabilistic process. *Advances in Civil Engineering*, 2017.
- [10] Sunitha, V., Veeraragavan, A., Srinivasan, K. K., & Mathew, S. Cluster-based pavement deterioration models for low-volume rural roads. International Scholarly Research Notices, 2012.
- [11] Karimzadeh A, Shoghli O, Sabeti S, Tabkhi H. Multi-Asset Defect Hotspot Prediction for Highway Maintenance Management: A Risk-Based Machine Learning Approach. Sustainability. 2022 Apr 21;14(9):4979.
- [12] Sandra, A. K., & Sarkar, A. K. Application of fuzzy logic and clustering techniques for pavement maintenance. Transportation Infrastructure Geotechnology, 2(3), 103-119, 2015.
- [13] Alashwal, H., El Halaby, M., Crouse, J. J., Abdalla, A., & Moustafa, A. A. The application of unsupervised clustering methods to Alzheimer's Disease. Frontiers in computational neuroscience, 13, 31, 2019.
- [14] Zhang, Z., Wu, Z., Martinez, M., & Gaspard, K. Pavement structures damage caused by Hurricane Katrina flooding. Journal of geotechnical and geoenvironmental engineering, 134(5), 633-643, 2008.
- [15] Chowdhury, S., Sultan, M., Chai, G., & Martin, T. A Study on the Flood Affected Flexible Pavements in Australia, 2016.
- [16] Karimzadeh A, Sabeti S, Shoghli O. Optimal clustering of pavement segments using K-prototype algorithm in a high-dimensional mixed feature space. Journal of Management in Engineering. Jul 1;37(4):04021022, 2021.
- [17] Gagniuc, P. A. Markov chains: from theory to implementation and experimentation: John Wiley & Sons, 2017.
- [18] Sayers, M. W. The little book of profiling: basic information about measuring and interpreting road profiles, 1998. Retrieved from
- [19] Wang, K. C., Zaniewski, J., & Way, G. Probabilistic behavior of pavements. Journal of Transportation Engineering, 120(3), 358-375, 1994.
- [20] Jiang, Y., Saito, M., & Sinha, K. C. Bridge performance prediction model using the Markov chain, 1988.
- [21] Ortiz-García, J. J., Costello, S. B., & Snaith, M. S. Derivation of transition probability matrices for pavement deterioration modeling. Journal of Transportation Engineering, 132(2), 141-161, 2006.

Predicting construction productivity with machine learning approaches

L. Florez-Perez^a, Z. Song^b and J.C. Cortissoz^c

^aBartlett School of Sustainable Construction, University College London, UK

^bDepartment of Computer Science, University of Cambridge, UK

^cDepartment of Mathematics, Universidad de los Andes, Colombia

E-mail: l.florez@ucl.ac.uk, zs393@cam.ac.uk, jcortiss@uniandes.edu.co

Abstract –

Machine learning (ML) is a purpose technology already starting to transform the global economy and has the potential to transform the construction industry with the use of data-driven solutions to improve the way projects are delivered. Unrealistic productivity predictions cause increased delivery cost and time. This study shows the application of supervised ML algorithms on a database including 1,977 productivity measures that were used to train, test, and validate the approach. Deep neural network (DNN), k-nearest neighbours (KNN), support vector machine (SVM), logistic regression, and Bayesian networks are used for predicting productivity by using a subjective measure (compatibility of personality), together with external and site conditions and other workforce characteristics. A case study of a masonry project is discussed to analyse and predict task productivity.

Keywords –

Machine learning; Labour productivity; Construction; Crew management

1. Introduction

There are numerous factors that impact task productivity of construction crews such as external conditions, site conditions, and workers characteristics. The interrelationships between the factors and the factors' effects need to be considered by site managers when planning work to better predict task productivity and determine which factors will have a negative impact, so that they can take actions such as identifying which workers will be part of a crew, determining optimal crew size, and allocating workers and crews to the proper tasks. Existing modeling approaches have not considered a subjective and essential characteristic of the workforce (compatibility of personality) and the interrelationships between workforce characteristics, site and external conditions to accurately predict task productivity of construction crews. Therefore, a model to better

understand the interactions between the factors and their combined effects to better predict task productivity needs to be developed. Additionally, the developed models make predictions without considering various levels of productivity that are useful for planning future work and establishing acceptable levels of production. In this work, levels of productivity are classified in 3 classes (high, medium and low) and the class definition was based on the number of standard deviations from the empirical mean productivity.

2. Machine learning in construction

Machine learning (ML) as a major area of interest within the field of artificial intelligence (AI) has been applied to construction and the built environment research for more than two decades [3], [5-7]. Supervised learning, including logistic regression, support vector machine (SVM), and random forest among others, are the most widely used type of ML algorithms in the construction field. Supervised learning is a ML method that learns a function that maps the input to output based on example input-output pairs and infers a function from labelled training data consisting of a set of training examples [10]. In most construction applications, supervised learning is used for data classification [8]. Unsupervised learning that focuses on data reduction and clustering problems is a method of machine learning. No pre-labelled training examples are given, and the input data is automatically classified or grouped [9-10]. Because in the construction industry the unlabelled data has many limitations, relatively speaking, the information that can be extracted is less than the labelled data, so unsupervised learning is less used in the construction industry than supervised learning.

Numerous ML applications have been developed for the construction industry. Examples include supervised learning such as logistic regression [15], SVM [13,16], AdaBoost [17], Random Forest [13,18], Bayesian Network [19], and KNN [20]; unsupervised learning examples such as Principal Component Analysis (PCA) [21], and K-means [22]; Deep learning examples such as

Convolutional Neural Networks (CNNs) [23], and Recurrent Neural Networks (RNN) [24]. In the construction industry, the common experience of applying ML methods is that these models often fail when dealing with real-world problems. One possible reason is that computer engineers currently establish many ML methods, and they lack the knowledge and experience of the corresponding industry application scenarios, which leads to the existence of many errors and biases in the process of model building [7]. The challenges to the application of ML in the construction industry are threefold. Significant challenges are the lack of data, the accuracy of some ML algorithms, and the complexity of the site environment among others [8]. When dealing with these challenges, it is essential to note that it is not just about building predictive models when solving real-life construction problems. In addition to training reliable ML models, it is also necessary to consider how to integrate the experience and knowledge of construction industry experts into the model building process as a comprehensive framework [7].

ML has created new opportunities for revealing, quantifying, and understanding labour productivity in the construction process. Determining the factors that affect the productivity of construction labour is often the first step in establishing research models. The performance of these models greatly depends on the input factors. Two research gaps need to be addressed urgently. One is the identification of comprehensive factors, and the other is the weight and relationship of these factors. In identifying the factors, some studies have not considered essential factors such as subjective factors related to the workforce (compatibility of personality). In determining the relation between the factors, many studies ignore the correlation between different types of productivity factors and only consider these as independent and isolated factors or they simplify the correlation between the factors [11-13]. Future models can obtain information from real data to export behaviour or data-rules, identify critical factors, and predict productivity performance. The more fundamental factors will play a more accurate role in forecasting productivity, which requires the model to first clarify the hierarchical relationship between the factors to develop sensible strategies to better predict labour productivity. By combining ML, construction site realities, and the builder's understanding of actual engineering problems, new models can be developed to correctly represent construction scenarios.

3. Case study

To illustrate the application of ML in construction, let us consider a real-life masonry project in Atlanta, GA in the United States. The project consisted of two main buildings with an approximate area of 950,000 ft².

Building A was mixed used space for upscale commercial stores and residential apartments. Building C had only upscale commercial stores. Up until the first storey, the floor use for both buildings were identical as well as the masonry units used. The second underground floor was used for parking, the first underground was used for storage for commercial clients, and the first floor was for commercial stores. Building A had 12 more floors of residential apartments. A data set was collected during the construction phase of this project, and it was used to determine the relationships between factors and factors' effects on task productivity. The reader is referred to [5] for an extended description of the factors.

Determining the factors that affect productivity was the first step for establishing the ML models. Various construction studies have used external conditions (temperature, humidity, wind speed, precipitation) [27, 28]; site conditions (floor level, work type, workload, complexity of task, congestion) [27,28,29]; and workers characteristics (age, experience, skill, crew size, personality [30, 31, 32], [26]). Note that there are numerous factors that affect productivity. When applying ML models, there is a game between being too narrow and too broad with the amount of information used to make predictions. If too narrow, some interrelationships might be lost and if being too broad the algorithms do not learn, and accuracy is low. In some cases, it seems that eliminating information might be more useful so that the levels of accuracy are acceptable, while allowing the models to learn from the information at hand. The best approach is often trial and error.

Productivity refers to the measure of the full utilisation of inputs to achieve an expected output [2]. In the field, productivity is measured at the task level, for practical considerations. Since masonry is one of the most labour-intensive trades in construction, the task-level model will be used in this study as single-factor productivity, which is expressed as the unit of work per labour hour [14]. To detail the factors three sections, namely, external conditions, site conditions, and workers characteristics describe typical attributes of masonry jobsites.

3.1 External conditions

The external conditions refer to the temperature regarding the building the crews were working at the specific time the data were collected. The temperature, both low and high temperature, was recorded for the day at the time the data were collected.

3.2 Conditions in masonry sites

Extensive site observations and interviews with masonry practitioners [6] were used to collect information of typical site conditions related to crews and walls. The crew size is determined by the length of the wall. A rule of thumb used by masonry practitioners is one mason for every 15-20 ft of wall [6] and it varies on site (depending on wall lengths) between one to five masons. The masonry tasks (walls) were classified in three different levels, namely: easy (difficulty = 1), normal (difficulty = 2), and difficult (difficulty = 3). This system considers crew sizes of one to five masons, but it was trained for two to three masons since it was the typical number of masons in this case study and dataset at hand.

3.3 Workers' characteristics

Masons have different ages and length of experience in the field, which could have impact on their productivity together with other external factors and conditions in the construction sites. The size of crews was annotated as it happened on site, which is typically determined by the superintendents. Compatibility between masons, defined as a measure of the capability of a group to interact and work well together to attain higher productivity [6], was collected through extensive site visits and interviews with masonry practitioners.

3.4 Dataset

The dataset of masonry work contains 1,977 data samples with 14 dimensions for training and prediction. Each of which includes the following features: low temperature of the day; high temperature of the day; level of difficulty of the masonry task; number of masons; compatibility of mason 1; compatibility (mason 1 & mason 2); compatibility (mason 1 & mason 3); compatibility (mason 2 & mason 3); age (mason 1,2&3); experience (mason 1,2&3). Productivity was measured by the number of blocks built in 5-minute time intervals. The dataset was divided into training and testing data sets and input data labelled by their corresponding productivity, which is measured by the number of blocks built per minute per mason. In the experiments, the level of productivity was classified as high (≥ 0.6), medium ($(0.2, 0.6]$), and low (< 0.2), considering that the average productivity of the whole data set is 0.433 and the standard deviation is 0.182. To ensure the input data was internally consistent, standardisation was implemented using Scikit-learn to pre-process the data. The dataset was balanced so that each class had approximately the same amount of data samples. To prevent the trained model from overfitting on certain classes while underfitting on other classes, enough duplication of the data in the minority classes were added to the dataset.

Then, the dataset was shuffled and divided into training, validation and testing sets in the ratio 2400:700:711. Further details of data processing can be found in [5].

3.5 Experiments

KNN [8] is a simple, supervised machine learning algorithm that can be used to solve both classification and regression problems. KNN classifier determines the class of a data point by majority voting principle. For example, if K is set to 5, the classes of 5 closest points are checked, and the prediction is done according to the majority class. To determine how close between the data points, Euclidean distance is one of the most used distance measurements. In this case, a KNN model where $K = 10$ is built using the Scikit-learn library and achieved the classification accuracy of 97.5% with grid search method. Different values of K have been explored and when $K = 10$, the model achieves the highest accuracy. The confusion matrix is plotted in Figure 1.

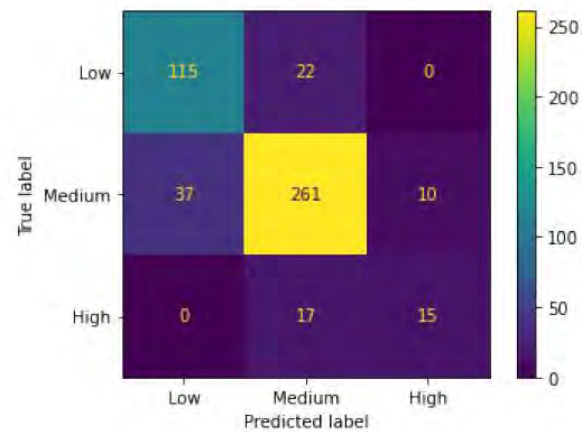


Figure 1. Confusion matrix of KNN

A logistic regression model is built using the Scikit-learn library and Liblinear [2] and achieved the classification accuracy of 85.2%. Logistic regression [3] is a statistical learning technique categorised in supervised machine learning methods for classification tasks. Logistic regression uses the sigmoid function 2, which takes any real value between zero and one. The logistic regression algorithm becomes a classification technique only when a decision threshold (default = 0.5) is brought into the picture.

A DNN is a deep learning model that is focused on emulating the learning approach that humans use to gain certain types of knowledge. Like biological neurons, which are present in the brain, DNN also contains several artificial neurons, and uses them to identify and store information, then transform the input into classification or regression results. In the experiments, rectified linear unit (ReLU) was chosen as it is commonly used and has

a well-performing activation function. In the output layer, log softmax was chosen to predict the class of the productivity level. Cross entropy loss was selected as the loss function with a two-layer architecture (14-8-3), that is, there are 14 neurons in the input layer, 8 in the hidden layer and 3 in the output layer, and a three-layer architecture (14-10-5-3). Then, the trained model was tested on the testing dataset. The best classification accuracy obtained, after probing with different architectures was 97.5%. A Support vector machine (SVM) [26] is a machine learning technique to find a hyperplane in an N-dimensional space (N – the number of features) that distinctly classifies the data points. In this task, SVM classifiers with Sigmoid kernels as expressed by the formula stated in (1):

$$K(X, Y) = \tanh(\gamma \cdot X^T Y + \gamma) \quad (1)$$

These were deployed to classify the level of productivity and the result, and the accuracy obtained was 95.8%. A Bayesian network is a type of the probabilistic graphical modelling technique that is used to compute uncertainties by using the concept of probability. Bayesian networks can take an observed event and forecast the likelihood that any of numerous known causes played a role. In this task, a Bayesian network was developed to classify the level of productivity and the accuracy obtained was 81.2%.

The confusion matrix is a performance measurement for ML classification problems to check the performance of a classification model on a set of test data for which the true values are known. The column represents the ground truth of the classification, and the row stands for the predicted classification results. The confusion matrix for logistic regression is shown in Figure 2. For instance, in the prediction of the “low productivity” tasks, the KNN model correctly classified 273 of the samples, while misclassified 45 of the “low productivity” as “medium productivity.”

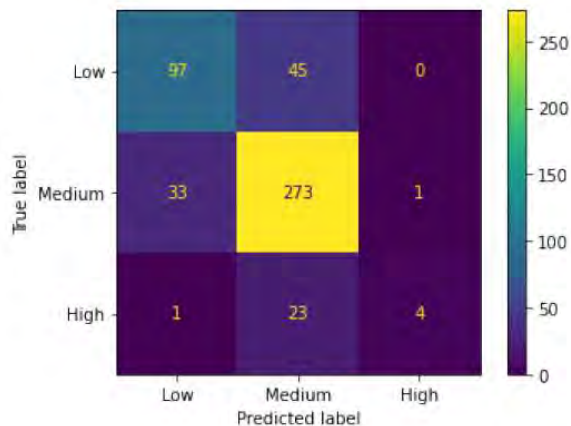


Figure 2. Confusion matrix of logistic regression

The KNN model achieved the highest accuracy (97.5%) on predicting the level of productivity of the construction project (see Table 1).

Table 1. Performance comparison of ML models

ML model	Classification	F1
	accuracy	Score
DNN with 2 layers	92.6%	0.903
DNN with 3 layers	88.2%	0.882
KNN (k=10)	97.5%	0.986
KNN (k=100)	81.4%	0.794
Logistic regression	85.2%	0.802
Sigmoid SVM	95.8%	0.965
Bayesian Network	81.2%	0.761

By predicting the level of productivity of the masons, the project manager can hence make decisions on how to group the masons based on their suitability and thus achieve maximum productivity and efficiency.

By removing all compatibility features (compatibility of mason1, compatibility (mason1 & mason2), compatibility (mason1 & mason3), compatibility (mason2 & mason3), the necessity of the compatibility feature could be determined. The classification results on the dataset without compatibility features are shown in Table 2. As shown in Table 2, removing the compatibility features from the input dataset results in a slight degradation on the accuracy of the classification (97.4%).

Table 2. Performance comparison of ML models (without compatibility)

ML model	Classification	F1 Score
	accuracy	
DNN with 2 layers	91.9%	0.891
DNN with 3 layers	89.8%	0.868
KNN (k=10)	97.4%	0.981
KNN (k=100)	81.5%	0.774
Logistic regression	85.4%	0.865
Sigmoid SVM	96.0%	0.944
Bayesian Network	83.6%	0.802

Figure 3 shows the feature importance graph for KNN (k=10). The features are in order: 0=low temperature, 1=high temperature, 2= level of difficulty, 3=number of masons, 4= compatibility of mason1, 5= compatibility (mason1&mason2), 6= compatibility (mason1&mason3), 7= compatibility (mason 2&mason3), 8= age (mason 1), 9=age (mason 2), 10= age (mason 3), 11=experience (mason 1), 12= experience (mason 2), 13= experience (mason 3). As shown in Figure 3, the lowest temperature (feature 0) and highest temperature (feature 1) and the difficulty of the tasks (feature 2) have the greatest impact

on predicting the productivity. Additionally, it is shown that other factors such as the experience, crew size, and compatibility play a role in the prediction power of the model.

The fact that the environmental conditions and difficulty of the task at hand prevail can be somehow explained by the fact that masonry is physically and labour intensive. Results also show that by removing compatibility gives mixed results regarding the prediction power of the model. Accuracy slightly improved in some cases (1-2%) and in other cases it was lower (1-2%). These results somehow show that compatibility might not impact productivity. However, a careful consideration in this study should be taken, as the measure for compatibility is largely subjective. In this case, there is no personality test and compatibility was determined via a subjective measure given by the foreman. While subjective, the foreman's opinion was used because it is based on the long time and careful observations of the masons and crews, she/he has managed on site. To make the metric more precise, personality tests could be done similarly to this study [25].

It might be appropriate to add here that the findings of this study contrast with [25] where it was found that personality compatibility does impact productivity and it has a positive correlation. Perhaps this is because in this study the crews were working in larger walls that were often divided by construction joints so that the foreman on site could have a better control. In the previous study [25], crews were working in residential projects that have smaller walls and crews thus required more interactions between the masons. It is interesting to look closely at the feature importance (age and wall difficulty) and how it impacts productivity. A thorough analysis of these results suggest that looking at the age of the masons might be more important to form productive teams. Perhaps pairing a young mason with a more experienced mason is better than always pairing experienced masons [6]. An additional consideration might be considering experience and compatibility depending on the type of project as well. Compatibility might be a more important factor to consider when forming crews for residential projects and easy walls, while experience might be a more important factor to consider when forming crews for commercial projects and difficult walls. These of course require further investigations.

Note that while the results might be different for other construction work, the methodology could be replicated in a similar way using ML and can be reused for projects that involve intensive labour works such as dry wall activities. Additionally, productivity measures vary depending on the type of work (structural consulting work might measure drawings per month). It might be interesting to normalise the productivity data in other activities to be able to compare with other type of works

than those considered in this paper. As the definition of low or high productivity depends on the standard deviation, it will not depend on the type of work. By normalising it, it allows to classify compatibility as well and apply the methods of this study across different activities.

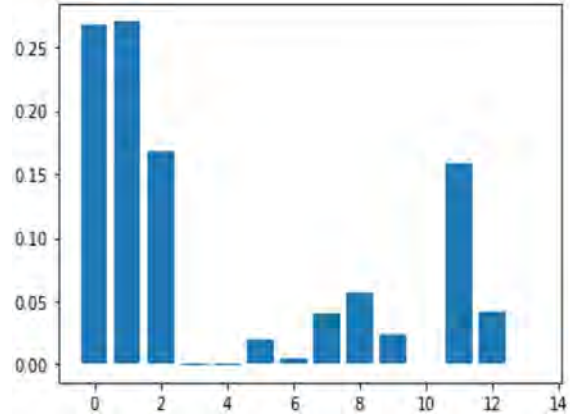


Figure 3. KNN feature importance K=10

4. Applications to optimization

Our work, and in fact any work focused on classifying and predicting productivity, can in principle be used to increase productivity. A first approach would be the following. Assume that we have a pool of N masons, from which we must select k crews of $n_1, n_2, n_3, \dots, n_k$ masons respectively. Then, we can form an exhaustive list of all possible teams of crews of the aforementioned sizes out of the total pool of masons. When N and the n_j 's are small this is quite feasible. Indeed, if $N = 12$, $k = 4$ and $n_j = 3$ for all j , the total number of teams made of four crews of three masons is given by:

$$\binom{12}{3} \binom{9}{3} \binom{6}{3} = 369,600$$

We measured the time taken by some of the algorithms we used, and once trained, the DNNs took around 0.2136 seconds to classify 697 data samples; hence, for the hypothetical example introduced above, it would take around 2 minutes to find an optimum in a single machine. The reader must consider that if the accuracy is at 85% (which is what we have now), then the probability of obtaining an optimum via this approach in the case described above is about 0.52, the colloquial toss of a coin. One way of improving these numbers is to obtain accuracies of at least 98%: in this case we would have that the probability of obtaining an optimum from exhaustive evaluation is around 0.90. Is it possible to reach these accuracy levels? Theoretically speaking, the answer is yes. For instance, in the case of DNNs by the

Universal Approximation Theorem (of course trying to avoid overfitting) and training with more data, this could be achieved or perhaps considering other factors that affect productivity, but this might have to wait some time.

Although the previous paragraph presents a simple, and perhaps obvious way of finding an optimum using classification methods in the case of larger N , say for instance $N = 21$, $k = 7$ and $n_j = 3$ for all j , an exhaustive approach becomes impracticable: in this case, the number of possible teams is 1.825×10^{14} . In this case, even if it only takes 10^{-6} seconds to predict the productivity of a crew, doing an exhaustive evaluation of all possible teams would take around 50,000 hours (around 5 years). Besides, if the model requires a prediction of the weather, more than 24 hours is too much to run the optimisation procedure.

Still, hope is not to be lost as something can be done. Assuming a high accuracy of prediction, we may proceed as follows: take a sample large enough but manageable and obtain an optimum from the sample. Using this approach (and to simplify our reasoning even further, assuming perfect accuracy), we can, for instance, if all we want is to get a team in the first decile of productivity, take a sample of 500 teams, and the probability of not obtaining a sample in the (real not estimated) first decile is 1.4×10^{-23} , so it is quite probable, and we could even say almost certain, that the team with the best productivity in the sample is in the real first decile of productivity. This sampling procedure gives us in turn a way to estimate the i -th decile of a distribution, and in this sense estimate the highest and lowest levels of productivity of the assembled teams.

5. Limitations and conclusions

The convergence of ML into the construction management domain provides the capability to learn the highly nonlinear, complex relationships between task characteristics, site conditions, and the characteristics of workers. This study leverages the power of ML and the existing wealth of real-life scenarios in construction projects to make productivity predictions. The experiments show moderate (logistic regression and Bayesian network) to very good (DNN with 2 layers, KNN with $K=10$ and SVM) accuracy when using ML models to classify and predict productivity with the data collected. The models were trained with only 1,700 data points. Given that the network is relatively small compared with the total number of data points, 700 data points were used to determine the accuracy. It is unknown if the network's accuracy will improve with more data. Adding more perceptrons to the DNN can theoretically improve accuracy, but it must be done with care so that the size of the network is still small compared

with the training data. If this can be successfully done, then this model can be used to plan a construction project to improve performance.

This confirms the appreciation that ML methods can be used as decision-making tools in managing crews in building construction and its use deserves to be more widely known. This can reduce the reliance on empirical estimates and computationally expensive analytical evaluations and better estimate productivity of construction crews. If these classification algorithms can be paired with some optimisation strategies as proposed in this paper, this would also confirm the appreciation that ML methods can be used as decision-making tools in managing crews in building construction, and its use deserves to be more widely known. If simple strategies for evaluating the performance of teams (groups of crews) with classification algorithms (again, we suggest using probabilistic approaches as the number of teams can be huge for exhaustive methods), this can reduce the reliance on empirical estimates and computationally expensive analytical evaluations.

There are some limitations of this study. Note that the models have not been validated yet, but we expect to do this in the future, hopefully with the support and collaboration from interested industry partners that could benefit from using the techniques developed in this study. Additionally, it would be interesting to run the experiments by removing factors to determine how much the accuracy is reduced by looking at the dynamics of the teams. For instance, determining if removing the size of the crews alongside compatibility impacts the performance of the models. These studies can also be refined using more productivity classes (in this study we used three). However, the amount of data at hand did not allow us to establish more refined productivity classes so that the algorithms would have reasonable accuracies. One of the challenges behind ML models is to obtain reliable data. We did run experiments with four classes of productivity (high, medium-high, medium-low and low) and obtained an accuracy of around 70%.

Acknowledgments

The authors thank University College London's Bartlett Research Grants Scheme 2020-2021 for funding this study. Data can be shared upon request to the first author.

References

- [1] Chan, P., & Kaka, A. (2004). Construction productivity measurement: A comparison of two case studies. Paper presented at the 20th Annual ARCOM Conference, Edinburgh, Scotland.
- [2] Crawford, P., & Vogl, B. (2006). Measuring

- productivity in the construction industry. *Building Research & Information*, 34(3), 208-219.
- [3] Dissanayake, M., Robinson Fayek, A., Russell, A. D., & Pedrycz, W. (2005). A hybrid neural network for predicting construction labour productivity. In *Computing in Civil Engineering* (2005) (pp. 1-12).
 - [4] Li, Y., & Liu, C. (2012). Labour productivity measurement with variable returns to scale in Australia's construction industry. *Architectural Science Review*, 55(2), 110-118.
 - [5] Florez-Perez, L., Song, Z. and Cortissoz, J. C. (2022). Using machine learning to analyze and predict construction task productivity. *Computer-Aided Civil and Infrastructure Engineering*, pp. 1–15.
 - [6] Florez, L. (2017). Crew Allocation System for the Masonry Industry. *Computer-Aided Civil and Infrastructure Engineering*, 32(10), pp. 874–889. Doi: 10.1111/mice.12301.
 - [7] Bilal, M., and Oyedele, L. O. (2020). Guidelines for applied machine learning in construction industry—A case of profit margins estimation. *Advanced Engineering Informatics*, 43, 101013.
 - [8] Xu, Y., Zhou, Y., Sekula, P., and Ding, L. (2021). Machine learning in construction: From shallow to deep learning. *Developments in the Built Environment*, 6, 100045.
 - [9] Celebi, M. E., and Aydin, K. (2016). Unsupervised learning algorithms: Springer.
 - [10] Hastie, T., Tibshirani, R., & Friedman, J. (2009b). Unsupervised learning. In *The elements of statistical learning* (pp. 485-585): Springer.
 - [11] Alaloul, W. S., Liew, M. S., Wan Zawawi, N. A., Mohammed, B. S., and Adamu, M. (2018). An Artificial neural networks (ANN) model for evaluating construction project performance based on coordination factors. *Cogent Engineering*, 5(1), 1507657.
 - [12] Ebrahimi, S., Fayek, A. R., and Sumati, V. (2021). Hybrid Artificial Intelligence HFS-RF-PSO Model for Construction Labor Productivity Prediction and Optimization. *Algorithms*, 14(7), 214.
 - [13] Momade, M. H., Shahid, S., Hainin, M. R., Nashwan, M. S., and Tahir Umar, A. (2020). Modelling labour productivity using SVM and RF: a comparative study on classifiers performance. *International Journal of Construction Management*, 1-11.
 - [14] Shehata, M. E., and El-Gohary, K. M. (2011). Towards improving construction labor productivity and projects' performance. *Alexandria Engineering Journal*, 50(4), 321-330.
 - [15] Wilson, O., Sharpe, K., and Kenley, R. (1987). Estimates given and tenders received: a comparison. *Construction Management and Economics*, 5(3), 211-226.
 - [16] Shu-quan, L., Xin-li, Z., Zhi-qiang, L., Li-xia, F., Lan, M., and Qiu-li, G. (2006). *Dynamic Monitoring on Construction Safety Based on Support Vector Machine*. Paper presented at the 2006 International Conference on Management Science and Engineering.
 - [17] Teizer, J., and Vela, P. A. (2009). Personnel tracking on construction sites using video cameras. *Advanced Engineering Informatics*, 23(4), 452-462.
 - [18] Liu, Z., Sadiq, R., Rajani, B., and Najjaran, H. (2010). Exploring the relationship between soil properties and deterioration of metallic pipes using predictive data mining methods. *Journal of Computing in Civil Engineering*, 24(3), 289-301.
 - [19] McCabe, B., AbouRizk, S. M., and Goebel, R. (1998). Belief networks for construction performance diagnostics. *Journal of Computing in Civil Engineering*, 12(2), 93-100.
 - [20] Wang, J., and Ashuri, B. (2016). *Predicting ENR'S construction cost index using the modified K nearest neighbors (KNN) algorithm*. Paper presented at the Construction Research Congress 2016
 - [21] Dogbegah, R., Owusu-Manu, D., and Omoteso, K. (2011). A principal component analysis of project management competencies for the Ghanaian construction industry. *Australasian Journal of Construction Economics and Building*, The, 11(1), 26-40.
 - [22] Fang, Q., Li, H., Luo, X., Ding, L., Luo, H., Rose, T. M., and An, W. (2018). Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. *Automation in Construction*, 85, 1-9.
 - [23] Gao, X., Shi, M., Song, X., Zhang, C., and Zhang, H. (2019). Recurrent neural networks for real-time prediction of TBM operating parameters. *Automation in Construction*, 98, 225-235.
 - [24] Florez, L., Armstrong, P., and Cortissoz, J. C. (2020). Does compatibility of personality affect productivity? Exploratory study with construction crews. *Journal of Management in Engineering*, 36(5), 04020049.
 - [25] Hearst, M. A. (1998). Support vector machines. *IEEE Intelligent Systems*, 13(4), 18–28.
 - [26] Dissanayake, M., Fayek, A. R., Russell, A. D. and Pedrycz, W. A hybrid neural network for predicting construction labour productivity. In *Proceedings of the International Conference of Computing in Civil Engineering*, American Society of Civil Engineers. Eds Lucio Soibelman and Feniosky Peña-Mora (pp 1–12).
 - [27] Golnaraghi, S., Zangenehmadar, Z., Moselhi, O., and Alkass, S. (2019). Application of artificial

- neural network(s) in predicting formwork labour productivity. *Advances in Civil Engineering*, 2019, 1–11.
- [28] Ebrahimi, S., Fayek, A. R., and Sumati, V. (2021). Hybrid artificial intelligence HFS-RF-PSO model for construction labor productivity prediction and optimization. *Algorithms*, 14(7), 214.
- [29] Alaloul, W. S., Liew, M. S., Zawawi, N. A. W., Mohammed, B. S., and Adamu, M. (2018). An artificial neural networks (ANN) model for evaluating construction project performance based on coordination factors. *Cogent Engineering*, 5(1), 1507657.
- [30] Jassmi, H.A., Ahmed, S., Philip, B., Mughairbi, F. A., and Ahmad, M.A. (2019). E-happiness physiological indicators of construction workers' productivity: A machine learning approach. *Journal of Asian Architecture and Building Engineering*, 18(6), 517–526.
- [31] Oral, E. L., Oral, M., and Andaç, M. (2016). Construction crew productivity prediction: Application of two novel methods. *International Journal of Civil Engineering*, 14(3), 181–186.

Automated Checking of Scaffold Safety Regulations using Multi-Class 3D Segmentation

J. Kim^a, J. Kim^a, N. Koo^a and H. Kim^a

^aDepartment of Civil and Environmental Engineering, Yonsei University, Korea

E-mail: john101010@yonsei.ac.kr, kah5125@yonsei.ac.kr, ahappyto@gmail.com, hyoungkwan@yonsei.ac.kr

Abstract

Scaffolds, one of the most widely used temporary structures, are prone to safety-related accidents. Despite the fact, checking regulations for a scaffold is manually being conducted, which is inefficient, especially for a large construction site. This paper proposes an automated method to check safety regulations regarding scaffolds on sites. 3D point cloud data obtained from Terrestrial Laser Scanning (TLS) is first processed by a deep learning-based 3D segmentation to automatically identify major entities. Then, a simple rule-based algorithm is applied to the segmented data to check for three types of major safety-related regulations. The result of our experiment shows potential for successfully automating scaffold safety checking at a construction site.

Keywords –

Deep learning; Scaffold; Point cloud; Semantic Segmentation; Safety regulation checking; Terrestrial Laser Scanning (TLS)

1 Introduction

Korea Occupational Safety and Health Agency's 2021 report on industrial accidents shows that the construction industry has the highest number of occupational deaths, constituting 50.6% of those of all industries [1].

One of the major reasons for those accidents are caused by temporary structures such as scaffolds [2]. Scaffolds are, due to their temporary nature, often not seriously considered and are prone to safety-related accidents [2]. It is, thus, necessary to check scaffolds on sites for violations of safety regulations. However, manual observation can be time-consuming and inaccurate, especially for large-scale construction sites.

Point cloud data acquired by laser scanning contain rich 3D geometric information of a site or an object. Pioneering studies, such as [3], demonstrated how TLS data can be processed in relation to CAD data. TLS data were also proven to have potential for safety regulation checking of scaffolds [4]. Recent advancement of deep

learning technology on point clouds such as [5] allowed for a more effective segmentation of scaffold point cloud data from a large-scale construction site [6].

The proposed methodology fully automates the safety-related regulation checking process of scaffolds on construction sites. Thanks to deep learning-based point cloud segmentation and rule-based algorithm, multi-class segmentation and safety regulation checking of scaffolds are successfully conducted.

2 Methodology

The proposed methodology is divided into two parts: multi-class segmentation and regulation checking. For multi-class segmentation, point cloud data of a construction site acquired by a terrestrial laser scanner (FARO m70) are used as the input of RandLA-Net [5]. RandLA-Net used in this study is trained to classify each point into one of six classes: 'stairs,' 'work platform,' 'uprights,' 'guard rail,' 'bracing,' and 'background.' For regulation checking, a simple and robust rule-based algorithm is used to check if the scaffold violates safety regulations. The safety regulations to be used were selected based on the Korea Occupational Safety & Health Agency safety work guidelines on steel pipe scaffold (KOSHA Guide C-30-2020 [7]) and modular scaffold (KOSHA Guide C-32-2020 [8]). Details of the regulations are shown in Table 1.

Table 1. Considered Regulations

#	Regulations
I	Attachment status of working platforms
II	Attachment status of stairs
III	Attachment status of guard rail

2.1 Multi-Class 3D Segmentation

RandLA-Net is a neural architecture structured for efficient 3D semantic segmentation on large-scale point clouds by using random sampling instead of complex point selection approaches. By using a local feature

aggregation module, RandLA-Net can capture complex local structures [5]. RandLA-Net has proven to effectively extract small entities from large scenes by showing high performance [5] on the Semantic3D dataset [9], a dataset of terrestrial laser scans of outdoor scenes. Previous study also proves the high performance of RandLA-Net when capturing features of scaffolds in large construction sites [6].

The RandLA-Net used in this study is composed of five sets of encoding and decoding layers. For the transfer learning, pre-trained parameters trained with the Semantic3D dataset were used as initial parameter values. Then, the network was fine-tuned by re-training the parameters of the inner six layers of the encoder and decoder.

2.2 Regulation checking algorithm

To check regulations of Table 1 with labeled outputs of RandLA-net, a representative 'uprights' coordinate (x and y) is first determined for each 'uprights' of the scaffold by peak finding based on the point density. The height (z-value) of each 'work platform' is also determined as the floor height from the data distribution. Then, potential fields of 'work platform' and 'guard rail' are calculated on the x-y plane based on the 'upright' coordinates by using the standard width of scaffolding entities. Figure 1 shows the potential fields of 'work platform' and 'guard rail'

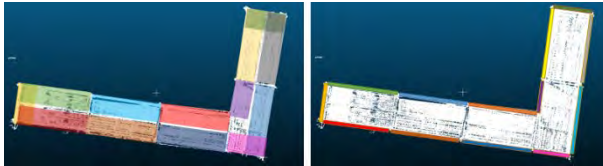


Figure 1. Potential fields of 'work platform'(left) and 'guard rail'(right); Different color shows different potential field instances.

2.2.1 Checking for Regulations I & II

Using the floor heights, 3D bounding boxes of potential fields for platforms are defined on each floor. To check regulation I (attachment status of working platforms), the 'work platform' points need to be extracted to see if there are enough points within each box. If a bounding box turns out to have no work platform, it is now time to check regulation II (attachment status of stairs). That is, the same bounding box is searched to see if there exists 'stairs' class points. The checking flow is shown in Figure 2.

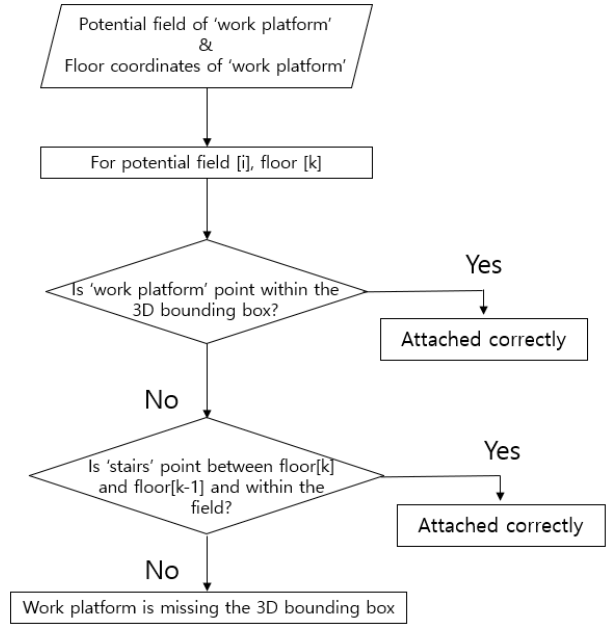


Figure 2. The checking flow for regulations I & II in a 3D bounding box.

2.2.2 Checking for Regulation III

A 3D bounding box for 'guard rail' is generated using potential fields of 'guard rail' and the z-values between two floors. On each bounding box, the presence of 'guard rail' class points is checked for regulation III (attachment status of guard rail). The checking flow is shown in Figure3.

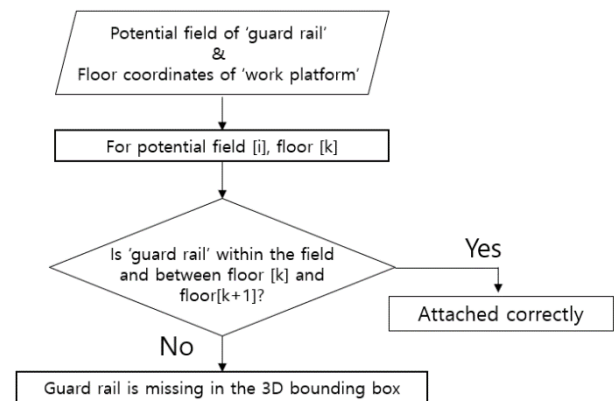


Figure 3. The checking flow for regulation III in a 3D bounding box.

3 Experiments and Results

3.1 Dataset preparation

The dataset used in the experiments were acquired using FARO m70 from four different construction sites as shown in Figure 4. A total of fifteen registered point clouds acquired from Sites A, B, and C were used to train RandLA-Net. Each point cloud data were labeled to represent a total of six classes: ‘stairs,’ ‘work platform,’ ‘guard rail,’ ‘uprights,’ ‘bracing,’ and ‘background.’ Data acquired from Site D were used for the testing. Total number of points in the training and testing data were 103,223,397 14,649,928 points, respectively.



Figure 4. Upper left; Site A, Upper right; Site B, Lower left; Site C, Lower right; Site D.

3.2 Evaluation metrics

To effectively calculate the performance of each class, Precision, Recall, and F1-score were used. As shown in Equations (1) ~ (3), Precision is a metric that calculates the percentage of ground truth labels within the predicted truth labels. Recall is a metric that calculates the percentage of predicted truth labels within the ground truth labels. Most importantly, F1-score is a metric that calculates the harmonic average of Precision and Recall.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (3)$$

3.3 Segmentation performance

All parameters of RandLA-Net, except for the training epoch, followed the setting of [6]. The training epoch was set as 50 considering the trend of training loss.

The results of multi-class segmentation by RandLA-Net are shown in Table 2. The results were evaluated taking an average of eight experiments with the same dataset and parameters. Class ‘background’ had the best F1-score of 98.83% followed by other successful segmentation results from ‘uprights’ 69.10%, ‘guard rail’ 68.95%, ‘work platform’ 61.14%, and ‘stairs’ 61.06%. Most of the false predictions of those five classes were found on the lower part of the scaffold. The ‘bracing’ segmentation results showed the poor performance with an 16.90% F1-score. This performance indicated a need for further studies if a need exists for regulation checking regarding bracings. Figure 5 shows a segmentation result.

Table 2. Segmentation results of RandLA-Net

Class	Precision (%) Recall (%)	F1 score (%)
‘stairs’	66.56 62.18	61.06
‘work platform’	74.87 51.92	61.14
‘guard rail’	93.11 55.89	68.95
‘uprights’	77.09 63.37	69.10
‘bracing’	14.23 29.80	16.90
‘background’	98.19 99.48	98.83

3.4 Results and discussions

According to the regulation checking algorithms, both ‘work platform’ and ‘guard rail’ had thirteen potential fields on the x-y plane (shown in Figure 1), and the number of floors was one. The results for the three regulations can be summarized as shown in Table 3. The two violations of regulation III are shown as yellow line in the last picture of Figure 5. They were all accurately predicted by the proposed method.

The regulation checking algorithm was focused on scaffold entities containing ‘stairs,’ ‘work platform,’ ‘guard rail,’ and ‘uprights.’ The misclassified points of ‘bracing’ could be filtered by following the steps of the regulation checking algorithm. This allowed the poor segmentation result of ‘bracing’ to not affect the final performance of the regulation checking process.

The scaffold of Site D was an L-shaped scaffold, which was not contained in Sites A, B, and C. This shows a robust performance of our model regarding the shape

of a scaffold. However, there is still a need to enrich the dataset to improve segmentation results and generalize the proposed methods at other construction sites. Data acquisition of this study was limited especially because of the temporary nature of scaffolds. Ablation studies of generating synthetic data could help to address this problem and improve the model generalization.

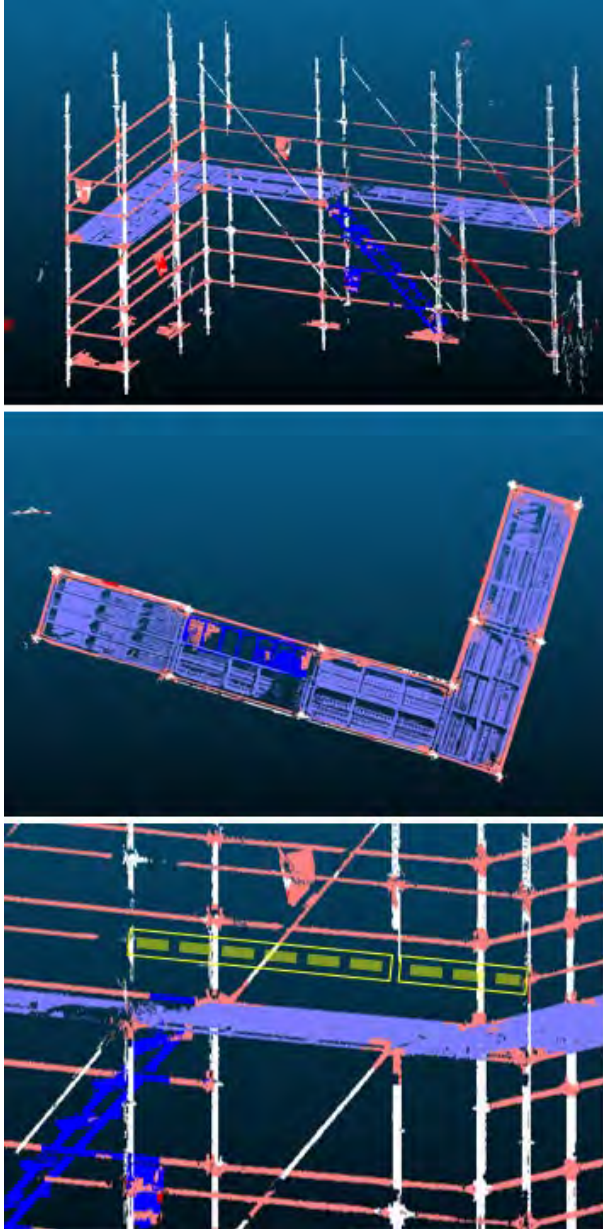


Figure 5. Visualization of segmentation results; {blue: 'stairs', purple: 'work platform', pink: 'guard rail', white: 'upright', red: 'bracing'}

4 Conclusion

This study presented a fully automated methodology to accurately check three major safety scaffold-related regulations specified in the KOSHA Guide to scaffolds [7, 8]. The proposed methodology is composed of a deep learning-based point cloud segmentation (RandLA-Net) and rule-based algorithms. The segmentation F1 scores were 98.83% for 'background,' 69.10% for 'uprights,' 68.95% for 'guard rail,' 61.14% for 'work platform,' 61.06% for 'stairs,' and 16.90% for 'bracing,' respectively. They were sufficient to successfully check all three major regulations considered on this study. These results indicate that this methodology has high potential to fully automate safety monitoring of scaffolds which will lead to a significant reduction in accidents and deaths in the construction industry.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT (No. 2021R1A2C2004308) and the Ministry of Education (No. 2018R1A6A1A08025348).

References

- [1] KOSHA (Korea Occupational Safety and Health Agency). Industrial Accident Status June 2021. On-line: <https://www.kosha.or.kr/kosha/data/industrialAccidentStatus.do?mode=view&articleNo=425349&article.offset=0&articleLimit=10>, Accessed: 26/02/2022.
- [2] Busan Metropolitan Corporation. Construction scaffolding work safety practice guide. On-line: <https://www.bmc.busan.kr/bmc/bbs/view.do?bIdx=703852&ptIdx=817&mId=0505060100>, Accessed: 26/02/2022.
- [3] Bosche, F., and Haas, C. T. Automated retrieval of 3D CAD model objects in construction range images. *Automation in Construction*, 17(4):499-512, 2008.
- [4] Wang, Q. Automatic checks from 3D point cloud data for safety regulation compliance for scaffold work platforms. *Automation in Construction*, 104:38-51, 2019.
- [5] Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., ... and Markham, A. Randla-net: Efficient semantic segmentation of large-scale point clouds.

Table 3. Results for the three regulations (G.T.: ground truth, O: normal, X: violated, -: not to be considered)

	Potential field	1	2	3	4	5	6	7	8	9	10	11	12	13
Regulation I	Prediction	O	O	X	O	O	O	O	O	O	O	O	O	O
	G.T.	O	O	X	O	O	O	O	O	O	O	O	O	O
Regulation II	Prediction	-	-	O	-	-	-	-	-	-	-	-	-	-
	G.T.	-	-	O	-	-	-	-	-	-	-	-	-	-
Regulation III	Prediction	O	O	O	O	O	O	O	O	O	X	X	O	O
	G.T.	O	O	O	O	O	O	O	O	O	X	X	O	O

In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, page 11108-11117, Seattle, USA, 2020.

- [6] Kim, J., Chung, D., Kim, Y., and Kim, H. Deep learning-based 3D reconstruction of scaffolds using a robot dog. *Automation in Construction*, 134:104092, 2022.
- [7] KOSHA (Korea Occupational Safety and Health Agency). Safety Work Guidelines for Steel Pipe Scaffold (KOSHA Guide C-30-2020). On-line: <https://www.kosha.or.kr/kosha/data/guidanceDetail.do>, Accessed: 26/02/2022.
- [8] KOSHA (Korea Occupational Safety and Health Agency). Safety Work Guidelines for Modular Scaffold (KOSHA Guide C-32-2020). On-line: <https://www.kosha.or.kr/kosha/data/guidanceDetail.do>, Accessed: 26/02/2022.
- [9] Hackel, T., Savinov, N., Ladicky, L., Wegner, J. D., Schindler, K., and Pollefeys, M. Semantic3d. net: A new large-scale point cloud classification benchmark. arXiv preprint arXiv:1704.03847, 2017.

A Systematic Classification and Evaluation of Automated Progress Monitoring Technologies in Construction

V.K. Reja ^{a, b}, M.S. Pradeep ^a, K. Varghese ^a

^a Department of Civil Engineering, IIT Madras, India

^b Faculty of Engineering & Information Technology, UTS, Australia

E-mail: varunreja7@gmail.com, meghaspradeep64@gmail.com, koshy@iitm.ac.in

Abstract-

Progress monitoring is one of the essential tasks while executing a construction project. Effective monitoring will lead to an accurate and timely analysis of the project's progress which is required to make vital decisions for project control. On the other hand, inefficient and delayed updates regarding the project's progress, which is estimated by comparing the as-built status with the as-planned status, will lead to time and cost overruns. Automated progress monitoring techniques are preferred over the conventional manual data entry method as the latter is time-consuming and complex, especially if the project scope is vast. Numerous tools and technologies are being used for progress monitoring of construction projects. Therefore, it is necessary to systematically classify and evaluate them based on their advantages and limitations for successful and appropriate implementation. Hence, this article identifies several progress monitoring methods and classifies them based on the technology they use to support progress monitoring. Then they are evaluated by highlighting their advantages and limitations. Several qualitative and quantitative factors affecting the selection of these technologies for implementation have also been identified. In future, a framework for objectively identifying the project-specific technology will be developed.

Keywords –

Progress monitoring technologies; Data acquisition; BIM; Internet of Things (IoT); Sensors; Computer vision; Extended reality; Literature review; Challenges; Limitations

1 Introduction

A project life cycle in a construction industry involves several stages, like designing, planning, scheduling, execution, monitoring, controlling, and demolition. Monitoring and control to minimize time and cost overruns are crucial for a construction project. Accurate progress monitoring is an essential step for achieving quality and safety parameters.

Progress monitoring also plays a vital role in avoiding unexpected circumstances and eliminating disputes and legal challenges among the stakeholders. Automating various tasks in monitoring and controlling will reduce the complexity involved in manual documentation and calculations in a project to a considerable extent. Hence, selecting a prompt and feasible automated progress monitoring technology is essential in the present-day construction sector [1][2].

Automation in progress monitoring has evolved over the past two decades, with several technologies with varying levels of automation used in projects. With several technologies available, there is not enough clarity on the type of technology that will be appropriate to be used for a specific case or project.

Existing papers have focused on the specific technology of progress monitoring, for example, specifically, vision-based [3] or IoT (sensor) based [4]. For a robust implementation, firstly, there is a critical need to identify and classify these technologies and, secondly, evaluate them based on their advantages and limitations. Hence this paper aims to:

1. Identify, classify, and evaluate technologies available for progress monitoring of construction projects.
2. To list and categorize factors that enable appropriate technology selection for the project-specific use case.

The paper is structured as follows. The review methodology has been discussed in Section 2. This is followed by a detailed review of the technologies in Section 3. The factors affecting progress monitoring technology selection are presented in Section 4. Section 5 is a discussion followed by conclusions in Section 6.

2 Review methodology

The reference literature for the review was collected from Scopus and Web of Science databases using a keyword search-based method, followed by snowballing technique. A total of 61 papers with 49 journal articles and 12 conference papers were identified from the databases, and an exhaustive review with analysis was

performed. The chronological distribution of the selected papers is shown in Figure 1.

The search attributes used in the review with the keywords used and search scope is as shown in Table 1. The relevant articles for the construction domain were filtered after reading the abstracts. The filtered articles were considered for meta-analysis.

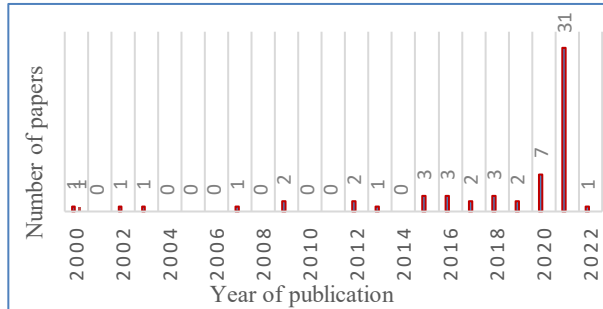


Figure 1 Chronological distribution of the selected papers

Table 1 Search attributes

Search attributes	Values used in the search
Databases	Web of Science, Scopus
Language	English
Duration	2000-2022
Type	Journal & conference articles
Keywords	Progress monitoring technologies, Automated progress monitoring.

3 Automated progress monitoring technologies

The selected papers contained case studies, challenges, and benefits of various automated techniques. Based on the meta-analysis, seventeen state-of-the-art progress monitoring techniques were identified. As shown in Table 2, these techniques were classified into six major categories based on the technology they use to support progress monitoring. These are conventional Information & Communication Technologies (ICT), tag-based methods, geospatial technologies, building information modelling and associated commercial software, computer vision-based approaches, and extended reality and are discussed below:

3.1. Conventional ICT: These include handheld computing devices (Personal Digital Assistants or PDAs, handheld personal computers), Interactive Voice Response or IVR, multimedia tools, and e-mails. These are the most basic techniques, which are information technology-based communication tools. These

technologies are primarily of lower cost with a limited level of automation but can increase the chances of communication between stakeholders, thereby helping in information tracking [2].

3.2. Tag-based techniques: These involve using tags and codes that can be attached to various resources on-site and are primarily used for material tracking and inventory, employee badge scanning, and equipment tracking. These include barcodes, quick-response or QR codes, radio-frequency identification or RFID tags, and ultra-wideband or UWB tags. Each tag has its working principle based on Automatic Identification and Data Capture (AIDC). It must be noted that a tag-based technology cannot directly extract spatial element information, visually represent the site changes, and collaborate with other vision-based techniques [5][6][7].

3.3 Geospatial techniques: These include fundamental technologies based on location-based sensors like Geographic Information System (GIS) and Global Positioning System (GPS). These techniques are used for geo-referenced data capture, analysis, and modelling. GIS can be used in large infrastructure projects where there is a need to store and handle huge amounts of data. It can be a useful geospatial tool, which uses location as the primary focus in database management, whereas GPS aids in the spatial analysis and navigation of different activities on the site [8].

3.4. Building Information Modelling (BIM) Based: BIM is a process involving different tools, technologies, and contracts, which aids in better visualization of construction sites for accurate project management. BIM also aids in stakeholder management practices of the construction industry for different aspects of communication, collaboration, engagement, and satisfaction. This can be used along with commercial scheduling software like MS Project so that progress monitoring can be done efficiently [9].

3.5. Computer Vision-Based Construction Progress Monitoring (CV-CPM): It is an emerging field that focuses on information retrieval through visual inputs [3]. These inputs can be digital images, videos, thermal images, as-built point clouds, panoramas, and photospheres. These techniques involve fixed surveillance, photogrammetry, videogrammetry, range imaging, and 3D laser scanning. Computer vision sub-domains include learning, 3D scene modelling, video tracking, 3D pose estimation, object recognition, scene reconstruction, object detection, and event detection, which can be used for progress monitoring [3][10][11]. For this, digital twin technology are also being explored for effective real time monitoring of projects [12].

Table 2 Detailed review of the progress monitoring technologies

No	Technology	Advantages	Limitations	Ref
1.	Conventional ICT			
	Hand-held computing	<ul style="list-style-type: none"> Small-sized, portable, handy, flexible devices with several features, that can be integrated with other technologies. 	<ul style="list-style-type: none"> Some devices are costly, and suitable applications need to be developed so that integration can be made more efficient. 	[2]
	Interactive Voice Response or IVR	<ul style="list-style-type: none"> Efficient and quick means of sending information from sites. 	<ul style="list-style-type: none"> Manual errors might occur while responding to multiple choices. Difficulty in retracing the already answered correct messages. 	
	Multimedia	<ul style="list-style-type: none"> Flexible tool to aid remote progress monitoring by safe documentation and visualization of project information. 	<ul style="list-style-type: none"> Manual site data capture results in errors. 	
	Electronic Mails or E-Mails	<ul style="list-style-type: none"> E-procurement tool for quality supply chain management. Attached with images, documents, videos, forms, etc, where the site personnel respond with the easy retracing of answered questions. 	<ul style="list-style-type: none"> Improper internet connection, and difficulty in responding from small devices might cause a delay in information exchange. 	
2.	Tag based			
	Bar-codes	<ul style="list-style-type: none"> Cost-effective, accurate, easy to use, portable, and flexible. No need for an external device to read the codes. 	<ul style="list-style-type: none"> Direct line of sight required for data capture, time taking in item tracking. Labels can get destroyed or lost due to adverse weather conditions. 	[2][15]
	Quick Response or QR codes	<ul style="list-style-type: none"> Portable and flexible technology with a better storage capacity. QR code reading applications can be installed on devices easily. Lightweight wireless QR code pocket printers can be used in sites. 	<ul style="list-style-type: none"> Might get affected by harsh environmental conditions. More effective in indoor tracking compared to outdoors. Monitoring of reinforcement, concreting, and those which are not easily accessible are difficult to be monitored using QR codes. 	[16]
	Radio Frequency Identification or RFID tags	<ul style="list-style-type: none"> Use radio waves that can be read accurately outside the line of sight as well, without direct contact compared to light waves. Reliable, portable, flexible, reusable, technique that can withstand harsh environmental site conditions. Supports indoor tracking of materials, facility and building component management, and information flow in large projects. Capable of identifying individual items and can read multiple items in an instant simultaneously. Storage capacity is higher than that of a barcode and is unaffected by differences in illumination. 	<ul style="list-style-type: none"> Costlier than barcodes. If there are metals or liquids or moisture in the nearby area, the results can be erroneous. Using RFID can be time-consuming, and costly if a single tag is used to track each one among several materials and equipment. Limitation of the battery operation time, and there is insufficient accuracy in location identification if it is not depending on a fixed network. There are insufficient international standards, multi-protocol tags and readers, and also a concern on the return of investment. 	[1][6][7][17]
	Ultra-wide band or UWB tags	<ul style="list-style-type: none"> More accurate than RFID with strong signals even in obstructions. Provide real-time resource tracking, 3-D coordinates for position sensing and consume low energy. 	<ul style="list-style-type: none"> UWB tags are not cost-effective compared to RFID. No daily necessity embedded tool or mini-device. 	[7]

No	Technology	Advantages	Limitations	Ref
3.	Geospatial	<ul style="list-style-type: none"> Optimal location for construction equipment can be found, with efficient capture, storage, and analysis of georeferenced information with minimum redundancy. Creation of geographical maps of high quality, by visually representing the construction schedule to monitor the plant and equipment, that can be provided to the clients. Can be used as a forecasting tool for early identification of time and space conflicts, and better safety regarding worksite considerations. 	<ul style="list-style-type: none"> Difficult to use in indoors. However there have been studies on the use of radio based (Wi-Fi, IoT, Bluetooth, UWB) and vision based signals like (markers/QR codes) to assist with obtaining the geospatial data. 	[4] [5][8]
	Global Positioning System or GPS	<ul style="list-style-type: none"> Accurate location of positions while material tracking in construction supply chain management can be facilitated by using the Global Positioning System or GPS. 	<ul style="list-style-type: none"> Difficult to use in indoors. Tagging several construction elements using GPS tags is very expensive. 	[18]
4.	BIM and commercial software	<ul style="list-style-type: none"> Visualization- clash detection & information management. Schedule updates- automatic quantity take off & cost-estimation. Enhanced collaboration & information exchange among stakeholders. Integrated with supply chain for product design & material delivery. Integration with LPS or Last Planner System for lean construction. Knowledge based systems that are active, along with various simulations in BIM contribute to data analysis. Choosing inputs using a hybrid video and laser scans, as-built BIM can be generated automatically. 	<ul style="list-style-type: none"> Limited in monitoring, scheduling, and decommissioning phases. Errors resulting from the manual navigation of BIM model and need for constant automated updates, especially for fast-tracked projects. Limited interoperability, even with other data acquisition techniques. Commercial software like MS Project, Primavera, etc cannot provide digital drawings and visualisation for construction. 	[1][4][17][19][20]
5.	Computer Vision (CV-CPM)	<ul style="list-style-type: none"> Fixed surveillance (crane cameras, closed circuit television cameras, etc) 	<ul style="list-style-type: none"> Crane camera images may result in noisy as-built point clouds and may get affected by heavy winds. Mounting the camera may require extra effort, and there is limited flexibility due to the motion range of the cranes. 3-D point clouds may be fragmented if there is incomplete site coverage. As the position of cameras is fixed, there is a limitation in the application of CCTV cameras in huge projects. There is a need to arrange several cameras, and the data clashes between different cameras have to be fixed. 	[1][6][7] [3][12][17][18][19]

No	Technology	Advantages	Limitations	Ref
Computer Vision (CV-CPM)		<ul style="list-style-type: none"> Automatic identification of objects using cameras and image processing algorithms by integration with n-D BIM. Lesser equipment cost and technical requirement, along with portability for the image capturing devices and improved flexibility. High resolution compared to satellite imaging, for representing the geometric attributes and high texture representation. The recorded images can be analyzed using software packages with computer vision techniques and machine learning algorithms for automatic updates with as-built -3D models by reconstruction. 	<ul style="list-style-type: none"> There is a limitation due to difference in lighting conditions, which may affect the resolution. Thermal images along with wireless sensors and BIM can be used to overcome this problem. Object edge detection might not be proper, occlusions, noisy images and presence of shadows will affect the accuracy of progress estimation. The location from which the photos are taken has to be matched with the check-points in the drawings, which can be a difficult problem. 	[3][18]
	Photogrammetry			
	Videogrammetry	<ul style="list-style-type: none"> Can be used for both indoor and outdoor progress tracking. Moving equipment can be tracked. 	<ul style="list-style-type: none"> Less accurate than laser scanning and photogrammetry, may get affected due to occlusions. 	[7] [3][20]
	Range or depth cameras	<ul style="list-style-type: none"> Easier as-built point cloud generation directly, as it contains depth information. Higher resolution compared to normal digital cameras, and higher portability with lower technical pre-requisite. 	<ul style="list-style-type: none"> Cost is lower than that for laser scanners but will be generally higher than that of normal digital cameras. Range of shooting is limited and is mostly used in automated indoor construction progress monitoring. 	[3][18]
	3D Laser scanners	<ul style="list-style-type: none"> High-resolution, precise, and accurate progress monitoring technology unaffected by illumination, and is used for quality control, structural health monitoring, condition assessment of structures, and tracking of components, along with active collaboration between the stakeholder teams. Automatic comparison being done between as-built and as-planned point clouds so that progress deviation detection becomes easier, and the schedule is updated accordingly. 	<ul style="list-style-type: none"> Highly expensive equipment with low portability, limited texture information, time-consuming data acquisition, mixed pixel restoration, need for sensor calibrations regularly, greater warm-up time. Operation requires larger technical knowledge, and might not be suitable for progress monitoring continuously. Accuracy of the data acquisition using laser scanning might be affected due to occlusions and in the site. 	[6] [3][7][17][18][19][21][22]
6.	Extended reality (XR)	<ul style="list-style-type: none"> Enables accurate visualization of the construction site from various angles. Worksite planning in construction, visualization of equipment operation for inspection, comparison between as-planned and as-built images can be done. They can be used easily in both interior and exterior locations, under different construction phases, and is cheap with the requirement of minimal training and set-up time. 	<ul style="list-style-type: none"> Automation quality depends on the technology in the device used. Stationary methods are limited in portability, have less cost-effectiveness, and need additional time for setting up the equipment when compared to mobile methods. Installation of fiducial markers requires additional investment in time and cost. 	[1] [3][13] [18][23]

3.6. Extended Reality (XR) Based: The relatively newer technology allows a combined real and virtual environment, supporting human-machine interactions. These techniques can be further classified into augmented reality (AR), virtual reality (VR), and mixed reality (MR), based on the difference in visualization. These techniques can be employed for the collection of digital data and can handle computing and network technologies in progress monitoring [13].

Each of these technologies has its unique advantages and limitations in construction site monitoring. Hence, there is a need for a detailed analysis to identify these to be used effectively. It is also to be noted that efficient integration of suitable technologies in sites has been done in various case studies, and it was found that time and cost overruns have been reduced to a great extent. A detailed and systematic review of each of the above-mentioned technologies is presented in Table 2, along with the relevant references.

4 Factors affecting progress monitoring technology selection

As shown in Table 2, each technology is characterized by its advantages and limitations. Apart from these, some key factors should be considered before choosing the appropriate technology for progress monitoring in a construction project. Some of these parameters have been identified by several authors through their research [1][14].

In our recent study [3], we conducted a systematic literature survey using PRISMA methodology and identified factors which affect progress monitoring technologies in construction.

Figure 2 shows the key factors to be considered while selecting the appropriate technology for progress

monitoring. These factors are categorized into quantitative and qualitative factors based on how they will be evaluated for consideration. The context of these factors for selection is described as follows:

A - Quantitative Factors

1. **Time efficiency:** The speed of data acquisitions as well as data processing.
2. **Operating range:** The distance up to which the employed technology works.
3. **Accuracy:** The reliability of the collected data along with precision.
4. **Cost:** The amount of financial cost as well as the computational cost incurred to adopt and implement the technology.

B - Qualitative Factors

1. **Utility:** Adaptability of the technology to be used both in interior and exterior construction progress monitoring. In other words, whether the technology is a general case solution.
2. **Level of Automation:** The extent of manual effort required while using the technique.
3. **Preparation required:** The level of preparation required while setting up the equipment or process at the deployment stage.
4. **Training required:** The amount of training or knowledge a user requires prior to using a particular technique.
5. **Susceptibility in adverse weather:** The extent of use of the technology in harsh environmental conditions like low visibility.
6. **Compatibility for use:** The level by which a particular technology can be integrated with other technologies or the existing Enterprise Resource

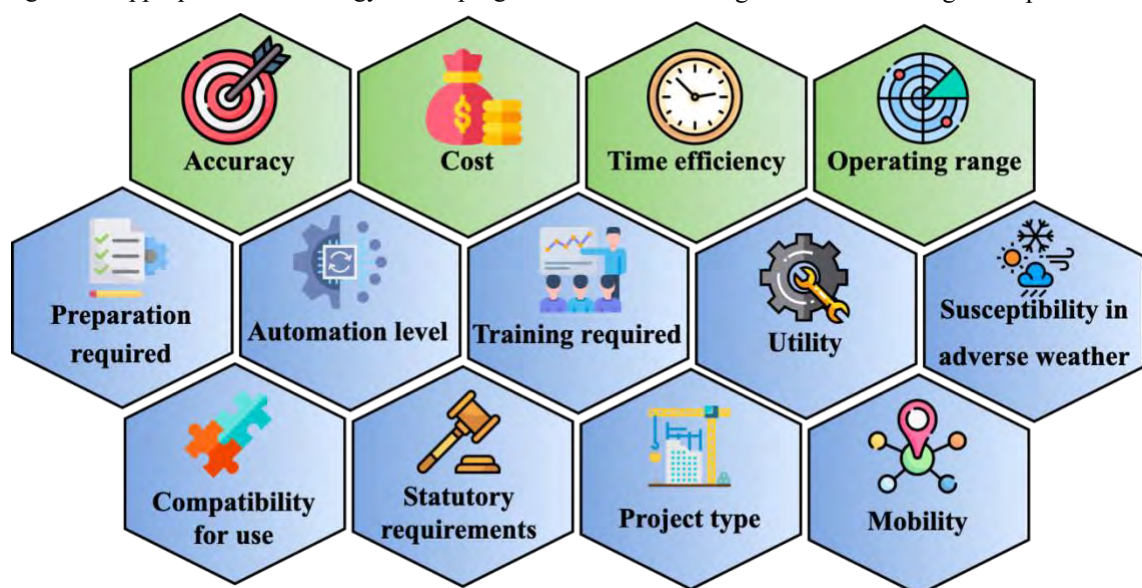


Figure 2 Factors affecting selection of technology: Quantitative (in green) and Qualitative factors (in blue)

- Management (ERP) system.
7. **Statutory requirements:** The legal codes and procedures to be followed while using a technique enforced by the authorities.
 8. **Mobility:** The ease, flexibility, and portability of the related equipment.
 9. **Project type & characteristics:** The type and characteristics of a particular construction project where the technology can be used.

Utilization of these factors for technology selection

To utilize these factors for appropriate technology selection, it is recommended to provide weightage to factors based on the project's constraints on cost, quality, and time. The qualitative and quantitative evaluation of categorized factors should be performed. All the six technologies mentioned in Table 2 can then be evaluated based on these factors for appropriate selection. This demonstration is not in the scope of this paper but will be part of future research.

5 Discussions

Progress monitoring is crucial for accurate project control. Choosing the appropriate automated technology based on required parameters is vital in the monitoring stage of a project. Automated technologies can be integrated based on the requirements and can be highly efficient in reducing project overruns compared to manual methods. The technology must be chosen without overselling, such that the investment returns from the project can be made higher.

An idealized situation would enable a higher efficiency in all these parameters, which is not practically possible in a single technology. Therefore, selecting a suitable technology that produces the maximum output based on these parameters would be the goal in the monitoring phase of a construction project.

Another important consideration is that newer technologies might face challenges about their widespread acceptance, as the construction companies might tend to reject the pilot integrated automation technology proposals. This happens mainly due to a lack of technical knowledge of automated technologies and the tendency to continue adopting conventional techniques. So integrated proposals through research should be added with proper inspection and maintenance guidelines, followed by proper incentives to the enterprises, so that widespread adoption can be facilitated. Moreover, data collection by a single resource tracking is never sufficient for accurate progress monitoring. Hence, applying data fusion techniques is vital to tracking multiple resources on a construction site.

6 Conclusions

This paper provides a systematic review of various automated progress monitoring technologies from 61 relevant publications to understand the state-of-the-art to guide future research. The paper also identifies the benefits and limitations associated with each technology (in Table 2), along with the factors affecting their selection (in section 4). It is to be noted that each construction project is unique and has its specific characteristics. As all the technologies have their advantages and shortcomings, selecting a technology that suits a particular project is extremely important. The technologies can be combined and integrated to minimize cost overruns. This review provides a basis for this selection, as it systematically identifies the scope for each automated technology. In addition, more review efforts are recommended to identify suitable mounting methods that can be used in combination with the techniques. Digital twin based progress monitoring is a potential area of future research [12].

Acknowledgements

The support first author received from the 'Prime Minister Research Fellowship (PMRF)' Scheme from the 'Ministry of Education, Government of India' is gratefully acknowledged.

7 References

- [1] M. Kopsida, I. Brilakis, P.A. Vela, A review of automated construction progress monitoring and inspection methods, in: Proc. 32nd CIB W78 Conf. Constr. IT, 2015.
- [2] T. Omar, M.L. Nehdi, Data acquisition technologies for construction progress tracking, *Autom. Constr.* 70 (2016) 143–155. <https://doi.org/10.1016/j.autcon.2016.06.016>.
- [3] V.K. Reja, K. Varghese, Q.P. Ha, Computer vision-based construction progress monitoring, *Autom. Constr.* 138 (2022) 104245. <https://doi.org/10.1016/j.autcon.2022.104245>.
- [4] V.K. Reja, K. Varghese, Impact of 5G Technology on IoT Applications in Construction Project Management, in: Proc. 36th Int. Symp. Autom. Robot. Constr. (ISARC 2019), Banff, Canada, 2019: pp. 209–217. <https://doi.org/10.22260/ISARC2019/0029>.
- [5] T.K. Geok, K.Z. Aung, M.S. Aung, M.T. Soe, A. Abdaziz, C.P. Liew, F. Hossain, C.P. Tso, W.H. Yong, Review of indoor positioning: Radio wave technology, *Appl. Sci.* 11 (2021) 1–44.

- <https://doi.org/10.3390/app11010279>.
- [6] G. Guven, E. Ergen, Tracking major resources for automated progress monitoring of construction activities: masonry work case, *Constr. Innov.* 21 (2021) 648–667. <https://doi.org/10.1108/CI-05-2020-0081>.
- [7] W.S. Alaloul, A.H. Qureshi, M.A. Musarat, S. Saad, Evolution of close-range detection and data acquisition technologies towards automation in construction progress monitoring, *J. Build. Eng.* 43 (2021) 102877. <https://doi.org/10.1016/j.jobbe.2021.102877>.
- [8] V. Thellakula, V.K. Reja, K. Varghese, A Web-Based GIS Tool for Progress Monitoring of Linear Construction Projects, in: *Proc. 38th Int. Symp. Autom. Robot. Constr., International Association for Automation and Robotics in Construction (IAARC)*, 2021. <https://doi.org/10.22260/isarc2021/0007>.
- [9] Y. Deng, V.J.L. Gan, M. Das, J.C.P. Cheng, C. Anumba, Integrating 4D BIM and GIS for Construction Supply Chain Management, *J. Constr. Eng. Manag.* 145 (2019) 04019016. [https://doi.org/10.1061/\(asce\)co.1943-7862.0001633](https://doi.org/10.1061/(asce)co.1943-7862.0001633).
- [10] B. Ekanayake, J.K.W. Wong, A.A.F. Fini, P. Smith, Computer vision-based interior construction progress monitoring: A literature review and future research directions, *Autom. Constr.* 127 (2021) 103705. <https://doi.org/10.1016/j.autcon.2021.103705>.
- [11] S. Paneru, I. Jeelani, Computer vision applications in construction: Current state, opportunities & challenges, *Autom. Constr.* 132 (2021) 103940. <https://doi.org/10.1016/j.autcon.2021.103940>.
- [12] V.K. Reja, K. Varghese, Digital Twin Applications for Construction Project Management, in: *Jt. Indo-Japanese Smart City Conf., Japan*, 2022. https://www.researchgate.net/publication/359992822_Digital_Twin_Applications_for_Construction_Project_Management
- [13] P. Bhadaniya, V.K. Reja, K. Varghese, Mixed Reality-Based Dataset Generation for Learning-Based Scan-to-BIM, in: *Lect. Notes Comput. Sci., Springer Science and Business Media Deutschland GmbH*, 2021: pp. 389–403. https://doi.org/10.1007/978-3-030-68787-8_29.
- [14] B. Ekanayake, J.K.-W. Wong, A.A.F. Fini, P. Smith, Computer vision-based interior construction progress monitoring: A literature review and future research directions, *Autom. Constr.* 127 (2021) 103705. <https://doi.org/10.1016/j.autcon.2021.103705>.
- [15] A. Keyvanfar, A. Shafaghat, M.A. Awanghamat, Optimization and Trajectory Analysis of Drone's Flying and Environmental Variables for 3D Modelling the Construction Progress Monitoring, *Int. J. Civ. Eng.* (2021). <https://doi.org/10.1007/s40999-021-00665-1>.
- [16] Z. Wang, Q. Zhang, B. Yang, T. Wu, K. Lei, B. Zhang, T. Fang, Vision-Based Framework for Automatic Progress Monitoring of Precast Walls by Using Surveillance Videos during the Construction Phase, *J. Comput. Civ. Eng.* 35 (2021) 04020056.
- [17] S. Alizadehsalehi, I. Yitmen, The Impact of Field Data Capturing Technologies on Automated Construction Project Progress Monitoring, *Procedia Eng.* 161 (2016) 97–103. <https://doi.org/10.1016/j.proeng.2016.08.504>.
- [18] J. Xue, X. Hou, Y. Zeng, Review of Image-Based 3D Reconstruction of Building for Automated Construction Progress Monitoring, *Appl. Sci.* 11 (2021) 7840. <https://doi.org/10.3390/app11177840>.
- [19] M.K. Masood, A. Aikala, O. Seppänen, V. Singh, Multi-Building Extraction and Alignment for As-Built Point Clouds: A Case Study With Crane Cameras, *Front. Built Environ.* 6 (2020). <https://doi.org/10.3389/fbuil.2020.581295>.
- [20] I. Mutis, V.A. Joshi, A. Singh, Object Detectors for Construction Resources Using Unmanned Aerial Vehicles, *Pract. Period. Struct. Des. Constr.* 26 (2021) 04021035. [https://doi.org/10.1061/\(ASCE\)SC.1943-5576.0000598](https://doi.org/10.1061/(ASCE)SC.1943-5576.0000598).
- [21] F. Arif, W.A. Khan, Smart Progress Monitoring Framework for Building Construction Elements Using Videography–MATLAB–BIM Integration, *Int. J. Civ. Eng.* (2021). <https://doi.org/10.1007/s40999-021-00601-3>.
- [22] V.K. Reja, P. Bhadaniya, K. Varghese, Q.P. Ha, Vision-Based Progress Monitoring of Building Structures Using Point-Intensity Approach, in: 2021. <https://doi.org/10.22260/ISARC2021/0049>.
- [23] A.K. Ali, O.J. Lee, D. Lee, C. Park, Remote indoor construction progress monitoring using extended reality, *Sustain.* 13 (2021) 1–24. <https://doi.org/10.3390/su13042290>.

Reducing spatial error in mobile laser scanning by real-time uncertainty visualization and human-machine interaction

M. Trzeciak, C.J. Burgoyne and I. Brilakis

Department of Engineering, University of Cambridge, United Kingdom

E-mail: mpt35@cam.ac.uk, cjb19@cam.ac.uk, ib340@cam.ac.uk

Abstract –

Scanning is a key element for many applications in the AECO industry. It provides point clouds used for construction quality assurance, scan-to-BIM workflows and construction surveys. However, data acquisition using static laser scanners or photogrammetry methods is lengthy and requires even lengthier subsequent processing. A quick and apparent escape from this problem might be mobile mapping solutions mainly based on lidars. However, current hand-held scanners suffer from drift, skewing point clouds and thus, increasing their spatial error. In this paper, we present a novel, real-time and fully explainable method exploiting human-machine interaction to increase the correctness of produced point clouds. Our method progressively reconstructs the scanned scene and predicts the regions of a potentially high error with a 95% confidence level. The user can then revisit these parts of the scene, which adds additional constraints on the underlying probabilistic graphical model, thus reducing the drift and increasing the confidence in the correctness of these regions. We build a prototypic lidar-based mobile scanner, implement our method and test it in a case study. The results show that the areas identified with a relatively high spatial error indeed suffer from it, while predicted areas with relatively higher correctness do have a smaller spatial error.

Keywords –

Mobile mapping; SLAM; Digitization; Pose graph; Uncertainty visualization; Uncertainty propagation; Human-machine interaction

1 Introduction

In this paper, we propose a novel mobile scanning technique aiming at reducing a spatial error in point clouds based on the user-scanner interaction. As the user traverses a scene, our real-time method visualizes the uncertainty related to the correct position of points in a

progressively-built point cloud, hence informing the user on the potentially increasing spatial error. The user can then take corrective actions on-the-fly by revisiting places with a potentially higher spatial error, thus imposing additional constraints on the underlying optimization problem and hence reducing the error.

Digitizing the geometry of existing assets is a key element for many use cases in the Architectural, Engineering, Construction and Operation (AECO) industry. However, data acquisition using static laser scanners or photogrammetry methods is lengthy and requires even lengthier subsequent processing [1], [2]. A quick and apparent escape from this problem might be mobile mapping solutions, mainly based on lidars.

However, the current state-of-practice mapping devices do not allow for scanning with high accuracy [3], [4]. Therefore, it is not uncommon that there is a mix of static and mobile scanners on construction sites depending on the requirements of the use cases at hand [3].

We define a spatial error as a distribution of distances between points by a mobile scanner and their corresponding ground truth. This corresponds with the geodetic “correctness” of point clouds [5], “absolute accuracy” [6] or, simply, “accuracy” [1]. We will use these notions interchangeably in this paper. With that in mind, the problem statement is that point clouds produced by current mobile scanners suffer from relatively higher spatial error because of drift increasing over time in Simultaneous Localization and Mapping (SLAM) systems [7], [8].

We propose a novel and fully explainable real-time scanning method based on the user-scanner interaction. As the user traverses the scene, our system propagates the uncertainty of odometric inter key-pose constraints in the underlying pose graph. Then, with a high confidence level, it computes the largest variability related to each of the key poses and visualizes it on the progressively-built point cloud using colours. Since the colours displayed on the 3D points are based on comparing the variability to accuracy levels/bands from surveying standards, our method is dedicated specifically to construction use cases.

The user can thus see which regions of the point cloud are likely to suffer from higher spatial error and can revisit them, which adds additional constraints onto the pose graph, thus reducing the uncertainty in the key poses.

The proposed system can be then integrated into the practical mobile scanning procedures in the following way. Before scanning starts, the user chooses the top and bottom levels of point cloud correctness they are willing to accept. As the user traverses the scene, the progressively-built point cloud gradually changes colour from green to red, indicating that the spatial error of the red regions of the scene is beyond the set lower level. After the desired parts of the scene have been scanned, the user revisits these parts of the scene that are coloured red, ideally linking them with green areas.

We build a prototypic scanner consisting of a lidar and a laptop, implement our method and test it in a case study. With a 95% confidence level, the results show that our system correctly predicts the regions of both relatively high correctness as well as those suffering from high spatial errors.

Before we proceed to the specifics of our method, however, the subsequent section introduces the reader to the state of the art (SOTA) pose-graph SLAM systems and uncertainty propagation in their underlying models.

2 Background

2.1 Pose-graph SLAM

According to [9], there are two probabilistic SLAM paradigms: 1) filtering and 2) smoothing. The former focuses on the estimation of the most current pose of the scanner given all the measurements of its sensors. It is useful in the case of robots that must determine their position in real-time as accurately as possible. The latter, in turn, focuses on the estimation of all the key poses comprising the trajectory of the scanner. Given the fact that the correctness of the produced point cloud is a function of the trajectory, its wrong estimation will yield a skewed point cloud resulting in its higher spatial error. Therefore, the smoothing SLAM paradigm is of interest in this paper because it focuses on the estimation of the whole trajectory, and hence the correctness of the produced point cloud as a whole.

There exist a number of SLAM methods under the smoothing paradigm, some of which are described by [9], [10] and [11]. Although these authors name their methods differently, their approaches share the same core. They model this problem using a probabilistic graph and then turn it into a problem involving minimizing non-linear least squares. Even with a relatively good initialization, there is no guarantee that such a problem can be executed in constant time due to iterations during the optimization and is generally considered

computationally expensive [12]. However, in the light of increasing computing power of mobile devices and recent scientific advancements in effective factorization methods [13], solving graph-based SLAM in near real-time has become increasingly possible.

The type of sensors used for SLAM also affects how the problem should be modelled. In cases where lidar is involved, it is not uncommon to avoid modelling an explicit map of the scanned scene and instead focus on the trajectory of the scanner only [14] since a stream of lidar points can yield relatively good odometric constraints between key poses [9]. Modelling the problem this way is otherwise known as pose-graph SLAM and involves only the mentioned odometric constraints and loop closures. The latter is important since they provide constraints that allow creation of a globally consistent trajectory, and therefore a globally consistent map.

2.2 Uncertainty propagation in non-linear least squares

On one hand, pose-graph SLAM under the smoothing paradigm still suffers from drift, which cannot be eliminated [15]. On the other hand, information about uncertainty of the key poses can help to localize those parts of the trajectory that suffer from drift. In this vein, the uncertainty/error of the inter-pose constraints can be propagated through the pose-graph so that the joint probability of the key poses can be computed [16]. From there, uncertainty on the individual poses can be calculated through the means of marginal covariances. Since the smoothing paradigm of SLAM can be viewed as a more general non-linear least-squares minimization problem, it is of interest to investigate the methods that propagate uncertainty through such systems. Some of them are presented in [17], [18] and [19].

2.3 Gaps in knowledge

On one hand, SLAM frameworks still suffer from growing drift, resulting in worsening spatial error of point clouds. They are not as accurate as the workflows based on static scanners and surveying, and there is no way to remove the drift in a user-aware manner. On the other hand, there are also ways to propagate the uncertainty in SLAM systems; however, they are either (1) rather theoretical, yielding faster and faster approximate methods of recovering marginal covariances, or (2) have applications in the detection of loop closures where more exact marginal covariances allow for fewer candidates among key-poses to be searched for to find the best match.

2.4 Research objectives

The point of this paper is to provide a real-time method for uncertainty visualization in pose-graph SLAM so that the user can be aware of regions of potentially higher spatial error and they can take corrective actions on-the-fly (during scanning).

3 Proposed solution

3.1 Scope & assumptions

Our solution is designed for SLAM systems working on pose-graphs. We assume that our inter key-pose and loop-closure constraints do not fail in any way during the execution of our system. The spatial error predictions are based on the translational part of marginal covariance matrices Σ_i .

3.2 Overview

The core of our idea is to pass the uncertainty encoded in marginal covariance matrices Σ_i of the estimated trajectory X onto the lidar points PC_i^{lidar} that are associated with the key poses x_i , and scene-reference the points. See Figure 1 for reference. Each key pose comprises a 3D rotation matrix (3×3) and a translation vector (3×1) defined in the coordinate system of the scene. The user can then intuitively see what areas of the progressively-built scan suffer from a potentially high spatial error after scene-referencing the lidar points. As explained in section 2.2, there are SOTA real-time methods to propagate the uncertainties $\{\Omega_1, \Omega_2, \dots\}$ of individual inter key-pose transformations $\{z_1, z_2, \dots\}$ in such a way that the joint probability of all the key poses can be computed. There are also real-time methods to recover the marginal covariances of the key poses after the propagation. Our idea is to pass these marginal covariances Σ_i in real-time onto the scene-referenced lidar points.

More specifically, our method is presented in Figure 2. As the trajectory of a mobile scanner grows, we progressively compute the greatest variance for each marginal covariance matrix Σ_i associated with the new key pose x_i . This will ensure that we find the maximal translational error for the new key pose. Given that the errors along each of the three axes might be correlated, we compute eigenvalues λ_k of the covariance matrix according to Equation (1) and pick the largest, marked as λ_L in processes (a) and (b) in Figure 2 respectively. v_k is an eigenvector associated with the corresponding eigenvalue.

$$\Sigma v_k = \lambda_k v_k \quad (1)$$

However, such computed variance itself is not a

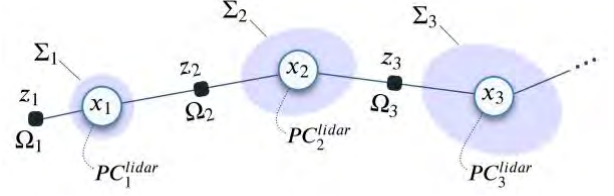


Figure 1. Part of progressively-built trajectory X consisting of a set of key poses $\{x_1, x_2, x_3, \dots\}$ represented as a pose graph. Black rectangles represent constraints Z including inter key-pose odometric transformations $\{z_1, z_2, z_3, \dots\}$. Each odometric transformation comes with an information matrix Ω_i representing uncertainty on these transformations. Pose uncertainties are marked in purple and are represented by marginal covariance matrices $\{\Sigma_1, \Sigma_2, \Sigma_3, \dots\}$

useful statistic in practice. Therefore, we convert it into a standard deviation and multiply it by 2. We thus obtain a 95% confidence level on the maximal translation error of the current pose. We choose a 95% confidence level to comply with land surveying guidance documents such as those by the Royal Institution of Chartered Surveyors [20] or specifications of Levels of Accuracy (LOA) [5] where this level of certainty is de facto a standard.

Next, we compare the two standard deviations to the accuracy levels or bands defined by USBSD or RICS respectively. Before, scanning starts, the user chooses which guidance document they want to comply with. Our system then compares $2\sqrt{\lambda_L}$ for each key pose against the accuracy levels according to the chosen standard.

We propose to colour-code the scene-referenced lidar points according to the accuracy levels. The most restrictive accuracy level (for example LOA 10 by [5]) is coloured in green while the bottom level (for example LOA 50) is red. All levels in between are then colorized according to hues in between these two, such as yellow, amber and orange. It is likely that current mobile devices will be unable to meet the requirements of the most stringent accuracy levels. Hence, we propose to cap the highest level to the one picked by the user, for example, LOA 30. In such a case, all the poses whose $2\sqrt{\lambda_L}$ are smaller than the maximal error associated with LOA 30, will remain green. A similar cap can be imposed for the bottom level. For example, if $2\sqrt{\lambda_L}$ is greater than the error associated with LOA 40, then the points related to this pose will be red.

Finally, for each new key pose, we transform lidar points in the lidar coordinate system PC_i^{lidar} to the scene coordinate system PC_i^{scene} according to Equation (2) and process (e) in Figure 2, with x_i , PC_i^{lidar} and PC_i^{scene} stored in homogenous coordinates. This way, the user scanning a scene can see what the predicted correctness

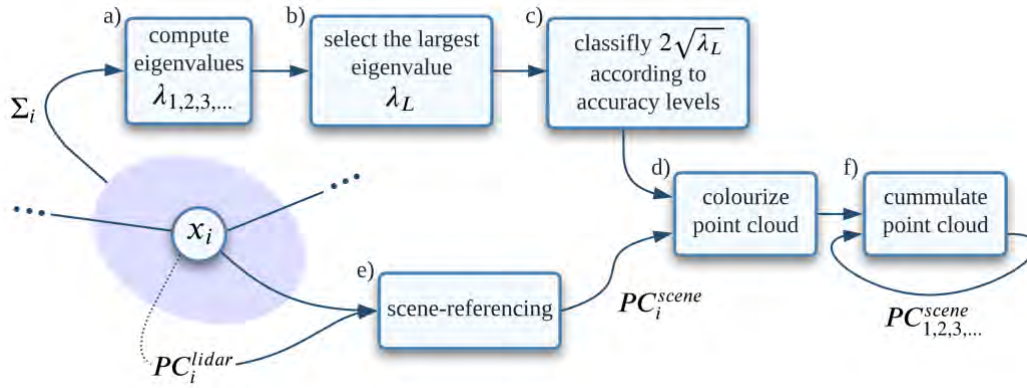


Figure 2. For each key pose x_i the following processes presented in this figure are executed so that the marginal covariance Σ_i is passed onto a scene-referenced pointcloud PC_i^{scene} and visualized to the user in real time.

of the progressively growing point cloud is in real-time.

$$PC_i^{scene} = x_i PC_i^{lidar} \quad (2)$$

The visualization system presented above can then be used in the following way. Before scanning starts, the user chooses the top and bottom levels of point cloud correctness they are willing to accept. As the user traverses new parts of the scene, the progressively-built point cloud gradually changes colour from green to red, indicating that there is a 95% chance that the spatial error of the red regions of the scene is beyond the set lower level. After the desired parts of the scene have been scanned, the user revisits these parts of the scene that are coloured red. By revisiting them, the user closes loops in the underlying pose graph representing the trajectory of the scanner, hence adding additional constraints to the graph. The underlying SLAM optimization process then shifts the trajectory, and hence the lidar points to a more correct position thanks to these additional constraints. This increases the confidence level in the correctness of the key poses and scene-referenced lidar points.

3.3 Hypothesis

For hand-held lidar-based scanners with pose-graph SLAM, real-time predictions on the correctness of the produced point cloud can enable the user to make informed corrective actions during the scanning process, thus increasing the correctness of the point cloud. The corrective actions are revisiting previous parts of the scene in an informed way.

4 Research methodology

4.1 Data collection

We built a prototypic scanner consisting of a

Velodyne VLP-16 lidar and a MacBook Pro laptop, both plugged into a portable power unit as shown in Figure 3. We coded up software for the scanner on Linux Ubuntu 20.04 with Robot Operating System [21] (version Noetic) using many own and publicly available repositories and frameworks.



Figure 3. Our prototypic scanner connected to a laptop during scanning.

Next, we went to one of the colleges and scanned its Front Court with the scanner to test our method. During scanning, we walked along the four walls so that the façade was captured, and we returned to the place we started scanning to close a loop. The estimated trajectory of the scanner can be seen in Figure 4. In addition, we used a FARO Focus 3D terrestrial scanner [22] to provide a ground truth scan. The scanner was placed in the middle of the court so that it could cover all four façades.

4.2 Methods

We follow a Maximum A Posteriori (MAP) incremental non-linear pose-graph optimization approach by [13], to find such key poses of trajectory X that their errors for the odometric inter key-pose and loop closure transformations z_i are the smallest (Equation (3)).

$$X^{MAP} = \min_X \sum_i \|h_i(x_i) - z_i\|_{\Omega_i^{-1}}^2 \quad (3)$$

In Equation (3), $h_i(\cdot)$ are non-linear functions transforming key poses stored in the coordinate system of the scene to the coordinate system of the previous pose in which z_i is measured. Computation of the odometric constraints is based on the Lidar Odometry And Mapping technique by [7].

We estimated the information matrices Ω_i in such a way that after the loop closure event shown in image d) of Figure 4, the most uncertain part of the point cloud (around the top corner) is coloured red and the other parts are green for LOA 10 to 50.

5 Results & Discussion

5.1 Raw results

In Figure 4, we present the progressively created point cloud with overlaid uncertainty information. At the start of the reconstruction (image a), all the points in the scene are green, indicating prediction in their relatively correct position. As we traverse the scene, the drift increases and so does the estimated spatial error in the newly accumulated points. This is shown first in amber and then in red in image (b), predicting that the relative error in these places is around 3-4 times higher than in the bottom part of the scene. Having this information and following our method, we decided to return to the green area to close the loop and potentially increase the correctness of these red points. However, as we go towards the place we started the scanning, the newly reconstructed parts of the scene on the right side of the scene turn even more red (image c), indicating that the error there might be 5 times higher than in the green areas. Finally, we close the loop in image (d), and the points around this area turn green. The system predicts that their spatial error decreased significantly. However, the points in the corner at the top part of the scene remain amber.

We will work on the point cloud shown in image (d) of Figure 4 and investigate what the actual spatial error in the amber region (top corner) is and contrast it with the spatial error in the top right corner, which should have a relatively lower error according to the predictions of our system.

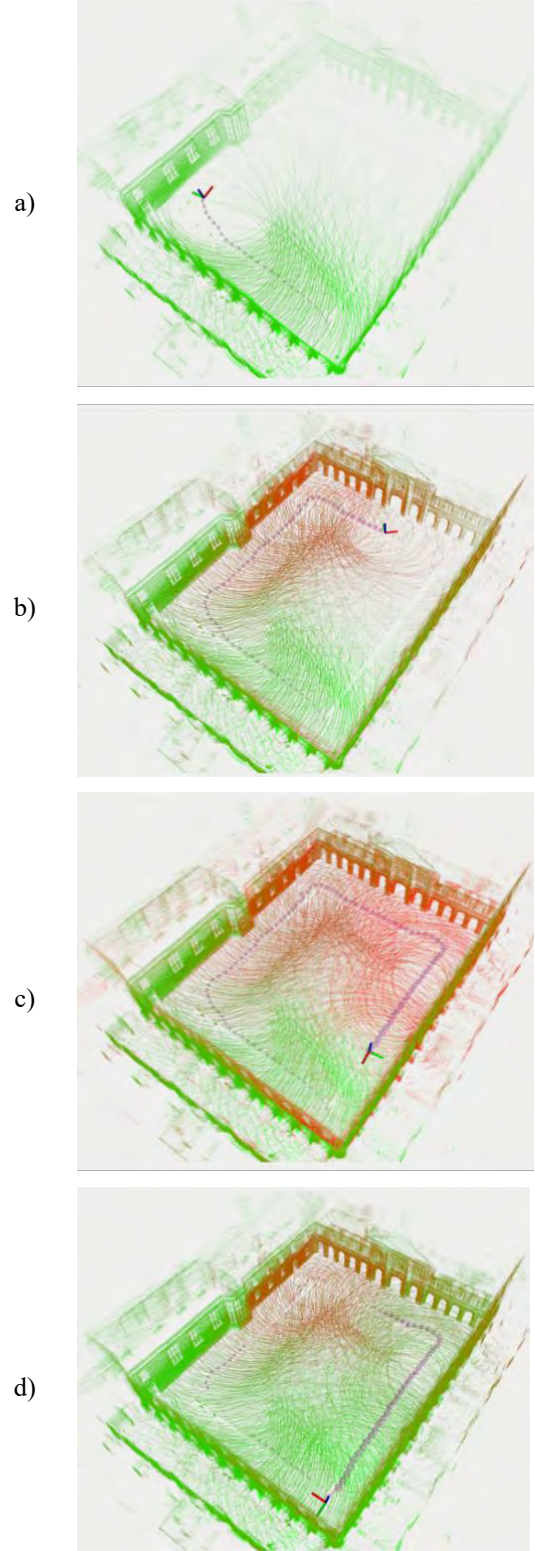


Figure 4. Progressively-built point cloud in real-time with overlaid uncertainty. Green indicates high point cloud correctness while red stands for predictions for a relatively higher spatial error.

5.2 Analysis

We segmented the point cloud by our scanner into three regions (all marked in blue in Figure 6): the one at the bottom of the figure was used to register the whole point cloud by our mobile scanner to the point cloud by the FARO Focus 3D scanner; the one at the top of the figure was predicted by our system as having a bigger spatial error (see Figure 4 d); finally, the region in the top right of the figure was predicted for a relatively lower spatial error than the previous region after the loop closure (again, see the corresponding Figure 4 d), although initially, the error there was even higher than that at the top corner (Figure 4 c).

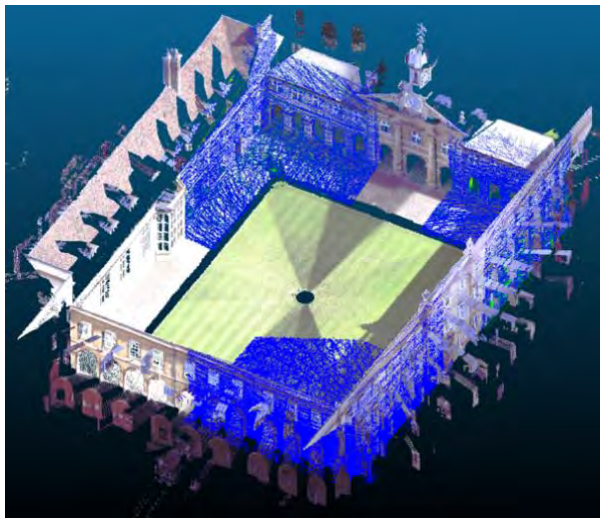


Figure 6. A scan from a terrestrial FARO Focus 3D scanner was used to provide ground truth (colourful point cloud). The segmented blue point cloud comes from our mobile scanner.

For the top two regions, we computed the distances to the ground truth scan using Cloud Compare. Next, we binned them and present them as histograms in Figure 5. On these distances, we also computed three statistics: 1) a mode (most likely value), a mean (average value) and a 95-percentile (see Table 1).

Judging by the modes in Table 1, the blue region at the top of Figure 6 has almost 5 times higher most likely spatial error than the point cloud in the top right corner (47 and 10 mm respectively). It also has around 39% higher mean error, and its 95-percentile is larger by 9 mm. These three statistics seem to confirm that our system correctly predicted the regions of relatively higher spatial error.

In addition to measuring the spatial error above, we also took a closer look at the loop closure event that occurred right before image (d) in Figure 4 was taken. In Figure 7, the initial key poses comprising the odometric trajectory (marked with a dashed light blue line) had very

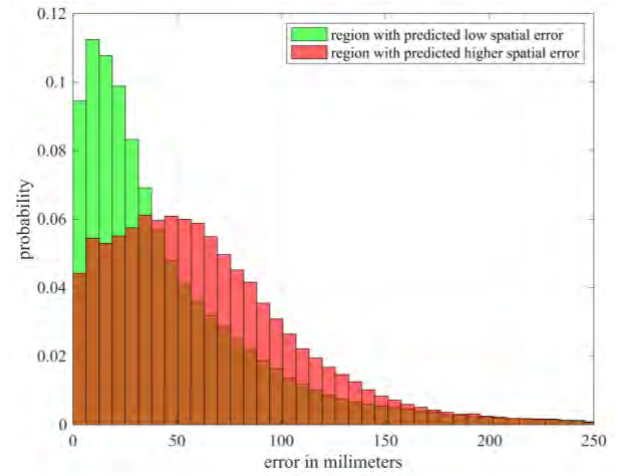


Figure 5. Distribution of error between the ground truth and the regions of a point cloud identified by our scanner as having a low spatial error (green bins) and a high one (red).

large marginal covariances (purple spheres), indicating that the position of the related point clouds was significantly off. After computing the relative transformation between the two key poses connected by the red line, the trajectory has been recomputed which resulted in an updated trajectory marked with a green line. The updated key poses now have significantly smaller marginal covariances indicated by the much smaller spheres. In fact, these spheres are as small as those at the start of scanning (Figure 4 a). The updated trajectory and its marginal covariances on the key poses seem to be in line with other studies on this topic such as those by [23]. This shows that we have implemented our method correctly.

Table 1. The actual spatial errors corresponding to the two segmented regions of our point cloud shown in Figure 6.

Statistics	Region with predicted low spatial error [mm]	Region with predicted higher spatial error [mm]
Mode	10	47
Mean	46	64
95-percentile	140	149

In this case study, the top corner marked in red in Figure 4 (d) still has lower accuracy compared to the green part. The question then is how to improve it? The informed user could walk up to the corner while scanning, diagonally through the lawn, and close the second loop. This would further reduce the uncertainty, especially in

that region, hence decreasing the error. We are planning to show this and demonstrate the impact of closing the second loop (and maybe more loops) in the following paper in the future. The reason for not demonstrating this here is the fact that more loop closures demand even tighter integration of odometry and SLAM systems than we have currently implemented. However, it is a problem related to the architecture of our software and the method described in this paper will still hold.

5.3 Our contributions

We presented a novel fully-explainable method for real-time predictions of areas with potentially high spatial error in progressively created point clouds by lidar-based mobile scanners. The method exploits the human-machine interaction and is suitable for SLAM systems based on pose graphs. We implemented it and tested its applicability in a case study which confirmed that the identified areas of relatively higher spatial error, indeed suffer from it. To the best of our knowledge, it is the first such system in the world.

Our system can help the construction industry to reduce the effort (time and money) put generally into scanning construction sites. Since the user is aware of the regions of potentially high spatial error during scanning, they can take corrective actions on-the-fly, and not after the data collection and processing have been completed. This, in turn, might eliminate the potential risk of re-scanning.

6 Conclusions

In this paper, we presented a real-time fully-explainable method exploiting human-machine interaction to increase the correctness of point clouds produced by lidar-based mobile scanners. With a 95% confidence level, our method predicts areas of relatively higher spatial error in a progressively created point cloud. We built a mobile scanner, implemented our method, and tested it in a case study. The results show that the areas identified with a relatively high spatial error indeed suffer from it, while areas with relatively higher correctness predicted, do have a smaller spatial error in reality.

The method presented here has the potential to increase the accuracy of point clouds produced by lidar-based SLAM systems operating on pose-graphs. This, in turn, might unlock demanding use-cases such as engineering surveying or high accuracy measured building surveys, which so far have been unreachable by mobile scanners due to systematic spatial errors in point clouds caused by drift. Moreover, our method might contribute to reducing the effort put into scanning by giving real-time predictions on the quality of produced point clouds, hence eliminating the need for potential rescanning.

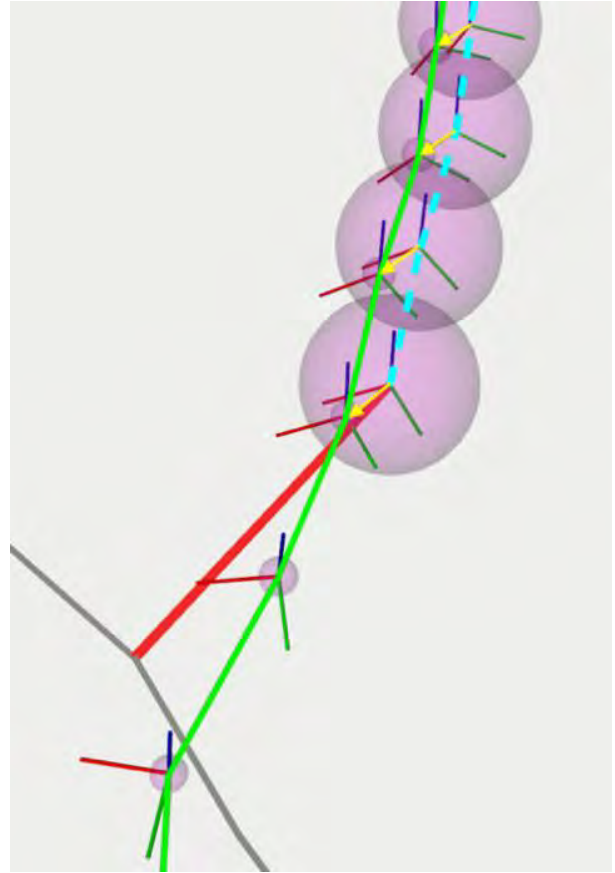


Figure 7. Correction of the odometric trajectory (dashed blue line) resulting in the SLAM trajectory (green line) in the 3D space of the scene. The red line connects the two key poses used for computing the loop closure with yellow arrows showing how the corresponding key poses shifted. Purple spheres represent marginal covariances.

In the future, we are planning to test our method in more case studies and extend it so that it yields predictions in cases where many loops have been closed.

Acknowledgements

The research leading to this paper received funding from BP, GeoSLAM, Laing O'Rourke, Topcon and Trimble. We would like to thank these companies for making our research possible. We also thank Emmanuel College for allowing access to their grounds. We gratefully acknowledge the collaboration of all the industrial partners. Any opinions, findings, conclusions or recommendations expressed in this material are ours and do not necessarily reflect the views of the companies mentioned above.

References

- [1] BIM task group, "Client guide to 3d scanning and data capturing," 2013.
- [2] T. S. Kalyan, P. A. Zadeh, S. Staub-French, and T. M. Froese, "Construction Quality Assessment Using 3D as-built Models Generated with Project Tango," *Procedia Engineering*, vol. 145, pp. 1416–1423, Jan. 2016, doi: 10.1016/j.proeng.2016.04.178.
- [3] NavVis, BIM+, Geo Week News, Lidar News, and GIM International, "State of Mobile Mapping: Survey," 2021. [Online]. Available: <https://f.hubspotusercontent40.net/hubfs/3339696/State%20of%20Mobile%20Mapping%202021.pdf?hsCtaTracking=2cf01e48-9803-4273-a9cb-28a5a72bfb88%7C1003557b-fe59-47bb-a928-e17d0470e054>
- [4] M. Trzeciak and I. Brilakis, "Comparison of accuracy and density of static and mobile laser scanners," Apr. 2021, doi: 10.17863/CAM.66857.
- [5] USIBD, "USIBD Level of Accuracy (LOA) Specification Guide." U.S. Institute of Building Documentation, 2016.
- [6] L. Lachauer, "Get real answers about mobile scanning accuracy." <https://www.navvis.com/blog/get-real-answers-about-mobile-scanning-accuracy> (accessed Feb. 25, 2022).
- [7] J. Zhang and S. Singh, "LOAM: Lidar Odometry and Mapping in Real-time," 2014. doi: 10.15607/RSS.2014.X.007.
- [8] C. Cadena *et al.*, "Past, Present, and Future of Simultaneous Localization And Mapping: Towards the Robust-Perception Age," *IEEE Trans. Robot.*, vol. 32, no. 6, pp. 1309–1332, Dec. 2016, doi: 10.1109/TRO.2016.2624754.
- [9] M. Kaess and F. Dellaert, "Factor Graphs for Robot Perception," 2017. <https://ieeexplore.ieee.org/document/8187520?bkn=8187520> (accessed Feb. 02, 2022).
- [10] G. Grisetti, R. Kümmerle, C. Stachniss, and W. Burgard, "A Tutorial on Graph-Based SLAM," *IEEE Intelligent Transportation Systems Magazine*, vol. 2, no. 4, pp. 31–43, winter 2010, doi: 10.1109/MITS.2010.939925.
- [11] Cyrill Stachniss, *Robust Least Squares for Graph-Based SLAM (Cyrill Stachniss)*, (Apr. 06, 2020). Accessed: Feb. 25, 2022. [Online Video]. Available: <https://www.youtube.com/watch?v=z60RbiY18I8>
- [12] T. D. Barfoot, *State Estimation for Robotics*. Cambridge: Cambridge University Press, 2017. doi: 10.1017/9781316671528.
- [13] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert, "iSAM2: Incremental smoothing and mapping with fluid relinearization and incremental variable reordering," in *2011 IEEE International Conference on Robotics and Automation*, May 2011, pp. 3281–3288. doi: 10.1109/ICRA.2011.5979641.
- [14] A. Jurić, F. Kendeš, I. Marković, and I. Petrović, "A Comparison of Graph Optimization Approaches for Pose Estimation in SLAM," in *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, Sep. 2021, pp. 1113–1118. doi: 10.23919/MIPRO52101.2021.9596721.
- [15] L. Carlone, *The Future of Robot Perception: Recent Progress and Opportunities Beyond SLAM* | Luca Carlone, (Jul. 15, 2021). Accessed: Feb. 02, 2022. [Online Video]. Available: <https://www.youtube.com/watch?v=j5g3efgdjRg>
- [16] "Factor Graphs and GTSAM," *GTSAM*, May 22, 2019. <http://gtsam.org/tutorials/intro.html> (accessed Feb. 25, 2022).
- [17] G. D. Tipaldi, G. Grisetti, and W. Burgard, "Approximate Covariance Estimation in Graphical Approaches to SLAM," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct. 2007, pp. 3460–3465. doi: 10.1109/IROS.2007.4399258.
- [18] M. Kaess and F. Dellaert, "Covariance recovery from a square root information matrix for data association," *Robot. Auton. Syst.*, vol. 57, no. 12, pp. 1198–1210, Grudzie 2009, doi: 10.1016/j.robot.2009.06.008.
- [19] V. Ila, L. Polok, M. Solony, P. Smrz, and P. Zemcik, "Fast covariance recovery in incremental nonlinear least square solvers," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, May 2015, pp. 4636–4643. doi: 10.1109/ICRA.2015.7139841.
- [20] RICS, "Measured surveys of land, buildings and utilities. RICS guidance note, global. 3rd edition." Royal Institution of Chartered Surveyors, 2014.
- [21] M. Quigley *et al.*, "ROS: an open-source Robot Operating System," in *ICRA workshop on open source software*, 2009, vol. 3, no. 3.2, p. 5.
- [22] FARO, "Technical Specification Sheet for the Focus3D X 30/130/330 and X 130/330 HDR," *FARO® Knowledge Base*, Dec. 20, 2016. https://knowledge.faro.com/Hardware/3D_Scanners/Focus/Technical_Specification_Sheet_for_the_Focus3D_X_30-130-330_and_X_130-330_HDR (accessed Jan. 19, 2021).
- [23] D. M. Cole and P. M. Newman, "Using laser range data for 3D SLAM in outdoor environments," in *Proceedings 2006 IEEE International Conference on Robotics and Automation*, 2006. *ICRA 2006.*, May 2006, pp. 1556–1563. doi: 10.1109/ROBOT.2006.1641929.

Toward Automation in Crack Detection and Measurements: Benchmarking of CNN-based Algorithms

D. Ji^a, Y. Turkan^a and P. Calvi^b

^aSchool of Civil and Construction Engineering, Oregon State University, United States of America

^bDepartment of Civil and Environmental Engineering, University of Washington, United States of America

E-mail: jid@oregonstate.edu, yelda.turkan@oregonstate.edu, Pmc85@uw.edu

Abstract

Cracks are one of the main defect features on concrete surfaces, and indicators of concrete structures' condition regarding their state of health. Since traditional methods to identify and assess cracks rely on manual measurements, a significant number of studies to date have focused on identifying ways to automate this process. Accordingly, this study proposes to combine Convolutional Neural Network (CNN)-based algorithms and traditional image morphological operation for crack detection and measurements with a specific goal to benchmark the proposed methodology on data of three images obtained from laboratory experiments in a controlled environment. The proposed methodology performed well, with 92.10% and 90.11% accuracy in crack length and width measurements, respectively. Future research will focus on fine-tuning the proposed crack detection and measurement methodology and evaluate them on a set of images acquired from a real full-scale structure such as a bridge or a building.

Keywords

Crack Detection; Crack Measurement; Image Morphological Operation; Convolutional Neural Network (CNN); Skeletonization

1 Introduction

According to the recent infrastructure report from the American Society of Civil Engineers (ASCE), approximately 231,000 bridges across the United States, or 37.4% of all bridges in the nation, need repair and preservation work, and 46,154, or 7.5% of the nation's bridges, are considered structurally deficient [1]. To prevent infrastructure failures, periodic inspections are needed to determine their condition. One of the most important procedures during periodic inspections are visual inspections carried out to identify and assess surface defects.

Cracks are one of the main defect features on concrete surfaces, and indicators of concrete structures' condition

regarding their stability and durability [2]. Thus, the identification and assessment of cracks plays a paramount role in concrete structures' inspections. However, since the traditional methods to identify and assess cracks mostly rely on manual measurements, which are labor intensive and error prone, several recent research efforts have focused on leveraging computer vision and machine learning techniques to automate this process.

Several image-based crack detection and measurement techniques, including deep learning-based algorithms, have been proposed to date to automate crack identification and assessment processes. These studies proposed various image-based crack detection methods that use edge detection algorithms [4]-[6], automatic threshold image segmentation [7], and image morphological operations [8][9]. In recent years, the focus has been on utilizing deep learning-based algorithms due to their robustness and accuracy. These studies utilized convolutional neural networks (CNN) architectures [10]-[13], and performed better compared to previously proposed computer vision-based methods [14]-[20].

Even though image-based methods have been proven to be effective in detecting cracks, most of them perform well when applied to images with well-distinguishable cracks, i.e., cracks with prominent edge-gradient changes, and none to minimal number of non-crack objects such as stains. In addition, deep learning-based crack detection algorithms require large number of training dataset of images and heavy computing resources. Therefore, it can be concluded that the currently available computer vision-based methods to identify and assess surface defects are not mature enough to be used in real-life scenarios, e.g., on data collected from bridges or buildings.

Several studies have focused on fine-tuning the previously proposed algorithms for crack detection and measurement to overcome the limitations summarized above. Moreover, to achieve an acceptable level of accuracy and sensitivity on real-case scenarios, new approaches that combine different methods (e.g., deep

learning and 3D point cloud) have been proposed.

Accordingly, this study aims to benchmark CNN based deep learning algorithms for crack detection and measurements on data obtained from a laboratory experiment in a controlled environment. A methodology that combines a CNN-based algorithm and the traditional image morphological operation is presented for crack detection. Next, skeletonization and orthogonal projection algorithms were applied to measure the crack length and width and their overall performance were evaluated.

2 Literature Review

In this section, previous studies on crack detection and measurements are reviewed. First, crack detection methods based on conventional image morphological operations, deep learning algorithms, and the approaches that combine the first two are examined and their strengths and weaknesses are discussed. Next, the studies that focused on crack measurements are analyzed. Finally, the methods selected for crack detection and measurement in this study are introduced.

2.1 Crack Detection

The majority of studies that focused on automating crack detection used digital images as input data. Thus, conventional object detection and segmentation algorithms are very important and play a major role in image-based crack detection. Initially, researchers have focused on detecting crack edges since they are the most distinguishable crack features, and various edge detection methods, such as Sobel, Prewitt, and Canny have been utilized for this purpose [4]-[6]. These methods detect crack edges based on prominent gradient changes in the image, especially the changes in intensity in certain directions along the feature of interest. If the gradient changes are not significant, meaning that the crack is not clearly distinguishable, the overall crack detection performance decreases drastically. Hence, several studies focused on extracting crack edges by applying various image processing algorithms.

To extract cracks, i.e., separate them from the background details in each image, Otsu [7] proposed a segmentation method based on image thresholding approach that uses the maximum grayscale intensity values in the image. Nguyen et al. [8] proposed the Free-Form Anisotropy method, which simultaneously considers various crack features including intensity values, texture, and others to detect cracks more accurately. Following these two approaches, several research studies applied morphological operations to pre/post process images for feature extraction purposes. For example, Xu and Turkan [9] presented an approach for automated crack detection using images of a bridge

acquired from an unmanned aircraft system (UAS). They applied Otsu's image gradient segmentation method and additional image pre-processing steps including contrast adjustment and noise reduction for crack detection and reported precision and recall values of 74.6% and 86.2%, respectively.

Although most crack detection methods based on image morphological operations have performed well and yielded meaningful results, issues such as the presence of shadows or stains, features similar to cracks, impact the performance of these methods negatively.

Meanwhile, several studies focused on utilizing deep learning-based object detection methods. The major CNN-based vision architectures, such as AlexNet [10], VGG-Net [11], Inception Network [12], ResNet [13], which performed well in detecting and classifying various objects, were utilized for image-based crack detection in several studies.

Dorafshan et al. [21] compared the performance of conventional edge detectors (Roberts, Prewitt, Sobel, Laplacian of Gaussian, Butterworth, and Gaussian) and a CNN-based crack detector. They tested AlexNet-based crack detector using fully trained, partially re-trained, and pre-trained datasets to determine its performance under different circumstances. They used images from SDNET dataset that include images with various surface defects [22]. The CNN-based crack detector performed well when a fully trained dataset used, with precision value of 99% and recall value of 66%. The precision values for partially trained and pre-trained datasets were 92% and 80% while the recall values were 86% and 84% respectively. In the meantime, the conventional edge detector based on the Laplacian of Gaussian algorithm achieved the precision and recall values of 60%, and 79%, respectively, which was the best performance among conventional methods. Wang et al. [23] tested the crack detection accuracy of six existing CNNs (VGG 16, Inception V2, V3, V4, Inception-ResNet-v2, and ResNet V1 50), using an original image dataset collected from the inspection of a slab element and the highest accuracy was obtained when using the Inception-ResNet-v2 network, with 80.08% of accuracy when utilizing pre-processed dataset.

Recent studies that focused on automatic crack detection proposed to combine different methods to improve the accuracy and reliability of the results. Several studies focused on augmenting digital images with three-dimensional (3D) data such as point clouds acquired using a laser scanner (i.e., lidar) to overcome the issues faced when using image-based methods such as loss of feature (crack) details due to shadows. Chen et al. [2] utilized point cloud data to obtain a depth image, which illustrates features in the image based on their measured depth. The depth image was then combined with the pre-processed image, and by applying Otsu's

crack detection algorithm, the proposed approach achieved, on average, 89.0%, 84.8%, and 86.7%, precision, recall and F1, respectively. Yan et al. [3] also utilized point cloud data to identify cracks, which enabled them to exclude background and other non-target segments in data processing, which is one of the most challenging issues for images obtained in real life cases, e.g., images of a bridge. Simultaneously, they processed images including the same crack features, obtained from the three columns of a bridge, using a VGG16-based crack detector. By combining the results obtained from point cloud and image-based methods, they achieved 93% crack detection accuracy on average, and obtained a precision and recall values of 93.9%, and 89.4%, respectively.

2.2 Crack Measurement

Recent studies on image-based crack measurements have focused on implementing a skeletonization algorithm to guarantee the accuracy of the measurements [3][23][24]. Using binary images, which contain the geometry of cracks, the skeletonization algorithm creates the centerline of the crack in one-pixel thickness. From this centerline, referred to as the crack skeleton, the length of a crack can be measured in pixel dimensions by counting the number of pixels along centerline. To measure crack width, the most accurate method proposed in previous studies calculates the continuous width of each pixel on the crack skeleton using the orthogonal projection algorithm [3][23][24], which is explained next. First, the orientation of each crack pixel is computed by fitting a line to the target pixel and its neighboring pixels on the crack skeleton. Then, an orthogonal line, which is perpendicular to the orientation of the crack in the target pixel, is projected to obtain two intersecting points between the orthogonal line and crack boundaries. In the final step, the width of the crack is calculated as the distance between these two intersecting points.

Qiu et al. [24] validated the performance of width measurement based on the skeletonization and the orthogonal projection algorithms. Using images that contained ten different cracks, with widths ranged between 18.1mm and 66.3mm (with an average of 36.4 mm), they calculated 1.4 mm difference, on average, between the ground truth values and their results. A study from Yan et al. [3], which utilized both CNN-based crack detector and point cloud data, presented, and tested their crack measurement approach based on skeletonization and orthogonal projection algorithms. The test datasets they used included cracks with lengths ranged between 44.5mm to 559.0mm and widths ranged between 1.0mm to 5.0mm. They obtained average error rates of $\pm 3\%$ and $\pm 8\%$ for length and width measurements, respectively. Wang et al. [23], who also adopted skeletonization and orthogonal projection methods, mainly focused on

classifying crack images into three classes based on their severity levels, i.e., average crack widths. They used an image dataset that contains cracks with an average width of less than 1.0mm and classified the dataset with the average accuracy of 97.41%.

One of the most essential steps in crack measurements is the conversion of measurements into real scale. This process converts the measured pixel dimensions into units such as mm or inches. For example, Wang et al. [23] utilized the width of railhead in the image as a reference. Since the actual geometric width of the railhead can be identified from its specifications document, the conversion factor can be easily calculated. This is a straightforward approach for conversion, but the images must contain a certain target feature with known dimensions in real world units. Also, during data collection an appropriate camera angle must be maintained. Another scale conversion approach is to utilize data from other sources. For example, Yan et al. [3] used values of depth and focal length of the lidar data to compute the scale factor. Kalfarisi et al. [25] utilized dimensions from the 3D mesh model, which was reconstructed from 2D images. The 3D mesh model provided the dimensions in both pixels and metric units.

3 Methodology

3.1 Crack Detection

In this study, a CNN-based crack segmentation algorithm called DeepCrack, which is proposed by Liu et al. [14], and a traditional image morphological operation based on Otsu's image segmentation algorithm [7] are combined for image-based crack detection. DeepCrack is a CNN-based crack segmentation algorithm that uses the VGG-16 network, one of the major vision architectures in computer vision. In this study, a pre-trained DeepCrack algorithm model, which was trained with 4,800 images and 3,792 images, was used for testing. The overall performance of the pre-trained DeepCrack model is as follows: the global prediction accuracy is 98.73%, while precision, recall and F1 scores are 85.82%, 84.56%, and 85.18%, respectively [14]. In the next step, the image morphological operation, which is based on Otsu's image segmentation algorithm along with additional pre-processing (e.g., contrast adjustment, smoothing) and post-processing (e.g., area filtering) steps [9] were implemented.

The detailed crack detection procedure followed in this study is provided in Figure 1.

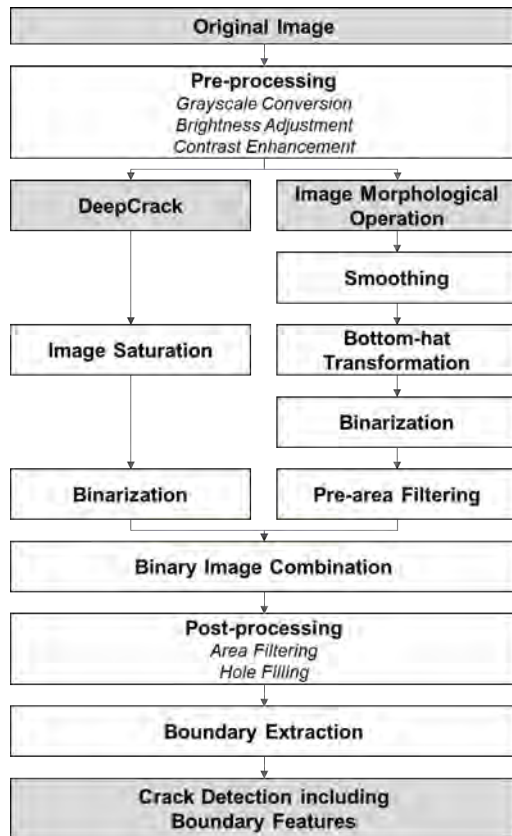


Figure 1. Crack Detection Procedure

First, the original image is pre-processed using grayscale conversion, brightness adjustment, and contrast enhancement techniques, respectively. In the next step, two different methods are applied to the pre-processed images in parallel:

1. The first method involves using the pre-processed image as an input for the pre-trained DeepCrack model. Next, the result from the DeepCrack algorithm is post-processed using image saturation parameters, and binarized to obtain a binary image.
2. The second method used is a conventional image morphological operation process that consists of procedures that include smoothing and discarding unnecessary details of non-crack features, bottom-hat transformation to extract dark regions that are not cracks, and pre-area filtering that uses a certain threshold value to extract detailed features of cracks, which may have been excluded from DeepCrack segmentation results.

After the two-track process based on DeepCrack and image morphological operations, the preliminary results obtained from each step are integrated. Next, the image was post-processed using second area filtering and a hole-filling operation. Lastly, the crack boundaries are extracted using image gradient thresholding technique.

3.2 Crack Measurement

In this study, two prominent crack measurement methods, the skeletonization and orthogonal projection algorithms, are used (Figure 2). The binary image obtained at the end of the crack detection process, is skeletonized. Since the original crack skeleton contains several branches that are not related to the crack length or the main orientation, a post-processing method was used to filter those branches (i.e., pruning). Next, the orthogonal projection algorithm is applied to the crack skeleton. The orientation of the crack skeleton is computed in each pixel, and the target pixel's orthogonal line is projected. By merging these projected orthogonal lines and the binary image of crack boundary from the crack detection step, two intersecting points between the orthogonal line and crack boundaries are obtained. As detailed in section 2, the crack length is measured by counting the number of pixels along the crack skeleton, and the width of the crack is measured by calculating the distance between the orthogonal line and crack boundaries. Finally, the crack length and width dimensions, which were calculated in numbers of pixels are converted into metric dimensions using the dimensions of the specimen that is used in this study.

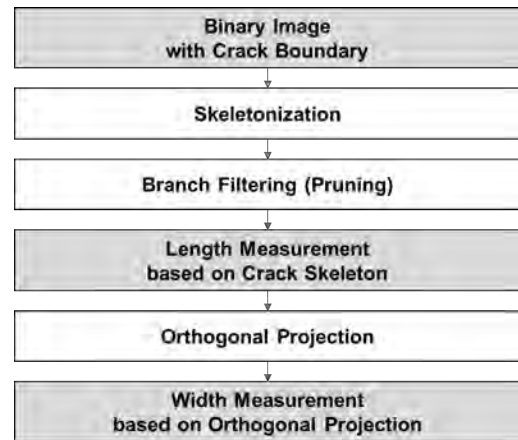


Figure 2. Crack Measurement Procedure

4 Experiment

The crack detection and measurement methods detailed in section 3 were applied to three images taken during a series of shear strength tests conducted at the University of Washington Structures laboratory and the performance of both methods are evaluated. The test data and the experimental procedure are detailed below.

4.1 Test Data

The test data used this study are three images obtained

from the shear strength tests on ultra-high-performance concrete (UHPC) [26]. For each test, an $890 \times 890 \times 70$ mm UHPC panel was tested under shear loads using the UW Panel Element Tester. The experimental setup was such that a major crack would form at a pre-selected location. The crack obtained at the end of this test was used as the target crack to be detected and measured in this study (Figure 3, Top).

The original test images were manually cropped and trimmed to extract regions of interests, which contain the target crack (Figure 3, Bottom). The size of the manually cropped test images that contains the regions of interests are 1808×1748 , 1350×1394 , and 2044×2011 pixels, respectively (left to right, i.e., images 1-3).

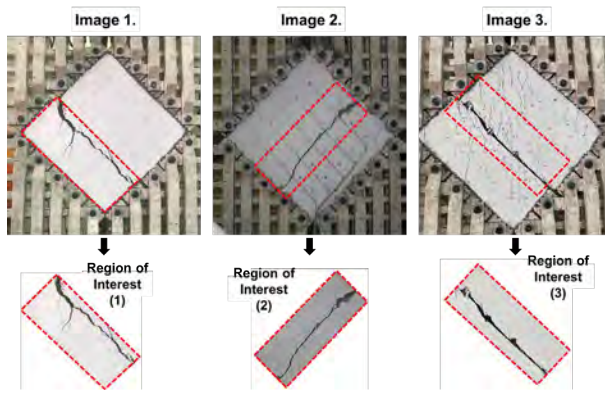


Figure 3. Original Image (Top) and Manually Cropped Test Image (Bottom)

4.2 Experiment Procedure

The images obtained from the laboratory experiments in a controlled environment are processed using the crack detection and measurement methods detailed in section 3. First, the performance of the crack detection method is evaluated based on feature segmentation accuracy that is determined using both the continuity of the feature, and the sensitivity such as preserving the details of the crack. Next, the measured values of length and width are compared with the manually annotated ground truths. More specifically, the length measurements obtained automatically following the procedure described in section 3.1 are directly compared to the manually annotated length from the image. For crack width comparison, five checkpoints along the crack are selected in each image, and the automatically measured widths in the checkpoints are compared with manually annotated widths. This experiment is designed to fine-tune the methodology described in section 3. Once the fine-tuning is achieved, the overall goal of this study is to apply this methodology to real case scenarios, e.g., images collected from bridges.

5 Results

5.1 Crack Detection

The crack detection results based on the methodology proposed in this study are presented in Figures 4 and 5. As can be seen, the cracks in all three images were detected, and their boundaries were extracted successfully. One issue that needs to be discussed here is that there is a loss in details where the crack branches are very thin (Figure 6). These narrow cracks are negligible when assessing reinforced concrete bridges for severe conditions according to North America standards [27], since this type of damage does not affect the integrity of the structure. Note that the average width of the cracks that were lost at the end of the detection process using the proposed methodology is 0.76 mm, which is manually measured at thirty random points and averaged.

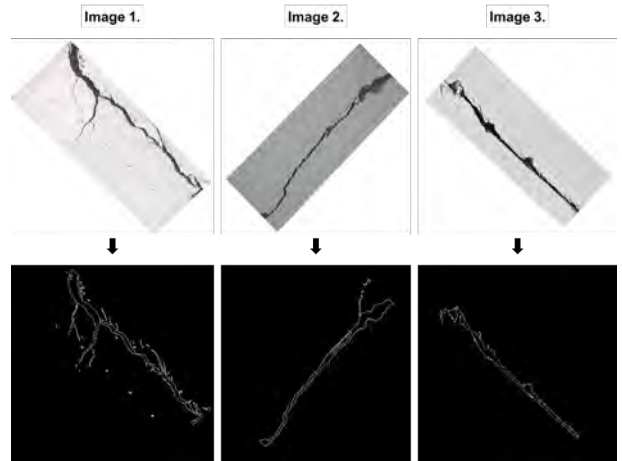


Figure 4. Test Images (Top) and the Crack Detection Results (Bottom)

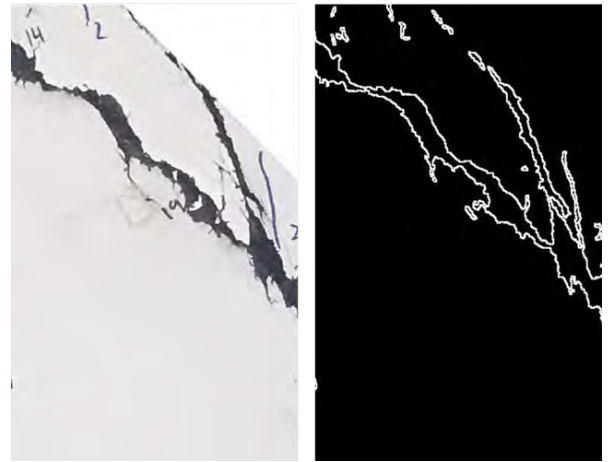


Figure 5. Detailed Detection Results obtained for Image 1

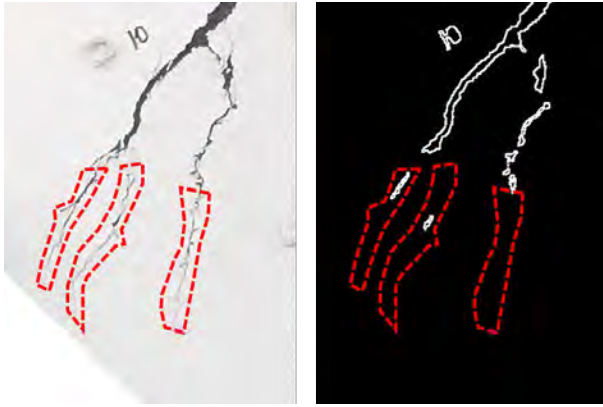


Figure 6. Detail Losses from Crack Detection

5.2 Crack Measurement

First, a real-scale conversion factor is calculated to convert pixel-dimensions into metric dimensions (in millimeter). The widths of the concrete panel in each image were manually measured directly, and the measured values of images 1, 2, and 3 are 1,770.6, 1375.4, and 1939.2 pixels respectively, which equaled to 890 mm. Using these values, the conversion factors for three images were calculated as 0.5027, 0.6471, and 0.4590 mm/pixel, respectively.

Second, the performance of skeletonization and branch filtering algorithms is evaluated. After branch filtering based on the threshold to obtain the proper crack skeleton in a 1-pixel thickness, the losses of pixels on either end of each of the cracks are observed (see Figure 7). The total losses in length along the main crack due to the branch filtering were 28.51, 34.78, and 30.29 mm in images 1 through 3, respectively.

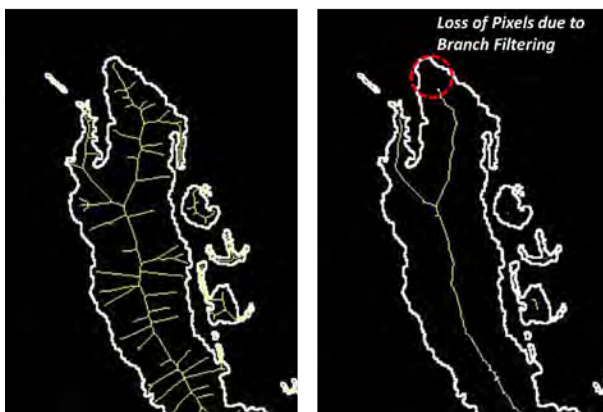


Figure 7. Skeletonization Result and Pixel Loss due to Branch Filtering

Third, the length measurement results are evaluated (Table 1). The measured length along the main crack skeleton in image 1 is 1019.48 mm, which contains 2028

pixels. The main crack skeleton in image 2 is measured as 798.52 mm, or 1234 pixels, and 772.42 mm, or 1683 pixels for image 3. The actual crack length (ground truth) was calculated by dividing each crack into ten segments since the crack is not straight, and then manually measuring each segment and adding them up. At the end of this process, it was determined that the total crack length (ground truth) of each crack in three images are 1071.94, 914.07, and 823.15 mm for images 1 through 3: corresponding to 95.11%, 87.36%, and 93.84% accuracy in length measurement respectively.

Table 1. Length Measurement Results

Image	Proposed method (mm)	Ground Truth (mm)	Accuracy (%)
1	1019.48	1071.94	95.11
2	798.52	914.07	87.36
3	772.42	823.15	93.84
Average			92.10

Lastly, the crack width measurement result is evaluated by measuring the width from five random checkpoints along each crack in the three images. Using the coordinates of the checkpoints, the width was measured using the method detailed in section 3.2. The actual crack widths (ground truths) at each checkpoint are manually measured directly from the original image. The automatic width measurement results (based on orthogonal projection) as well as the ground truth values for all check points are provided in Table 2. As can be seen, the average width measurement accuracy for each image is 90.54%, 90.78%, and 89.01%, respectively.



Figure 8. Representation of Width Measurement

Table 2. Width Measurement Results

Checkpoint	Automated Measurement (mm)	Ground Truth (mm)	Accuracy (%)
1.1	13.51	15.29	88.36
1.2	3.83	4.32	88.86
1.3	15.78	16.41	96.16
1.4	12.42	14.00	88.71
1.5	13.05	14.40	90.63
Average - 1			90.54
2.1	29.77	27.00	89.75
2.2	11.65	13.28	87.73
2.3	10.11	11.45	88.30
2.4	9.15	9.71	94.23
2.5	53.55	50.47	93.88
Average - 2			90.78
3.1	14.69	15.42	95.27
3.2	12.75	12.22	95.65
3.3	22.03	17.66	75.27
3.4	30.78	37.73	81.58
3.5	25.04	25.74	97.28
Average - 3			89.01
Total Average (1, 2, 3)			90.11

6 Conclusions

Visual inspections to identify and assess surface defects are important to ensure structural safety. In particular, dimensions of surface cracks are used to determine concrete structures' condition in terms of their stability and durability. However, the assessment of concrete cracks still relies on manual identification and assessment. Therefore, recent research has focused on leveraging computer vision and machine learning techniques to automate this process.

This study presented a methodology that combined CNN-based algorithms and the traditional image morphological operation for crack detection, and implemented a crack measurement approach based on skeletonization and orthogonal projection algorithms. This approach was tested on using data from three images obtained from the shear strength test on UHPC, which was conducted in a controlled laboratory environment. This was done to evaluate the proposed approach and fine tune it for future applications, including data collected from real-life structures, e.g., bridges, building, etc.

The results showed that the proposed crack detection method detected the crack and its boundary accurately. However, some of the crack details, with an average width of 0.76 mm, were lost. This was considered to be insignificant since those details do not have a major impact on the structural integrity compared to the larger cracks. Next, the results from skeletonization and branch filtering algorithms were tested. Both algorithms could

extract the crack skeleton with a thickness of 1-pixel, which is essential for the following measurement procedures. Lastly, the accuracy of the proposed crack measurement method was evaluated with manually measured ground truth values from the original images. The proposed crack measurement method was found to be promising as it measured the crack length with 92.10% accuracy and the crack width with 90.11% accuracy.

To summarize the results, the proposed methodology for crack detection and measurement performs well and is capable of achieving the level of accuracy reported in previous studies. However, since the ground truths used in this study was obtained through manual annotations in the digital image, it is recommended that more precise evaluation based on real crack dimensions (measured physically using a tape) should be performed in future studies.

The study presented in this paper represents an important step in an overall research framework that aims to utilize unmanned aircraft system (UAS) images to automatically identify and measure surface defects of large structures such as bridges and map that information to Building Information Models (BIM), so that inspection reports are integrated with structural drawings and models in a single database. Accordingly, after fine-tuning the crack detection and measurement methods proposed in this paper, they will be evaluated on a set of images acquired from a bridge using an UAS and the results will be mapped to the corresponding BIM.

References

- [1] American Society of Civil Engineers (ASCE). *2021 Report card for America's infrastructure*. ASCE, Reston, VA, United States, 2021.
- [2] Chen, X., Li, J., Huang, S., Cui, H., Liu, P. and Sun, Q. An Automatic Concrete Crack-Detection Method Fusing Point Clouds and Images Based on Improved Otsu's Algorithm. *Sensors*, 21(5):1581, 2021.
- [3] Yan, Y., Mao, Z., Wu, J., Padir, T. and Hajjar, J. F. Towards automated detection and quantification of concrete cracks using integrated images and lidar data from unmanned aerial vehicles. *Structural Control Health Monitoring*, 28(8): e2757, 2021.
- [4] Kanopoulos, N., Vasanthavada, N. and Baker, R.L. Design of an image edge detection filter using the Sobel operator, *IEEE Journal of Solid State Circuits*, 23(2):358-367, 2002.
- [5] Wang, D. and Zhou, S. Color Image Recognition Method Based on the Prewitt Operator. In *Proceedings of the International Conference on Computer Science & Software Engineering*, pages 170-173, Wuhan, China, 2008.
- [6] Li, E.-S., Zhu, S.-L., Zhu, B.-S., Zhao, Y., Xia, C.-

- G. and Song, L.-H. An Adaptive Edge-detection Method Based on the Canny Operator. In *Proceedings of the International Conference on Environmental Science & Information Application Technology*, pages 465-469, Wuhan, China, 2009.
- [7] Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62-66, 1979.
- [8] Nguyen, T. S., Begot, S., Duculty, F. and Avila, M. Free-form anisotropy: A new method for crack detection on pavement surface images. In *Proceedings of the IEEE International Conference on Image Processing*, pages 1069-1072, Brussels, Belgium, 2011.
- [9] Xu, Y. and Turkan, Y. BrIM and UAS for bridge inspections and management. *Engineering, Construction and Architectural Management*, 27(3):785-807, 2020.
- [10] Krizhevsky, A., Sutskever, I. and Hinton, G.E. ImageNet classification with deep convolutional neural networks. In *Advanced in Neural Information Processing Systems 25 (NIPS 2012)*, Stateline, Nevada, United States, 2012.
- [11] Simonyan, K. and Zisserman, A. Very deep convolutional networks for large-scale image recognition. In *Proceedings of the International Conference on Learning Representations (ICLR 2015)*, San Diego, California, United States, 2015.
- [12] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1-9, Boston, Massachusetts, United States, 2015.
- [13] He, K., Zhang, X., Ren, S. and Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770-778, Las Vegas, Nevada, United States, 2016.
- [14] Liu, Y., Yao, J., Lu, X., Xie, R. and Li, L. DeepCrack: A deep hierarchical feature learning architecture for crack segmentation. *Neurocomputing*, 338:139-153, 2019.
- [15] Yang, X., Li, H., Yu, Y., Luo, X., Huang, T. and Yang, X. Automatic Pixel-Level Crack Detection and Measurement Using Fully Convolutional Network. *Comput. Aided Civ. Infrastruct. Eng.*, 33:1090-1109, 2018.
- [16] Gopalakrishnan, K., Khaitan, S.K., Choudhary, A. and Agrawal, A. Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection. *Constr. Build. Mater.*, 157:322-330, 2017.
- [17] Cha, Y.-J., Choi, W. and Buyukozturk, O. Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. *Comput. Aided Civ. Infrastruct. Eng.*, 32:361-378, 2017.
- [18] Lee, J., Kim, H.-S., Kim, N., Ryu, E.-M. and Kang, J.-W. Learning to Detect Cracks on Damaged Concrete Surfaces Using Two-Branched Convolutional Neural Network. *Sensors*, 19:4796, 2019.
- [19] Islam, M.M.M. and Kim, J.-M. Vision-Based Autonomous Crack Detection of Concrete Structures Using a Fully Convolutional Encoder—Decoder Network. *Sensors*, 19:4251, 2019.
- [20] Bhowmick, S., Nagarajaiah, S. and Veeraraghavan, A. Vision and Deep Learning-Based Algorithms to Detect and Quantify Cracks on Concrete Surfaces from UAV Videos. *Sensors*, 20:6229, 2020.
- [21] Dorafshan, S., Thomas, R.J. and Maguire, M. Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. *Construction and Building Materials*, 186:1031-1045, 2018.
- [22] Maguire, M., Dorafshan, S. and Thomas, R. SDNET2018: A Concrete Crack Image Dataset for Machine Learning Applications, Utah State University, Logan, 2018.
- [23] Wang, W., Hu, W., Wang, W., Xu, X., Wang, M., Shi, Y., Qiu, S. and Tutumluer, E. Automated crack severity level detection and classification for ballastless track slab using deep convolutional neural network. *Automation in Construction*, 124:103484, 2021.
- [24] Qiu, S., Wang, W., Wang, S. and Wang, K.C.P. Methodology for accurate AASHTO PP67-10-based cracking quantification using 1-mm 3D pavement images. *Journal of Computing in Civil Engineering*, 31(2):04016056, 2017.
- [25] Kalfarisi, R., Wu, Z.Y. and Soh, K. Crack Detection and Segmentation Using Deep Learning with 3D Reality Mesh Model for Quantitative Assessment and Integrated Visualization. *Journal of Computing in Civil Engineering*, 34(3):04020010, 2020.
- [26] Voytko, D., Calvi, P.M. and Stanton, J. Shear strength of ultra-high-performance concrete. *Engineering Structures*, 255:113961, 2022.
- [27] Calvi, P.M., Proestos, G.T. and Ruggiero, D.M. Toward the development of direct crack-based assessment of structures. *ACI Structural Journal*, 328:9.1-9.20, 2018.

Innovative Model for Forecasting Trailer Usage for Prefabricated Exterior Wall Panels

A. Rener ^{a,b}, A. Karatas ^c, M. Cole ^d

^aDepartment Civil and Architectural Engineering, Lawrence Technological University, US

^b Centerline Prefab, LLC, US

^cDepartment of Civil, Materials and Environmental Engineering, University of Illinois at Chicago, US

^d College of Business and Information Technology, Lawrence Technological University, US

E-mail: arener@centerlineprefab.com, akaratas@uic.edu, mcole@ltu.edu

Abstract –

With growing demand for large scale exterior envelope prefabrication solutions beyond precast concrete, Engineer-to-Order (ETO) prefabrication firms must develop reliable methodologies to manage finish goods inventory of large components. The imbalance between fabrication time and installation requires ETOs to forecast likely consumption of transportation resources at the proposal stage with only a conceptual understanding of the final project. Since 2019 US domestic trucking costs have increased due to demand. Therefore, ETOs need a reliable forecasting trailer usage model for estimating transportation costs at the proposal phase and mitigate the risks of an arbitrary approach. The proposed model was developed and evaluated by using a supervised machine learning algorithm on a large data set collected from completed exterior prefabricated panel projects in the US. The model was then tested and compared to estimations completed by a professional prefabrication estimator. The model can assist ETOs with projecting the quantity of trailers likely necessary for a prefabricated panel project with less variance at the proposal phase with limited information. The increased accuracy can reduce the financial exposure of the ETOs.

Keywords –

Prefabrication; Offsite Construction; Exterior Wall Panels; Logistics; Estimating; Supply Chain

1 Introduction

Industrialized Construction (IC) and Modern Methods of Construction (MMC) are evolving applications of manufacturing methodology and lean practices to improve the productivity and final project outcomes in the construction industry. This is achieved through the decoupling of sub-assembly operations from

the traditional construction site and fabricating these building components at facilities located off-site [1]. An accepted overarching term utilized throughout the construction industry is *prefabrication*. Adoption of prefabrication and offsite construction continues to increase to address challenges with schedules, quality, sustainability and manpower [2]–[4]. To meet this demand, most prefabrication companies deliver their products in an Engineer-to-Order (ETO) model [5], [6].

Prefabrication, as a concept, is not a new idea. The United States government utilized prefabrication techniques and planning during the Manhattan Project to rapidly construct communities that now are part of Oak Ridge National Labs [7]. There was a period in the 20th Century where companies like Sears and Roebuck sold “kitted” residential homes that were a form of prefabrication. However, the complexity of systems and components of buildings that are now being utilized are much more intricate for onsite prefabrication and offsite prefabrication. Onsite prefabrication can include constructing subassemblies while on the jobsite prior to the permanent installation. Offsite prefabrication is completed in a production facility focused on a specific building element, and requires consideration of new challenges in transportation, supply chain and logistics. Therefore, ETO’s must examine other products and solutions that have evolved to address these challenges in the supply chain.

This study specifically focuses on examining prefabricated exterior systems, such as the panel systems fabricated by the ETO companies, that are multi-layered high-energy performance cladding systems that layer various building materials to create the panels [1]. For decades, precast concrete has been available to building owners as an exterior cladding option for their facilities. Unitized glazing systems have offered building owners an exterior system that can be preassembled on or offsite and installed as large units on the façade, improving overall productivity.

Whether its precast, unitized glazing, or prefabricated exterior panels, these units are large and take up substantial space while being staged at either an offsite facility or onsite [8]. Depending on the jobsite or the prefabrication facility location, staging areas are often limited. Some jobsites may constrain the subcontracting companies to follow a Just-In-Time (JIT) delivery approach because there are insufficient staging areas within the confines of the site [9].

ETO companies develop project specific solutions to meet the needs of a given project. Often, this work is procured at the schematic or design development stage of the design process. As such, granular details about the panel products necessary to complete the project are not yet designed. However, ETO's regularly must provide a fixed price or Guaranteed Maximum Price to their clients that have to be cost competitive to warrant the inclusion of their products in the project. This requires preconstruction and estimating teams of the ETO to utilize their experiential knowledge to develop their proposal. At the point of proposal submission, the ETO generally has information regarding the orientation of the panels, the square footage of the project, a preliminary panel count, framing style and, aesthetic finishes. The experiential knowledge of the estimators is subjective and limited to their personal experience or the institutional knowledge of their organization. One area that shows large fluctuations in accuracy in the estimate is the estimation of the number of trailers necessary to transport the finished panels from the manufacturing facility to the jobsite. Since the onset of the COVID pandemic, over-the-road trucking costs have increased four to five-fold compared to 2019 for subcontracted over-the-road hauling services in the United States. This has magnified the financial impact of errors in the forecasting of trailers at the proposal stage. In the event the estimator errors too low, the costs quickly compound and erode profitability for the ETO. Conversely, overly conservative estimating of trailers may cause the ETO's proposal to be viewed as too expensive. To address this challenge, ETO companies must assess multiple strategies for addressing management of their finished goods inventories to align both installer and customer demands.

This study presents an automated forecasting model that ETO fabricators can utilize at the estimate stage of a project to forecast the quantity of over-the-road trailers necessary to transport a prefabricated panel project from its place of fabrication to the jobsite for final installation on the building. By drawing on a large data set compiled from completed projects of twelve ETOs the model can forecast likely trailer counts with less variance than a professional estimator. ETOs are generally limited to their own experiential knowledge from their data set of projects.

The paper is organized to present the objective, followed by a literature review, then an overview of the model creation including data collection and validation. Conclusions drawn from the study, limitations of the current model and future research complete the paper.

1.1 Objective

The objective of this study is to develop an innovative automated model that is capable of forecasting trailer usage for prefabricated exterior wall panels with greater accuracy than current methods. The model utilizes supervised machine learning algorithms on a large data set collected from completed projects from multiple companies in the US. The model can be used by ETO fabricators at the estimate stage of a project to forecast trailers necessary to transport a project from its origin place of fabrication to the final destination (i.e., jobsite). The development of the tool can reduce the potential for financial risk associated with poorly predicting the number of trailers necessary for a project.

2 Literature Review

Studies have identified multiple factors that contribute to the complexity of the supply chain challenges in the construction industry [10]. The research determined there were four overarching categories that the challenges could be organized within. Those are: material flow, company communication, project communication, and complexity. The challenges identified by their respondents are wide ranging and require substantial management and planning effort to overcome successfully.

Communication between the project site and ETO companies can be challenging relative to demand needs. This creates a disruption in the production process. Panova and Hilletoft [11] used dynamic modelling to attempt to manage construction supply chain risks caused by delays. Their research recommends that suppliers implement safety stocks as a method of minimizing disruptions. Utilizing buffer space for safety stock also minimized disruption of the production sequence and onsite activities, ultimately reducing the bullwhip effect [12], [13]. While the managerial approach of creating safety stock to address the fluctuating demand prevents the potential site disruption caused by delayed deliveries may appear to address the problem, it creates a secondary problem of storage for large construction components and assemblies at the factory or intermediate staging site.

The importance of thorough planning in construction supply chain is a critical step increasing the likelihood of a successful project and it becomes even more critical when utilizing ETO prefabricated components [6], [10]. There is positive impact on the project costs and nonvalue add process of Zero Inventory compared to the benefits to the project utilizing a Smart

Manufacturing Zero-Warehousing approach that relies on communication and feedback between the ETO companies and construction site [8]. The importance of communication between the onsite installation and the ETO company is critical to the successful outcome of the project. Inter-organizational coordination, cooperation and learning to form an overall project team focused on executing a successful project versus multiple independent teams can be achieved [14].

Demand variability on construction sites for construction materials, such as precast products, can be an impediment to success due the demand of on-time delivery to keep the project on schedule [15]. The importance of thorough planning in construction supply chain is a critical step to increase the likelihood of a successful project and it becomes even more critical when supply is constricted [10]. A research study was conducted to attempt to optimize transportation costs and the quantity of trucks [16]. Although this research has some fundamental application, it considered weight and volume of prefabricated components as part of the characteristics of the products. However, there were no multi-dimensional attributes considered. It did not consider specific important transportation aspects inherent in building façade panels, such as precast, and the likelihood for oversized loads. This research also did not account for large variability in sizes that may result from custom nontypical building products.

On time deliveries of products to the jobsite are critical to maintain a project's flow. One approach to mitigate late deliveries caused by manufacturing challenges can be addressed through the managed incorporation of a buffer or safety stock. Implementation of a safety stock either at a permanent or temporary location can help to minimize disruptions [11]. Further research through surveys have examined preferences for buffer stock levels to mitigate disruptions [9]. Therefore, it is important to consider the costs associated with large buffer stocks which can become costly if there is not a contractual vehicle for billing for that material in a timely manner.

For full realization of the schedule benefits of prefabrication, the prefabricated components must arrive at the project site when the schedule demands. Late or early deliveries of prefabricated assemblies can cause disruption to the project site and can result in double handling [12]. The expectation of the prime contractor is the entity responsible for transportation plans for storage space to deal with slow site installation or bad weather.

Through a survey conducted of ETO exterior panel fabricators it was found that all the respondents store their finished goods inventory of panels on trailers prior to shipment to the jobsite [17]. Nearly 50% of the respondents reported slowing or halting production due to issues with storing finished good inventory. To

address these aforementioned research gaps, this study focuses on the utilization of completed projects to develop a model capable of forecasting trailer resources for projects at the proposal stage.

3 Model Development

The automated forecasting model was developed in two steps. Data were collected on randomly selected prefabricated panel projects. In order to be included in the data set, the projects had to be completed so that characteristics of the project were actual and not estimated or theoretical. Data were analyzed, and validated using a supervised machine learning algorithm to assess the practical functionality of the model in forecasting trailer resources.

3.1 Project Data Collection

Data were collected from ETO prefabricators specializing in exterior wall panels for 107 completed projects that included characteristics of the panels along with the number of trailers utilized to transport the finished goods to the jobsites. This task needed an exhausting effort and meticulous organization to reach out to all ETO prefabricators in the US and create a dataset on the completed projects.

The project specific data spanned over a four-year period and were collected from multiple companies servicing different regions of US. It is noteworthy to mention that this research did not focus on transportation distance because it does not affect actual trailer trips to and from the factory and jobsite.

Data set created for utilization in this study included both macro and specific characteristics of the exterior panels fabricated for the project. Macro level characteristics included total square footage of panels and number of panels built for the project. Additionally, the respondents were followed up to gather more detailed information relative to the panels' structural configuration as part of the building skin as either a By-Pass, Infill or Load-Bearing configuration. Specific information on the panels fabricated for a project included panel finishes. The following variables were included:

- Cornice/Parapet Element
- Frame only
- Back-up
- EIFS
- Metal
- Thin Brick (with cast bed)
- Thin Brick (over foam)
- Fiber Cement Siding
- Aluminum Composite Panel (ACM)
- Acrylic Panel

- Other

The cornice/parapet element refers to a thickened architectural detail at the top of the panel. This has potential significance because it would create differing panel thickness that must be accounted for in the loading of panels. Frame only panels are comprised of only cold-formed metal studs and no other materials. This would make them, potentially, the thinnest cross-section of any of the panel types. Back-up panels are comprised of studs, sheathing, and an applied air/vapor barrier. Exterior Insulated Finishing System (EIFS) panels have foam and an applied finish beyond that of a back-up panel. Metal panel finishes could include insulated metal panel, corrugated metal siding or machine bent formed panel assemblies. Thin brick over cast bed is a multilayered approach applied to the back-up panel that includes a slip sheet, lath/mesh and a cementitious cast bed applied to the panel prior the installation of commercially available thin brick materials. In contrast to the thin brick over cast bed, the thin brick over foam allows for the adhesion of thin brick materials to the cementitious base coat applied to foam that is adhered to the back-up panel. Fiber cement siding, along with attendant insulation, is applied to the back-up panel assembly per the manufacturer's instructions. This is often constructed in a rain-screen configuration. The aluminum composite panel is often fabricated out of sheets and machined into desired geometries prior to being installed on the back up panel with the attendant carrier system. The "Other" finish allowed for the respondents to provide data on projects that did not have any of the finishes provided as options. Respondents also provided the actual number of trailers that were utilized to transport the finished panels from their fabrication facility to the project site.

3.2 Data Analysis and Validation

Dataset were analyzed using a supervised machine learning algorithm involving multiple data variables for analysis. The data was analyzed in Minitab V20.4 utilizing multivariate regression of the panel characteristics (predictors) regressed on the number of trailers (response) The results were evaluated along with the residuals to determine the reliability of the selected machine learning algorithm. To assist with evaluation and description of the resulting equation, the number of trailers was transformed to the log natural and the multivariate regression was then performed again [20].

To estimate the validity of the automated forecasting model, derived from the set of 107 completed projects, data from actual projects not in the data set were input into the equation and the forecasted trailer resources were then compared to the actual number of trailers utilized for

that project. Furthermore, the developed model was also tested on a professional estimator. The estimator had no prior knowledge of the actual projects and was engaged to forecast the number of trailers necessary on the four sample projects utilizing their experiential knowledge. The variances from actual were then evaluated for practicality of the developed model.

4 Results and Discussion

4.1 Numerical Results

4.1.1 Data Collection

Data from 107 randomly selected completed projects was gathered and analyzed as part of this study. Table 1 summarizes the total data set by the 14 predictor variables (x) and the outcome variable (y).

It is noteworthy to mention that there were no respondents that provided data for acrylic panel finishes so that predictor was not considered in the evaluation as it would add no value. The square footage and number of panels predictors are continuous variables; the remaining predictors are binary categorical variables.

Table 1. Summarized Data Set of Panel Projects

Variable	Total	% of Total
Square Footage (x₁)	5,020,976	100
Number of Panels (x₂)	43,722	100
Bypass (x₃)	71	66
Infill (x₄)	12	11
Load-Bearing (x₅)	26	24
Cornice/Parapet (x₆)	9	8
Frame Only (x₇)	26	24
Back-up (x₈)	22	21
EIFS (x₉)	44	41
Metal (x₉)	4	4
Thin Brick (cast bed) (x₁₀)	3	3
Thin Brick (over foam) (x₁₁)	3	3
Fiber Cement Siding (x₁₂)	2	2
Aluminum Composite (x₁₃)	1	1
Other (x₁₄)	6	6
Number of Trailers (y)	2,559	100

4.1.2 Data Analysis

To analyze the collected data, the forecasting model was developed by using a supervised machine learning algorithm. The panel characteristics (predictors presented in Table 1) were regressed on the number of trailers (outcome). Equation 1 presents results from the analysis, that represents the unstandardized regression coefficients ($R^2 = 79\%$):

$$\begin{aligned} Trailors = & -6.23 + 0.000411x_1 - 0.01204x_2 + \\ & 17.2x_3 + 18.11x_4 + 11.58x_5 + 0.4x_6 - 10.06x_7 - \\ & 5.45x_8 - 0.69x_9 + 25.73x_{10} - 1.09x_{11} - 2x_{12} + \\ & 44.95x_{13} - 1.08x_{14} \end{aligned} \quad (1)$$

Examination of residuals suggests a few data points are outliers. Figure 1 presents the Normal Probability Plot of the residuals and Figure 2 shows the Versus Order Plot of Observation Order and Residuals.

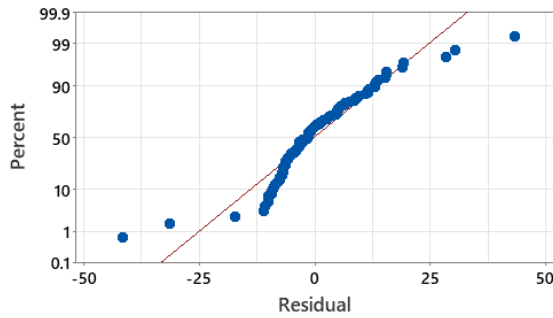


Figure 1. Normal probability plot of residuals versus percentage

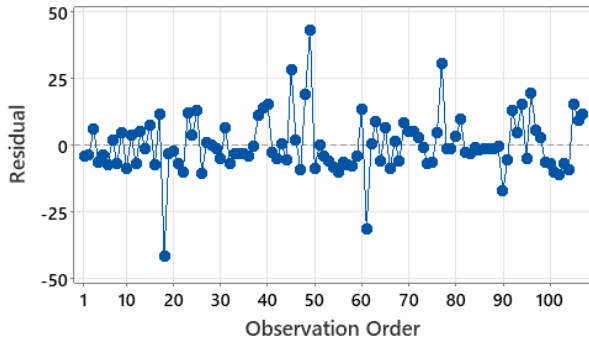


Figure 2. Versus order plot of observation order versus residuals

To help interpret the unstandardized regression coefficients from Equation 1 in terms of the estimated percent change in the number of trailers as a function of a one unit change of the predictors, trailers was transformed to its natural log (LN transformation) and the regression equation was recalculated. As an example, for every 1,000 square feet of panel (x_1) there is an estimated 1.3% increase in the number of trailers needed (see Equation 2):

$$\begin{aligned} LN Trailers = & 1.158 + 0.000013x_1 - 0.000118x_2 + \\ & 1.191x_3 + 0.835x_4 + 0.695x_5 + 0.066x_6 - \\ & 0.688x_7 - 0.233x_8 + 0.019x_9 + 0.775x_{10} + \\ & 0.066x_{11} + 0.104x_{12} + 1.358x_{13} - 0.133x_{14} \end{aligned} \quad (2)$$

To evaluate the functionality and accuracy of the applied machine learning algorithm, four randomly

selected projects were evaluated that were not part of the original data set. The characteristics for the model were gathered for each of the projects. Utilizing Equation 1, predictions were made relative to the number of trailers utilized. The predictors for the four sample projects are presented in Table 2. The square footage and panel count are presented as integer numbers. The balance of the predictors is presented as binary number because they either are part of the panels of the selected project or are not.

Utilizing Equation 1, project predictors were regressed on the number of trailers. The results of the application of the model equation are presented in Table 3 along with the actual number of trailers utilized in the sample projects.

As a practical matter, there are no partial trailers so the resultant prediction can either be rounded up for conservative purposes or the practitioner can choose to round down the predicted number and manage their efforts to meet that goal. Results across the four projects found the regression model produced reasonable results for three of the four sample projects. In the case of project 2, the results exceeded the actuals by approximately one-third.

Table 2. Sample Projects with Attendant Predictors

Predictor	#1	#2	#3	#4
Square Footage	10,942	11,507	40,658	92,500
Number of Panels	70	33	109	287
Bypass	1	1	1	1
Infill	0	0	0	0
Load-Bearing	0	0	0	0
Cornice/Parapet	0	0	1	1
Frame Only	0	0	0	0
Back-up	1	1	1	0
EIFS	0	0	1	0
Metal	0	0	0	1
T Brk (cast bed)	0	0	0	0
T Bck (ovr foam)	0	0	0	0
Fiber Cmt Siding	0	0	0	0
ACM	0	0	0	0
Other	0	1	0	0

To assess the practicality of the model an estimator in the prefabrication space was interviewed. The estimator stated that the challenge of trucking costs of finished panels has caused financial challenges on some projects due to the cost impact related to the estimated trailer quantities and the actuals necessary to complete the project. The estimator has started to utilize a rate of \$2,000USD per trailer load compared to \$500USD per trailer load 3 years ago. The estimator was then asked to evaluate the 4 sample projects to predict the trailer usage necessary for those projects and was provided the same project characteristics as presented in Table 2. The results of the estimator's forecast along with potential cost impacts, utilizing \$2,000USD per trailer load cost, are

presented in Table 3. Negative cost variance values represent a loss or expense to the ETO and positive variances are savings compared to estimate.

The forecasts for both the model and the estimator compared to actual resulted in a likely loss on transportation costs for the ETO. However, the utilization of the model with the same data set as provided the estimator improved the accuracy of the estimate by nearly 400% in terms of dollars saved. The estimated value for each trailer load can vary by location, region, and available resources. Total haul distance (mileage) can also have an impact on per load cost. However, the magnitude of the variance is significant enough to be a potentially desirable solution for ETO's as a risk mitigation tool for cost overruns relative to trailer usage on project compared to arbitrary means.

Table 3. Comparison of Model and Estimator Predictions to Actual with Attendant Cost Variances

Project	Model Prediction	Estimator Prediction	Actual	Variance by Estimator	Variance by Model	Effective Cost Variance (Estimator)	Effective Cost Variance (Model)
1	9.17	4	10	60%	8%	-\$12,000	\$0
2	8.77	5	6	17%	32%	-\$2,000	\$6,000
3	20.62	16	18	11%	12.8%	\$4,000	\$6,000
4	71.66	30	96	69%	25.4%	-\$132,000	\$48,000
Total Variance =						-\$142,000	-\$36,000

5 Conclusion

Statistical analysis of 107 completed prefabricated panel projects was conducted to evaluate whether specific characteristics of the project can be utilized to forecast the number of trailers necessary to transport the finished panels to the jobsite. A predetermined but broad set of predictors were analyzed using a supervised machine learning algorithm to estimate the number of trailers. The resulting equation can be utilized to forecast the number of trailers required to ship panels from the ETO's facility to the project site. Project specific information of four additional projects not included in the original data set were utilized to validate the model by comparing actual number of trailers to the estimated quantities of trailers determined by a professional estimator in the prefabrication industry. The comparison shows that the model can provide forecasts of necessary trailers with less variance to actual compared to an experienced estimator. ETO's are incentivized to utilize a data driven approach to forecasting compared to historical arbitrary approaches due to the high costs of transporting the finished panels and the potential for adverse financial outcomes.

Dimensional data of the panels utilized in the data set was not solicited due to the custom nature of the solutions provided by ETOs and the variability in modern architectural aesthetics coupled with structural systems of buildings to meet specific project requirements. Panel dimensions and packing methodologies of each ETO can vary resulting in more or less panels being loaded onto a given trailer for transport. By way of example some ETOs may prefer nominal 4" dimensional lumber compared to others that may use 2" high density foam for dunnage. Variation in different states over-the-road load restrictions may also affect the number of trailers necessary to complete a project.

No consideration was given to the availability of trailers for an ETO as a potential constraint. It is presumed that acquisition of the necessary quantity of trailers is possible.

5.1 Limitations

Results of this study may be limited by size and validity of the sample of 107 projects submitted by ETO's for the purposes of this study. For example, residuals plots suggest some data points were outliers, and some predictors only occurred a few times (e.g., metal, thin brick, fiber cement siding, aluminum composite). Additionally, the regression analysis did not account for projects that have multiple finish characteristics such as EIFS, Metal or Fiber Cement to achieve the architectural aesthetic. Users should separate dissimilar finishes and utilize the supervised machine learning algorithm as if there was a separate project, with

correlating square footage, for each variation in finish type.

5.2 Future Research

Ongoing collection of completed project data from the original study participants will be utilized to further refine and optimize the forecasting model as a user-friendly tool to help practitioners in the panel prefabrication space use data-driven forecasting of trailer resources. Future research will also create a simple user interface to allow the practitioner to quickly input the know predictors and receive a response from the model.

5.3 Acknowledgements

The authors would like to thank the ETO companies who provided data for this project and the anonymous estimator who was willing to participate in this study. Any opinions, findings or conclusions of this paper are that of the authors and do not necessarily represent those of the companies participating in the research.

References

- [1] G. Correia Lopes, R. Vicente, M. Azenha, and T. M. Ferreira, "A systematic review of Prefabricated Enclosure Wall Panel Systems: Focus on technology driven for performance requirements," *Sustainable Cities and Society*, vol. 40, pp. 688–703, Jul. 2018, doi: 10.1016/j.scs.2017.12.027.
- [2] H. Khan and K. Jain, "Study on the Trends & Usage of Prefabrication and Modularization: Increasing Productivity in the Construction Industry," p. 10, 2017.
- [3] S. Das, "Global Research and Analytics Firm," Dec. 18, 2020. <https://www.aranca.com/knowledge-library/articles/business-research/construction-technology-trend-prefabrication> (accessed Oct. 04, 2021).
- [4] K. Davis, "Prefabrication + the New Frontier," *Construction Superintendent*, Jan. 22, 2020. Accessed: Oct. 04, 2021. [Online]. Available: <https://consupt.com/2020/01/prefabrication-the-new-frontier/>
- [5] D. T. Matt, P. Dallasega, and E. Rauch, "Synchronization of the Manufacturing Process and On-site Installation in ETO Companies," *Procedia CIRP*, vol. 17, pp. 457–462, Jan. 2014, doi: 10.1016/j.procir.2014.01.058.
- [6] E. Rauch, P. Dallasega, and D. T. Matt, "Complexity reduction in engineer-to-order industry through real-time capable production planning and control," *Prod. Eng. Res. Devel.*, vol. 12, no. 3, pp. 341–352, Jun. 2018, doi: 10.1007/s11740-018-0809-0.
- [7] B. Feijoo, R. Interiano, and D. Fisler, "The Secret City that Paved the Way for Modern Prefabrication," *ADL Ventures*, Nov. 18, 2020. <https://adlventures.com/the-secret-city-that-paved-the-way-for-modern-prefabrication/> (accessed Nov. 10, 2021).
- [8] Z. Lyu, P. Lin, D. Guo, and G. Q. Huang, "Towards Zero-Warehousing Smart Manufacturing from Zero-Inventory Just-In-Time production," *Robotics and Computer-Integrated Manufacturing*, vol. 64, p. 101932, Aug. 2020, doi: 10.1016/j.rcim.2020.101932.
- [9] L. Sui Pheng and C. Joo Chuan, "Just-in-time management in precast concrete construction: a survey of the readiness of main contractors in Singapore," *Integrated Manufacturing Systems*, vol. 12, no. 6, pp. 416–429, Jan. 2001, doi: 10.1108/EUM00000000006107.
- [10] M. Thunberg and A. Fredriksson, "Bringing planning back into the picture - How can supply chain planning aid in dealing with supply chain-related problems in construction?," *Construction Management & Economics*, vol. 36, no. 8, pp. 425–442, Aug. 2018, doi: 10.1080/01446193.2017.1394579.
- [11] Y. Panova and P. Hilletoft, "Managing supply chain risks and delays in construction project," *Industrial Management & Data Systems*, vol. 118, no. 7, pp. 1413–1431, 2018, doi: <http://dx.doi.org.ezproxy.ltu.edu:8080/10.1108/I-MDS-09-2017-0422>.
- [12] Y. Zhai, R. Y. Zhong, and G. Q. Huang, "Buffer space hedging and coordination in prefabricated construction supply chain management," *International Journal of Production Economics*, vol. 200, pp. 192–206, 2018, doi: <https://doi.org/10.1016/j.ijpe.2018.03.014>.
- [13] H. L. Lee, "Taming the bullwhip," *Journal of Supply Chain Management*, vol. 46, no. 1, pp. 7–8, Jan. 2010.
- [14] P. E. D. Love, Z. Irani, and D. J. Edwards, "A seamless supply chain management model for construction," *Supply Chain Management: An International Journal*, vol. 9, no. 1, pp. 43–56, Jan. 2004, doi: 10.1108/13598540410517575.
- [15] C.-H. Ko, "Production control in precast fabrication: considering demand variability in production schedules," *Canadian Journal of Civil Engineering*, vol. 38, no. 2, pp. 191–199, Feb. 2011, doi: 10.1139/L10-123.
- [16] W. Yi, S. Wang, and A. Zhang, "Optimal transportation planning for prefabricated products in construction," *Computer-Aided Civil &*

- Infrastructure Engineering*, vol. 35, no. 4, pp. 342–353, Apr. 2020, doi: 10.1111/mice.12504.
- [17] A. T. Rener, “Examination of Supply Chain Challenges in Large Scale Exterior Building Panels from the Perspective of an ETO Fabricator,” Harare, Zimbabwe, 2020, p. 13.
 - [18] A. Sahay, *Applied Regression and Modeling: A Computer Integrated Approach*. New York, UNITED STATES: Business Expert Press, 2016. Accessed: Jan. 17, 2022. [Online]. Available: <http://ebookcentral.proquest.com/lib/lawrencetu-ebooks/detail.action?docID=4560113>
 - [19] R. Silhavy, P. Silhavy, and Z. Prokopova, “Analysis and selection of a regression model for the Use Case Points method using a stepwise approach,” *Journal of Systems and Software*, vol. 125, pp. 1–14, Mar. 2017, doi: 10.1016/j.jss.2016.11.029.
 - [20] T. J. Cole, “Sympercents: symmetric percentage differences on the 100 loge scale simplify the presentation of log transformed data,” *Statistics in Medicine*, vol. 19, no. 22, pp. 3109–3125, 2000, doi: 10.1002/1097-0258(20001130)19:22<3109::AID-SIM558>3.0.CO;2-F.

Automated Wall Detection in 2D CAD Drawings to Create Digital 3D Models

C. Wei^a, M. Gupta^a and T. Czerniawski^a

^aSchool of Sustainable Engineering and the Built Environment, Arizona State University, USA

E-mail: cwei32@asu.edu, mgupta70@asu.edu, Thomas.Czerniawski@asu.edu

Abstract –

Generating digital 3D buildings models from scratch is time consuming and labor intensive. In this paper, we present an automated detection process leveraging computer vision and the information available in 2D drawings to reduce 3D modeling time. The recognition system is limited to walls and has two parts: (1) Image classification on walls by ResNet-50 model, (2) Object Detection on walls by YOLOv3 model. The system accepts new 2D drawings and outputs parameters of recognized walls. The parameters are input into Dynamo for 3D model reconstruction. We anticipate these types of systems, which rely on 2D drawings as recognition priors, will be pivotal to the industry's transition from 2D to 3D information modalities.

Keywords –

Deep Learning; Image Classification; Object Detection; 2D Drawings; 3D Building Models

1 Introduction

Building Information Modeling (BIM) plays an important role during the entire life cycle of a building. First, Cloud hosted BIMs enhance communication by ensuring all views into the model are synchronized. Secondly, achieving better visualization allows clients to have a good understanding at each stage reducing the possibility of design changes in the future. Thirdly, performing clash detection tasks on BIM models can be cost effective and improve safety. Lastly, environmental analysis and simulation on BIM models accomplish sustainable and AI-based architectural design [1].

For projects during preconstruction phase, it is quite beneficial if 3D modeling process is efficient since it helps push construction start date forward by enhancing the collaboration among entities early on. Other projects are 3D reconstruction for as-built buildings which can assist in reconstructing historical buildings and better visualization through digital representation. In detail, doing structural analysis on 3D models of historical

buildings can reach better maintenance. Moreover, taking this paper application as an example, constructing campus 3D representation can let online students have experience of looking around campus through virtual tours and even have interaction, such as reading posters on bulletin boards. To speed up the 3D modeling process with high accuracy, computer vision is a promising approach.

Computer vision is a prevailing implementation of artificial intelligence. It can achieve pattern recognition, object detection, image classification, and instance segmentation tasks on images, such as 2D drawings. This paper is experimenting with computer vision techniques in the automatic 3D model generation process to solve time-consuming manual modeling conditions. The scale and distribution of the dataset, quality of labeling and model choosing would all affect the model performance.

In our study, we proposed an object detection model to localize the vertical and horizontal walls on 2D drawings by computer vision technique. Some researchers were using semantic segmentation models to achieve this process [2, 3, 4, 8, 10], however, to make the process more efficient, we chose an object detection model which can reduce much labeling time and still achieve good performance. After we expand this work to multi-classes detection on 2D drawings, the performance metrics would include a confusion matrix and mean Average Precision (mAP).

2 Related Work

We focus on reviewing related work on how deep learning assists the process of 3D model reconstruction from 2D drawings, the benefits and applications of later usage.

2.1 Deep Learning Approach on 2D Drawings

There are lots of researchers doing recognition and segmentation of components on 2D drawings by deep learning approaches. Xiao et al. [2] cropped the original

2D drawings into smaller image dimensions for feeding into a neural network model. They manually did pixel-labeling on 300 2D drawings and implemented a transfer learning from ResNet-152. This model is pretrained on the ImageNet dataset and then executed recognition and localization on five architectural components: wall, window, door, column, stairs. Liu et al. [3] transformed rasterized images to vector-graphics representation since vector images can make further 3D reconstruction models having better visualization, being easier to manipulate and do computational analysis. They first did deep representation learning through Convolutional Neural Network (CNN) to convert a raster image to a set of junctions and pixel-wise semantics and then assembled junctions to lines and boxes by setting constraints through integer programming with a straightforward post-processing to achieve vectorization. Dodge et al. [4] utilized Fully convolutional networks (FCN) to do wall segmentation after trying different pixel strides. They conducted Faster R-CNN and Optical Character Recognition (OCR) to estimate room sizes. Mishra et al. [5] did furniture and architectural components detection on floorplans by Cascade Mask R-CNN network with CNN and Deformable Convolutional Networks (DCN) separately.

The recognition and segmentation tasks are not limited to architectural floor plans and can be more generalized to different types of buildings. Zhao et al. [6] detected structural components and grid reference by YOLO model. The structural elements include columns, horizontal beams, vertical beams and sloped beams. Kalervo et al. [7] presents a large-scale floor plan dataset, CubiCasa5K, which is carefully annotated by applying the Quality Assessment process and including 5000 samples with over 80 object classes.

In this paper, we adopt ResNet-50 and YOLOv3 in our methodology for wall recognition and segmentation tasks since both algorithms have their prevalence in pretrained models and can be easily deployed by industries' practitioners. Moreover, implementing the ResNet-50 model for classification tasks can be trained easily without increasing the training error with a large number of layers and this algorithm could help vanish gradient problems during the backpropagation process [8]. The YOLOv3 model for object detection has high speed and comprehends generalized object representation [9].

2.2 3D Model Generation and Application

Kalervo et al. [7] mentioned that the application of 3D models includes 3D real estate virtual tours and AR/VR technology. Jang et al. [8] created CityGML (City Geography Markup Language) and IndoorGML

(Indoor Geography Markup Language) 3D data models by extracting indoor spatial information from floorplan images. These 3D models support a structure for 3D geospatial data integration, storage and exchange. Kippers et al. [9] claimed that 3D models can support better decision making, data analysis and scenario simulation. They integrated CityJSON and floor plan images to reconstruct the 3D model. Some researchers indicated that CityJSON is easier to use than CityGML.

Seo et al. [10] stated the application of the 3D model reconstruction process at each stage. Architectural component recognition can contribute to evacuation paths generation and evacuation distances calculation. Analysis of building energy ratings with window area ratios could be executed by automatically calculating the window and wall areas. Moreover, Using Generative Adversarial Networks (GAN) can generate much more virtual drawing images to integrate into AI-based architectural design in the future.

For indoor furniture fitting, Dodge et al. [4] performed Optical Character Recognition (OCR) to extract text information from 2D drawings which can measure each room size and compute the pixel density for fitting interior components.

3 Methodology

This section presents data generation process, training, inference pipelines and 3D model reconstruction process in Revit.

3.1 Data Creation

In this study, Arizona State University Tempe campus 2D Drawings were assembled by the University Facilities Management department. We utilized 29 sheets of size 3400 x 2200 pixels as our original dataset.

We then used the LabelMe annotation tool [11] to manually label walls' location. We drew a rectangle as a bounding box surrounding each wall so that we can get the location of each wall which includes the coordinates of the upper-left point, width and height of each bounding box.

To feed into the CNN architecture, we randomly sampled each sheet with 100 crops of 256 x 256 pixels images, and computed the labeled information to the relative position for the small crop images. In the meantime, we programmed the small cropped images to two folders depending on whether or not the image contained the wall for further training.

Ultimately, we assembled all labeled information into a csv file and got 1250 non-wall images and 1650 wall-contained images individually.

3.2 Model Training

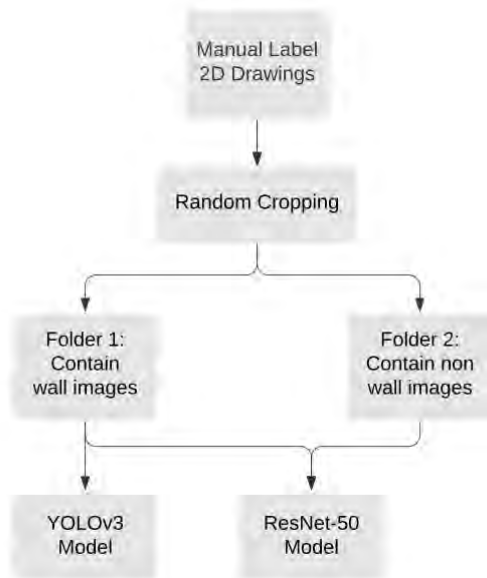


Figure 1. The training pipeline from labeling to Models' Generation

3.2.1 Wall Classification

The first step of doing model training is determining programming languages, deep learning libraries, programming environment and referred neural network models. We used Python with Tensorflow and Keras libraries on Google Colaboratory which is an online cloud-based Jupyter notebook environment. We also referred to the YOLOv3 model architecture from PyLessons [12] website which posts deep learning applications and tutorials for customized data. We chose their architecture and did further modification since the descriptions and explanations are clear, code is precise and the results they have shown are promising.

To classify wall objects, we utilized the ResNet-50 model to train a binary classification task. The training dataset was collected during the preprocessing step and a 70/30 ratio to split training and test dataset was determined. We utilized all training set as our input for the training pipeline shown in Figure 1, and test set for inference pipeline shown in Figure 3, separately.

We first programmed the ImageNet pretrained model for transfer learning. The setting of the pretrained model includes average pooling, 256 x 256 x 3 pixels input shape, 2 classes classification, and all layers inside the pretrained model are not trainable, which could reduce time complexity during the training process.

We flattened the output of the pretrained model into one dimension and added a fully connected layer with

512 nodes and a ReLU activation function. Finally, using a sigmoid function to determine the probability of wall existence.

During the training process, we set a 32 batch size, 100 epochs, and Adam optimizer with a 0.001 learning rate. The result shows that the training and test accuracy converged to 98% and 94%, respectively, as shown in Figure 2.

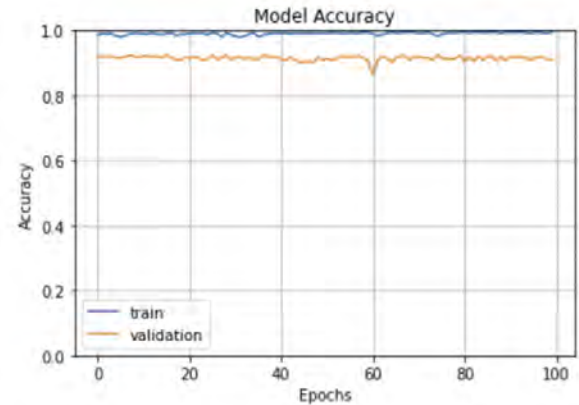


Figure 2. ResNet-50 classification training and test accuracy versus number of Epochs

3.2.2 Wall Detection

This section describes the wall detection task using the wall-contained images as our training set and YOLOv3 model for training. For this supervised learning training, we parsed the images and loaded the labeled information csv file with image filename and corresponding bounding box of wall information. We used an 80/20 ratio to split the training and test dataset, set the batch size as 4, Adam optimizer and 30 epochs in total. For learning rate, we used two warmup epochs first and then decayed linearly from 1e-4 to 1e-6.

3.3 Model Inference

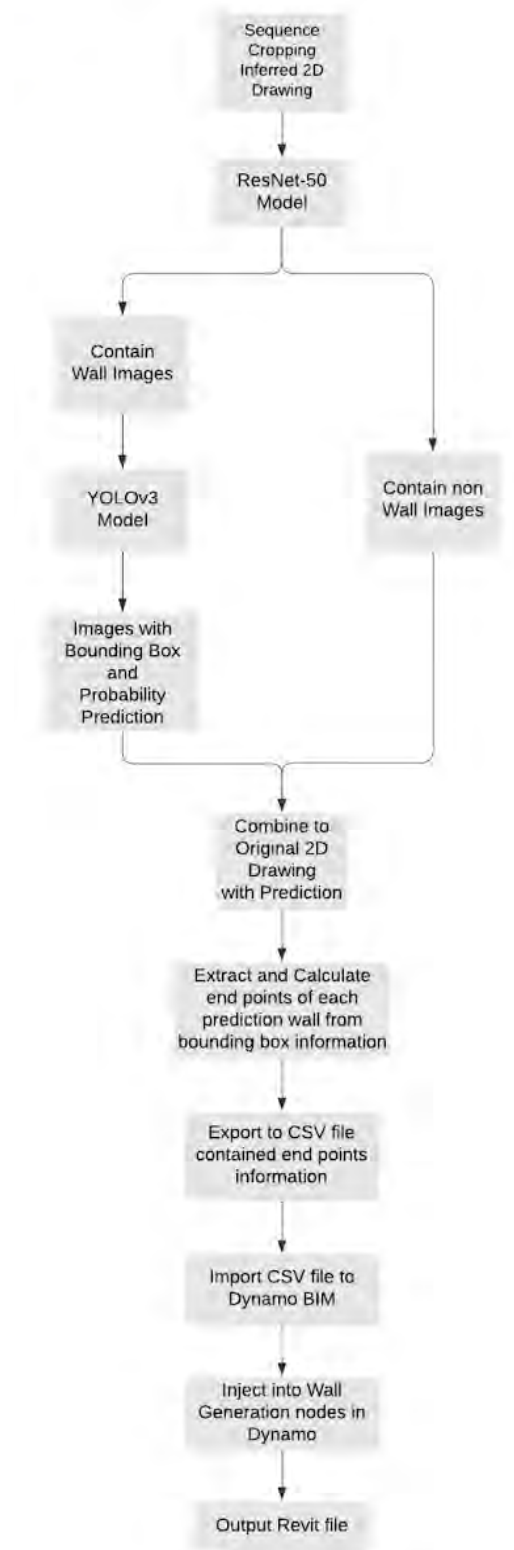


Figure 3. The inference pipeline for 3D wall reconstruction in Revit

To do model inference on new 2D drawings, we first cropped each drawing into 256 x 256 pixels size in a sliding windows sequence since we need to match the input size of deep learning algorithms that we set during the training process.

We feed each crop into our ResNet-50 model to classify if the crop contains walls or not. The inferred non-wall images stay the same since there are no walls to detect the location, however, we parsed the inferred wall images into our YOLOv3 model to infer bounding box information and probability of walls. The inferred rectangle bounding box surrounding walls and probability would be shown on each crop. Eventually, we can combine each crop to the original sheet size in a sequence after the inference process. The overall inference process is shown as Figure 3.

3.4 3D Model Reconstruction in Revit

After the model inference process described in section 3.3, we can get the prediction value of the wall location which represents the bounding boxes information. We extract and calculate the end points of each predicted wall location by their predicted bounding boxes. These parameters are saved in a csv file with height information on the elevation plan and then imported to Dynamo BIM. In the Dynamo environment, we inject the csv file into wall generation nodes and connect to the wall family type. Ultimately, we reconstruct the walls in Revit through Dynamo with our csv file and the family type we specify. The overall procedure is shown in Figure 4.

4 Results and Discussion

The results are shown in Figure 5 which are zooming a portion of the new 2D drawing. The red rectangles indicate the bounding box and the number on the upper-left corner above the rectangles is the neural network's confidence probability of the wall.

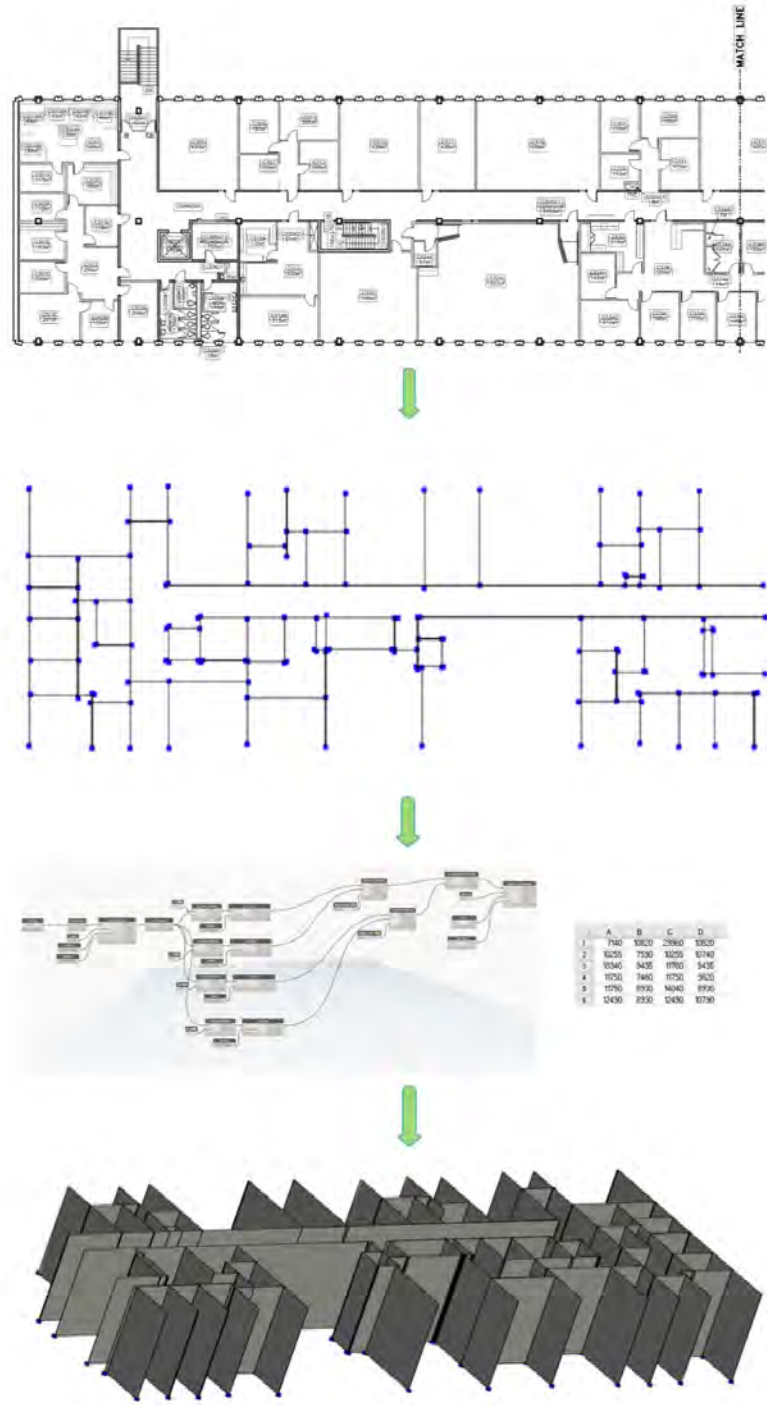


Figure 4. The procedure of creating a 3D model from extraction wall location through Dynamo and CSV file

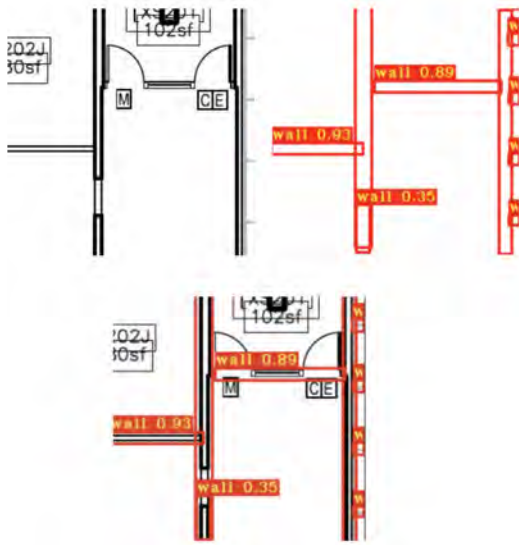


Figure 5. Original small portion of floorplan (upper-left), bounding boxes (upper-right) and floorplan with bounding boxes (lower)

We raised three failure cases on our inferred drawings as shown in Figure 6. First, the rectangle surrounding “MATCH LINE” on the drawing causes the CNN model to possibly be misinterpreted as a wall appearance. Secondly, we discovered that “curved walls” and “diagonal walls” are not suitable for object detection tasks during manual labeling process since the bounding boxes are axis aligned which could not tightly delineate the curved and diagonal lines. Therefore, we did not label the curved and diagonal walls during the training process so that our model could not recognize them on the inferred dataset.

Future work will include, first, we can remove legend, title and match line symbol before doing the

inference process so that we can avoid detecting the rectangle surrounding the match line symbol as walls. Secondly, we will make a list of all failed inference cases and explore the potential improvement. Fourthly, we would like to explore different methods to do “diagonal walls” and “curved walls” detection since the original labeled method would easily make a bounding box of walls covering other objects. The most promising method we would attempt first is an instance segmentation algorithm. Lastly, expanding the automation process to more architectural components, doors and windows, would be a critical contribution for this research and our strategy would be doing multiclass classification and detection, creating precise interface connection among all components.

5 Conclusion

BIM has become an essential methodology in the AEC industry in the past few decades. To facilitate BIM applications, we hope that we could make the BIM model generation effective and precise.

In the paper, we demonstrated that using computer vision techniques to detect the location of walls on the 2D drawings so that we could automatically generate 3D models later on. We achieve good results on the inferred dataset and plan to expand the scale to other components inference and experimental instance segmentation for non-straight walls.

Deep learning related research is very active and advanced, lots of machine learning models are being optimized and well-performed on many projects. Therefore, using those techniques on 2D drawings to 3D model generation could be of practical use in the AEC industry in the near future. These automated systems could reduce budget, improve safety, enhance communication among entities and promote more advanced implementation for researchers and expertise.



Figure 6. Failure cases: “MATCH LINE” (left), “diagonal wall” (middle) and “curved wall” (right)

References

- [1] Michael T. Benefits of Building Information Modeling. On-line: <https://www.ny-engineers.com/blog/benefits-of-building-information-modeling>, Accessed: 24/01/2022
- [2] Xiao Y., Chen S., Ikeda Y. and Hotta K. Automatic recognition and segmentation of architectural elements from 2D drawings by convolutional neural network. In *Proceedings of the 25th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA)*, pages 843–852, Hong Kong, 2020.
- [3] Liu C., Wu J., Kohli P. and Furukawa Y. Raster-to-vector: Revisiting floorplan transformation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2195–2203, Venice, Italy, 2017.
- [4] Dodge S., Xu J. and Stenger B. Parsing floor plan images. In *Proceedings of the 15th IAPR International Conference on Machine Vision Applications (MVA)*, pages 358–361, Nagoya, Japan, 2017.
- [5] Mishra S., Hashmi K., Pagani A., Liwicki M., Stricker D and Afzal M. Towards robust object detection in floor plan images: A data augmentation approach. *Applied Sciences*. 11(23):11174, 2021.
- [6] Zhao Y., Deng X. and Lai H. A Deep Learning-Based Method to Detect Components from Scanned Structural Drawings for Reconstructing 3D Models. *Applied Sciences*. 10(6):2066, 2020.
- [7] Kalervo A., Ylioinas J., Häikiö M., Karhu A. and Kannala J. CubiCasa5k: A dataset and an improved multi-task model for floorplan image analysis. In *Proceedings of the Scandinavian Conference on Image Analysis*, pages 28–40, 2019.
- [8] Jang H., Yu K. and Yang J. Indoor Reconstruction from Floorplan images with a Deep Learning Approach. *ISPRS International Journal of Geo-Information*. 9(2), 2020.
- [9] Kippers R., Koeva M., Van Keulen M. and Oude Elberink S. Automatic 3d Building Model Generation Using Deep Learning Methods Based on Cityjson and 2d Floor Plans. In *Proceedings of International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 49–54, 2021.
- [10] Seo J., Park H. and Choo S. Inference of drawing elements and space usage on architectural drawings using semantic segmentation. *Applied Sciences*. 10(20):7347, 2020.
- [11] MIT, Computer Science and Artificial Intelligence Laboratory. LabelMe. On-line: <http://labelme.csail.mit.edu/Release3.0/>, Accessed: 04/11/2021.
- [12] Rokas B. PyLessons. On-line: <https://pylessons.com/>, Accessed: 22/11/2021.

IFC-based Information Extraction and Analysis of HVAC Objects to Support Building Energy Modeling

H. Li^a and J. Zhang^a

^a*Automation and Intelligent Construction (AutoIC) Lab, School of Construction Management Technology, Purdue University, West Lafayette, IN 47907, USA*

E-mail: li4016@purdue.edu, zhan3062@purdue.edu

Abstract –

The heating, ventilation, and air conditioning (HVAC) system is a highly complex part of a building that requires high specialty and expertise to understand and analyze for energy modelling and simulation purposes. Significant manual effort is needed for information extraction from the mechanical designs, to support the creation of an energy model, including information such as HVAC system type, cooling/heating load, pressure drop, and thermal zones, etc. However, such information can be readily available in Building Information Modeling (BIM)-based mechanical, electrical, and plumbing (MEP) models. In this paper, data analysis and information extraction were conducted on HVAC systems of industry foundation classes (IFC)-based MEP models. By following the state-of-the-art Data-driven Reverse Engineering Algorithm Development (D-READ) method, an algorithm was developed to automatically parse and extract HVAC information from the IFC models. The algorithm was tested on a commercial building with 1 hot water boiler and 19 radiators, which achieved error-free information parsing and extraction. This is expected to reduce the manual effort in information extraction of HVAC systems for building energy modeling (BEM). It also is built upon and supports the open and neutral IFC-based information workflow, which could be a solid step towards automation and interoperability between BIM and BEM in the HVAC domain.

Keywords –

Building Information Modeling (BIM); Building Energy Modelling (BEM); Heating, Ventilation, and Air Conditioning (HVAC); Industry Foundation Classes (IFC); Automation; Information Extraction; Interoperability

1 Introduction

In both commercial and residential building sectors, the heating, ventilation, and air conditioning (HVAC) component has been a top contributor to energy

consumption, among all building components. According to a report from the U.S. Environmental Protection Agency, the HVAC systems in commercial and residential buildings contribute up to 15.6% of the electricity use in the US. [1]. Considering such a significant energy consumption from the HVAC systems and the increasing emphasis on energy conservation in the society [2], there is an urgent need of an efficient method for evaluating the energy consumption of HVAC systems, which is expected to help better inform owners and professionals of the energy performance and potential energy savings throughout their building's lifecycle. However, the traditional energy evaluation method is a time-consuming process. Mechanical engineers/designers and energy modelers still heavily rely on 2D mechanical drawings for HVAC information interpretation and analysis as the standard industry practice. It requires a large number of labor hours to manually retrieve HVAC information from 2D mechanical drawings. With such consideration, researchers nowadays are seeking digital alternatives to the manual approach used in the state of the practice, to automatically extract and analyze HVAC information from Building Information Modeling (BIM) [3].

BIM can provide an object-orientated data repository for managing information of building projects including architectural, structural, and mechanical, electrical, and plumbing (MEP) systems, among others. As an important part of the MEP system, the HVAC system can also be modeled in BIM including information such as system type, cooling/heating load, pressure drop, and associated thermal zones. The information-rich nature of BIM has been shown successful in supporting many tasks in the architectural, engineering, and construction (AEC) industry such as quantity takeoff [4] and structural analysis [5]. The extensive adoption and associated benefits of BIM have also increased the demand to leverage BIM for energy analysis. The importance of BIM to energy analysis is further supported by the increasing levels of BIM adoption by architects and the need to improve the energy modeling processes to better fit into the architectural design workflow [6].

Industry Foundation Classes (IFC) is an international

open standard that is independent of proprietary BIM authoring tools. The neutral and transparent feature of IFC can provide flexibility for data sharing and collaboration across platforms and stakeholders. In this paper, the authors developed a new systematic methodology for extracting essential information of HVAC systems from IFC-based BIM by following the state-of-the-art Data-driven Reverse Engineering Algorithm Development (D-READ) method [4]. The developed information extraction algorithm following the proposed methodology was also tested on a commercial building model and achieved error-free performance.

2 Model Description

Two IFC models were collected for HVAC information analysis and extraction. The first model was a 2-story duplex apartment building [7]. Figure 1 and Figure 2 show the architectural model and MEP model of the duplex apartment building, respectively. There were 926 entity instances in total in the IFC-based MEP model of the duplex apartment building. Essentially, the building consisted of 3 levels: Level 1, Level 2, and Roof Level. Each level contained multiple types of IFC entities such as *IfcFlowControllers*, *IfcFlowSegments*, and *IfcSpaces*, etc. It was found there were 21 *IfcSpace* entity instances in total.

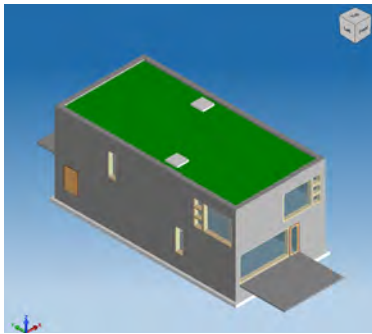


Figure 1. Architectural model for the duplex apartment building



Figure 2. MEP model for the duplex apartment building

The second model was an office building with a different HVAC system type from the first model [7]. Figure 3 and Figure 4 show the architectural model and MEP model of the office building, respectively. There were 5697 entity instances in the IFC-based MEP model. The data structure of the office building model was the same as the duplex apartment building model.

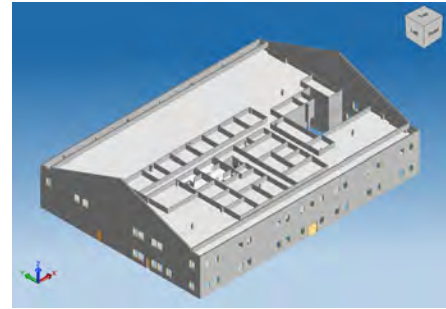


Figure 3. Architectural model for the office building



Figure 4. MEP model for the office building

3 Preliminary Analysis

By analyzing the MEP models of the two buildings in a data-driven manner, it was found that the main HVAC components such as boilers and chillers were defined as *IfcEnergyConversionDevice* in IFC. And all the pipes/ducts and elbows for constructing and connecting the HVAC loops were defined as *IfcFlowSegments* and *IfcFlowFitting*, respectively. The pumps or fans for circulating the flow through the loops were defined as *IfcFlowMovingDevice*. The diffusers for distributing the conditioned air to the targeted thermal zones were defined as *IfcFlowTerminal*. Table 1 summarizes all the corresponding IFC entities for HVAC components.

Table 1. Corresponding IFC entities for HVAC components

HVAC components	Corresponding IFC entity
Boiler	<i>IfcEnergyConversionDevice</i>
Radiator	<i>IfcEnergyConversionDevice</i>
Chiller	<i>IfcEnergyConversionDevice</i>
Pump/fan	<i>IfcFlowMovingDevice</i>
Variable air volume (VAV) unit	<i>IfcFlowMovingDevice</i>
Air handling unit (AHU)	<i>IfcFlowMovingDevice</i>
Pipe	<i>IfcFlowSegment</i>
Duct	<i>IfcFlowSegment</i>
Elbow	<i>IfcFlowFitting</i>
Diffuser	<i>IfcFlowTerminal</i>

There were 16 *IfcEnergyConversionDevice* entity instances (i.e., 14 radiators and 2 hot water boilers) in the duplex apartment model. Table 2 and Table 3 summarize the number of IFC entity instances corresponding to each HVAC component for the duplex apartment model and the office model, respectively.

Table 2. Number of HVAC related entity instances for the duplex apartment model

HVAC component	Number of entity instances
Boiler	2
Radiator	14
Pump	2
Pipe	417

Table 3. Number of HVAC related entity instances for the office model

HVAC component	Number of entity instances
Screw chiller	1
AHU	2
VAV unit	21
Pump	1
Duct	643
Diffuser	245

4 Proposed Methodology

Figure 5 illustrates the four-step methodology of the HVAC information extraction the authors proposed. Firstly, the HVAC system type is identified by analyzing the main HVAC components represented by *IfcEnergyConversionDevice* entity in the IFC-based MEP model. For example, a chiller with VAV units can help determine the system type to be a VAV system. Secondly, the HVAC plant loops are extracted (e.g., hot water loops, the chilled water loop, and the air loop) and

corresponding components in each loop are identified. In the third step, detailed parameters for HVAC components are parsed such as the heating capacity of the boiler, volume flow rate of the pump, etc. In the final step, the related thermal zones served by each HVAC terminal are identified.

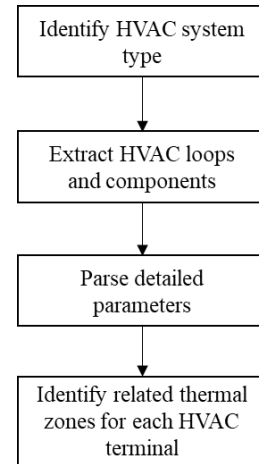


Figure 5. The proposed four-step methodology of HVAC information extraction

5 Application of the Proposed Methodology

5.1 Identify HVAC System Type

Based on the preliminary analysis results as shown in Table 2, the HVAC system type for the duplex apartment model was identified as a hot water boiler radiator system. A standard hot water boiler radiator system consists of supply equipment and demand equipment. The supply equipment in this case included a hot water boiler for generating hot water and a pump for circulating the water through the loop. On the demand side, several radiators were connected to targeted thermal zones. The hot water flew through the radiators and provided heating to the thermal zones.

For the office model, a VAV system consisting of a chilled water loop and an air loop was identified. The chilled water loop included a chiller and a pump for creating cooled water for the air handling units. And the air loop consisted of an air handling unit, an electric heating coil, and a fan for providing conditioned air through the loop. The VAV boxes on the demand side received the conditioned air and distributed it to the targeted thermal zones.

5.2 Extract HVAC Loops and Components

Through the Data-driven Reverse Engineering

Algorithm Development (D-READ) method [4], the authors developed an algorithm to extract the HVAC loops and corresponding components. For example, the system name and system type were two features for distinguishing the two hot water loops of the duplex apartment model. System names “Unit A Hydronic Supply In” and “Unit B Hydronic Supply In” were used to distinguish and extract all the *IfcFlowSegment*, *IfcFlowFitting*, *IfcFlowMovingDevice*, and *IfcEnergyConversionDevice* entity instances in Hot Water Loops A and B, respectively. The extraction results for Hot Water Loops A and B are shown in Figure 6 and Figure 7, respectively. Similarly, system types “supply air” and “return air” were used to extract all the related HVAC entity instances in the air loop of the VAV system from the office model (Figure 8 and Figure 9).

The loop extraction is important in HVAC system analysis because it defines the relationships of components and the functionality of each loop. For example, Figure 6 and Figure 7 show the two hot water plant loops extracted from the IFC-based MEP model of the duplex apartment building, respectively. It can be seen that each loop consisted of a boiler, with multiple radiators serving corresponding thermal zones. By extracting and separating the two hot water loops, information such as the service relationship between boiler A and radiator 1, and the number of radiators served by boiler A in the hot water loop A, were obtained (as shown in Figure 6). In addition, The pressure drop along the pipe in each loop configuration can also be calculated based on the pipe length.

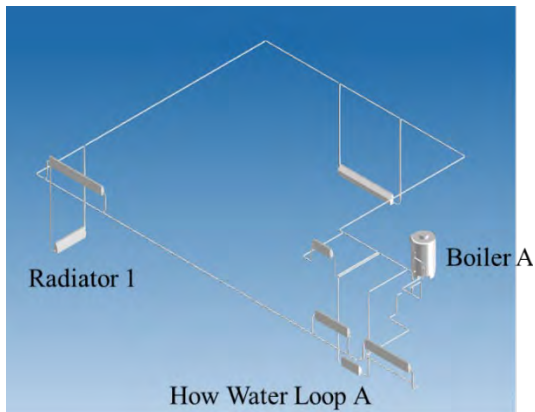


Figure 6. Extracted hot water plant loop A from the duplex apartment MEP model.

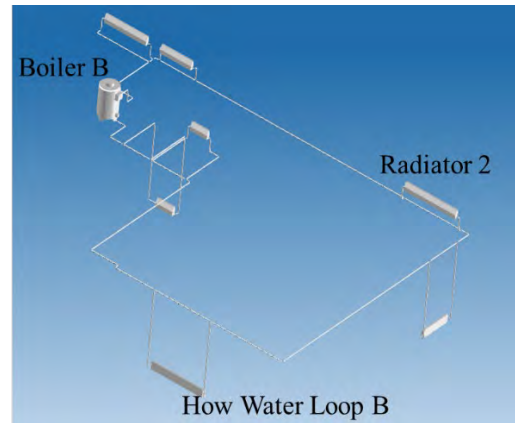


Figure 7. Extracted hot water plant loop B from the duplex apartment MEP model.

Figure 8 and Figure 9 show the extracted chilled water loop and air loop for the office building, respectively. The chilled water loop consisted of a screw chiller, a chilled water pump, air handling units, and cooling coils. A chilled water pump pushes chilled water through the chiller and the chilled water line around the building. The AHU or cooling coil serves as the heat transformation media that connects the chilled water loop to the air loop. The air is cooled by the chilled water in the AHU and delivered to the thermal zones through the HVAC terminals (i.e., diffusers). The air loop consisted of VAV boxes, air handling units, ductwork, fans, and diffusers.

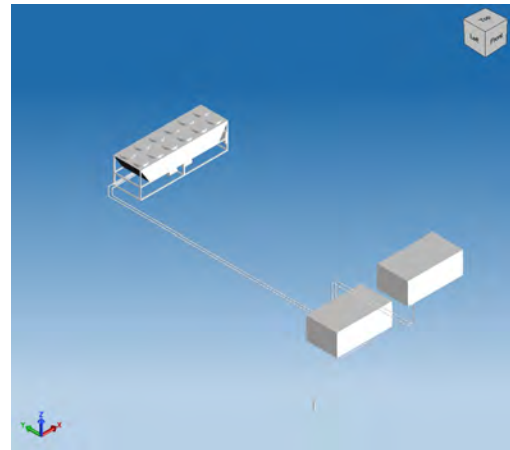


Figure 8. Extracted chilled water loop from the office building MEP model.

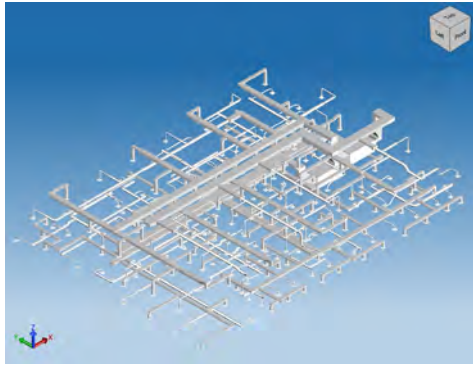


Figure 9. Extracted air loop from the office building MEP model.

5.3 Parse Detailed Parameters

The next step was to retrieve detailed parameters for each HVAC component. Some of the basic parameters such as the system type (e.g., natural gas boiler or electric boiler) were contained in the attributes of the corresponding IFC object. The mechanical properties were extracted through the *IfcPropertySet* entity instance related to each HVAC component. For example, the most essential mechanical parameters of the boiler include heating capacity, efficiency, etc. By retrieving the output heat and input heat as 2949311 Watts and 3465979 Watts, respectively, the thermal efficiency of the boiler was calculated by Equation (1).

$$\text{boiler efficiency} = \frac{\text{output heat}}{\text{input heat}} = \frac{2949311W}{3465979W} = 85\% \quad (1)$$

Table 4 shows the summary of detailed parameters of the HVAC components of the boiler radiator system in the duplex apartment model.

5.4 Identify Related Thermal Zone for HVAC Terminals

One of the most important parameters for an HVAC terminal (e.g., a radiator or diffuser) is the thermal zone that it serves. In the IFC data, this relationship was defined by the *IfcRelContainedInSpatialStructure* entity (as shown in Figure 10), which consisted of *RelatedElements* and *RelatingStructure*. The *RelatedElements* attribute specified the HVAC terminal such as radiators and diffusers. The *RelatingStructure* attribute specified the spaces or thermal zones that the related equipment served. By iteratively parsing through the *IfcRelContainedInSpatialStructure* entity instance, the related thermal zones for each HVAC terminal were extracted.

6 Case Study

The information extraction algorithm developed by

following the proposed methodology was then tested on the testing model - a commercial building with 4 stories as shown in Figure 11. The IFC-based MEP model was parsed through our developed algorithm. The HVAC system was identified as the boiler-radiator system. The hot water loop was extracted correctly, and the related HVAC components including 1 boiler and 19 radiators were successfully parsed including their detailed parameters. And the 19 thermal zones served by corresponding radiators were identified. Table 5 summarized the testing results of the information extraction algorithm, 396 entity instances were successfully extracted without error.

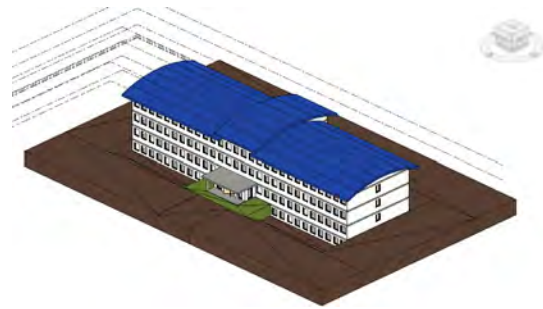


Figure 10. Testing model visualized in Revit

7 Conclusion

In this paper, the authors proposed a four-step methodology for parsing and extracting HVAC information from the IFC-based building information models. Two types of HVAC systems from two BIM models were analyzed following the proposed methodology, namely the VAV system and the boiler-radiator system, respectively. Information including HVAC system type, HVAC loops, detailed component parameters, and related thermal zones for HVAC terminals were identified and parsed systematically. The HVAC information extraction algorithm following the proposed methodology was tested on a commercial building model with the boiler-radiator system and achieved error-free extraction. This is expected to help reduce the time and effort in information extraction of HVAC systems from BIM compared to that by manual extraction from 2D mechanical drawings.

7.1 Contributions to the Body of Knowledge

This research contributes to the body of knowledge in three main ways: First, the authors proposed a new methodology for extracting essential HVAC information from IFC-based BIMs. The proposed methodology was used to develop an automated HVAC information extraction algorithm based on existing IFC-based BIM.

Table 4. Summary of parameters of the boiler radiator system in the duplex apartment model

HVAC components	Number of entity instances	System type	Capacity	Pressure drop	Efficiency
Boiler	2	Natural gas	147kW	99974.4 Pa	85%
Radiator	14	Baseboard	2744.47 Btu/Feet	0.28 meters H ₂ O	N/A
Pump	2	Centrifugal	3.9 Liters/Second	0.8 meters H ₂ O	Motor: 90%

```
{'id': 195486,
'type': 'IfcRelContainedInSpatialStructure',
'GlobalId': '2rU3W39sT9n9Bbw02tJD0I',
'OwnerHistory': #1=IfcOwnerHistory(#136873,#57792,$,.NOCHANGE.,$,$,$,0),
'Name': None,
'Description': None,
'RelatedElements': (#703=IfcEnergyConversionDevice('2W7U6tKFT4vf9lgsJlEr9H',#1,'M_Radiator - Hosted:Radiator - 25:Radiator - 25:535063',$,.Radiator - 25',#135150,#134152,'535063'),
#544=IfcFlowTerminal('1K7eM1Qof1d0c9$mY9I4C8',#1,'M_Pendant Light - Hemisphere:150W - 120V:150W - 120V:575385',$,'150W - 120 V',#135311,#133855,'575385'),
#794=IfcFlowTerminal('1K7eM1Qof1d0c9$mY9I4C8',#1,'M_Lighting Switches:Single Pole:Single Pole:575402',$,'Single Pole',#135180,#133837,'575402'),
#1755=IfcFlowTerminal('1K7eM1Qof1d0c9$mY9I4De',#1,'M_Duplex Receptacle:Duplex Receptacle:Duplex Receptacle:575434',$,'Duplex Receptacle',#135146,#134427,'575434'),
#1745=IfcFlowTerminal('1K7eM1Qof1d0c9$mY9I4Df',#1,'M_Duplex Receptacle:Duplex Receptacle:Duplex Receptacle:575435',$,'Duplex Receptacle',#135381,#134428,'575435'),
#1739=IfcFlowTerminal('1K7eM1Qof1d0c9$mY9I4Do',#1,'M_Duplex Receptacle:Duplex Receptacle:Duplex Receptacle:575440',$,'Duplex Receptacle',#135439,#134432,'575440'),
#2229=IfcFlowController('112cg02D5FS9t6u8i4Cve3',#1,'M_Smoke Detector:Smoke Detector:Smoke Detector:610319',$,'Smoke Detector',#135484,#137631,'610319')),
'RelatingStructure': #2839=IfcSpace('0CRPz_SEr94Ah74P8LhwKx',#1,'A203','',$,#3633,#134081,'Bedroom 2',.ELEMENT,.INTERNAL.,$)}
```

Figure 11. An example of *IfcRelContainedInSpatialStructure* entity instance

Table 5. Testing results of the developed information extraction algorithm

HVAC components	Corresponding IFC entity	Number of entity instances in IFC-based MEP model	Number of objects correctly parsed by developed algorithm	Accuracy
Water Loop	-	1	1	100%
Boiler	<i>IfcEnergyConversionDevice</i>	1	1	100%
Radiator	<i>IfcEnergyConversionDevice</i>	19	19	100%
Pump	<i>IfcFlowMovingDevice</i>	1	1	100%
Pipe	<i>IfcFlowSegment</i>	355	355	100%
Thermal zones	<i>IfcSpace</i>	19	19	100%
Overall	-	396	396	100%

data. In contrast to manually extracting HVAC information from 2D mechanical drawings, the developed algorithm is expected to extract HVAC information more efficiently and objectively. Second, this work facilitates the transition from 2D drawings to 3D digital representations by supporting the automated information extraction from IFC-based BIMs. It also builds a solid step towards automation and interoperability between BIM and BEM in HVAC information exchange. Third, some HVAC components were found missing from the IFC-based BIMs. For example, in the office building model, no cooling coil was found in the chilled water loop. The cooling coil can be defined as *IfcCoil* according to BuildingSMART [8]. Regarding the missing elements, the extraction algorithm can be converted/adapted to one for model design verification, i.e., checking if there are any missing HVAC components in the IFC model.

7.2 Limitations, and Future Work

Two main limitations of this study are acknowledged. First, the lack of comprehensive industry standards for classifying IFC entity instances could render the extraction error-prone in some circumstances. For example, as it is shown in Figure 12, four types of HVAC components were classified as the *IfcFlowMovingDevice* entity for the office building model (as highlighted in different colors). However, they were on different loops. The centrifugal fan, air handling units, and VAV units were on the air loop, whereas the pump was on the chilled water loop. Although it was correct to regard both air and water as flow (as defined by *IfcFlowMovingDevice*). This could make the loop extraction process confusing. To address this issue, future work should investigate a more thorough industry standard for creating the IFC-based MEP model so that components can be classified more rigorously. Second, the scope of this study currently focuses on the extraction of high-level information in four essential categories: (1) HVAC system type, (2) HVAC loops and components, (3) detailed parameters for components, and (4) related thermal zone for HVAC terminals. The HVAC system types used in the experiment were limited to the boiler-radiator system and VAV system. More advanced HVAC system types were not investigated in this paper, such as the variable refrigerant flow (VRF) system, dedicated outdoor air system, heat pump, etc. In addition, air distribution components (e.g., diffusers and grilles for distributing conditioned air) and detailed parameters (e.g., size of ducts/pipes), as well as the control strategies such as operation schedules/modes required by BEM, have not been extracted yet, due to limitations of our experimental IFC model. A more comprehensive and robust algorithm can be developed in future work to cover a broader set of IFC-based MEP models with more detailed parameters

by following the proposed methodology. In addition, automated methods/approaches in semantic enrichment of IFC-based BIM and model validation/checking could help facilitate the IFC-based BIM interoperability with BEM, which should be investigated in future work as well.

References

- [1] US EPA, "Electricity Customers", *US EPA*, 2022. [Online]. Available: <https://www.epa.gov/energy/electricity-customers>. [Accessed: 02- Jan- 2022].
- [2] C. de la Cruz-Lovera, A. Perea-Moreno, J. de la Cruz-Fernández, J. Alvarez-Bermejo, and F. Manzano-Agugliaro, "Worldwide Research on Energy Efficiency and Sustainability in Public Buildings", *Sustainability*, vol. 9, no. 8, p. 1294, 2017. Available: 10.3390/su9081294.
- [3] E. Kamel and A. Memari, "Review of BIM's application in energy simulation: Tools, issues, and solutions", *Automation in Construction*, vol. 97, pp. 164-180, 2019. Available: 10.1016/j.autcon.2018.11.008.
- [4] T. Akanbi, J. Zhang and Y. Lee, "Data-Driven Reverse Engineering Algorithm Development Method for Developing Interoperable Quantity Takeoff Algorithms Using IFC-Based BIM", *Journal of Computing in Civil Engineering*, vol. 34, no. 5, p. 04020036, 2020. Available: 10.1061/(asce)cp.1943-5487.0000909.
- [5] J. Wu, H. Sadraddin, R. Ren, J. Zhang and X. Shao, "Invariant Signatures of Architecture, Engineering, and Construction Objects to Support BIM Interoperability between Architectural Design and Structural Analysis", *Journal of Construction Engineering and Management*, vol. 147, no. 1, p. 04020148, 2021. Available: 10.1061/(asce)co.1943-7862.0001943.
- [6] McGraw Hill Construction, "The business value of BIM in North America: multi-year trend analysis and user ratings (2007-2012)." Smart Market Report, McGraw Hill Construction, Bedford, MA, USA, 2012
- [7] Whole Building Design Guide (WBDG). "Common Building Information Model Files And Tools", *Wbdg.org*, 2021. [Online]. Available: <https://www.wbdg.org/bim/cobic/comm-on-bim-files>. [Accessed: 02- Jan- 2022].
- [8] BuildingSMART, "Industry Foundation Classes Version 4.2 bSI Draft Standard", *Standards.buildingsmart.org*, 2022. [Online]. Available: https://standards.buildingsmart.org/IFC/DEV/IFC4_2/FINAL/HTML/. [Accessed: 02- Jan- 2022].

```

ifc_office.by_type('IfcFlowMovingDevice')

[#2647=IfcFlowMovingDevice('2HvpIy7czF08IJRzJYrFl0',#1,M_Centrifugal Fan - Rooftop - Upblast:991-1905 LPS:991-1905 LPS:71213
8',$,'991-1905 LPS',#448840,#463851,'712138'),
#3114=IfcFlowMovingDevice('0e33T6sVz0_wmX66ZhWKYh',#1,M_Air Handling Unit - Split System - Horizontal:63300000 J:63300000 J:5
67396',$,'63300000 J',#452920,#464456,'567396'),
#3115=IfcFlowMovingDevice('3Y4wfEOIXFGQc5PzV5wNdi',#1,M_Air Handling Unit - Split System - Horizontal:63300000 J:63300000 J:5
83145',$,'63300000 J',#448524,#464457,'583145'),
#3913=IfcFlowMovingDevice('0uWa1aOkT9xe4KLkxi5gyz',#1,M Inline Pump - Circulator:3.9 LPS - 0.8 Meter Head:3.9 LPS - 0.8 Meter
Head:715905',$,'3.9 LPS - 0.8 Meter Head',#448071,#441597,'715905'),
#9435=IfcFlowMovingDevice('3HmORFD6X1KPVGYdFqdiJ4',#1,M VAV Unit - Single Duct:400 mm:400 mm:597363',$,'400 mm',#449414,#4445
93,'597363'),
#9436=IfcFlowMovingDevice('1mZDDjjUH3wPVg5W8w_Ana',#1,M VAV Unit - Single Duct:400 mm:400 mm:596002',$,'400 mm',#451276,#4463
73,'596002'),
#9437=IfcFlowMovingDevice('1mZDDjjUH3wPVg5W8w_Awf',#1,M VAV Unit - Single Duct:400 mm:400 mm:596719',$,'400 mm',#448841,#4463
74,'596719'),
#9438=IfcFlowMovingDevice('2XhMvYvGPEzOprAHjdikpM',#1,M VAV Unit - Single Duct:300 mm:300 mm:582290',$,'300 mm',#453269,#4440
94,'582290'),

```

Figure 12. *IfcFlowMovingDevice* entity in the office building

Systematic Literature Review of Building Information Modelling and Green Building Certification Systems

O.I. Olanrewaju^a, W.I. Enegbuma^a and M. Donn^a

^aWellington School of Architecture, Victoria University of Wellington, New Zealand

E-mail: Oludolapo.olanrewaju@vuw.ac.nz, wallace.enegbuma@vuw.ac.nz, michael.donn@vuw.ac.nz

Abstract –

The rapid pace of technological transformation in sustainability assessment in the construction industry has directed the development of tools and policies. Building information modelling (BIM)-based documentation processes for green building certification systems (GBCS) credits continually require re-assessments. The relationship between BIM and GBCS is minimal in the current literature, which motivates the need for this systematic literature review. This study aims to map the synergies and potentials of BIM and GBCS integration for improvements in the sustainability assessment process. A systematic literature review was adopted to map existing gaps, potentials, and future research areas. A total of 84 papers between 2009 and 2020 from top indexed built environment journals. Energy possessed the highest representation of 71% in the environmental sustainability dimension while economic and social had 15% and 11% respectively. LEED possessed the highest representation of 35% in multi-criteria GBCS. The findings revealed predominant neglect of social and economic dimensions of sustainability credits. Regenerative credits such as biodiversity, water, land use and ecology, socio-economic and acoustics are less incorporated into sustainability assessment models. These findings have direct implications on sustainability assessment policy improvements to implement the use of emerging technologies such as internet of things (IoT) and Blockchain.

Keywords –

Building information modelling; BIM; Green building certification system; Net-zero buildings; Systematic literature review; Sustainability

1 Introduction

Sustainability problem solving is complex and requires an integrated approach to effectively address inherent environmental issues [1]. Consequently, the high rate of global resource utilisation in such forms as

excessive water, energy, and forest use and raw materials has encouraged the concept of sustainable development with the view to meet the current needs without adverse effects on the future [2, 3]. Buildings are one of the most dominant sources of resource usage and environmental emissions, using about 50% of raw materials, consuming 71% of electricity and 16% of water usage, and producing 40% waste disposed of in landfills [4].

The global Architecture, Engineering and Construction (AEC) industry is currently undergoing a dynamic transformation with the introduction of green building (GB) [5]. This has led to the development of several GB certification systems (GBCS) such as Green Star (Australia, New Zealand, and South Africa), Leadership in Energy and Environmental Design [LEED] (the United States of America and Canada), Building Research Establishment Environmental Assessment Method [BREEAM] (United Kingdom), SBTool (Portugal and Czech), Chinese evaluation standard of green building [ESGB] (China), Building Environmental Assessment Method [BEAM] Plus (Hong Kong), Green Mark (Singapore), Green Building Index [GBI] (Malaysia), Deutsche Gesellschaft für Nachhaltiges Bauen [DGNB] (Germany), and Living Building Challenge (LBC). In addition, researchers have also developed GBCS for developing countries by adapting existing GBCS. Some prominent examples of newly developed GBCS are as follows: BSAM - Sub-Saharan Africa [6, 7], SABA - Jordan [8], and GB tool for existing buildings [9].

The AEC industry is more focused on applying emerging technologies to improve environmental sustainability [10]. Technologies such as artificial intelligence (AI), prefabrication, BIM, business process reengineering (BPR), and total quality management (TQM), among others, have been deployed to facilitate the integration of construction processes and achieve value for money [11].

McGraw-Hill Construction conducted an online survey to investigate a wide range of industry professionals who use BIM tools to deliver green buildings. The survey indicated that BIM could significantly facilitate green construction, and it is

expected to be extensively used in the future if relevant challenges can be identified and effectively tackled [12]. BIM surfaced as a solution to facilitate the integration and management of information throughout the building life cycle [12], thus, enabling the effective use of design data for sustainability evaluations [13].

Several functions of BIM have been studied, such as lighting analysis, energy performance simulation, and construction and demolition waste analysis. In addition, different management and incentive aspects associated with BIM adoption have been highlighted, such as its economic benefits, organisational challenges, and motivational challenges [14-16]. A number of BIM applications have been proposed and developed to seamlessly integrate sustainability analysis into traditional design, construction, and operation processes, but there are limited reviews on the connections between BIM and green rating systems [12].

The successful integration of BIM and GB assessment requires a lot of data. Quantitative and qualitative data are the backbone of sustainability assessment [13, 17]. However, Zhao *et al.* [18] noted that one of the fundamental gaps in green building is information technology applications. Lu *et al.* [12] revealed that contemporary BIM software packages are still inadequate in delivering a unified analytical solution for a distinct green building assessment, as they cannot simultaneously analyse all green aspects of buildings.

Furthermore, Gandhi and Jupp [19] findings revealed one of the problems of BIM and green integration as a lack of alignment between the activities, processes, and tasks that encourage modelling and analysis and those that encourage the realisation of green building certification criteria. Also, Ansah *et al.* [13] identified interoperability issues as a major setback in implementing BIM for green building assessment due to the loss of quality information. Xu *et al.* [10] further acknowledged that one of the major challenges with the green construction process is an inaccurate assessment.

There have been several studies on integrating BIM and GBCS [18-25] in recent years. Examples include LEED [15, 20-22], SBTool [23-25], and ESGB [26]. Also, Jalaei and Mohammadi [21] developed a BIM-LEED Revit plugin including a data mining method (K-Nearest Neighbour) that integrated location and transportation, sustainable sites, energy and atmosphere, materials and resources, indoor environmental quality, innovation in design, and regional priority accounting for 6 of 8 credits in LEED v4. Similarly, Kang [22] created a rule-based LEED evaluation method with BIM to facilitate the sustainability assessment process and save time.

Although, many BIM applications have been proposed and developed to seamlessly integrate BIM and GBCS into the traditional design, construction, and

operation processes. Nonetheless, there is still a need for comprehensive integration of BIM and GBCS throughout the project lifecycle [12]. As a result, a systematic literature review (SLR) was carried out. Relevant publications between 2009 and 2020 were identified and synthesised accordingly. The year range was based on literature search output after removing papers that were not relevant to the study. The outcome of this research will provide a roadmap and future directions for BIM and GBCS integration.

2 Research Methodology

Kitchenham and Charters [27] defined a systematic literature review (SLR) as:

“a methodology used to identify, evaluate and interpret research relevant to a determined topic area, research question or phenomenon of interest.”

Generally, there are five major reasons for executing a literature review as stipulated by Paré *et al.* [28] and Paré and Kitsiou [29]. These reasons include:

- establishing the current state of knowledge on the subject or topic;
- resolving the scope to which certain research area divulges any logical patterns;
- gathering practical findings in line with a research question to support its validity;
- creating new frameworks and theories; and
- ascertaining research topics or areas for future research.

The main focus of this SLR is to provide research topics or areas for future research. The review process is divided into two steps namely; SLR protocol development (Step I) and SLR execution (Step II). The SLR protocol development has four sub-steps (purpose definition, research question formulation, keywords and database selection), while the second step has six sub-steps (search strings refinement and article search, inclusion and exclusion criteria, title and abstract scanning, quality assessment, coding, and data extraction pattern). Carvalho *et al.* [24] highlighted the need for a more elaborate systematic literature review to include more keywords and green building rating tools such as Green Star, DGNB, BEAM Plus, etc. The purpose of the review is to provide answers to the following questions:

- What is the level of BIM implementation in GBCS sustainability areas?
- What are the future research directions in the BIM-GBCS domain?

The keywords for the literature search are shown in Table 1. Based on the keywords identified, a search string was developed and implemented on the Scopus search engine. As seen in Figure 1, the initial search string

revealed a total of 400 research papers, while eighty-four (84) were selected for in-depth analysis. The findings of each of the papers are briefly summarised in Table 2.

Table 1. Keywords for systematic literature review

BIM	Green building and assessment tools	Method
BIM OR "Building information modeling" OR "Building Information Modeling (BIM)" OR "Building information model" OR "Building an information-model" OR "BIM (building information modeling)" OR "Building Information Modelling"	"green building" OR "green construction" OR "sustainable building" OR "high performance building" OR "building environmental performance" OR "ecological building" OR sustaina* OR "Green building concept" OR LEED OR CASBEE OR GBI OR DGNB OR BREEAM OR "Green Star" OR "Green Mark" OR "Green Globes" OR ESGB OR GBL OR ECOEFFECT OR ECOPROFILE OR ESCALE OR HK-BEAM OR "BEAM Plus" OR "GB Tool" OR "SB Tool" OR "Home Star" OR "Energy Star"	"Assessment Method" OR "Assessment Rating Method" OR "Certification" OR "Green Building Certification" OR "Evaluation" OR "Labeling Method or System" OR "Guideline" OR "Benchmark" OR "Assessment Standard" OR "Green Building Rating Tools"

3 Analysis and Discussions

3.1 Descriptive Analysis

This section of the paper presents a descriptive analysis of the papers considered in this SLR. The presentation was based on the frequency of papers, GB sustainability area, and GBCS considered.

Frequency of Papers between 2009 and 2020

The frequency of publications between 2009 and 2020 is presented in Figure 2. It is seen that the year 2020 has the highest number of publications accounting for 29% of the total papers used in this study while no publication was recorded in 2012. Overall, it can be deduced that annual publications have grown haphazardly over the years. Also, it can be said that the increasing interest in GB is due to government policies and the coronavirus pandemic [30-32].

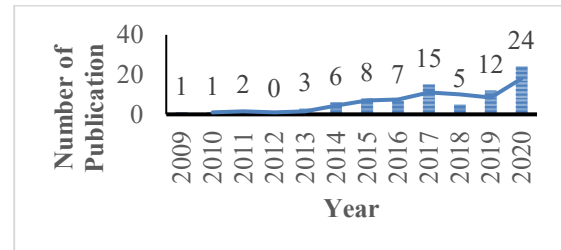


Figure 2. Frequency of Publication

Classification of Papers by GBCS Considered

It is important to note that many studies are mainly focused on the multi-criteria tool while only one paper by Acampa *et al.* [17] mentioned some of the existing lifecycle-based assessment tools (BEES, BEAT, and EcoQuantum). Forty percent (40%) of the studies did not report any GBCS, 35% used/mentioned LEED as a basis for their research while 10% used/mentioned BREEAM. Other GBCS such as Green Star (7%), SBTool (5%), GBI (5%), etc. has been seldom considered in the existing literature. Furthermore, some of the existing studies by Khoshdelnezhamiha *et al.* [33], Olawumi and Chan [34], Mahmoud *et al.* [9], and Ahmad and Thaheem [35] have developed new rating tools (5%) to complement the shortcomings of the existing GBCS.

Journal of Selected Papers

Table 3 shows the top 10 journals in which the selected papers were published. Nearly 40% were published in *Sustainability [Switzerland]* (15.48%), *Automation in Construction* (9.52%), *Sustainable Cities and Society* (7.14%), *Journal of Green Building* (4.76%), and *Applied Sciences [Switzerland]* (2.38%). The table

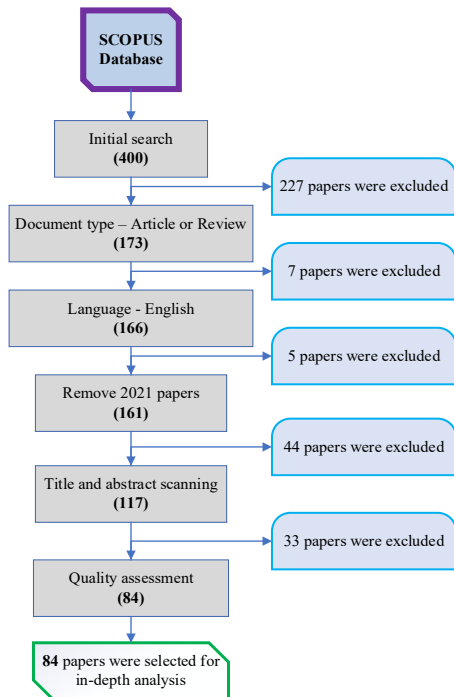


Figure 1. Overview of the paper elimination process

contains highly ranked journal papers that guarantee the reliability of data to be gathered during the SLR.

Table 3. Top 10 Journals of Selected Papers

S/N	Journal	Frequency	*Percentage (%)
1	Sustainability	13	15.48%
2	Automation in Construction	8	9.52%
3	Sustainable Cities and Society	6	7.14%
4	Journal of Green Building	4	4.76%
5	Applied Sciences (Switzerland)	2	2.38%
6	Architectural Engineering and Design Management	2	2.38%
7	ARPJ Journal of Engineering and Applied Sciences	2	2.38%
8	Buildings	2	2.38%
9	Energy and Buildings	2	2.38%
10	International Journal of Sustainable Building Technology and Urban Development	2	2.38%

*Percentage is based on the overall contribution of the journal based on the 84 papers

3.2 Thematic Analysis

Level of BIM implementation in GB Sustainability Area

Using the Green Star NZ as a yardstick, a large portion of the papers focused on the environmental aspect of sustainability which includes; management (4%), indoor environment quality (35%), energy (71%), transport (20%), water (26%), materials (54%), land use and ecology (35%), emissions (23%), and innovation (18%). On the other side, the social (11%) and economic (15%) aspects of sustainability were not often researched. These findings connote that the management, innovation, water, and transport areas under the environmental dimension of sustainability need more research.

Lu *et al.* [12] emphasised that BIM is still not capable of assessing buildings' environmental and social sustainability in a holistic manner. Furthermore, Zanni *et al.* [36] and Lu *et al.* [12] indicated four sustainability areas not supported by BIM software packages: management, ecological issues, innovative techniques and performance, and transportation conditions. The discussion in this part of the paper is structured based on Green Star Design and As-Built New Zealand v1.0 (Green Star NZ) certification system. It includes nine credit categories namely, management, indoor environment quality, energy, transport, water, materials, land use and ecology, emissions, and innovation [37]. In addition to the credit categories of Green Star NZ, the social and economic dimension of sustainability that is

lacking in the GBCS [7] will also be discussed concerning BIM integration.

The management and innovation criteria are often given less attention in most of the existing GBCS despite their presence in many sustainability research [38, 39]. This is evident in Green Mark GBCS, where less than 2% of the entire credits is allocated to “*Management*” [7]. Gandhi and Jupp [19] mentioned that assessing management and innovation components of green building assessment tools is almost unfeasible with BIM. Hence, future studies could concentrate on developing a qualitative oriented and adaptable approach for credits not achievable with BIM [40].

There have been few studies on integrating BIM and IEQ components in GBCS over the years. Al-Sulaihi *et al.* [41] developed a framework to integrate IEQ data into the BIM model in order to detect and track IEQ problems. Marzouk and Abdelaty [42] integrated BIM and wireless sensor network (WSN) for effective monitoring of thermal comfort in subways, an essential ingredient of IEQ. The study revealed that BIM better visualises the building components and IEQ metrics. It also enables efficient control of HVAC systems in an energy-efficient manner. Furthermore, HSE [43] revealed the six factors that influence the thermal comfort of occupants as 1) air temperature, 2) radiant temperature, 3) air velocity, 4) humidity, 5) clothing insulation, and 6) metabolic heat.

Pučko *et al.* [44] introduced a systematic approach for energy and cost analysis of building envelop using ArchiCAD, DesignBuilder and Vico Office. Venkatraj *et al.* [45] used Kieran Timberlake's Tally and Autodesk Green Building Studio (GBS) to compute embodied and operating energy respectively. The study quantitatively revealed that altering the amount of insulation, glazing type, window to wall ratio and depth of external shades to reduce operating energy may lead to the use of energy-consuming. Rodrigues *et al.* [46] compared the performance of GBS and ECO.AP for energy analysis. It was revealed that GBS is suitable at the early design stage of a building with poor performance. Galiano-Garrigós *et al.* [47] revealed that GBS, Herramienta Unificada Lider-Calener (HULC), Sefaira and DesignBuilder perform better than other energy supply tools because their values were close to the real values. Galiano-Garrigós *et al.* [47] suggested that incorporating crossed ventilation or thermal energy inertia could help minimise the differences in values between simulations and real energy and CO₂ emissions values.

It is observed that most of the studies were aimed at developing a plugin, while only Li *et al.* [48] developed a methodology by integrating Dynamo and “Amap” WMS. Furthermore, LEED GBCS is the only one considered in all the four studies that have attempted the automation location analysis credits. Over 70% of existing studies use Google Map API due to its versatility

and flexibility. However, there is a need to explore other WMS to explore strengths and weaknesses. It has been established that the “Amap” WMS has geographical and language limitations, influencing its adoption in other locations.

Water is an essential part of GB, which account for a significant portion of most existing GBCS. There is a lack of research regarding automating BIM for water efficiency analysis [49], although some existing BIM-GB applications such as EnergyPlus, EcoDesigner and Ecotect offer metrics on water usage analysis. Recently, Liu *et al.* [50] developed a water efficiency framework for sustainable building design and construction management, which is yet to be validated. The framework would act as a road map for potential software development in future. However, there is a need to develop an adaptable plugin or framework for GBCS based analysis.

Al-Ghamdi and Bilec [51] conducted a comparative assessment of three common LCA tools (SimaPro, Athena Environmental Impact Estimator, and Kieran Timberlake’s Tally). The study revealed discrepancies in the selected LCA tools due to different methodologies and databases used for calculations. As a result, there is 10% and 17% variation in terms of embedded and operational impacts, respectively. Similarly, Schultz *et al.* [52] conducted a comparative analysis between Athena Impact Estimator and Kieran Timberlake’s Tally. It was stressed that there is a disconnection between design and sustainability assessment at the early stage of buildings. Furthermore, it was revealed that there is a lack of an elaborate database for the LCA of materials. Also, there is a need for existing and new sustainability tools to incorporate the concept of transparency and to benchmark in LCA evaluation.

Ahmad and Thaheem [35] stressed that BIM still has a long way to go to integrate the social dimension of sustainability into BIM. Also, the economic aspect of sustainability is not left out of this trend. Incorporating corporate social responsibility (CSR) and BIM into GBCS gives it the ability to act as a “social integrative system” [53]. Reychav *et al.* [53] established that existing BIM guidelines only focused on stakeholder management, teamwork, and participation. It was revealed that tenants are not involved in a project’s design and construction stage, which negates the equity aspect of social sustainability. The research also presented CSR as a basis for BIM and sociocultural sustainability integration.

Additionally, Zanni *et al.* [54] argued that a clear line of communication aid an effective, sustainable building design process. This emphasises the need to keep project stakeholders, including end-users (tenants), abreast with up-to-date information regarding the project. Similarly, Kim *et al.* [55] used the number of workers needed to

represent the social aspect of a large-scale development master plan sustainability assessment. Subsequently, Ahmad and Thaheem [35] developed a conceptual framework for social sustainability assessment BIM plugin applicable to residential buildings.

Future Research Directions in BIM-GBCS Domain

Firstly, it is evident that there is still disagreement among authors regarding the GBCS credits assessable with BIM, which has led to gaps between BIM technologies and GBCS. However, different GBCS used in various countries have their differences, one of the major reasons for this disagreement. Nonetheless, it was observed that most of the studies that stated certain percentages have not thoroughly validated their speculations. This will help automate GBCS processes, which are mostly manual at the moment and reduce waiting time for GBCS certificates. Consequently, a study is required to validate the speculations in previous research regarding the number of credits assessable with BIM. Secondly, several authors have revealed different percentage differences between the results of BIM-GB software packages with no guideline in benchmarking the outputs for these software packages. Furthermore, the lack of transparency between the computation processes of different software packages has also fuelled the differences in their outputs.

As a result, the following questions have emerged based on the observations of the disagreement obvious in the existing literature:

- How to develop an improved GBCS encompassing critical GB sustainability criteria that can be used to validate the actual credits assessable with BIM technologies?
- How to minimise the gaps between the outputs of existing BIM-GB analysis tools and develop a benchmark for their outputs?

Sustainability consists of three critical pillars, which include social, economic, and environmental dimensions. Nonetheless, the social and economic dimensions of sustainability have little or no points in most of the existing GBCS. This questions the overall aim of sustainability assessments in buildings when all dimensions of sustainability are not considered. Although, studies have attempted to design new GBCS that will include more points for social and economic dimensions [7-9]. However, most of these new GBCSs are still yet to encompass the social and economic dimensions to a satisfactory level. Studies by Ahmad and Thaheem [35] explored the concept of social sustainability in the BIM-GBCS domain without integrating the developed indicators and frameworks into

any of the existing GBCS. Furthermore, a framework to satisfy the economic evaluation of buildings in GBCS and the need to integrate these social and economic indicator frameworks into the existing GBCS is required.

Conclusively, the neglect spotting revealed critical areas that have been overlooked, under-researched, and lack empirical support. The neglect gaps are summarised below:

- Other sustainability areas such as biodiversity, water, land use and ecology, socio-economic and acoustics have been potentially overlooked by researchers in the BIM-GBCS domain.
- The need to extend further research to include other prominent GBCS which have been neglected because most existing studies focused on the LEED GBCS.
- Lack of framework or guideline for evaluating software packages to be used in GBCS evaluations.
- There is a lack of empirical support for ontology-based BIM-GBCS and current literature focus only on the Chinese ESGB.

Based on the extensive literature review, gap-spotting and consideration of different perspectives outlined, the following key areas are essential for future research in the BIM-GBCS domain as summarised below:

- **Development of a holistic framework for BIM-GBCS integration:** there is scanty literature on the holistic application of BIM for GBCS. Different GBCS have unique credit categories as a result of geographical differences.
- **Development of framework or guideline for evaluating software packages to be used for GBCS evaluations:** there is a need to develop a framework or guideline that will assist GB professionals in selecting the best software packages to be used for the building evaluation.
- **Minimising the gap between simulation outputs from GB analysis tools and real-world values:** It is observed that there are still differences between GB analysis tools and real values. In some scenarios, these differences are large, and it questions the integrity of existing GB analysis tools.
- **Integrating BIM and ontology for GBCS credits evaluation:** Currently, there are few pieces of literature on ontology in the BIM-GBCS domain despite the benefit ontology offers that have been reported in previous studies.
- **Integrating BDA, blockchain and IoT for BIM-based GBCS evaluation:** Combining the potentials of big data analytics and the internet

of things for BIM-based GBCS evaluation will enhance the GB evaluation process and as well reduce manpower requirements and process costs.

4 Conclusion

BIM has grown to be a useful tool in the AEC industry. The use of BIM to support GBCS has received huge attention from researchers. Particularly, between 2019 and 2020, the number of published works increased significantly by 50%. This can be attributed to the effect of government policies and the coronavirus pandemic. This study presents a systemic literature review of the relationship between BIM and GBCS. Owing to the review of journal articles ($n = 84$) between 2009 and 2020, this study summarised the: (1) GBCS currently BIM-enabled, (2) level of BIM implementation in GBCS sustainability areas, and (3) areas for further research in the BIM-GBCS domain.

The main findings of this study include the following; Firstly, the current body of research has solely concentrated on the environmental dimension of sustainability while the social and economic dimensions are neglected in the BIM-GBCS. There is a need to maintain a balance across the dimensions of sustainability. Secondly, the study revealed that biodiversity, water, land use and ecology, socio-economic and acoustics had been potentially neglected in the BIM-GBCS domain. As a result, there is a need for thorough research to address these gaps. Additionally, the study will serve as a background for future research works in the BIM-GBCS domain as it highlights the gaps for future studies.

References

- [1] 1. Horn, R., et al., The BIM2LCA Approach: An Industry Foundation Classes (IFC)-Based Interface to Integrate Life Cycle Assessment in Integral Planning. Sustainability, 2020. **12**(16).
- [2] 2. Hussin, J.M., I.A. Rahman, and A.H. Memon, The way forward in sustainable construction: issues and challenges. International Journal of Advances in Applied Sciences, 2013. **2**(1): p. 15-24.
- [3] 3. Zhang, J., et al., A successful delivery process of green buildings: the project owners' view, motivation and commitment. Renewable energy, 2019. **138**: p. 651-658.
- [4] 4. Oduyemi, O., M.I. Okoroh, and O.S. Fajana, The application and barriers of BIM in sustainable building design. Journal of Facilities Management, 2017. **15**(1): p. 15-34.
- [5] 5. Chen, P.H. and T.C. Nguyen, A BIM-WMS integrated decision support tool for supply chain

- management in construction. *Automation in Construction*, 2019. **98**: p. 289-301.
- [6] 6. Olawumi, T.O. and D.W. Chan, Application of Generalized Choquet Fuzzy Integral Method in the Sustainability Rating of Green Buildings based on the BSAM Scheme. *Sustainable Cities and Society*, 2020: p. 102147.
- [7] 7. Olawumi, T.O., et al., Development of a building sustainability assessment method (BSAM) for developing countries in sub-Saharan Africa. *Journal of Cleaner Production*, 2020: p. 121514.
- [8] 8. Ali, H.H. and S.F. Al Nsairat, Developing a green building assessment tool for developing countries—Case of Jordan. *Building and environment*, 2009. **44**(5): p. 1053-1064.
- [9] 9. Mahmoud, S., T. Zayed, and M. Fahmy, Development of sustainability assessment tool for existing buildings. *Sustainable Cities and Society*, 2019. **44**: p. 99-119.
- [10] 10. Xu, Z., et al., Study on the Evaluation Method of Green Construction Based on Ontology and BIM. *Advances in Civil Engineering*, 2019. **2019**.
- [11] 11. Olanrewaju, O.I., et al., Investigating the barriers to building information modeling (BIM) implementation within the Nigerian construction industry. *Engineering, Construction and Architectural Management*, 2020. **27**(10): p. 2931-2958.
- [12] 12. Lu, Y., et al., Building Information Modeling (BIM) for green buildings: A critical review and future directions. *Automation in Construction*, 2017. **83**: p. 134-148.
- [13] 13. Ansah, M.K., et al., A review and outlook for integrated BIM application in green building assessment. *Sustainable Cities and Society*, 2019. **48**.
- [14] 14. Azhar, S. and J. Brown, Bim for sustainability analyses. *International Journal of Construction Education and Research*, 2009. **5**(4): p. 276-292.
- [15] 15. Azhar, S., et al., Building information modeling for sustainable design and LEED ® rating analysis. *Automation in Construction*, 2011. **20**(2): p. 217-224.
- [16] 16. Wu, W. and R.R.A. Issa, BIM execution planning in green building projects: LEED as a use case. *Journal of Management in Engineering*, 2014. **31**(1).
- [17] 17. Acampa, G., et al., Project sustainability: Criteria to be introduced in BIM. *Valori e Valutazioni*, 2019. **2019**(23): p. 119-128.
- [18] 18. Zhao, X., et al., A bibliometric review of green building research 2000–2016. *Architectural Science Review*, 2019. **62**(1): p. 74-88.
- [19] 19. Gandhi, S. and J. Jupp, BIM and Australian green star building certification, in *Computing in Civil and Building Engineering* (2014). 2014. p. 275-282.
- [20] 20. Jalaei, F. and A. Jrade, Integrating building information modeling (BIM) and LEED system at the conceptual design stage of sustainable buildings. *Sustainable Cities and Society*, 2015. **18**: p. 95-107.
- [21] 21. Jalaei, F. and S. Mohammadi, An integrated BIM-LEED application to automate sustainable design assessment framework at the conceptual stage of building projects. *Sustainable Cities and Society*, 2020. **53**.
- [22] 22. Kang, T., Rule-Based LEED Evaluation Method considering BIM Linkage and Variability. *KSCE Journal of Civil Engineering*, 2020. **24**(1): p. 110-121.
- [23] 23. Carvalho, J.P., L. Bragança, and R. Mateus, Optimising building sustainability assessment using BIM. *Automation in Construction*, 2019. **102**: p. 170-182.
- [24] 24. Carvalho, J.P., L. Bragança, and R. Mateus, A systematic review of the role of BIM in building sustainability assessment methods. *Applied Sciences (Switzerland)*, 2020. **10**(13).
- [25] 25. Sanhudo, L.P.N. and J.P.D.S.P. Martins, Building information modelling for an automated building sustainability assessment. *Civil Engineering and Environmental Systems*, 2018. **35**(1-4): p. 99-116.
- [26] 26. Zhang, D., et al., A semantic and social approach for real-time green building rating in BIM-based design. *Sustainability (Switzerland)*, 2019. **11**(14).
- [27] 27. Kitchenham, B. and S. Charters, Guidelines for performing systematic literature reviews in software engineering. 2007, Technical report, Ver. 2.3 EBSE Technical Report. EBSE.
- [28] 28. Paré, G., et al., Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 2015. **52**(2): p. 183-199.
- [29] 29. Paré, G. and S. Kitsiou, Methods for literature reviews, in *Handbook of eHealth Evaluation: An Evidence-based Approach [Internet]*. 2017, University of Victoria.
- [30] 30. Fezi, B.A., Health engaged architecture in the context of COVID-19. *Journal of Green Building*, 2020. **15**(2): p. 185-212.
- [31] 31. The City of Alexandria, Green Building Policy 2019, Planning, Editor. 2019: The City of Alexandria.
- [32] 32. The U.S. Green Building Council. Public Policies. 2019 [cited 2021 5/10/2021]; Available from: <https://public-policies.usgbc.org/>.
- [33] 33. Khoshdelnezhamiha, G., et al., Evaluation of bim application for water efficiency assessment. *Journal*

- of Green Building, 2020. **15**(4): p. 91-115.
- [34] 34. Olawumi, T.O. and D.W. Chan, Green-building information modelling (Green-BIM) assessment framework for evaluating sustainability performance of building projects: a case of Nigeria. *Architectural Engineering and Design Management*, 2020: p. 1-20.
- [35] 35. Ahmad, T. and M.J. Thaheem, Developing a residential building-related social sustainability assessment framework and its implications for BIM. *Sustainable Cities and Society*, 2017. **28**: p. 1-15.
- [36] 36. Zanni, M.A., R. Soetanto, and K. Ruikar, Defining the sustainable building design process: Methods for BIM execution planning in the UK. *International Journal of Energy Sector Management*, 2014. **8**(4): p. 562-587.
- [37] 37. NZGBC, Green Star Design and As Built New Zealand v1.0 Submission Guidelines. 2019.
- [38] 38. Illankoon, I.C.S., et al., Key credit criteria among international green building rating tools. *Journal of cleaner production*, 2017. **164**: p. 209-220.
- [39] 39. Sev, A., How can the construction industry contribute to sustainable development? A conceptual framework. *Sustainable Development*, 2009. **17**(3): p. 161-173.
- [40] 40. Raouf, A.M. and S.G. Al-Ghamdi, Building information modelling and green buildings: challenges and opportunities. *Architectural Engineering and Design Management*, 2019. **15**(1): p. 1-28.
- [41] 41. Al-Sulaihi, I., et al., Assessing indoor environmental quality of educational buildings using BIM. *Journal of Environmental Science and Engineering B*, 2015. **4**(8): p. 451-458.
- [42] 42. Marzouk, M. and A. Abdelaty, Monitoring thermal comfort in subways using building information modeling. *Energy and buildings*, 2014. **84**: p. 252-257.
- [43] 43. HSE. Thermal comfort. 2020 [cited 2020 14/11/2020]; Available from: <https://www.hse.gov.uk/temperature/thermal/>.
- [44] 44. Pučko, Z., D. Maučec, and N. Šuman, Energy and cost analysis of building envelope components using BIM: A systematic approach. *Energies*, 2020. **13**(10).
- [45] 45. Venkatraj, V., et al., Evaluating the impact of operating energy reduction measures on embodied energy. *Energy and Buildings*, 2020. **226**.
- [46] 46. Rodrigues, F., et al., Energy efficiency assessment of a public building resourcing a BIM model. *Innovative Infrastructure Solutions*, 2020. **5**(2).
- [47] 47. Galiano-Garrigós, A., et al., Evaluation of BIM energy performance and CO2 emissions assessment tools: a case study in warm weather. *Building Research and Information*, 2019. **47**(7): p. 787-812.
- [48] 48. Li, J., et al., Integration of Building Information Modeling and Web Service Application Programming Interface for assessing building surroundings in early design stages. *Building and Environment*, 2019. **153**: p. 91-100.
- [49] 49. Chang, Y.-T. and S.-H. Hsieh, A review of Building Information Modeling research for green building design through building performance analysis. *ITcon*, 2020. **25**: p. 1-40.
- [50] 50. Liu, Z., et al., A Building Information Modelling (BIM) based Water Efficiency (BWe) Framework for Sustainable Building Design and Construction Management. *Electronics*, 2019. **8**(6): p. 599.
- [51] 51. Al-Ghamdi, S.G. and M.M. Bilec, Green Building Rating Systems and Whole-Building Life Cycle Assessment: Comparative Study of the Existing Assessment Tools. *Journal of Architectural Engineering*, 2017. **23**(1).
- [52] 52. Schultz, J., et al., A benchmark study of BIM-based whole-building life-cycle assessment tools and processes. *International Journal of Sustainable Building Technology and Urban Development*, 2016. **7**(3-4): p. 219-229.
- [53] 53. Reychav, I., R. Maskil Leitan, and R. McHaney, Sociocultural sustainability in green building information modeling. *Clean Technologies and Environmental Policy*, 2017. **19**(9): p. 2245-2254.
- [54] 54. Zanni, M.A., R. Soetanto, and K. Ruikar, Towards a BIM-enabled sustainable building design process: roles, responsibilities, and requirements. *Architectural Engineering and Design Management*, 2017. **13**(2): p. 101-129.
- [55] 55. Kim, J.I., et al., BIM-based decision-support method for master planning of sustainable large-scale developments. *Automation in Construction*, 2015. **58**: p. 95-108.

Information Extraction from Text Documents for the Semantic Enrichment of Building Information Models of Bridges

P. Schönfelder¹, T. Al-Wesabi², A. Bach² and M. König¹

¹Department of Civil and Environmental Engineering, Ruhr University Bochum, Germany

²Schüssler-Plan Ingenieurgesellschaft mbH, Düsseldorf, Germany

phillip.schoenfelder@rub.de, talwesabi@schuessler-plan.de, abach@schuessler-plan.de, koenig@inf.bi.rub.de

Abstract -

The majority of innovative approaches in the realm of the retrospective generation of building information models for existing structures deal with geometry extraction from point clouds or engineering drawings. However, many building-specific or object-specific attributes for the enrichment of building models cannot be inferred from these geometric and visual data sources, and thus their acquisition requires the analysis of textual building documentation. One type of such documents are structural bridge records, which include specifications regarding used material, location, structural health, modifications, and administrative data. The documents are semi-structured and hardly allow a robust information extraction based on traditional programming, since the implementation of such an approach would result in a complex nesting of conditional clauses, which is not guaranteed to remain effective for future versions of the document structure. Therefore, a data-driven approach is adopted for the information extraction. This paper demonstrates an end-to-end semantic enrichment method, taking a bridge status report as input and feeding structured object parameters directly to the building information modeling software for the enrichment of the model. The proposed method requires little user interaction and achieves production-ready accuracy. It is tested on an as-built model of an actual bridge and shows promising results.

Keywords -

Building information modeling; Information extraction; Semantic enrichment; Natural language processing; Named entity recognition; Machine learning

1 Introduction

In the operations and maintenance phase of a structure, building information modeling (BIM) can facilitate the access to important condition data. [1] For the majority of the bridges currently in use, however, an as-is model does not exist. [2] Since the manual and retrospective modeling of these structures is a demanding task even for expert engineers [3], current research aims at the full or partial

automation of the task. The geometrical aspect of the model is mostly dealt with by the acquisition and processing of point clouds, images or construction plans. [4, 5] In any case, to produce a useful BIM model, not only the structure's geometry, but also semantic information about the covered parts has to be provided and included into the model. [6] This sort of information is hardly present in 2D drawings and not at all in point clouds or images. Therefore, additional sources of information have to be utilized. For bridges specifically, these can be documents like structural records or inspection reports.

In particular, this study deals with the automatic extraction of the needed information from structural bridge records and the integration of the discovered information into bridge models with BIM software. The source documents are text-based and semi-structured in the sense that the data is represented in an almost tabular manner, which is however not encoded as a formal table in the document. This is to facilitate the common exchange and storage of the documents in widely available formats such as PDFs. For that purpose, multiple strategies are pursued to extract the textual data, and to fill in pre-defined building part attributes. After the relevant information is extracted, it is fed into the modeling software Autodesk Revit by means of an import tool developed with the open source visual programming language Dynamo. The proposed method therefore represents an end-to-end semantic enrichment procedure for bridge digital twins. Both labor-intensive tasks, namely the extraction of information from text documents and its integration into BIM models, are addressed by the method. This study therefore contributes to automating the process of semantic enrichment.

The paper is organized as follows: In Section 2, an overview of existing semantic enrichment approaches for digital twins is provided. Also, prior works towards information extraction from bridge documents are listed. Section 3 gives deeper insight into the structure of the source document type for this study and deals with the underlying implementation of the proposed information extraction method. It also includes an overview of the model enrichment procedure. In Section 4, the information extraction

method is tested by using a test set of admissible bridge records. To give an impression of the practical value provided by the methodology, it is demonstrated in Section 5 by the example of a real world bridge, along with its respective BIM geometry and the carried out enrichment. Section 6 concludes the paper by discussing some of the methodology limitations and an outlook to future works.

2 Related Works

Though the number of scientific contributions on the subject of semantic enrichment of BIM models is limited, the studies [7, 8, 9] provide elaborate literature reviews. Previous works have focused mainly on inferring semantic information from the BIM model geometry itself. Belsky et al. [10] presented the SeeBIM software, which allows engineers to manually define rules to create new object relations in BIM models according to the defined conditions. For example, adjacent objects can be aggregated if certain requirements are fulfilled, e.g., matching object types and near proximity. An extension of the software is shown in [11], which also includes the possibility to add alphanumerical information from external sources (e.g., bridge management systems provided by highway agencies) to the model. Bloch and Sacks [8] presented an approach to enrich objects with a semantic specification by inferring it from their geometry with the help of machine learning. In particular, in this example, the room types of a building story are inferred from the floor plan geometry. In comparison with a rule-based approach, the machine learning approach shows superior results in this example. In an approach by Isailović et al. [12], damage information is gathered by means of point cloud acquisition and processing and is later inserted into the BIM model of a bridge. In this example, the model is enriched with detailed geometric and with semantic information.

A difference of the method proposed in this paper to most of the previous semantic enrichment approaches is the source of the additional semantic information. Instead of inferring the semantic information directly from the model's parts (e.g., by their shape and adjacencies), in this approach, external documents are consulted and processed.

On the subject of information extraction, there have been some publications with application in the built environment as well: Liu and El-Gohary [13] applied a CRF-based named entity recognition approach to extract information about deficiencies, their causes and other related entities from bridge status reports. Using their developed bridge deterioration ontology [14], they made use of domain-specific semantics to improve the NER classifier. The corpus used in their study consists of the reports for multiple bridges, each including many years of bridge condition records, and is written in natural language. Moon

et al. [15] followed a similar method, while choosing a Bi-LSTM architecture instead of the CRF model. Also, they made use of the active learning concept to reduce the annotation effort. Capitalizing on the US National Bridge Inventory database, Li and Harris [16] created a text corpus from Virginia bridge status reports. They trained a Bi-LSTM-CRF classifier to recognize damage types, damage severity and the respective locations in the bridges. It was shown that the model outperforms alternative models for the task at hand [17]. Li et al. [18] propose an innovative recognition method for both flat and nested named entities in bridge inspection reports. It is based on the BERT language model, a Bi-LSTM neural network and uses lexicon augmented word embeddings. Their approach mainly differs from previous approaches by the novelty that the question-answering technique is used to find the desired entities (name, structural elements, location, defects, descriptions) in the text. Liu and El-Gohary [19] developed a dependency parsing method to extract the relations between found entities in bridge reports. They used semantic and syntactic features of the text to train a neural network ensemble classifier and achieved promising results, both for the entity-level recognition and the relation extraction. Inspired by the literature, multiple machine learning models are tested in this study.

3 Methodology

For the understanding of the method presented in the following sub-sections, it is useful to have a basic idea of the source documents' structure. Each bridge record in the document corpus follows a strict table of contents and all reports include the same chapters. However, it is not guaranteed for each chapter to actually display content. If a certain piece of information is not provided, the respective section in the document might be empty. Also, even though the chapters themselves are organized in a very structured way, there exist certain exceptions to that structure, e.g., if special works have been carried out on site and are documented accordingly. After all, varying bridge designs also come with varying document structures – may it be for different numbers of parts or even whole part-structures.

Provided by the regional highway agencies, the records list bridge details such as administrative data (owner, operator), geometric data (position, coordinate system, alignment), mechanical information (material, steel grade, coating) and various other types of information. For a human reader, the documents appear to have tabular structure. However, encoded in the PDF files is only line-by-line text with lots of white space and formatting, which makes the documents *appear* like tables. To recreate the tabular structure, one cannot rely on the occurrence of specific keywords or line-breaks, as the included keys vary from

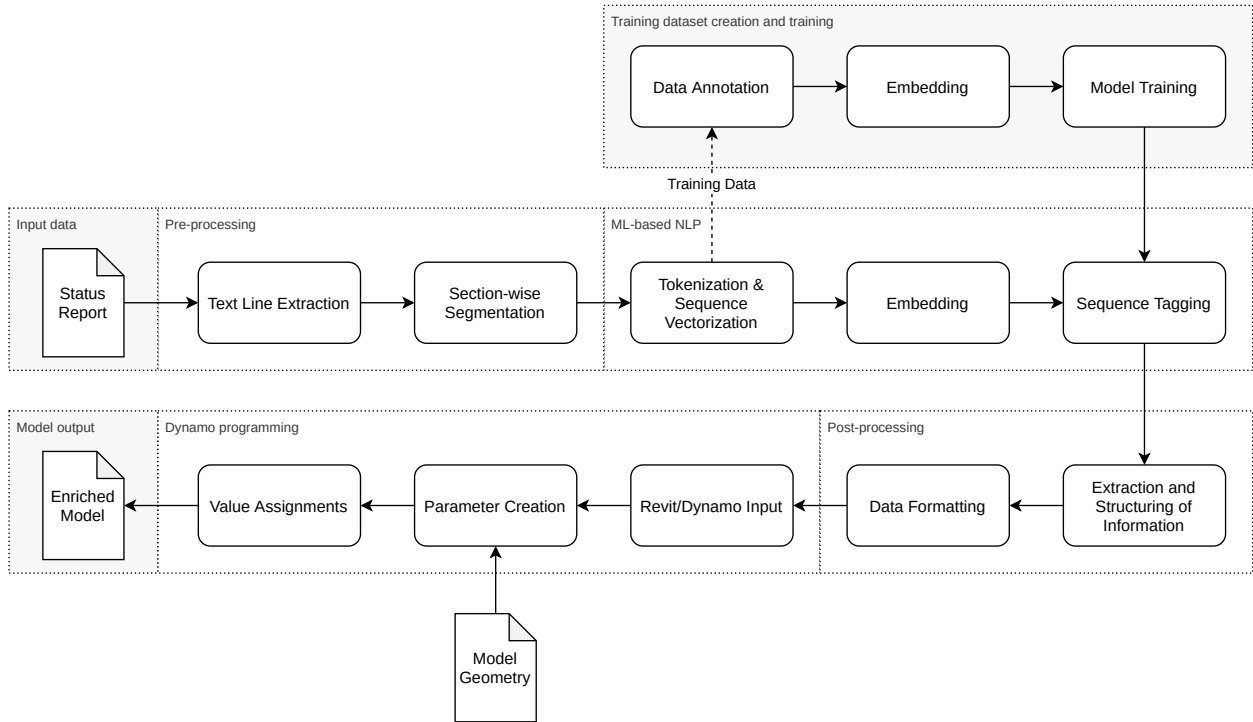


Figure 1. Process diagram of the presented end-to-end extraction and enrichment process.

document to document and as slight imperfections in the formatting are an issue. A traditional programming approach based on regular expression will therefore require an infeasible number of rules and rule exceptions to extract keys and values. Thus, this particular document structure emphasizes the need for a flexible text interpreter.

3.1 Training dataset

To train the machine learning model discussed in Section 3.3, a sufficiently large training dataset is necessary. Therefore, a total of 38 status reports are gathered, pre-processed and annotated manually to be suited for training the neural network and validating the method. The reports cover road and highway bridges in the state of North-Rhine Westfalia in Germany. In total, 98 distinct labels are to be differentiated. Each label represents a parameter name to be extracted with its respective value. Within this study, it is assumed that the documents are free of contradictions, i.e., each parameter is assigned a single value. It is noted that, depending on the types of bridges involved (e.g., truss or beam bridge) and the used construction material (e.g., steel or concrete), a suitable set of training data files has to be selected to include training cases for all the desired pieces information from the reports.

3.2 Pre-processing

Since the documents originally are in PDF format, the textual data is first detached from the typeset document with the use of a PDF manipulation library. Since there is no inherent table structure in the documents, it is converted from PDF to TXT format line-by-line. Excess white space is removed. Having each report present in TXT format, they are segmented in accordance with the highest hierarchy level of the table of contents, which is the only structure of the document relied upon. This segmentation shortens the length of the individual sequences being fed to the model, and with it the number of unique labels a single model has to differentiate. This is because separate models are trained for each of the document sections (c.f. Section 3.3). After the the segmentation, the training text files are annotated manually, labeling the values of the key-value pairs with the name of the key (see Fig. 2). This label choice facilitates the post-processing, as a mapping from a predicted label to a word sequence will directly represent a key-value pair. Moreover, the labels are given prefixes according to the CoNLL 2002 benchmark format [20], i.e., the label of the first token in an entity is extended by the prefix “B-”, all the following tokens of the entity receive the prefix “I-”. This doubles the number of distinct labels to 196.

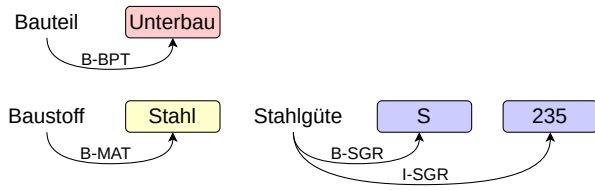


Figure 2. Example of the labeling scheme in a document excerpt. Key names are the label names for the values: Bauteil (building part) → BPT, Baustoff (material) → MAT, Stahlgüte (steel grade) → SGR. Prefixes are in accordance with CoNLL 2002 [20] format.

3.3 Extraction Algorithm

Without all the imperfections to the document structure mentioned in Section 3, a hard-coded approach based on regular expressions would be the ideal way to extract the needed information, however, it simply does not come with the flexibility needed in order to deal with the status reports at hand. Therefore, a more tolerant, data-driven approach is adopted, enabling the processing of documents with slightly different structure, and even for possible future changes to the standard structure.

A fitting candidate approach are recurrent neural networks (RNNs), which are well-suited for sequence processing and have proven useful in all sorts of natural language processing tasks. Hochreiter and Schmidhuber [21] introduced Long-Short-Term-Memory (LSTM) RNNs, which are, in contrast to simple RNNs, able to deal with much longer sequences. Since the bridge record chapters typically include hundreds of words, LSTMs are chosen for the task. Furthermore, bi-directional LSTMs (Bi-LSTMs) may provide an additional advantage over LSTMs, since they can infer a word's label not only with the knowledge of all words before it, but also those after the particular words. This may be of use especially in those situations, in which a certain *key* word does not have an associated *value* following it, if the information is missing. A one-directional LSTM may tend to falsely interpret the following word as a value whatsoever, where a Bi-LSTM may handle the situation better, supposedly. Another possible extension to the model is to append a conditional random field (CRF) layer, which may improve the performance as well [22]. A CRF infers token labels based on the conditional probability that a label occurs given the neighboring labels, thus, it takes the context into account. In the course of this study, multiple model architectures are tested for the sequence tagging: (1) a CRF, (2) an LSTM, (3) an LSTM + CRF layer, (4) a Bi-LSTM, (5) a Bi-LSTM + CRF layer, and (6) a Bi-LSTM followed by

another LSTM and a CRF layer. As a classification layer, a fully connected (FC) layer is inserted after the LSTM layers. It serves as a hidden layer in the models which include a CRF for classification. Also, to transform tokens to a vector representation, each model includes an embedding layer, mapping words to vectors of size d_e . All included LSTM layers have the hidden dimension h_d .

To find the set of best hyper-parameters for each of the model architectures, a Bayesian optimization was run with the help of the KerasTuner library [23] to find concrete values for the hidden dimension d_h and the embedding dimension d_e . The optimal hyper-parameters found for the models being applied to the dataset at hand are summarized in Table 1. All the listed models are trained and evaluated in Section 4.

Table 1. Model hyper-parameters found with Bayesian Optimization, by model design.

Model design	d_h	d_e
LSTM	256	928
Bi-LSTM	320	1024
LSTM + CRF	800	896
Bi-LSTM + CRF	512	1024
Bi-LSTM + LSTM + CRF	32	768

3.4 Post-processing

As the model's output is simply a tagged sequence, it needs to be cast into a tabular form for further use. Therefore, consecutive tokens with matching tags are simply concatenated and condensed to a single key-value entry in the output table. The output is saved as a CSV file to serve as input to the desired BIM software.

It is noted that incorrect predictions can lead to contradictory value assignments. Therefore, after the automatic processing of the documents, an engineer still has to examine the output file and resolve potential conflicts.

3.5 Model Enrichment

Dynamo for Revit is a Python-based visual programming tool that is available to Revit users to allow visually constructed scripts and logic. It allows communication with the Revit API in Python which is a great advantage for developers and enables the automatic enrichment of the BIM model as follows:

One step in the process of BIM modeling is the creation of parameters in Revit. However, it is also one of the most time-consuming processes since the creation of the parameters and assigning their values are typically done manually. The complexity goes to great heights when the

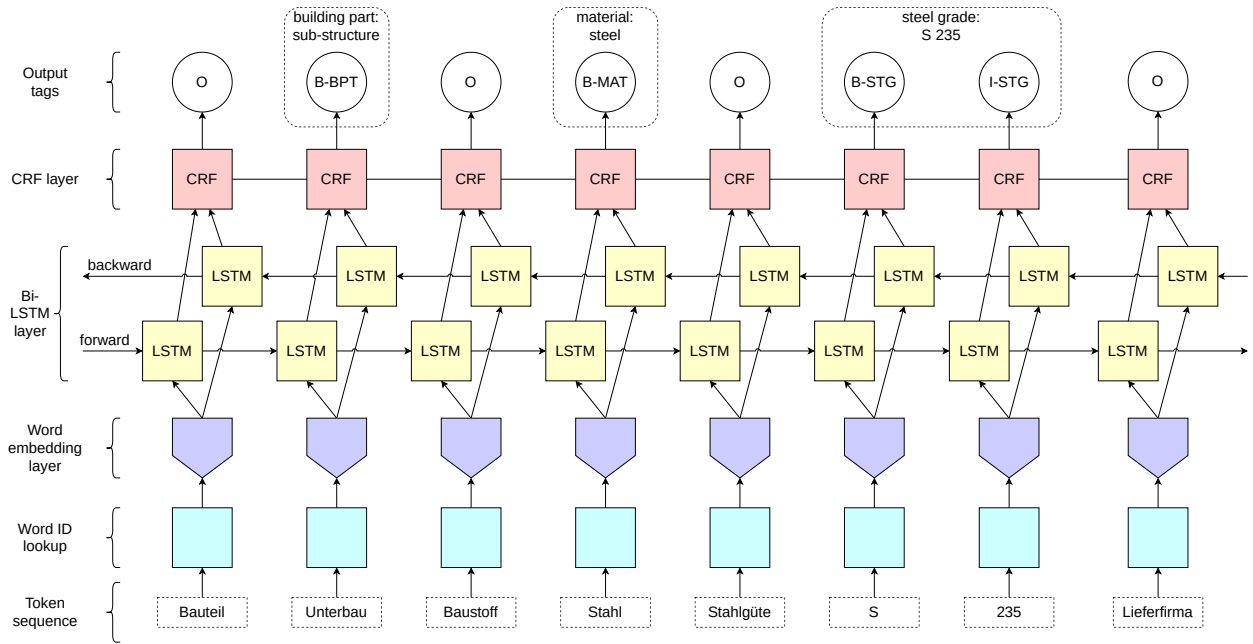


Figure 3. Schematic view of the best performing model design. The example sequence translates to: “Building part substructure building material steel steel grade S 235 supplier”

project has thousands of elements and many layers of classification. Whereas the Revit API provides a great value in automating processes, it does not contain methods of creating project parameters and for that, a workaround is performed through the use of shared parameters. The developed program reads the parameter names and values from the CSV output file of the RNN model. This data is further processed and grouped in Dynamo. The processing of parameters and values includes supplementing them with information to make them readable in Revit. These supplements include adding the data type of each value and choosing the element category for which the list of parameters will be created. After that, the program reads the values and links each parameter with the correct value in order to complete the process of semantic enrichment. In addition, a graphical user interface has been created for the program to be user-friendly and add a level of flexibility for the users. Moreover, the program is integrated with Revit as an add-in in the Revit ribbon to act as a complementary built-in tool to serve the automation, optimization, and ease of use.

4 Validation

To ensure the satisfactory performance of the chosen sequence tagging model, a 10-fold cross-validation experiment is carried out with the annotated corpus. In each of the 10 runs, the dataset is split into 10 parts, thereof 9 for

training and one for testing. The performances are averaged to compute the final performance scores summarized in Table 2. The precision P and the recall R are defined by

$$P = \frac{\# \text{ TP}}{\# \text{ TP} + \# \text{ FP}}$$

$$\text{and } R = \frac{\# \text{ TP}}{\# \text{ TP} + \# \text{ FN}},$$

where $\# \text{ TP}$ is the number of true positives, $\# \text{ FP}$ is the number of false positives and $\# \text{ FN}$ is the number of false negatives in the sequence tagger's predictions. The F1 score is the harmonic mean of precision and recall. Since the Bi-LSTM-CRF shows the best performance among the tested model designs, it is selected for the information extraction pipeline.

The models are implemented in TensorFlow 2.7 [24], and all computations are executed under Ubuntu 20.04.3 LTS using an NVIDIA A100 GPU.

5 Demonstration

To test the proposed method on a real-world example, the highway bridge *Hachmannbrücke*, located in Hamburg, Germany, is modeled in Autodesk Revit and enriched with information from its respective bridge record. The

Table 2. Performance scores of the model variations in the 10-fold cross-validation experiments. Each number represents the averaged result from testing the model on the test set of each of the 10 folds.

Model design	P	R	F1
CRF	0.430	0.437	0.416
LSTM	0.970	0.966	0.965
LSTM + CRF	0.959	0.959	0.957
Bi-LSTM	0.967	0.964	0.963
Bi-LSTM + CRF	0.970	0.969	0.967
Bi-LSTM + LSTM + CRF	0.958	0.957	0.955

model comprises 1481 elements.

The user interface is depicted in Figure 4 and aids the engineer as follows: The information extraction pipeline takes a bridge record as input and outputs a CSV file. The user imports the file via the GUI (1) and proceeds with the selection of parameters and their respective values to be imported (2). The GUI snippets (4) and (5) depict the created and the enriched element parameters of the model (3), respectively. In the shown example, new parameters are assigned to elements of the category *Fundamente* (foundations), e.g., *Hauptbaustoff* (material) and *Baujahr* (year of construction). Without additional user input, they take the values as extracted from the original bridge record.

In this example, the method achieves an accuracy of 86%, given the parameter names listed in Figure 4. All of them are extracted correctly, except the parameter *Bemerkungen* (i.e., remarks). Presumably, this is because remarks can have arbitrary content, contrarily to parameters like steel grade or year of construction, which have a limited set of possible values and are, thus, more likely to be recognized by the model.

Without the trained extraction method and the designed Revit GUI, the user has to read into the PDF file, find the desired values, create the needed parameters for the relevant objects and assign the values to them one-by-one. For the selected parameters to be imported, the proposed method greatly facilitates the process compared to the manual editing of model file. Nonetheless, this semi-automatic model enrichment process still depends on a proficient user and their understanding of the Revit software.

6 Conclusions

Few approaches in the research subject of retrospective BIM model creation deal with the semantic enrichment of bridge models, and if they do, their focus is mostly on inferring information from geometry. This paper, however, demonstrates the value of text documents as a source of information for semantic enrichment. The study's contribu-

tions are twofold: First, an ML-based information extraction algorithm is proposed for the processing of structural bridge records provided by German highway agencies. It shows promising performance in the cross-validation experiments and for the real-world example document. Second, an end-to-end pipeline is developed for the semantic enrichment of bridge BIM models, including an ML-based information extraction method and a Dynamo tool to incorporate the data into the model. The data integration tool is also available as a Revit add-in, which might promote industry acceptance.

However, the presented approach has three main drawbacks: First, it is limited to bridges and, moreover, to these very kind of documents, at least as far as no other training data is provided. Second, it has only been tested for Revit for now, but since the extracted data is stored in an open format, it can serve as input data for other BIM modeling software or for the enrichment of models saved in IFC format as well. Lastly, as the post-processing does not include any semantic processing to ensure the quality of the results, it is advisable for an engineer to check the exported file for errors before importing it to the model. In follow-up studies, it is anticipated to automate this check for the most apparent extraction errors.

Acknowledgements

This research is conducted as part of the BIMKIT project, funded by the German Federal Ministry for Economic Affairs and Climate Action. The authors would like to express their gratitude towards Schüßler-Plan Ingenieurgesellschaft, who generously provided the bridge model for the demonstration in Section 5, and towards Hamburg Port Authority, the owner of the bridge. Also, the authors thank both the Landesbetrieb Straßen NRW and the Autobahn GmbH for making available enough bridge status reports to train and test our developed algorithms thoroughly.

References

- [1] Jack C. P. Cheng, Qiqi Lu, and Yichuan Deng. Analytical review and evaluation of civil information modeling. *Automation in Construction*, 67:31–47, 2016. doi:10.1016/j.autcon.2016.02.006.
- [2] Rebekka Volk, Julian Stengel, and Frank Schultmann. Building Information Modeling (BIM) for existing buildings — Literature review and future needs. *Automation in Construction*, 38:109–127, 2014. doi:10.1016/j.autcon.2013.10.023.
- [3] André Borrmann, Markus König, Christian Koch, and Jakob Beetz. Building Information Modeling: Why? What? How? In André Borrmann,

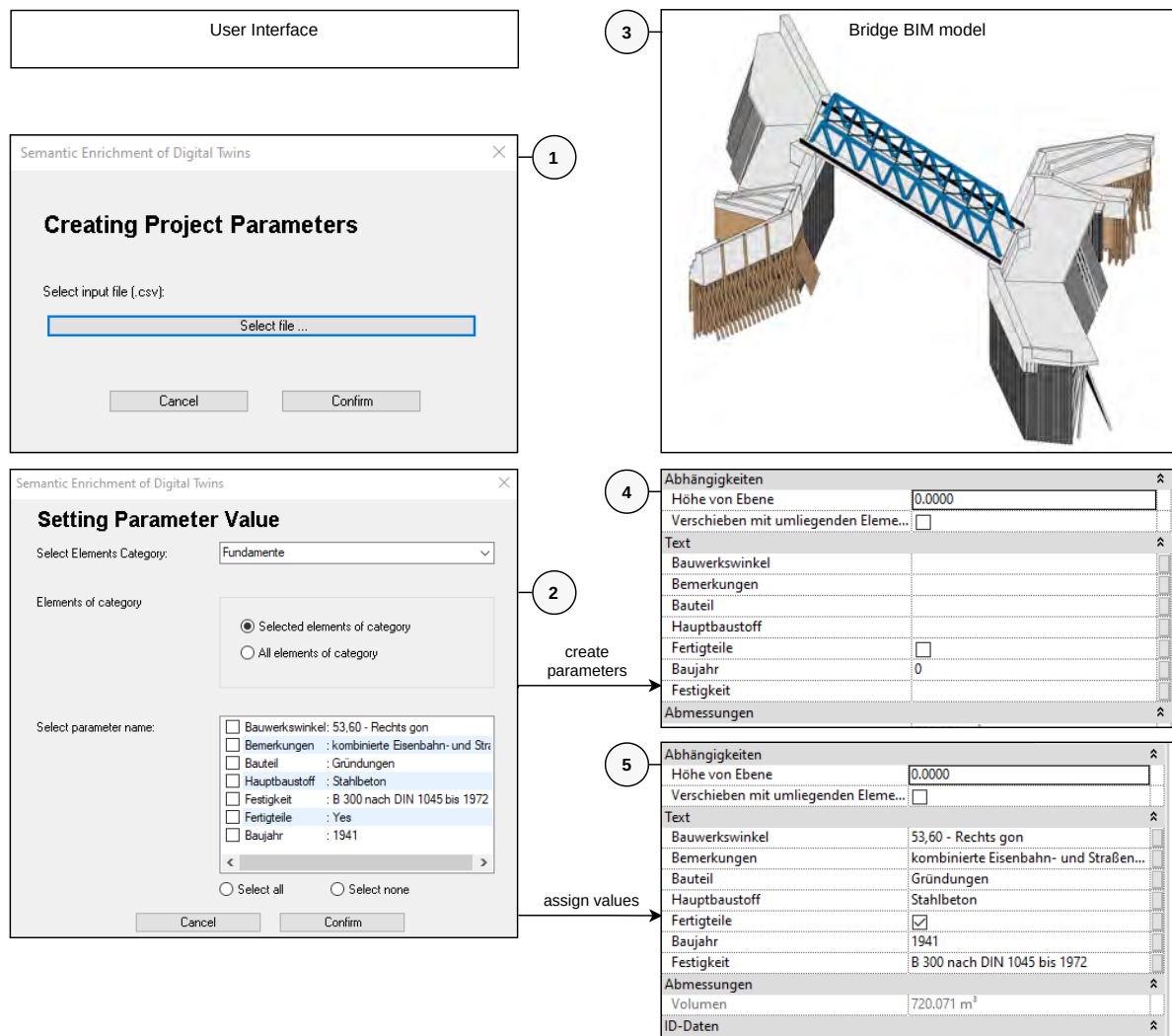


Figure 4. Illustration of the model enrichment user interface. 1: GUI to select the input CSV file, 2: GUI to accept user input to define which element categories should be assigned the respective extracted values, 3: BIM model image, 4: parameter table after parameter creation, 5: parameter table after value insertion.

- Markus König, Christian Koch, and Jakob Beetz, editors, *Building Information Modeling: Technology Foundations and Industry Practice*, pages 1–24. Springer International Publishing, Cham, 2018. doi:10.1007/978-3-319-92862-3_1.
- [4] Viorica Pătrăucean, Iro Armeni, Mohammad Nhang, Jamie Yeung, Ioannis Brilakis, and Carl Haas. State of research in automatic as-built modelling. *Advanced Engineering Informatics*, 29(2):162–171, 2015. doi:10.1016/j.aei.2015.01.001.
- [5] Lucile Gimenez, Jean-Laurent Hippolyte, Sylvain Robert, Frédéric Suard, and Khaldoun Zreik. Review: reconstruction of 3D building information models from 2D scanned plans. *Journal of Building Engineering*, 2:24–35, 2015. doi:10.1016/j.job.2015.04.002.
- [6] Christian Koch and Markus König. Data Modeling. In André Borrmann, Markus König, Christian Koch, and Jakob Beetz, editors, *Building Information Modeling: Technology Foundations and Industry Practice*, pages 43–62. Springer International Publishing, Cham, 2018. doi:10.1007/978-3-319-92862-3_3. URL https://doi.org/10.1007/978-3-319-92862-3_3.
- [7] Nouha Hichri, Chiara Stefani, Livio De Luca, and Philippe Véron. Review of the "as-built BIM" ap-

- proaches. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-5-W1, pages 107–112, Trento, Italy, 2013. doi:10.5194/isprsarchives-XL-5-W1-107-2013.
- [8] Tanya Bloch and Rafael Sacks. Comparing machine learning and rule-based inferencing for semantic enrichment of BIM models. *Automation in Construction*, 91:256–272, 2018. doi:10.1016/j.autcon.2018.03.018.
- [9] Fábio Matoseiro Dinis, João Poças Martins, Ana Sofia Guimarães, and Bárbara Rangel. BIM and Semantic Enrichment Methods and Applications: A Review of Recent Developments. *Archives of Computational Methods in Engineering*, 2021. doi:10.1007/s11831-021-09595-6.
- [10] Michael Belsky, Rafael Sacks, and Ioannis Brilakis. Semantic Enrichment for Building Information Modeling. *Computer-Aided Civil and Infrastructure Engineering*, 31(4):261–274, 2016. doi:10.1111/mice.12128.
- [11] Rafael Sacks, Ling Ma, Raz Yosef, Andre Borrmann, Simon Daum, and Uri Kattel. Semantic Enrichment for Building Information Modeling: Procedure for Compiling Inference Rules and Operators for Complex Geometry. *Journal of Computing in Civil Engineering*, 31(6):04017062, 2017. doi:10.1061/(ASCE)CP.1943-5487.0000705.
- [12] Dušan Isailović, Vladeta Stojanovic, Matthias Trapp, Rico Richter, Rade Hajdin, and Jürgen Döllner. Bridge damage: Detection, IFC-based semantic enrichment and visualization. *Automation in Construction*, 112:103088, 2020. doi:10.1016/j.autcon.2020.103088.
- [13] Kaijian Liu and Nora El-Gohary. Ontology-based semi-supervised conditional random fields for automated information extraction from bridge inspection reports. *Automation in Construction*, 81:313–327, 2017. doi:10.1016/j.autcon.2017.02.003.
- [14] Kaijian Liu and Nora El-Gohary. Semantic Modeling of Bridge Deterioration Knowledge for Supporting Big Bridge Data Analytics. In *Construction Research Congress 2016: Old and New Construction Technologies Converge in Historic San Juan*, pages 930–939, San Juan, Puerto Rico, 2016. doi:10.1061/9780784479827.094.
- [15] Seonghyeon Moon, Sehwan Chung, and Seokho Chi. Bridge Damage Recognition from Inspection Reports Using NER Based on Recurrent Neural Network with Active Learning. *Journal of Performance of Constructed Facilities*, 34(6):04020119, 2020. doi:10.1061/(ASCE)CF.1943-5509.0001530.
- [16] Tianshu Li and Devin Harris. Automated construction of bridge condition inventory using natural language processing and historical inspection reports. In *Nondestructive Characterization and Monitoring of Advanced Materials, Aerospace, Civil Infrastructure, and Transportation XIII*, volume 10971, pages 206–213, 2019. doi:10.1117/12.2514006.
- [17] Tianshu Li, Mohamad Alipour, and Devin K. Harris. Context-aware sequence labeling for condition information extraction from historical bridge inspection reports. *Advanced Engineering Informatics*, 49:101333, 2021. doi:10.1016/j.aei.2021.101333.
- [18] Ren Li, Tianjin Mo, Jianxi Yang, Dong Li, Shixin Jiang, and Di Wang. Bridge inspection named entity recognition via BERT and lexicon augmented machine reading comprehension neural model. *Advanced Engineering Informatics*, 50:101416, 2021. doi:10.1016/j.aei.2021.101416.
- [19] Kaijian Liu and Nora El-Gohary. Semantic Neural Network Ensemble for Automated Dependency Relation Extraction from Bridge Inspection Reports. *Journal of Computing in Civil Engineering*, 35(4):04021007, 2021. doi:10.1061/(ASCE)CP.1943-5487.0000961.
- [20] Erik F. Tjong Kim Sang. Introduction to the CoNLL-2002 Shared Task: Language-Independent Named Entity Recognition. In *COLING-02: The 6th Conference on Natural Language Learning*, pages 155–158, Taipei, Taiwan, 2002. URL <https://aclanthology.org/W02-2024>.
- [21] Sepp Hochreiter and Jürgen Schmidhuber. Long Short-term Memory. *Neural computation*, 9:1735–80, 1997. doi:10.1162/neco.1997.9.8.1735.
- [22] Zhiheng Huang, Wei Xu, and Kai Yu. Bidirectional LSTM-CRF Models for Sequence Tagging. 2015. arXiv: 1508.01991.
- [23] Tom O’Malley, Elie Bursztein, James Long, François Chollet, Haifeng Jin, Luca Invernizzi, and others. Keras Tuner, 2019. URL <https://github.com/keras-team/keras-tuner>.
- [24] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, and others. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, 2015. URL <https://www.tensorflow.org/>.

BIM-Integration of Light Construction Equipment

M. Schöberl^a, S. Offinger^a, T. Goldfuß^a, S. Kessler^a and J. Fottner^a

^aChair of Materials Handling, Material Flow, Logistics, Technical University of Munich, Germany
E-mail: max.schoeberl@tum.de

Abstract –

BIM (Building Information Modeling) has become increasingly important in recent years. At the center of BIM is the digital information model of a building. While such information models are already used intensively in building construction for small-scale robotic applications, digital models are only used for large-scale measures and machines in civil engineering. Light construction equipment has not yet been integrated into digital construction management in civil engineering despite its manifold potentials. This publication therefore takes a closer look at the integration of light construction equipment into BIM-supported digital civil engineering. For this purpose, the fundamentals and the state of the art are presented based on literature, and the linking of BIM and light construction equipment is conceptualized and validated by means of a compaction case study.

Keywords –

BIM; DTM; civil engineering; digital construction; light construction equipment; information models

1 Introduction

BIM (Building Information Modeling) has become increasingly important in recent years [1]. The digital model of a building is at the center of BIM. This contains both three-dimensional geometric information and non-geometrical information such as materials, costs, and technical properties, and is therefore characterized by a high level of information depth [2].

Due to this development, BIM is also increasingly entering the focus of construction equipment manufacturers. Construction equipment is to be integrated with BIM so that efficiency potentials can be exploited and new business models can be developed. This will enable construction equipment to be increasingly integrated into the construction value chain, reducing costs and increasing safety and efficiency [3].

While this integration is fully underway for newly developed construction robot concepts and heavy earth-moving machines (see section 2.2), light construction equipment distinguished by an operating weight of up to

1.5 metric tons [4,5], like compaction plates, walk-behind dumpsters, and trench rollers, is excluded from this development, leaving safety, efficiency, and cost reduction potentials undiscovered [6,7]. Therefore, the object of this paper is to *investigate BIM-integration of light construction equipment*, specifically by (2) depicting the current state of the art and requirements regarding BIM-integration of construction equipment, (3.1) examining the requirement gaps between heavy and light equipment integration, (3.2) introducing a conceptual framework for the integration and (4) conducting a compaction case study evaluating the framework.

2 State of the art

With ongoing digitalization efforts, the digital construction ecosystem is becoming more complex. It is therefore useful to depict task management before diving deeper into its components. Generally, tasks on a construction site pass three levels: project management, work instruction, and execution [8]. Depending on the type of construction (building or infrastructure), project management revolves around either a BIM (e. g. *.ifc) or digital terrain (e. g. Land*.xml) model that is enhanced through multiple dimensions (5D) and simulations [9]. Between project management and the actual execution on site, the respective information model must be transformed through a Construction Site Control System (CS²) in the work instruction level to form an executable task. The task is executed by semi-automated heavy equipment such as machine-controlled excavators or autonomous robots. Figure 1 depicts the digital task management on future construction sites.

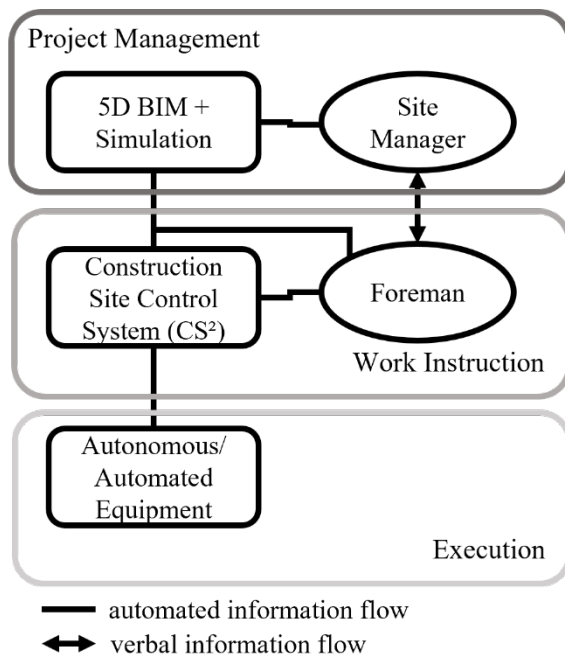


Figure 1: Task management ecosystem on construction sites [8]

A similar three-level structure is part of the ISO 15143-4, a norm under development focusing on topographical data exchange on mixed fleet worksites [10]. It standardizes the server-to-server data exchange between machine-specific Vendor Integration Systems (VIS), more commonly known as grade- or machine-control systems, and a general Site Management System (SMS) containing overall jobsite information and digital terrain models, as shown in Figure 2.

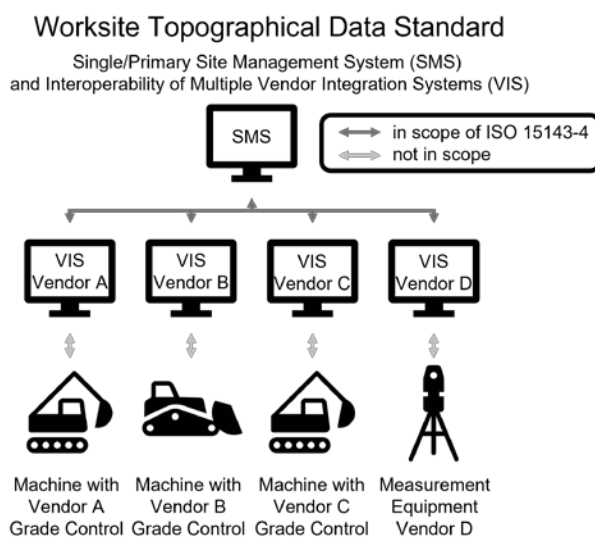


Figure 2: Structure of topographical data

exchange on construction sites according to ISO 15143-4 [10]

According to expert interviews on “building information modeling (BIM) in construction equipment scheduling and equipment-heavy operations” [11] with six equipment management professionals, BIM managers, and construction equipment developers conducted by the authors, five requirements must be met in order to integrate construction equipment into the BIM methodology:

The equipment

- must be able to receive, transform, and transmit data, and
- must align itself and data collected during operations with the information model.

The information model

- must contain information necessary for the task to be executed by the equipment, and
- must provide a sufficient information structure to include received information in a meaningful way.

The task

- must be sufficiently specified in both the information model and the software of the equipment.

After this process-oriented digression, the next section depicts the technological fundamentals.

2.1 Fundamentals

In the context of this work, building information modeling (BIM) is defined as the “use of a shared digital representation of a built asset to facilitate design, construction, and operation processes to form a reliable basis for decisions” [12]. The afore-mentioned information models (*.ifc and Land*.xml) comply with this definition. A BIM-model is an object-oriented representation of a building, and a digital terrain model (DTM) is a digital representation by means of a point cloud and a polygon mesh of existing or planned topographies.

In order to interact with these digital representations, automated equipment and robots must be able to align themselves with the information model in the digital environment. Therefore, automated equipment in open-field infrastructure construction use positioning systems such as differential global navigation satellite system (DGNSS) or real-time kinematics (RTK) [13,14], while autonomous robots or more automated equipment in building construction use localization algorithms such as simultaneous localization and mapping (SLAM) [15,16] or ultra-wideband (UWB) [17].

Construction equipment can be classified by size or purpose. This paper focuses on light construction equipment (LE). Light refers to equipment with an

operating weight of up to 1.5 metric tons. This includes mostly hand-held and non-ride-on equipment. Most of the newly developed construction robot concepts fall under this construction equipment segment.

2.2 State of research

The state of research covers two main areas of effort regarding BIM integration of construction equipment.

First, conventional heavy equipment is digitally enhanced to interact with information models.

Bouvet et al. [18] developed a real-time localization system for heavy compactors that maintains a positioning error lower than 0.2 m with a low-cost internal sensor set.

Heikkilä et al. [19] introduced eight modules to produce a fully autonomous compact excavator. These contain “open information modeling”, “positioning in accurate information system”, and “mission planning and work task creation” in regard to BIM-integration of construction equipment.

Yamamoto et al. [20] presented an autonomous hydraulic excavator that uses a simplified three-dimensional mesh of design data comparable to specific digital terrain models.

Halbach and Halme [21] developed 3D graphical job planning tools to support an autonomous wheel loader. Standard information models such as DTM or BIM are not supported by these tools.

Second, construction robotics researchers rely on information models for navigation and control purposes.

Follini et al. combined the Robot Operating System (ROS) with BIM for applications in construction logistics [22], specifically for navigation and task planning. Automatic updating of the construction process in the *.ifc-file is a desired future development goal [23].

Xu et al. [24] evaluated Hilti's Jaibot in a case study. The Jaibot collects design information from a modified BIM model. They found that characteristics of the information model (e. g., planning accuracy) influence the drilling performance.

Brosque et al. [25] compared manual and robotic concrete drilling for installation hangers. The deployed robot accesses BIM-models with a level of development (LOD) of 400 for design information.

Regarding light equipment, research projects and industrial applications are sparse. Light equipment vendors have upgraded sensor systems in recent years in order to obtain more data. An example is the coupling of telematics and compaction data, which allows the user to track the compaction progress in a web-application [26–28]. Since the obtained information is solely kept in the respective applications and cannot be aligned with an information model with a higher level of information or geometry, integration of light equipment and BIM is still missing.

3 BIM-Integration Concept

In order to counter the missing integration of light equipment and BIM, this chapter introduces a general BIM-integration concept in section 3.1. The concept is enhanced with light equipment specific requirements in section 3.2.

3.1 BIM-Integration of Light Equipment Concept

The concept for integrating BIM and construction equipment comprises five essential steps derived from the requirements mentioned in section 2 (see Figure 2).

1. First, the information necessary for the task must be included or generated in the information model. For this purpose, BIM or DTM software allows the user to check or edit the information model.
2. The information must then be exported from the information model, which can be done using the *.ifc or Land*.xml standard data formats. The data exported in this way should focus on data relevant to the construction equipment operation to keep data traffic low and must obey a predefined specification of the task so that the meaning of the data is conserved.
3. In order to work with the aforementioned information, the equipment must be technically able to receive and transmit information. Since light equipment moves freely and on different construction sites, a wireless network connection via tele- or radiocommunication is recommended.
4. Data transformation capabilities in the equipment ensure adequate interpretation of the received and the to-be transmitted data. In addition to this, light equipment must be able to align itself and to data generated in operation with the information model. Sensor systems along with localization and object detection algorithms provide this ability (see 2.1).
5. After transmitting data generated during operation from light equipment to the information model, it must be added to the information model for documentation or progress tracking purposes. In order to do so, the information model must follow an adequate information structure. BIM files in the *.ifc-format are object-oriented and allow information to be added to the respective objects. DTMs in the Land*.xml-format can either be updated (new file) or the information can be added to each geographical point. Another option for both information models is to include the information globally, e. g., in the file-header.

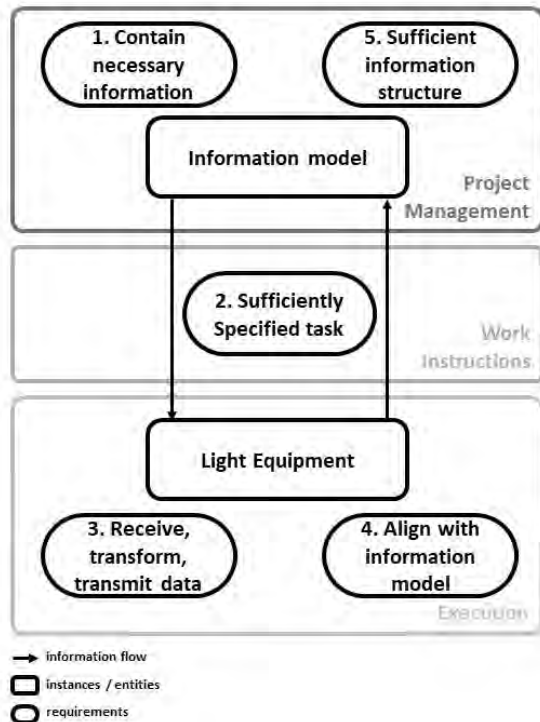


Figure 3: BIM-integration of light construction equipment concept

3.2 Light Equipment Requirements

Light equipment poses specific requirements for BIM integration. Since the mean price per machine in the light equipment segment is small compared to heavy equipment [29], expensive sensors increase the overall price. In order to keep a product compatible, sensor costs must be as low as possible.

Another characteristic of light equipment is its compact size and low operating weight. Therefore, the sensor system underlies strict space and weight limitations. All other requirements are similar to heavier equipment, e. g., robustness, easy commissioning, etc..

The presented concept for the integration of BIM and light construction equipment is evaluated in a compaction case study detailed in the next section. For better traceability, the respective numbers from chapter 3 are assigned to the individual steps below.

4 Case Study

To validate the concept described in section 3, the data exchange between an information model and a vibratory plate is implemented as an example. The objective is to specify a certain area to be compacted by the vibratory plate up to a certain compaction value with the help of an information model. This model combines the terrain information of a DTM with the property

parameters of the work task in a BIM model. This information is used to carry out the compaction with the vibratory plate and transmit information about the degree of compaction back to the BIM model.

4.1 Setup

The test area consists of a ballast bed framed by concrete blocks in which the vibratory plate can move freely.

The vibratory plate is transformed into a cyber-physical system that can process, acquire and document data using the hardware components shown in Figure 3, and the Robot Operating System (ROS) version 2 [30].

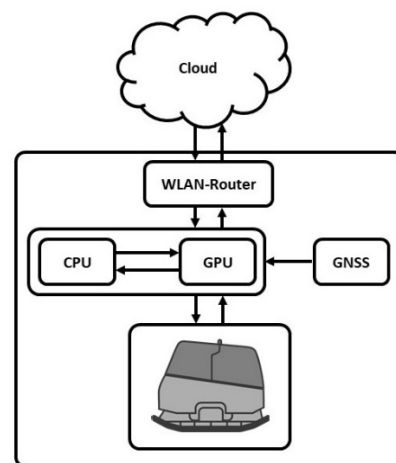


Figure 4: Hardware structure of the light equipment

A Central Processing Unit (CPU) and a Graphics Processing Unit (GPU) extend the vibratory plate via a machine interface, in order to be able to process the required information.

In addition, a GNSS module generates data for global localization. Since the accuracy of the GNSS position depends on external influences, position determination can be improved with DGNSS (meter accuracy) or Real Time Kinematics (RTK) (centimeter accuracy) [31]. This involves using a base station to provide correction data that attenuates the environmental influences on the GNSS position (see Fig. 4). The correction data can be obtained from a service provider as well, but this does not offer the same level of accuracy as a base station. In this study, RTK delivers positioning data with a mean accuracy of 1.5 to 3.0 cm.

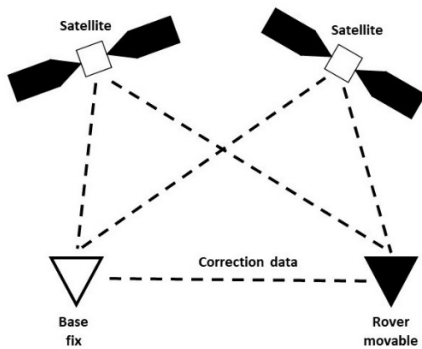


Figure 5: DGNSS and RTK concept

Via a WLAN interface and a router, the vibratory plate communicates with an external computing unit.

4.2 Setup and Execution

In accordance with Figure 2, Figure 6 shows the five necessary steps for the BIM-integration of light equipment. The information flows include the deployed data formats. The individual steps are described in detail below.

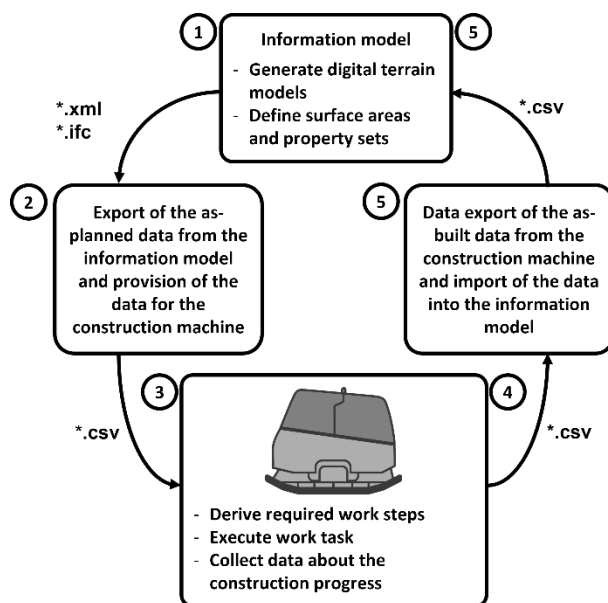


Figure 6: Overview Case Study Operation Setup

1. The first step in the realization of the case study was to create a digital terrain model of the construction area with the test area in (blue circle – see Fig. 7). This was done using the Autodesk Infrastructure and Civil 3D applications, which were used to generate the DTM from a point cloud.

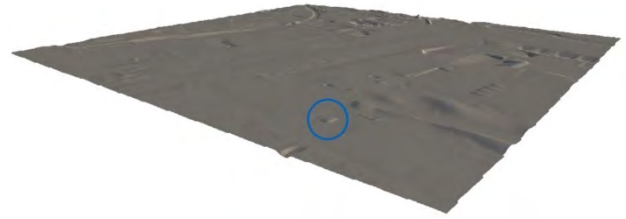


Figure 7: Digital terrain model of the construction site

Subsequently, a coordinate-based surface was defined in Civil 3D, to which a compaction value was assigned. The resulting surface is shown in Figure 7, where a bird's eye view of the construction site gives an overview of the entire test area.

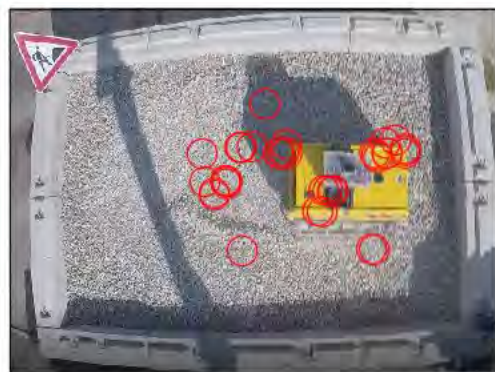


Figure 8: Digital terrain model with compaction factor as input

2. The data generated in this way was exported from Civil 3D using the *.ifc data format. The file obtained in this way was limited to the relevant information and converted into a data format that is readable by the machine. For this purpose, a python script was used. The resulting *.csv-file was then transferred to the vibratory plate.
3. As described in chapter 4.1, the vibratory plate is

able to process and use the information and to move independently. By measuring the superstructure acceleration of the vibrating plate, conclusions about the current compaction state can be drawn.

4. Based on these measurements, the vibratory plate executed the given work task. During execution, the compaction value was obtained locally by the vibratory plate. Matched with the GNSS position, the compaction value is stored locally in a documentation *.csv-file, which has the structure of a point cloud.
5. This file was then transmitted by the vibratory plate and imported directly back into Civil 3D. The result of the import, individual compaction points and the associated compaction values is shown in Fig. 8. The scattering of the compaction values is due to the incompatible soil used in the test bed.



Point #	Elevation	Northing	Easting	Compaction Value
5	477.00	5349101.26	698051.09	2.0
6	476.90	5349101.26	698051.09	6.0
7	476.90	5349101.24	698051.09	8.0
8	476.30	5349100.63	698051.04	5.0
9	476.60	5349099.43	698050.66	3.0
10	476.70	5349099.45	698050.70	4.0
11	476.60	5349099.47	698050.73	7.0
12	476.30	5349101.09	698051.08	4.0
13	476.50	5349101.78	698051.10	5.0
14	476.90	5349099.68	698049.44	3.0
15	476.80	5349099.72	698049.45	5.0
16	476.50	5349101.45	698049.71	8.0
17	476.80	5349101.80	698050.71	7.0
18	476.50	5349100.35	698050.08	5.0
19	476.50	5349100.30	698050.06	7.0
20	476.30	5349099.30	698050.85	2.0
21	476.20	5349099.16	698050.87	4.0
22	476.00	5349099.06	698050.67	3.0
23	476.00	5349099.06	698050.71	3.0
24	476.00	5349099.04	698050.73	6.0

Figure 8: Data points with position and

compaction factor as output

4.3 Evaluation of requirements specific to light equipment

The light equipment specific requirements mentioned in section 3.2 of robustness, easy commissioning, light weight, low cost, and compact size are evaluated for the case study. Table 1 shows an overview of all necessary additional components. The requirements of robustness and easy commissioning are fulfilled by these components for a prototyping/concept purpose, but could be improved in future works or in mass-production. Weight is negligible in the case of a vibratory plate (793 kg) due to the high weight of the base plate. The hardware system needed for the mentioned functionality has a total cost of less than 600 € (566,90 €) and requires a space of roughly 300 x 200 x 100 mm. The major cost drivers are the two needed development boards, which are RTK capable GNSS receivers.

Table 1. Hardware components size and cost

Component	Size [mm]	Cost [€]
CPU + GPU	85 x 56 x 16	77,90
Router	144 x 230 x 37	33
2 x RTK/GNSS receiver	110 x 55	456 (228 each)

The cost and size of the integration is small relative to the vibratory plate's dimensions (1.183 x 870 x 830 mm) and market price (10,000 €). Therefore, the requirements of low cost and compact size are met in this case study.

5 Conclusion

The paper at hand investigates BIM-integration of light construction equipment. Therefore, the state of the art and requirements of BIM-integration of construction equipment is depicted. With light equipment specific requirements in mind, an integration concept is introduced. Finally, the concept is successfully validated in a compaction case study.

The case study revealed that an integration of light equipment and BIM methodology can comply with the requirements in practice. Since the case study required adaptations to the information model and light equipment as well as introduction of a standardized task description, future research should focus on the introduction of transparent interfaces between the information model and light equipment, as well as standardized task descriptions. The ISO 15143-4 standard for the exchange of worksite topographical data, which is currently under development, could prove to be an adequate starting point [10]. The industry should evaluate serial

development of the solution to become part of mass-produced light construction equipment.

6 References

- [1] A.J. Spengler, J. Peter, *Die Methode Building Information Modeling*, Springer Fachmedien Wiesbaden, Wiesbaden, 2020.
- [2] A. Borrmann, M. König, C. Koch, J. Beetz, *Building Information Modeling*, Springer Fachmedien Wiesbaden, Wiesbaden, 2015.
- [3] C. Terol, *BIM and construction machines: What are their capabilities?*, 2020.
<https://www.globalcad.co.uk/bim-and-construction-machines-what-are-their-capabilities/> (accessed 28 January 2022).
- [4] Ammann Group, *Light Equipment*, 2022.
<https://www.ammann.com/en/machines/light-equipment> (accessed 28 April 2022).
- [5] Bomag GmbH, *Light equipment for soil and asphalt compaction*, 2022.
<https://www.bomag.com/ww-en/machinery/categories/light-equipment/> (accessed 28 April 2022).
- [6] Leica Geosystems, *Working Smarter, Not Harder: Using Technology to Level the Playing Field*, White Paper, 2021.
- [7] S. Benko, *You're Using Machine Control Technology. Now What?*, 2022.
<https://constructible.trimble.com/construction-industry/youre-using-machine-control-technology-now-what> (accessed 28 April 2022).
- [8] M. Schöberl, D. Bartmann, S. Kessler, J. Fottner, *Towards a Construction Site Control System ? Task Management in Construction Operations and Intralogistics*, in: *Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC)*, Dubai, UAE, International Association for Automation and Robotics in Construction (IAARC), 2021.
- [9] A. Fischer, Z. Li, S. Kessler, J. Fottner, *Importance of secondary processes in heavy equipment resource scheduling using hybrid simulation*, in: *Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC)*, Dubai, UAE, International Association for Automation and Robotics in Construction (IAARC), 2021.
- [10] T. Costlow, *New standard will ease grade-control data sharing between construction machines*, 2020. <https://www.sae.org/news/2020/02/new-iso-standard-for-construction-sites> (accessed 28 January 2022).
- [11] M. Schöberl, L. Herr, S. Kessler, J. Fottner, *Building Information Modeling (BIM) in Construction Equipment Scheduling and Equipment-heavy Operations*, in: *Proceedings of the 9. Fachtagung Baumaschinentechnik 2022*, Dresden, Germany, 2022 (to be published).
- [12] International Standardization Organization, *ISO 19650-1:2018-12: Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) - Information management using building information modelling - Part 1: Concepts and principles 35.240.67, 91.010.01*, 2018.
- [13] J. Mallela, A. Mitchell, J. Gustafson, M. Olsen, C. Parrish, D. Gillins, M. Kumpula, G. Roe, *Effective Use of Geospatial Tools in Highway Construction: Technical Report*, 2018.
- [14] D. Persson, *Surveying in the Construction Industry: A study of surveying and machine guidance systems in excavators on ground construction projects*. Master's Thesis, Gothenburg, 2018.
- [15] M. Immonen, I. Niskanen, L. Hallman, P. Keranen, M. Hiltunen, J. Kostamovaara, R. Heikkilä, *Fusion of 4D Point Clouds From a 2D Profilometer and a 3D Lidar on an Excavator*, *IEEE Sensors J.* 21 (2021) 17200–17206.
<https://doi.org/10.1109/JSEN.2021.3078301>.
- [16] Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, *Algorithm Toolbox for Autonomous Mobile Robots*, 2022.
<https://www.iosb.fraunhofer.de/en/projects-and-products/algorithm-toolbox.html> (accessed 27 January 2022).
- [17] F. Vahdatikhaki, A. Hammad, H. Siddiqui, *Optimization-based excavator pose estimation using real-time location systems*, *Automation in Construction* 56 (2015) 76–92.
<https://doi.org/10.1016/j.autcon.2015.03.006>.
- [18] D. Bouvet, M. Froumentin, G. Garcia, *A real-time localization system for compactors*, *Automation in Construction* 10 (2001) 417–428.
[https://doi.org/10.1016/S0926-5805\(00\)00077-7](https://doi.org/10.1016/S0926-5805(00)00077-7).
- [19] R. Heikkilä, T. Makkonen, I. Nishanen, M. Immonen, M. Hiltunen, T. Kolli, P. Tyni, *Development of an Earthmoving Machinery Autonomous Excavator Development Platform*, in: *Proceedings of the 36th International Symposium on Automation and Robotics in Construction (ISARC)*, Banff, AB, Canada, International Association for Automation and Robotics in Construction (IAARC), 2019.
- [20] H. Yamamoto, M. Moteki, H. Shao, T. Ootuki, H. Kanazawa, Y. Tanaka, *Basic Technology toward Autonomous Hydraulic Excavator*, in: *Proceedings of the 2009 International Symposium*

- on Automation and Robotics in Construction (ISARC 2009), Austin, TX, USA, International Association for Automation and Robotics in Construction (IAARC), 2009.
- [21] E. Halbach, A. Halme, Job planning and supervisory control for automated earthmoving using 3D graphical tools, *Automation in Construction* 32 (2013) 145–160. <https://doi.org/10.1016/j.autcon.2013.01.017>.
- [22] C. Follini, M. Terzer, C. Marcher, A. Giusti, D.T. Matt, Combining the Robot Operating System with Building Information Modeling for Robotic Applications in Construction Logistics, in: S. Zeghloul, M.A. Laribi, J.S. Sandoval Arevalo (Eds.), *Advances in Service and Industrial Robotics*, Springer International Publishing, Cham, 2020, pp. 245–253.
- [23] C. Follini, V. Magnago, K. Freitag, M. Terzer, C. Marcher, M. Riedl, A. Giusti, D.T. Matt, BIM-Integrated Collaborative Robotics for Application in Building Construction and Maintenance, *Robotics* 10 (2021) 2. <https://doi.org/10.3390/robotics10010002>.
- [24] X. Xu, T. Holgate, P. Coban, B. García de Soto, Implementation of a Robotic System for Overhead Drilling Operations: A Case Study of the Jaibot in the UAE, in: *Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC)*, Dubai, UAE, International Association for Automation and Robotics in Construction (IAARC), 2021.
- [25] C. Brosque, G. Skeie, M. Fischer, Comparative Analysis of Manual and Robotic Concrete Drilling for Installation Hangers, *J. Constr. Eng. Manage.* 147 (2021) 5021001. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0002002](https://doi.org/10.1061/(ASCE)CO.1943-7862.0002002).
- [26] BOMAG GmbH, BOMAP: Smart helper on construction sites, 2022. <https://www.bomag.com/us-en/technologies/overview/bomap/> (accessed 27 January 2022).
- [27] Wacker Neuson SE, DPU110 - Uncompromisingly functional, 2022. <https://www.wackerneuson.de/en/products/compaction/vibratory-plates/reversible-vibratory-plates/model/dpu110/>.
- [28] Ammann Switzerland Ltd, ACEecon Offered on Ammann Reversible Plates, 2017. <https://www.ammann.com/en/news-media/news/aceecon-offered-on-ammann-reversible-plates> (accessed 27 January 2022).
- [29] Bauverlag BV GmbH, Construction equipment register (BGL), 2020. <https://www.bgl-online.info/en/basic-information/> (accessed 28 April 2022).
- [30] Open Robotics, ROS 2 Documentation: Foxy, 2022. <https://docs.ros.org/en/foxy/index.html> (accessed 27 January 2022).
- [31] Institute of Flight System Dynamics, TUM, Satellite Navigation: Differential GNSS, Realtime Kinematics (RTK), 2022. <https://www.fsd.ed.tum.de/research/sensors-data-fusion-and-navigation/research-and-competence-areas/satellite-navigation/#DifferentialGNSS> (accessed 27 January 2022).

Visualisation and graph-based storage of customised changes in early design phases

D. Napps^a, A. Zahedi^b, M. König^a and F. Petzold^b

^aChair of Computing in Engineering, Ruhr University Bochum, Germany

^bChair of Architectural Informatics, Technical University of Munich, Germany

E-mail: daniel.napps@ruhr-uni-bochum.de, ata.zahedi@tum.de, markus.koenig@ruhr-uni-bochum.de, petzold@tum.de

Abstract

Building Information Modeling (BIM) was conceived as a working method for networked planning using software assistance. Multiple stakeholders in the construction industry are involved in the design and management of digital representations of the physical and functional characteristics of a facility in the BIM model.

Design variant management coupled with documentation and recording of design knowledge are the main goals of this paper. The use of a BIM-based plugin for design rationale documentation grants support for other architects in the early design phases, to retrieve suitable building designs and ideas, based on their needs and architectural concepts. Uniting Design Episodes with different Variant Types allows the architects to receive inspiration and freely select important parameters for their design of the building in advance. These input parameters are compared with other BIM-based design graphs, following the graph representation approach.

The relevant design options and variants corresponding to the new project are visualised and stored for both current and future users. This can lead to cost and time savings and allow the possibility for future extensibility in terms of other analysis parameters (such as the level of detail or energy efficiency).

Keywords –

Building Information Modeling (BIM); Industry Foundation Classes (IFC); Early Design Phases; Design Decisions; Graph Representation; Design Episodes; Variant Types; Plugin

1 Introduction

1.1 The problem statement

In the early design phases of buildings, the cooperation of interdisciplinary experts with architects and civil engineers is essential for considering the complex aims and requirements of a building. The architect attempts to find an efficient and aesthetic solution for the future project with the given specifications (e.g., sustainability) and evaluates a variety of designs for a building, which can result in a significant component of the project phase which is, on the one hand, only temporary but, on the other hand, can incur high expenses. These designs must be subsequently continuously coordinated and adapted with the other project participants [1]. In order to avoid the architects from designing basic concepts, such as floorplans, construction types or individual elements for each upcoming building, it is essential to provide a high-performance efficient solution to support design reuse.

Currently many architects are using Building Information Modeling (BIM) because of its benefits for exchanging models with other stakeholders. In particular, BIM includes all important information and visualisations of the building [2]. Nevertheless, for the different requirements and comparisons, several BIM models are made during the architecture process in order to evaluate the advantages and disadvantages of the various designs of the building [3]. In addition, the requirements for a new project may be vague or detailed, depending on the components of the construction, which in turn complicates the architect's work. An insufficient knowledge base for integrating previous experience is especially challenging for architects who have not been in practice for years [4], as they have either no or only limited experience with similar projects. Inadequate documentation of the early design process and an absence of transparency (in case of subjective

estimations) can result in mistakes [5], when adapting the processes of previous architects to a new project, and is thus problematic even for experienced architects.

1.2 Previous research and proposed solution

During the past years of research, a Variant Type concept was developed, which allows individual elements in a BIM model to be classified. A classification is made between three Variant Types, namely structural, functional and product variants [6]. In addition, a graph representation of the variants was established. This representation is based on the structure of the Industry Foundation Classes (IFC) according to IFC4 [7].

Recently, a concept for the retrieval process of similar building designs in the early design phase of a building was presented that incorporates this diversity of variants. This process is part of the case-based reasoning approach, where problems are solved using problems that have already been solved and proven. The findings provide an example of how an architect can select the most suitable graph for the upcoming building project from a graph database [8].

Design decisions are based on both objective and subjective criteria. Proper documentation of the design process demands recording qualitative and subjective assessments and decisions. To address this, Zahedi et al. [9] introduced the concept of Explanation Tags to describe and elaborate the subjective aspects of design decisions and enhance the design documentation and its future readability and reuse. Furthermore, Zahedi et al. [9] presented the concept of Design Episodes to store various pieces and sections of design in corresponding segmentations that enable the traceability and reuse of these design ideas and concepts to address future design tasks and problems.

Graph structures are useful in various disciplines to analyse and extract information. In the construction industry, they are used in the context of BIM because of their capacity to describe complex digital models and internal relationships [10].

1.3 Objectives and scope

The research presented in this paper aims to identify a possibility to make the architect's work more comprehensible in the early design phases of a building. In particular, the use of a plugin that can be implemented in a BIM-enabled environment which captures the knowledge of previous architects, as well as subjective design decisions related to existing buildings.

The objective in using the tool is to capture design requirements for the building and to store and manage the information, changes and options using IFC.

Focusing on the following two areas is of particular importance:

- Retrieval of individual Design Episodes and Variant Types that match the selected input parameters in order to search for inspiration regarding a new project.
- Ensure consistency of the entire design or individual Design Episodes through pattern matching of the individual BIM model graphs.

An efficient way to apply this graph-based method has its origins in the field of mathematics and has already been successfully used in the building sector as well [11].

The graph representation of BIM data models is generated based on the IFC data structure. Existing building graphs are compared with the design graphs described from the input parameters in terms of similarity and thus provide the architect with relevant solutions that give inspiration coming out of an existing case. As a result, the scope of this concept is particularly flexible, as the scale of view can be set in such a way that both entire buildings and individual design episodes, right down to variant types, can be compared using the plugin.

2 Background

In the following, the background information on which the methodological approach is based and which contributes to the understanding of the case study is specified. The importance of BIM and the IFC interface is briefly outlined.

The comprehensive digital representation of a building information model is widely used in the construction industry. In contrast to the outdated analogue form of building drafts on paper, the use of BIM allows, among other things, the capture of three-dimensional geometries, non-physical elements and semantic information. In particular, it enables the stakeholders involved in the project to use and exchange the same, consistent, up-to-date knowledge, and it covers the entire life cycle of a building (design, planning, construction, management, refurbishment) [2].

Industry Foundation Classes as an international standard (ISO 16739) is important for the exchange of the BIM data. It is an object-based file format developed and continuously updated by buildingSMART. IFC is described as an entity-relationship model that can represent all entities of a building that are organised and interconnected in an object-based inheritance hierarchy [12]. The IFC structure is relevant in this context because the graph representation of the BIM model is based on its internal architecture.

2.1 Graph representation

Existing graph-based publications for structured data can be divided into two principal categories. The first category covers Resource Description Framework (RDF) graph models. The second includes property graph models (PGM) [13].

The RDF approach is often used related to linked data and declares graphs as a triple (subject, predicate, object), see figure 1.

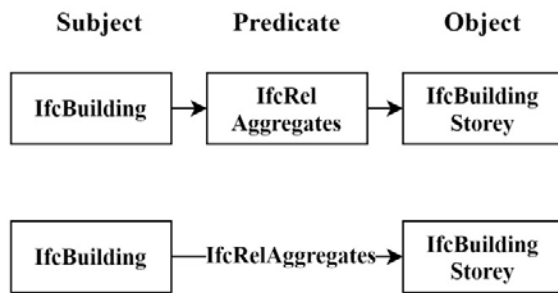


Figure 1. RDF representation of IFC Entities

Because of the Uniform Resource Identifier (URI), the graphs are unique but do not have an internal structure. As used in the construction industry when comparing entire BIM models with each other, the file size becomes problematic when exporting complex buildings [14].

In parallel the (labelled) property graph (LPG) or entity centred graph approach is used for object-oriented programming and overcomes the limitations of the triple centred approach. In each edge and node, information, attributes or properties can be stored (Figure 2).

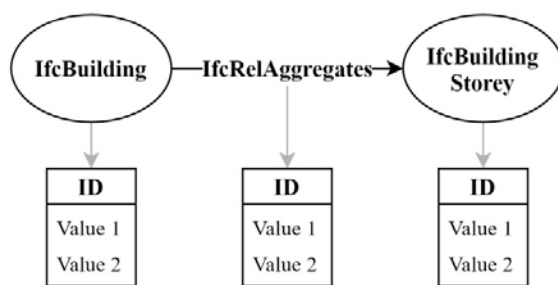


Figure 2. LPG representation of IFC Entities

Zhao et al. recently presented an approach, in which they described a method for the intact representation of relationships between IFC entities as an essential requirement for the correct classification of IFC entities [15]. Differences to the RFD approach are that LPGs are characterised by an internal, more reduced structure. Instances of relationships of the same type can also be

identified so that they can be qualified or attributed [16].

Using graph pattern matching, several graphs or subgraphs can be compared with each other.

2.2 Variant management and pattern matching

This work is intended to contribute to the previous knowledge in this topic, where a BIM-based solution for the retrieval process in the early design phase of a building was identified [6]. Searching for similar floor plans currently operates on the basis of filtering with respect to size, type of use, number of rooms and other parameters. A combination of Variant Types and defined Design Episodes has not been made. Structure variants offer options to the structure of the building, for example the geometry of the elements. Functional variants refer to the functions, e.g. a room layout or structural engineering elements. The product variants include individual objects (e.g. windows) that can be exchanged with similar objects or other property values.

Due to a similarity calculation, the upcoming project is matched with a similar building from a database of existing designs to provide inspiration for architects in early design phases [6]. Graph pattern matching is used for the matching method. This method has been widely accepted in understanding and accessing network data and works for both graph representations. Searching for a similar graph in a graph database, graph pattern matching finds the answers A as a set of graphs from D [17].

Based on a sample query, matching graphs are selected from a database that provide an answer to the existing query. There are two approaches for this process: exact pattern matching and inexact pattern matching [8]. The first type is characterised by a one-to-one isomorphism of two examined graphs. In the second approach, similar subgraphs are examined in the overall graph.

2.3 Explanation Tags and Design Episodes

To explain and elaborate on certain aspects of design, an open-ended collection of Explanation Tags is offered within the plugin that cover many design criteria and concepts and may be used to assign to various design elements or their specific properties to justify and clarify the rationale behind those details and decisions. A particular focus has been on sustainability requirements in planning competitions, while designing the first set of Explanation Tags [18].

Certain parts of design can be documented and stored as Design Episodes. Each Design Episode in the plugin consists of multiple design elements, e.g., building components or spaces, and a name and description explaining this design chapter or situation. A

Design Episode can contain a solution for a design problem or a template to address a design case or situation [18].

3 Methodology

3.1 Definition of property graphs

A labelled property graph consists of nodes or vertices (V), relationships, properties and labels. Relationships, that can also be described as edges (E), connect nodes and structure the graph, by directing from a start point to end point. Both, nodes and relationships can have properties. Labels (L) group nodes together and specify their role within the dataset [19].

Resulting from this, a graph can be described mathematically [20].

$$G = (V, E) \text{ with } V = (v_1, v_2, v_3, \dots, v_i) \text{ and } E = (e_1, e_2, e_3, \dots, e_i)$$

Generated subgraphs can be either labelled or attributed.

$$G = (V, E, L) \text{ with } L_v = (l_{v_1}, l_{v_2}, l_{v_3}, \dots, l_{v_i}) \text{ and } L_e = (l_{e_1}, l_{e_2}, l_{e_3}, \dots, l_{e_i})$$

The importance of the vertices and relations can be determined by weighting all elements of the respective set.

$$W = (v) \forall v \in V \text{ and } W = (e) \forall e$$

Definitions below refer to the elements described:

- Vertices (V) of the graph represent the IFC entities of the BIM model. All entities in a digital model (e.g. IfcBuilding or IfcWall) can be considered as virtual physical objects and differ from each other. The label of the vertex, stored in the nodes, is the entity's type.
- An edge (E) of a graph is a relation, i.e. an interaction between two entities. To distinguish between the two directions a relationship may have, we define so-called directed graph edges.
- Properties represent features of entities and relationships. In a LPG this information is stored in the vertices and edges.

3.2 Plugin

Zahedi and Petzold [18] developed a plugin for design documentation, that provides two main capabilities for the users. The first is to use Explanation Tags to enhance and elaborate on the argumentation and

rationale of design decisions while also providing the ability to explore and find various assigned tags in any given BIM model and to manipulate and change them if need be. Furthermore, is the ability to extend and add to the open-end collection of tags based on users' specific needs, while being able to import or export custom-designed tags as an archive. The second capability is the ability to create and manipulate Design Episodes. This plugin is designed for Autodesk Revit.

Furthermore, the plugin enables the export of Design Episodes to a graph database such as Neo4j. When exporting to Neo4j, the user has the choice to select the desired properties of various elements of a Design Episode to be exported or not. The result is a LPG that contains the Design Episode as the parent node, and Episode Elements, as child nodes connected via Episode Element edges. The Design Episode nodes have, as attributes

- a name and a description to explain the purpose of this episode using storytelling techniques,
- a model identifier for tracing the original BIM model that this episode originated from,
- and a Global Unique Identifier (GUID) to help distinguish between Design Episodes with similar names.

The design element nodes have, as attributes, an ID, a Unique ID, and an Object Name as the basis, as well as every other property from Revit that the user has decided to be exported to Neo4j. When the elements in Revit are directly connected to each other, the resulting design element nodes are also connected with an IsConnected edge. When an element is contained in another one in Revit, the corresponding design element nodes in Neo4j are connected via a ContainedIn edge. When the elements in Revit are adjacent to each other, the IsAdjacent edge is used to connect the exported element nodes in Neo4j.

3.3 Matching

Graph pattern matching and pattern construction are important tasks in the comparison of two graph representations. Based on the definition of property graphs, graphs assume characteristics from each other. Comparing graphs requires that the graphs are connected. Thus, a graph always consists of nodes that are connected by edges. (Figure 3).

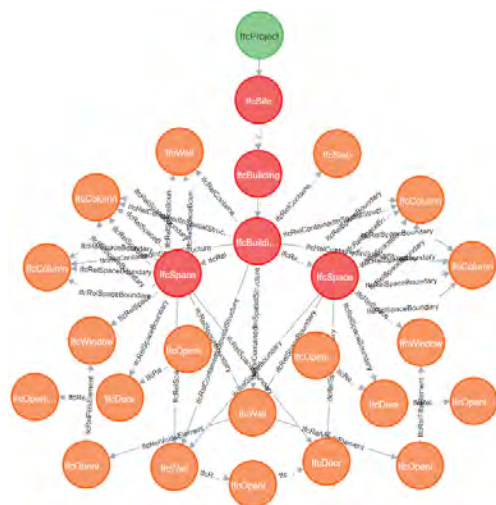


Figure 3. Graph representation of the IFC structure

The inexact pattern matching, mentioned at the beginning, accommodates the heterogeneity of the building industry, because two buildings are never exactly similar.

Using a pattern query, the substructures of a graph G can be analysed for matches using a query function. An inexact graph pattern match $I_{\text{gpm}}(G)$ of a pattern query I in a data graph G is a subgraph G^* of G . Both vertices V and edges E of the graph G are analysed during the matching process [18].

The subgraphs that are searched for can be simplistically found as the result of the following command:

Find all matches in $G = (V, E, L)$ with a given query $I = (V_I, E_I, L_I)$.

Combining Design Episodes and Variant Types enables querying a subgraph matching of different building graphs in various ways. This approach enables searching directly for defined Design Episodes or for individual optional Variant Types for specific rooms or storeys.

4 Case study

In the following, the two approaches of Zahedi et al. [9] and Napps et al. [8] are combined. In terms of further research, this results in a potential filtering for finding building designs. Furthermore, architects will be provided with a tool to define Variant Types and Design Episodes in a BIM model to save them transparently and comprehensible for other users. In order to illustrate this process, a sample Revit project file is used provided by Autodesk [21]. Three BIM models are created, each with different functional and structural variants that

belong to a defined Design Episode. Focussing only on the two Variant Types mentioned is due to the fact that the process for the product variant is comparable to that of the functional variant.

4.1 Define Variant Types and Design Episodes

The introduced plugin is used to define an exemplary Design Episode in all three BIM models. For this purpose, the Explanation Tags are used to define specific attributes of a Design Episode. The assignment is based on the subjective decisions of the architect.

Components within this episode receive additional Explanation Tags, defined by the architect in order to create a design concept. To add the variants to the model, the plugin is first extended by the three new Explanation Tags, each representing a Variant Type (Figure 4).

Design Documentation

Explanation tags | Design episodes

Tag	Topic	Synonyms	Description
	Aesthetics		Venustus as one of three foundations and essential components of successful architecture by Vitruvius (the ancient Roman architect), regarded as beauty or delight, is responsible for aesthetic quality, imparted style, proportion and visual beauty.
	Privacy		"To a highly general sense, the interior stands for privacy, possession and self-gathering, the exterior for the public sphere, availability and display." Privacy could be defined and integrated through other terms and concepts such as accessibility and exclusivity, protection, cell, facade, inside and outside, residence, screening, territory, view instead of, closure to extensive openness, the requirements of separation.
	Structure variant		Building elements, that effect the structure of the building and the outline of the building.
	Product variant		Small-scale allocation of the building elements that may be interchangeable but do not effect the structure and function of the building (e.g. a replacement of a door or window with the same dimension).
	Function variant		Building elements, that effect the function of the building.

Figure 4. Creating new Explanation Tags for Variant Types

During the next phase, different variants are implemented in the three models and identified with new tags. These identifications are saved, can be retrieved for further tasks and are available for other users.

Each Variant Type is identified by its colour (Figure 5-7). Included in the Design Episode are the dining room and the kitchen, bathroom and mechanical room. Individual elements that are located in this Design Episode are provided with tags, so that many options are generated.

The function variant changes from a wall with an open doorway separating the kitchen from the dining area (model one) to a glass wall including a door (model two) and in case of model three, to a load-bearing column without a separating effect. Equally, the structure variant is changed and thus varies between a glass curtain wall, an energy-efficient wall and an extension of the room structure.

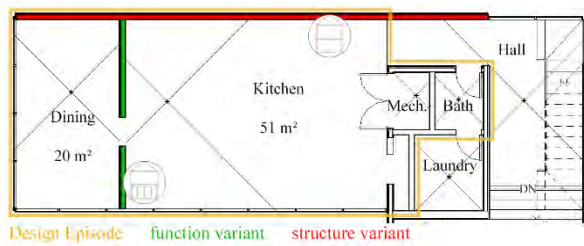


Figure 5. Defining Variant Types in the first BIM model

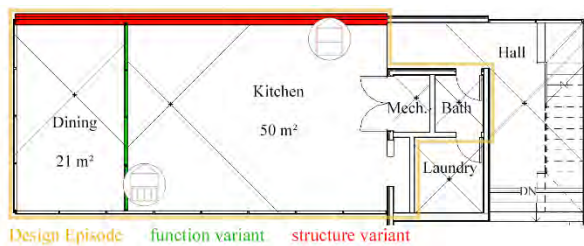


Figure 6. Defining Variant Types in the second BIM model

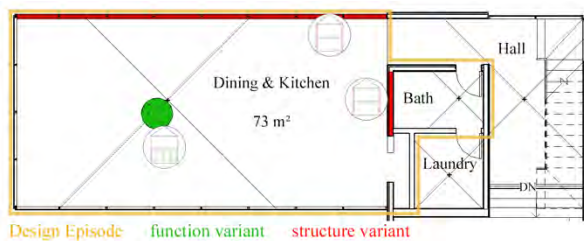


Figure 7. Defining Variant Types in the third BIM model

The arguments and justification for each variant case, whether it be functional or structural in the examples above, could be discussed and saved as an episode description in the Revit plugin.

Choosing a glass wall over a regular wall, for example, will provide openness and transparency, whereas the other choice will provide privacy.

4.2 Transfer BIM model to graphical representation

Neo4j is used to realise the graph-based representation of the BIM model. Exporting the entire BIM model as an IFC4 file is possible in order to import the data into Neo4j, but not necessary, as the plugin offers a more efficient way. Using this interface, the Design Episodes are displayed correctly. The user is able to directly export Design Episodes to Neo4j, since this feature is implemented as part of the plugin.

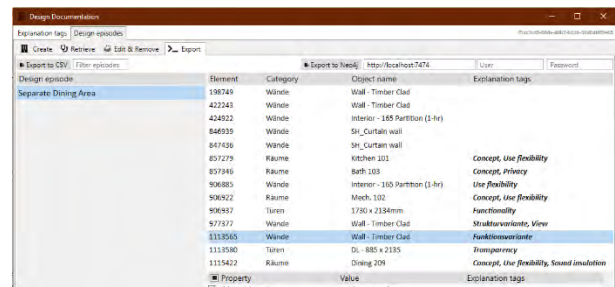


Figure 8. Set and export Design Episodes

Converting the model into a graph network allows an analysis of the relationships between the individual properties of a complex model. This is an effective feature for exporting subgraphs from the entire BIM model because the Design Episodes are part of the entire building. For a graphical comparison of the Design Episode, it is exported from all three BIM models, which ensures different Variant Types in the subgraphs. By the end of this process, three different subgraphs will have been loaded into Neo4j.

4.3 Visualisation and application

After a Design Episode with different Explanation Tags has been created and exported (Figure 8), the architect is able to choose possible matching variants, stored in a database, based on different requirements. Continuing, the Design Episode can be visualised by its graphical structure.

Users of the plugin are either able to create variants themselves, that contain other functions in the Design Episode or they can retrieve variants for this type of Design Episode that have already been stored in the database.

Within this case study, the basic assumption is that there are three different options stored in the database for a similar Design Episode.

In the case the architect requires a separation of the kitchen and the diner in terms of noise and smell and additionally a panoramic view to the outside, the Variant Types of model one are the most suitable alternative for the project.

The vertices of the graph that are a function variant are marked with a green circle. The nodes that are a structural variant are circled in red. The Explanation Tags can also be listed in the names of the nodes, but in this case it was decided to represent the entities in order to reflect the structure of the model.

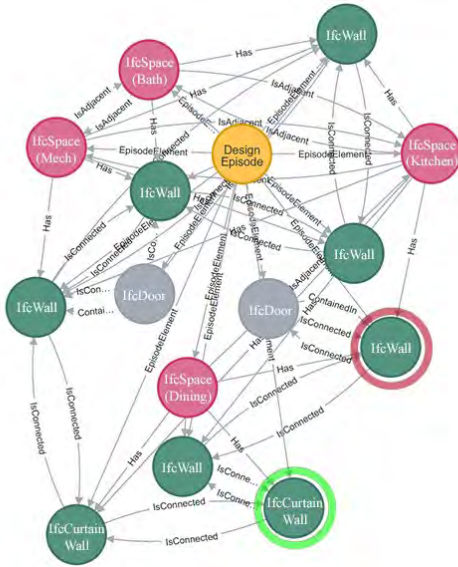


Figure 9. Graph representation of BIM model 1

If it is important for the client to have a separation between the kitchen and the diner, but desires a visual connection between the rooms and would like to build as energy-efficiently as possible, then the Variant Types from model two correspond to the ideas.



Figure 10. Graph representation of BIM model 2

A request for a shared kitchen and dining area can be supplemented by static adjustments (e.g. a column). Other room modifications that affect the structure can also be found, for example if a larger bathroom is desired and another room is only optional (Figure 7).

Adjustments made via the variant management can have large or small effects on the graph structure (Figure 9-11).

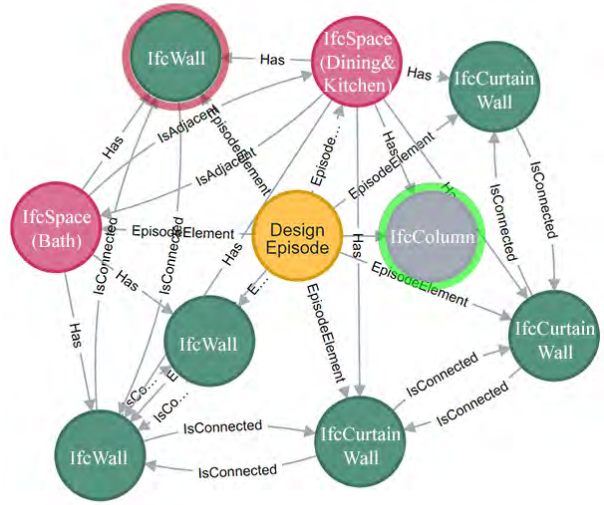


Figure 11. Graph representation of BIM model 3

There is a difference between the three graphical representations of the digital models. In a design selection process by a user, only the most relevant Design Episodes are selected according to the input conditions for the project or Design Episode.

Several Variant Types for an entity can be grouped into option nodes, which makes it easier for the user, as he or she is directly shown possible options on the basis of the Explanation Tags made for a Design Episode.

5 Conclusion and future work

In this paper a way for documenting design decisions by using the announced Revit plugin in the early design phases was introduced to retrieve suitable building designs and ideas, based on architectural needs and concepts. This workflow is based on the traceability of earlier design decisions, leading to time and cost savings for future similar projects.

Users of the plugin may specify a Design Episode, which can be assigned with various Explanation Tags. This results in important design decisions for the architect, which can be entered and stored directly in a BIM model. Based on these design decisions, architects are able to define Variant Types in the model. It has been demonstrated that depending on the Explanation Tags, different Design Episodes can be shown to the architect. This inspiration can be refined by different variations without changing the basic settings of the Design Episode, if these are stored in an option node.

Pattern matching allows similar design variants to be retrieved from a database that match the client's requirements. In addition, options from a variant pool are displayed, allowing the architect to make decisions based on the design decisions at the beginning of the

project. Thus, the user is inspired by different options, which fit the requirements of the new building or the defined Design Episode.

It is to be investigated whether the graphically simplified relationship descriptions are sufficient for more complex applications in the construction industry. Alternatively, the relationship notation of the IFC structure can be used. This is important for the similarity calculations in the retrieval process of the CBR approach.

The benefit of the tool for architects is limited by the voluntary use and implementation of the tool in Revit. A large overall benefit will only arise if many architects use the plugin regularly and save the designs in a database for other users. In addition, future work might take an approach that permits an automatic assignment of Variant Types, so that the user's effort is limited. Following on from this, the integration of additional area-specific topics such as structural design, detailing and energy efficiency is to be evaluated in order to optimise the model, as well as an appropriate visualisation of the design changes.

References

- [1] Darwich, A. K., Means, J. K. and Lawrence, T., ASHRAE GreenGuide. Design, construction and operation of sustainable buildings, Eds. 2018. ASHRAE, Atlanta, GA, 2018.
- [2] Borrmann, A., König, M., Koch C. and Beetz, J., Building Information Modeling. Technology Foundations and Industry Practice, Eds. 2018. Springer, Cham, Switzerland, 2018.
- [3] MITRE, System Engineering Guide. 2014.
- [4] Demian, P. and Fruchter, R. An ethnographic study of design knowledge reuse in the architecture, engineering, and construction industry. *Research in Engineering Design*, 16: 184-195, 2006.
- [5] Zima, K., Plebankiewicz, E., Wiczorek, D. A SWOT Analysis of the Use of BIM Technology in the Polish Construction Industry. *Buildings*, 10: 16-28, 2020.
- [6] Mattern, H. and König M. Concepts for formal modeling and management of building design options. In *Proceedings of the ICCET*, Seattle, Washington, 2017.
- [7] Mattern, H. and König, M. BIM-based modeling and management of design options at early planning phases. *Advanced Engineering Informatics*, 38: 316-329, 2018.
- [8] Napps, D., Pawlowski, D. and König M. BIM-based variant retrieval of building designs using case-based reasoning and pattern matching. In *Proceedings of the ISARC*, pages 435-442, Dubai, United Arab Emirates, 2021.
- [9] Zahedi, A., Abualdenien, J., Petzold, F., Borrmann, A. Documenting Design Decisions using Design Episodes, Explanation Tags and Constraints. Manuscript submitted for publication, 2021.
- [10] Isaac, S., Sadeghpour, F. and Navon, R. Analyzing Building Information Using Graph Theory. In *Proceedings of the ISARC*, pages 1013-1020, Montreal, Canada, 2013.
- [11] Wu, J. and Zhang, J. New Automated BIM Object Classification Method to Support BIM Interoperability. *Journal of Computing in Civil Engineering*, 33(5): 1-55, 2019.
- [12] ISO. ISO 16739-1: 2018. Online: <https://www.iso.org/standard/70303.html>. Accessed: 24.01.2022.
- [13] Hor, A. H., Jadidi, A. and Sohn G. BIM-GIS integrated geospatial information model using semantic web and RDF graphs. In *Proceedings of the ISPRS*, pages 73-79, Prague, Czech Republic, 2016.
- [14] Cao, H. and Connolly, D. Delta: an ontology for the distribution of differences between RDF graphs. In *Proceedings of the W3*, New York City, USA, 2004.
- [15] Zhao, Q., Li, Y., Hei, X. and Yang, M. A graph-based method for IFC data merging, *Advances in Civil Engineering*, 2020: 1-15, 2020.
- [16] Barrasa, J. and Howard, R. RDF triple stores vs. labeled property graphs: What's the difference? Online: <https://neo4j.com/blog/rdf-triple-store-vs-labeled-property-graph-difference/>. Accessed: 10.01.2020.
- [17] Wu, Y. and Khan, A. A graph pattern matching. In Sakr, S., Zomaya, A. (Publisher) *Encyclopedia of big data technologies*, 1-5, 2020.
- [18] Zahedi, A., Petzold, F. Revit Add-In For Documenting Design Decisions and Rationale. A BIM-based tool to capture tacit design knowledge and support its reuse. In *Proceedings of the CAADRIA 2022*, pages 557-566, Sydney, Australia, 2022.
- [19] Robinson, I. Graph databases. New opportunities for connected data, Eds. 2015. O'Reilly, Beijing, 2015.
- [20] Kriege, N. and Mutzel, P. Subgraph matching kernels for attributed graphs. In *Proceedings of the ICML*, pages 291-298, Edinburgh, Scotland, 2012.
- [21] Autodesk. Revit sample project files. Architecture. Online: <https://knowledge.autodesk.com/support/revit/getting-started/caas/CloudHelp/cloudhelp/2022/ENU/Revit-GetStarted/files/GUID-61EF2F22-3A1F-4317-B925-1E85F138BE88-htm.html>. Accessed: 03.01.2022.

Conceptual Modeling of Lifecycle Digital Twin Architecture for Bridges: A Data Structure Approach

I.M. Giorgadze^a, F. Vahdatikhaki^a, J.H. Voordijk^a

^a Department of Construction Management & Engineering, University of Twente, The Netherlands
E-mail: i.m.giorgadze@utwente.nl, f.vahdatikhaki@utwente.nl, j.t.voordijk@utwente.nl

Abstract

The concept of Digital Twin (DT) has emerged in recent years to facilitate the use of Building Information Modeling during the entire projects' lifecycle. In the DT concept, cyber-physical system theory is utilized to collect condition data about an existing asset and then integrate this data into the digital model. The major limitation though is that the current scope of DT is limited to the operation and maintenance phase. Nevertheless, the DT concept can be extended to the entire lifecycle of the asset if the relevant sensory and non-sensory data are incorporated into the digital model in an automated and systematic way. However, in the current literature, there is no clear insight about such a holistic and life-cycle DT concept for infrastructure projects. Especially, there is very little understanding about how various sensory and non-sensory data from construction and operation phases can be seamlessly integrated into the 3D BIM models. Therefore, this research aims to develop a conceptual model for the architecture of Lifecycle DT (LDT) focusing on bridges. To this end, an ontological modeling approach is adopted. The proposed ontology is validated through a workshop session where domain experts assessed the results with respect to some competency questions. The outcome of the session indicated that the proposed ontology scored sufficiently in all the criteria and succeeded in satisfying the information needs of the LDT. Overall, the proposed model offers an insight into a lifecycle modeling practice as well as automated data incorporation, enabling a smooth transition towards an upgraded modeling practice.

Keywords –

Digital twin; Lifecycle digital twin; Bridge information modelling; Ontological modelling

1 Introduction

Integrating the lifecycle information of a construction project in a centralized/federated model has

received much attention lately [1]. Building Information Modeling (BIM) attempts to store and represent all the relevant information of a project's lifecycle in an object-oriented 3D model. However, time-consuming and error-prone manual work is required for the generation, maintenance, and upkeep of the information in BIM [2, 3]. This problem can be potentially addressed by automating information acquisition and integration [3~7].

Digital Twin (DT) is an upcoming concept, which can address the need for automatic acquisition and integration of information. DT is a multi-physics model, fed with meaningful data about the asset and the environment around it [8]. The concept of DT was first introduced in 2012 for the verification and validation of aerospace vehicle models. As far as the construction industry is concerned, DT can be described as an extension of BIM, with the addendum of the dynamic and reactive aspect emerging from the use and integration of sensors, IoT architecture, and cyber-physical systems [10~11]. Such technologies make the physical asset smart and enable it to communicate with its digital counterpart about its health and condition. In other words, DT intend to fuse as-designed and as-is physical representations [3]. In this sense, DT is a system consisting of a physical twin, a digital twin, and a communication interface that connects the two [9]. Among the characteristics of DT, the real-time reflection of the physical space to the virtual one has been expressed as a distinguishing factor for the concept.

Nevertheless, current applications of DT is mostly limited to Operation and Maintenance (O&M) activities [12]. From literature, it is already known that design information/decision has significant impact on the performance and the maintenance of a project [12~15]. That is why it is important to take a life-cycle approach towards the development of DT models. To integrate the entire lifecycle information, the digital and physical counterparts should evolve in parallel from the design phase until the final demolition, creating what can be labelled as a Lifecycle Digital Twin (LDT). BIM can be potentially used as a platform for developing LDT, but this requires the expansion of its scope to real-time and automated data acquisition. Because of this, the BIM

should be developed in view of the requirements of the LDT. All the necessary lifecycle information and the respective sensors should be predefined, and the expected data pieces should be allocated to the relevant model components. This way, a LDT-ready BIM model, which is created in the design phase, can function as a LDT once the physical counterpart and the communication channels with the model are in place. In other words, future BIM models should be prepared from the design phase already in view of a full-scale LDT model.

Currently, there is a lack of a comprehensive approach towards building a LDT-ready BIM model that can be used for different purposes across the lifecycle of an asset. To address this problem, this research aims to offer a conceptual model for the architecture of the LDT, focusing on bridges. This model outlines requirements for the transformation of conventional BIM models to DT-ready models.

The remainder of the paper is structured as follows. Firstly the scope of the study and the methodology are briefly explained. Next, each step of the methodology is further elaborated. More specifically, the current modelling practices are explained and a data map of the current data structure is presented. Then, the information requirements are presented followed by proposed enriched ontology, which integrates the information requirement. A case study that explains how the proposed ontology can be of use is presented in the next section. This is followed by the validation of the results. Finally, a reflection upon the findings of the research is presented in the discussion and conclusions section

2 Research scope and methodology

To identify the missing elements for a smooth transition towards the DT concept, it is important to have a clear picture of the current modeling practices. To address that need, the first step of the research was an exploration of existing BIM models of bridges. A Dutch contractor was studied as a context. More specifically, Revit files of different bridges were investigated to extract the data scheme, i.e., how the bridge model is decomposed into different objects, what are the properties of the different objects and how they are stored. In the next step, a set of interviews with domain experts from disciplines covering the entire lifecycle were carried out to identify the information that the different disciplines desire to extract from a LDT. These information needs were then expressed as specific properties that LDT-ready model needs to incorporate. Some of the missing properties can be easily allocated to the existing elements, while some may require the introduction of new entities. The addition of new entities in the current data structure led to an enriched bridge LDT ontology. This ontology was further validated at the

final step of the research via the assistance of human domain expertise. More specifically, domain experts from disciplines covering the entire lifecycle, were asked to assess and rate the proposed ontology with respect to its correctness, completeness, conciseness and extensibility by answering a set of competency questions.

3 Requirements Analysis

In the following section, the different steps of the methodology are explained.

3.1 Current Modelling Practices

Figure 1 presents the high-level ontology that represents the dominant approach toward bridge information models. This model emerged from the exploration of several bridge BIM models. It was observed that these models consist of two main parts, the topographic view, and the bridge model itself. The latter consists of a set of digital entities representing the bridge objects, e.g., piles, pile caps, piers, headstock, girder, abutments, wing walls, etc. All these elements are characterized by a shared set of parameters like the area, volume, and year of construction. Apart from the geometric properties, all the objects of the model are characterized by a unique identifier code. This code is composed of a set of digits that are determined based on the location and decomposition of the objects. For example, a pile is characterized by 6 digits, where the first two digits indicate whether the pile is located in the western or eastern half of the bridge, the second two digits indicate that the pile is part of the foundation, and the last two digits designate the specific pile from the pile bundle that the foundation consists of. This naming approach aims to create an unambiguous naming policy for the elements across the different disciplines and is a prerequisite for applying filtering processes and logic rules, which enable the management of metadata in an automated way.

3.2 Information Requirements

During the interviews, the domain experts were asked about what information they would like to be able to extract from the LDT in order to assist their tasks. A set of example information requirements that emerged from the interviews are presented in Table 1.

Each information requirement was further linked to a target class, i.e., the ontology entity that should host that specific information. Some of the target classes already exist in the current ontology, while some new classes had to be introduced to allow a meaningful distribution of information. For example, the required information of the end of the lifespan (No2) refers to all the construction instances that already exist in the current

ontology. On the other hand, the required information of the location of the equipment (No10) demands the introduction of the equipment class, as well as a class for a technology that traces and registers the location of the equipment. Among others, some of the newly introduced target classes are the processes, equipment, the agent and process model, as well as some information collecting technologies like the laser scanners and Radio Frequency Identification (RFID) tags. The content of the new classes as well as the way they are incorporated to enrich the current ontology are further explained in the next section.

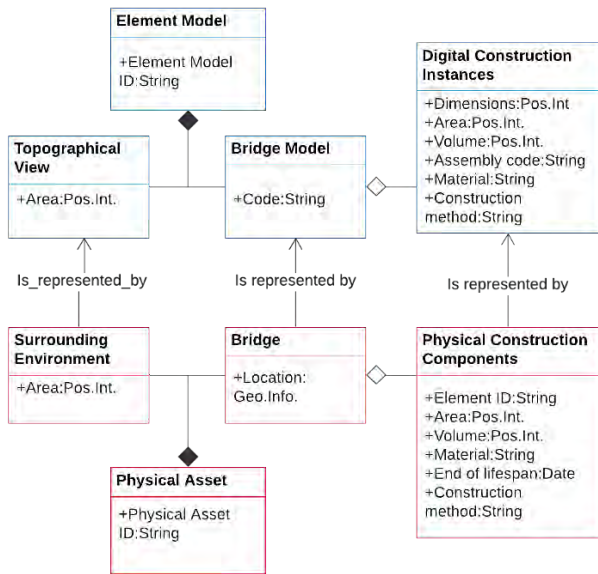


Figure 1. Current ontology

3.3 Proposed Ontology

The new information pieces and respective required classes that emerged from the requirements analysis are used to develop the proposed ontology for a DT-ready bridge information model. The proposed ontology consists of three main sections: the physical asset, the digital model, and the lifecycle, which is the main addition compared to the current ontology. The physical asset has a lifecycle that is proposed to be represented in the digital model, which in turn communicates bilaterally with the physical asset, as the concept of DT implies. This relationship between the three main sections is depicted in the higher-level proposed ontology shown in Figure 2.

Figure 3 presents a more detailed representation of the proposed bridge LDT. More specifically, the lifecycle section includes new aspects such as the resources, processes, and involved risks. The resources class includes the subclasses of equipment and material used in the construction activities, as well as the software and human agents, which may be designers, engineers or

workers. The processes refer to different kinds of construction-related activities like designing, concrete pouring, drilling and transferring.

Table 1. Information Requirements Taxonomy

No	Information Requirement	Target Class
1	To select and filter the elements in the model to highlight those satisfying a requirement	Construction Instances, Processes
2	End of lifespan for each element separately	Construction Instances
3	Environmental Impact Value in the model for each element	Construction Instances
4	Access Material Passport (MP) of each element	Construction Instances
5	Executed quality checks and potential non-conformances with the model	Construction Instances, Processes, LiDAR for as-built scanning
6	Duration and cost of the execution for the different work sets and objects separately	Construction Instances, Processes, Process Model
7	Usage loads bearing the deck	Deck, Weight-In-Motion sensors
8	Deviation between planned maintenance schedule and implemented maintenance record	Construction Instances, Processes
9	Concentration of dangerous substances in the air during drilling activities	Drilling, Surrounding Environment, Real-time respirable dust monitoring devices
10	Trace the location of equipment	Equipment, RFID for asset tracking
11	Duration and cost of rented equipment	Equipment, RFID for asset tracking
12	Deviation between actual and designed environmental variables	Construction Instances, Surrounding Environment, Hygrometer, Thermometer
14	Proximity between workers and identified danger (steep slope, height, moving vehicle, dangerous materials)	Crew, Surrounding Environment, Materials, Agent Model, RFID for people and asset tracking,
15	Proximity between moving vehicles and areas with unstable or humid ground.	Moving Vehicles, Surrounding Environment, Agent Model, RFID for asset tracking, Terrestrial laser scanner

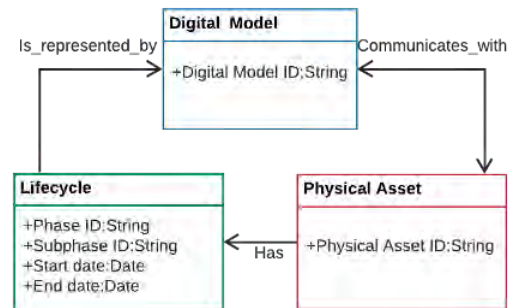


Figure 2. Higher-Level Proposed Ontology

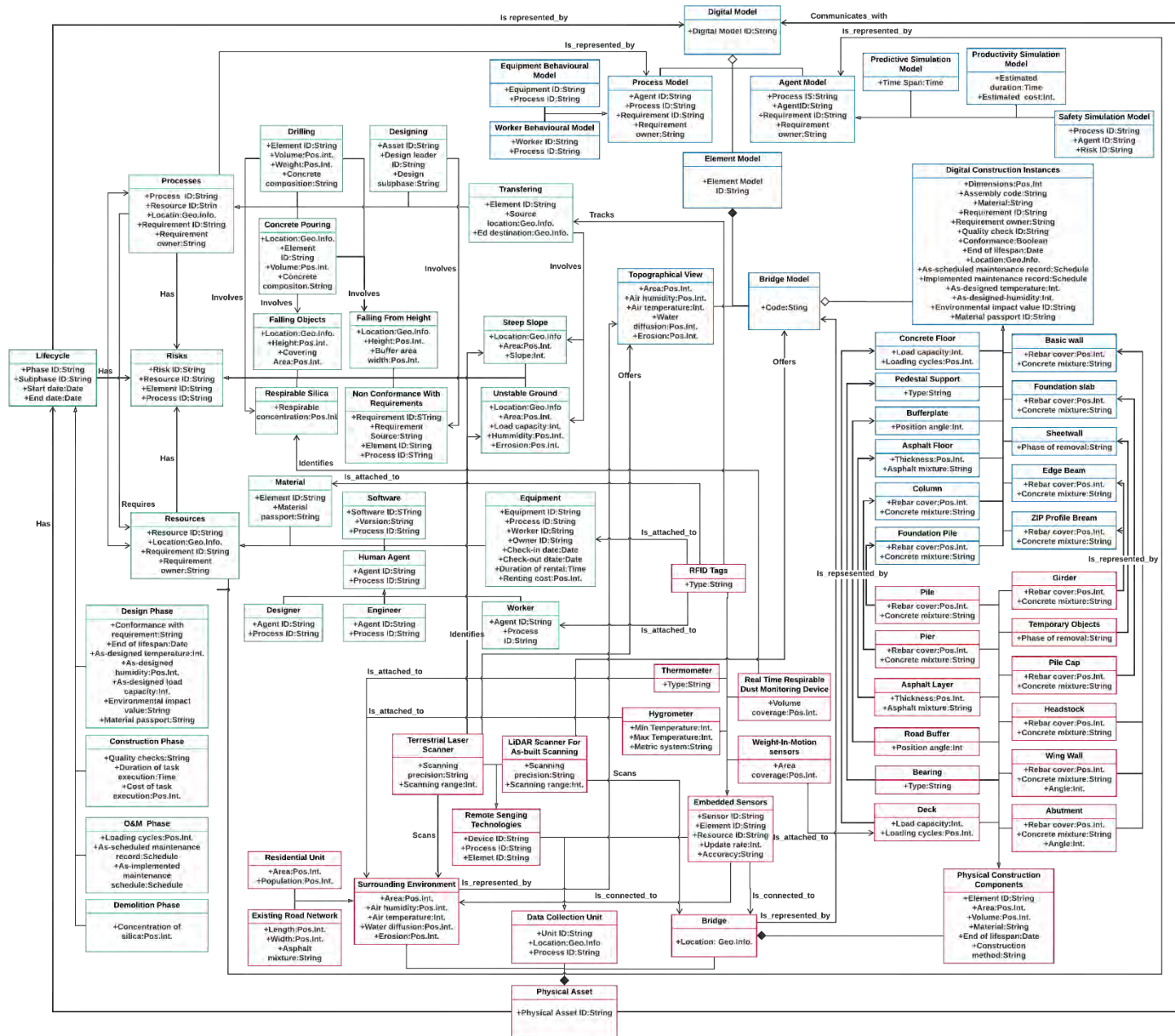


Figure 3. Proposed Enriched Ontology for the LDT

Furthermore, the risks concern potential accidents like falling from height or from steep slopes, breathing dangerous substances but also design related risks like insufficiency to meet the client's requirements. Finally, regarding the relationships between these classes, the different processes require resources while the also involving risks.

Regarding the digital model section, an agent and process model were added alongside the element model. While the element model represents physical components of the bridge, the agent and process models offer an insight in the behavior of the asset. An example of the behavior of an excavator is shown in the agent model in Figure 4 [16]. Process models, on the other hand, represent the sequence of activities that take place in different lifecycle operations, e.g., excavation, as shown in Figure 5 [17]. The combination of agent and process models allows to develop various types of simulation models that can support the asset management by predicting the asset behaviors, assessing safety, and estimating the productivity of different operations.

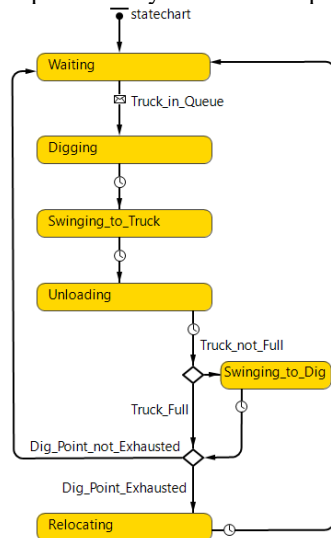


Figure 4. Example of an agent model (adapted from [16])

As far as the physical asset section is concerned, a set of collecting units is added, which collects data about the physical asset and feeds that data to the digital model. The data collecting units may either be (1) embedded sensors like thermometers and hygrometers attached to the bridge, RFID tags attached to the resources, and weight sensors embedded in the deck, or (2) contactless sensors like LiDAR scanners, which register the as-built condition, or terrestrial scanners, which offer a detailed depiction of the ground surface.

The incorporation of the new classes in the LDT ontology led to a large set of new relationships, which connect different interrelated entities. An exhaustive description of all the relationships and linking possibilities is beyond the scope of this paper, however

the following two examples aim to offer an insight about the information linking potential of the proposed ontology:

Example 1: During the construction phase, the transportation process takes place. The resource used in this process is a vehicle. This process involves the risk of the vehicle approaching a steep slope on the construction site. The location of the steep slope is acquired via the assistance of a terrestrial laser scanner and is registered in the topographic view of the element model. Furthermore, the location of the moving vehicle can be traced via the assistance of RFID tags. This information is registered to a safety simulation model. The simulation model and the topographic view are interlinked parts of the digital model, and this way the proximity of the moving vehicle to the steep slope can be calculated.

Example 2: The design phase involves designer agents. A risk associated with this process is failing to meet the client's requirements. Additionally, the designing process is assisted by the use of terrestrial scanners that offer a representation of the ground surface of the surrounding environment. Finally, once the asset is built, a LiDAR scanner for as-built scanning can identify non-conformances between the as-designed and as-built situation.

Overall, the proposed ontological model addresses the current limitations in multiple ways. Firstly, the combination of different model types, namely agent, process and element models, offers a multiscale modeling practice and a holistic modeling approach. Such a global modeling perception is in favor of managing and incorporating diverse lifecycle data generated through the lifespan of a construction project. Furthermore, the definition and inclusion of the different properties, as well as the information flow channels, increases the DT-readiness of the element model. Moreover, the new emerged relationships assist the combination and extraction of a big variety of information as explained in the examples. With the assistance of some algorithms and the application of logic rules, it is possible to deeply mine data and identify patterns, apply correlation analyses, and identify emergent and latent phenomena. These applications serve as a basis for machine learning and artificial intelligence, which are concepts closely related to the DT.

3.4 Hypothetical Validation

To further explain how the proposed ontology can assist the extraction of information, a hypothetical validation is developed based on a scenario extracted from the required information in Table 1.

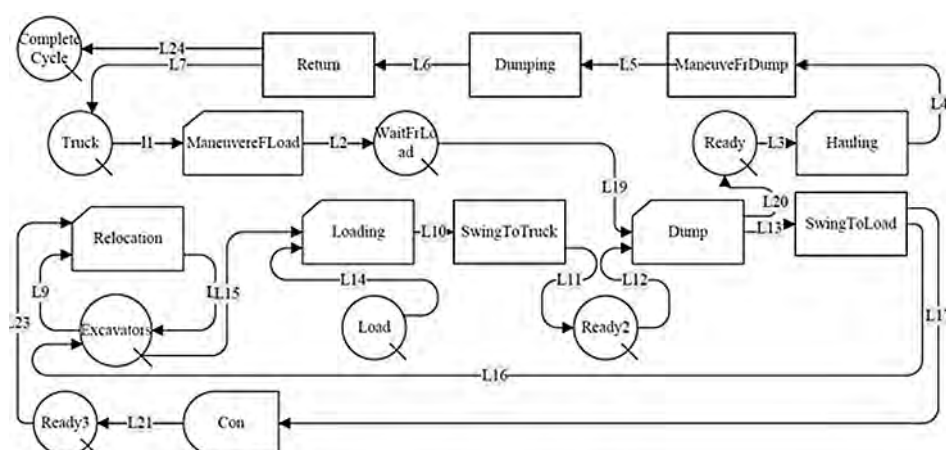


Figure 5. Example of a process model (adapted from [17])

It is required to extract the deviation between the programmed and implemented maintenance tasks. According to the proposed ontology, the as-planned maintenance frequency and the implemented maintenance record are properties assigned to all the construction instances. Therefore, it is possible to use a query to highlight the elements of the model that require more frequent than scheduled maintenance.

Similarly, it has been asked to extract the deviation between the designed temperature and humidity capacity of the elements and the actual ambient conditions. The as-designed values are included in all the different construction instances, and the actual ambient conditions are being retrieved via the use of embedded hygrometers and thermometers, and stored in the topographic view of the model. Therefore, it is possible to filter and isolate the elements that present a deviation between the as-designed and implemented ambient conditions.

Finally, it is possible to combine the aforementioned information and identify potential correlations between the climate change and the maintenance demand of the asset. This insight allows optimizing the maintenance schedule of the reference project, but also considering more resilient construction practices for new generations of the project. It can be hence observed that the proposed ontology combines different information to offer a deep insight and it can even provoke the genesis of new knowledge.

3.5 Validation

The proposed ontology was validated through a workshop session where the ontology was presented to a group of six domain experts from different phases of the entire lifecycle of a bridge. More specifically, the participants were a systems engineer, a designer, a site engineer, a maintenance engineer, a safety and a sustainability specialists. The participants were asked to

assess the proposed ontology with regard to a set of predefined criteria by answering a set of competency questions. The evaluation criteria used to assess the ontology were correctness, completeness, conciseness, and adaptability [18~19]. The criterion of correctness indicates whether the ontology correctly represents the real-world concept of the lifecycle of a bridge. Completeness measures whether the domain of interest is appropriately covered, while conciseness indicates that an ontology does not include unnecessary concepts or redundancies. The last criterion, adaptability or extendibility, designates whether an ontology is easily adaptable in the case of adding new definitions and new knowledge to existing ones.

The results of the session, which are presented in Table 2, showed that the proposed ontology scored well in all criteria. An improvement recommendation about the correctness of the proposed ontology concerns the introduction of tiers in classifying the elements, meaning libraries, objects and instances. Regarding the completeness, it was proposed to add the pre-design phase in the lifecycle to incorporate tendering processes and involved risks as well. Additionally, it was proposed to keep the ontology at a higher level of abstraction, and further develop it at a project level. This way the ontology can be adapted for several infrastructure assets apart from bridges. Lastly, it was suggested to align the object breakdown structure to the national standards to improve the adaptability potential of the ontology. Overall, the proposed ontology is considered to adequately cover the lifecycle information needs of a LDT of a bridge.

4 Discussion

Regarding the scientific contribution of this study, the proposed ontology incorporates and allocates sensory data to the elements of the BIM model, addressing the issue of seamless integration of the automated/sensory data into lifecycle information models. Furthermore, it

proposes a lifecycle implementation framework for the DT, introducing a holistic approach, which is missing from the existing fragmented application efforts. Moreover, the research provides an insight into how to transit towards DT-ready bridge information models by indicating how sensory data, different databases and simulation models can increase the automation potential of the BIM models.

Table 2. Validation scores and potential improvements

Criterion	Score (1-5)	Potential Improvement
Correctness	3.8	Structure of the classification of libraries, objects and instances
Completeness	4	Inclusion of pre-design phase
Conciseness	4.2	-
Adaptability/ extensibility	3.25	Keep ontology at a higher level of abstraction Align with national standards about OBS

The outcome of the study indicates that the DT is not necessarily only about sophisticated technological application, but also systematic data structure. The actual value of the DT concept is in the incorporation and linkage of various data and its continuous flow across the lifecycle.

Regarding the strengths of the research, the participation of actual domain experts in the development of the ontology adds to the implementability potential of the proposed ontology. The data pieces, which were used to enrich the current ontology, were derived from actual practitioners and reflect the real modeling needs. Additionally, apart from input, domain expertise was also used for validation. Another strong point of the proposed ontology is that it is conducted at a relatively high level of abstraction, allowing its easy adoption for other types of infrastructure assets without significant changes. On the other hand, a limitation of the research is that it is not an exhaustive study of neither all the lifecycle aspects that the DT model, nor all the potential data collection variety that can assist such a model. The proposed ontology serves rather as a high-level data framework for a bridge's DT modeling practice.

As far as the application of the ontology is concerned, there are some prerequisites that need to be met. From a technical point of view, the different platforms, databases, and data collecting technologies should be interoperable to apply the recommended data structure. Interoperability, in this sense, does not only refer to the data format but also the data granularity. In other words, the decomposition of objects, worksets and the naming policy should be aligned across different disciplines. Regarding the working routines, they will

also need to undergo some changes. New collaborations, roles and tasks will emerge, and people should be persuaded to embrace them in view of a greater lifecycle picture. Closer collaborations and more intense interdisciplinary communication are needed to communicate the information needs and find the best path for the information to flow. Finally, people should be adequately informed about the added value of the DT readiness practice to embrace enthusiasm about and avoid resistance regarding new working routines.

5 Conclusion

To summarize, this research aimed to bring the current state of art one step closer to the application of DT by identifying the requirements a BIM model should meet to allow a gradual transition towards LDT. To address the research objective, an ontological model was developed to map the distribution of various data pieces among elements of the model. This ontology describes what is the additional information that need to be included in the BIM model, what are the needed data collection technologies, how are the sensor measurements distributed in the model and what are the relationships between the different entities. The proposed data structure indicates how the models should be created in consideration of the LDT offering a smooth transition for the application of the concept. Overall, the results of the research indicate that the proposed ontology has great potentials to support a variety of activities throughout the entire lifecycle of a bridge.

References

- [1] "White paper: National digital twin | Bits & Pieces." <https://global.royalhaskoningdhv.com/digital/resources/publications/national-digital-twin> (accessed Feb. 14, 2022).
- [2] T. Borangiu, D. Trentesaux, P. Leitão, A. G. Boggino, and V. Botti, "Studies in Computational Intelligence 853 Service Oriented, Holonic and Multi-agent Manufacturing Systems for Industry of the Future." [Online]. Available: <http://www.springer.com/series/7092>
- [3] V. Stojanovic, M. Trapp, R. Richter, B. Hagedorn, and J. Dollner, "Semantic Enrichment of Indoor Point Clouds An Overview of Progress towards Digital Twinning."
- [4] Ü. Işıkdağ, "Enhanced Building Information Models Using IoT Services and Integration Patterns."
- [5] J. P. 1959- Wulfsberg, *I. interdisziplinäre Konferenz zur Zukunft der Wertschöpfung Konferenzband*.

- [6] B. Becerik-Gerber, F. Jazizadeh, N. Li, and G. Calis, "Application Areas and Data Requirements for BIM-Enabled Facilities Management," *Journal of Construction Engineering and Management*, vol. 138, no. 3, pp. 431–442, Mar. 2012, doi: 10.1061/(asce)co.1943-7862.0000433.
- [7] J. Chen, T. Bulbul, J. E. Taylor, and G. Olgun, "A Case Study of Embedding Real Time Infrastructure Sensor Data to BIM."
- [8] E. H. Glaessgen and D. S. Stargel, "The digital twin paradigm for future NASA and U.S. Air force vehicles," 2012. doi: 10.2514/6.2012-1818.
- [9] S. Haag and R. Anderl, "Digital twin – Proof of concept," *Manufacturing Letters*, vol. 15, pp. 64–66, Jan. 2018, doi: 10.1016/j.mfglet.2018.02.006.
- [10] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui, "Digital twin-driven product design, manufacturing and service with big data," *International Journal of Advanced Manufacturing Technology*, vol. 94, no. 9–12, pp. 3563–3576, Feb. 2018, doi: 10.1007/s00170-017-0233-1.
- [11] D.-G. J. Opoku, S. Perera, R. Osei-Kyei, M. Rashidi, T. Famakinwa, and K. Bamdad, "Drivers for Digital Twin Adoption in the Construction Industry: A Systematic Literature Review," *Buildings*, vol. 12, no. 2, p. 113, Jan. 2022, doi: 10.3390/buildings12020113.
- [12] D. G. J. Opoku, S. Perera, R. Osei-Kyei, and M. Rashidi, "Digital twin application in the construction industry: A literature review," *Journal of Building Engineering*, vol. 40, Elsevier Ltd, Aug. 01, 2021. doi: 10.1016/j.jobbe.2021.102726.
- [13] H. Gervasio, S. Dimova, and A. Pinto, "Benchmarking the life-cycle environmental performance of buildings," *Sustainability (Switzerland)*, vol. 10, no. 5, May 2018, doi: 10.3390/su10051454.
- [14] W. Hu, T. Zhang, X. Deng, Z. Liu, and J. Tan, "Digital twin: a state-of-the-art review of its enabling technologies, applications and challenges," *Journal of Intelligent Manufacturing and Special Equipment*, vol. 2, no. 1, pp. 1–34, Aug. 2021, doi: 10.1108/jimse-12-2020-010.
- [15] F. Jiang, L. Ma, T. Broyd, and K. Chen, "Digital twin and its implementations in the civil engineering sector," *Automation in Construction*, vol. 130, Elsevier B.V., Oct. 01, 2021. doi: 10.1016/j.autcon.2021.103838.
- [16] F. Vahdatikhaki, K. el Ammari, A. K. Langroodi, S. Miller, A. Hammad, and A. Doree, "Beyond data visualization: A context-realistic construction equipment training simulators," *Automation in Construction*, vol. 106, Oct. 2019, doi: 10.1016/j.autcon.2019.102853.
- [17] F. Vahdatikhaki, "TOWARDS SMART EARTHWORK SITES USING LOCATION-BASED GUIDANCE AND MULTI-AGENT SYSTEMS," 2015.
- [18] J. Raad, C. Cruz, and C. A. Cruz, "A Survey on Ontology Evaluation Methods," 2015, doi: 10.5220/0005591001790186.
- [19] P. Delir Haghighi, F. Burstein, A. Zaslavsky, and P. Arbon, "Development and evaluation of ontology for intelligent decision support in medical emergency management for mass gatherings," *Decision Support Systems*, vol. 54, no. 2, pp. 1192–1204, Jan. 2013, doi: 10.1016/j.dss.2012.11.013.

Towards the development of a digital twin for subsoil monitoring and stability against overturning of a mobile drilling rig

F.W. Riquer¹, D.A. Dao¹ and J. Grabe¹

¹Institute of Geotechnical Engineering and Construction Management, Hamburg University of Technology

francisco.williams@tuhh.de, duy.anh.dao@tuhh.de, grabe@tuhh.de

Abstract -

Mobile drilling rigs are particularly susceptible to overturning due to the high location of their centre of mass. In some cases, overturning occurs due to a failure in the subsoil. Until now, machine operators are solely responsible for monitoring the machines' stability, and assessment of dangerous conditions is based mainly on experience. This investigation aims to set the grounds for elaborating a digital twin to online monitor the machine's stability to prevent overturning. Stress transmitted to the ground, tracks' settlement, and bearing capacity are calculated from a multibody simulation (MBS) and following existing standards. Furthermore, soil-dependent stability diagrams are generated to describe the stable location of the machine's centre of mass and predict soil failure. Results offer the possibility to function as an online alarm system running parallel to the machine's operation and alerting the operator about dangerous conditions.

Keywords -

Digital Twin (DT); Multibody simulations (MBS); Stability of mobile drilling rigs; Settlement prediction; Monitoring of bearing capacity

1 Introduction

Every year, the overturning of construction machinery causes several accidents worldwide. From 2007 to 2008, there were 38 fatalities and 679 injuries in the UK caused by the overturning of construction machinery, including drilling rigs [1]. Similarly, in the USA, there were 323 fatalities from 1980 to 1992 [2]. In Germany, there were 9 fatalities and 21 serious injuries from 1993 to 2003 [3]. Among the causes of these accidents is ground failure due to poorly prepared working platforms. Satoshi and Tomohito [4] summarize the operations before the accident, equilibrium conditions, and ground properties for a real case study of the overturning of a drilling rig. In this case, a ground penetration of the tracks was observed.

This publication investigates the idea of developing a digital twin to monitor the stability of the subsoil un-

derneath the construction machine's tracks. A multibody simulation (MBS) of the mobile drilling rig is implemented in the Simscape environment of the software Matlab/Simulink. Signals simulating the reading of the internal machine's sensors are used to recreate working conditions. Following standards DIN EN 16228-1, DIN 4019 and DIN4017, the stress distribution, settlement and bearing capacity underneath the machine's tracks are estimated online for the MBS. For these estimations, local information about the soil properties is required. This work presents a mockup dashboard for an online monitoring system of the stability of the subsoil for limiting scenarios.

Additionally, stability diagrams are generated by analyzing the limiting conditions for the failure of the subsoil. Thus, stable areas are determined for the subsoil for all possible locations of the centre of mass. These could be helpful for on-site rapid stability checks or enhance current standards.

The company Liebherr has developed a system following similar industry standards [5]. However, this publication conceptually differs since the calculations of the location of the centre of mass follow an MBS. Furthermore, to the authors' best knowledge, the Liebherr system is based on the assumption of a rigid subsoil.

2 Theoretical framework

This section presents the theoretical tools used in this work. First, the idea of digital twins is briefly addressed. Secondly, the concept of MBS is mentioned, and then a review of the used standards is included.

2.1 Digital twins

In recent years, the fast development of simulation, data acquisition, data communication, and other technologies facilitated the interactions between physical and virtual spaces [6]. The digital twin's framework has emerged from this recent development as the natural consequence to merge simulations and physical space. A digital twin is defined as a comprehensive physical and functional de-

scription of a component, product, or system, including helpful information for present and future life cycle phases [7].

The idea of digital twins has been exploited at length in industries like product design, production prognostics, and health management [6]. Given that it is a new emerging trend, there is not a strict framework to define digital twins. However, most authors agree that a digital twin includes three parts: physical product, virtual product, and their connections ([8], [9]). Therefore, digital twins enable manufacturers and users to make accurate predictions and informed rational decisions.

The objective of the digital twin framework proposed in this work is the stress monitoring of the subsoil to avoid working conditions that could cause soil failure. The proposed simulation provides the backbone of an eventual one-to-one simulation with the possibility of intuitively extending it to include detailed machine internal mechanics.

2.2 Multibody simulations (MBS)

MBS are crucial for research areas like vehicle dynamics, robotics, biomechanics, etc. Thanks to the recent development in computational dynamics, MBS can run in real-time on digital computers and thus, are of great importance for the elaboration of digital twins. In principle, MBS are numerical simulations including rigid and flexible bodies with dynamics represented by their equations of motion [10].

One of the advantages of using the Matlab Simscape Multibody environment for MBS is the possibility of integrating a wide variety of elements from Simscape and Simulink into the model. This environment for 3D mechanical systems provides a block language including libraries to represent a variety of bodies, joints, constraints, force elements, and sensors [11]. Simscape internally solves the equations of motion for complete dynamical assemblies. Furthermore, Simulink libraries can be integrated to represent control systems, hydraulics, internal friction between elements, etc. Nonetheless, it is crucial to balance the objective of the created model and the complexity level since a higher complexity corresponds to longer simulation times. Thus, the real-time requirement to use it as a digital twin can be compromised.

Additionally, stand-alone applications designed to run in external hardware are supported by C-code generation in Simscape Multibody [11]. Models designed in this environment can be deployed in the internal computers of the mobile drilling rigs and work as digital twins to simulate and monitor the working conditions of the machine in real-time.

2.3 Review of industry standards

This section summarizes the standards used to estimate the stability of the drilling rig against overturning and the estimation of soil failure. These standards are incorporated in the MBS.

Stability calculations

The DIN EN 16228-1 standard [12] bases the stability standpoint of the machine by determining a maximum allowable tilting angle. Figure 1 shows the inclination of a drilling rig about a tilting edge located at the front of the tracks. The maximum allowable tilting angle α_{sr} about a given tilting edge is compared with the actual stability angle α_s . Therefore if $\alpha_s \geq \alpha_{sr}$, the analyzed operating condition is considered as stable. This industry standard proposes an offline calculation of this angle for all possible or expected operating conditions. After that, the smallest of all calculated α_{sr} is defined as the lower bound for α_s . This standard assumes a rigid subsoil for its calculations.

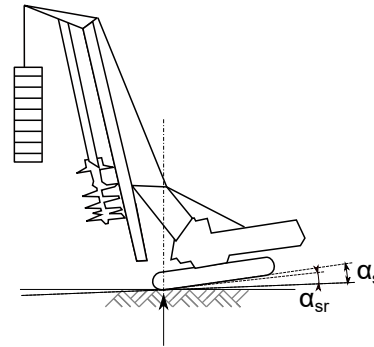


Figure 1: Determination of stability angles according to DIN EN 16228-1 [12]

Stress calculation

Several standards ([12], [13], [14]) describe the pressure transmitted to the subsoil underneath the drilling rig's tracks. These standards simplify the weight and forces acting on the machine using a resulting single-point load. This load represents the total drilling rig's weight, including lifted external weights. Then, the load is divided according to its proximity to each track, and an eccentricity is calculated. Figure 2 shows this simplification.

The stress distribution transmitted to the subsoil is calculated, assuming a trapezoidal distribution. Table 1 summarizes the different cases for different positions of the load along the tracks. Eccentricity e gives the position of the corresponding load acting on each track, and a and b represent the track's length and width correspondingly.

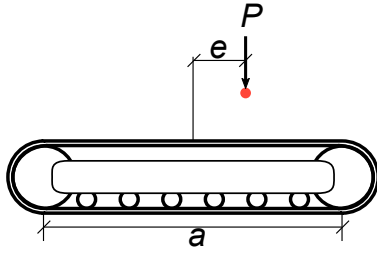


Figure 2: Simplification of a point load acting on one of the tracks

Table 1: Calculation of ground pressure transmitted to the soil

Diagram	Load position	Stress
	$e = 0$	$\sigma_1 = \sigma_2 = \frac{P}{ba}$
	$e < \frac{a}{6}$	$\sigma_1 = \frac{P}{ba} \left(1 - \frac{6e}{a}\right)$ $\sigma_2 = \frac{P}{ba} \left(1 + \frac{6e}{a}\right)$
	$e = \frac{a}{6}$	$\sigma_1 = 0$ $\sigma_2 = \frac{2P}{ba}$
	$e > \frac{a}{6}$ $c = \frac{a}{2} - e$	$\sigma = \frac{2P}{3cb}$

Settlement prediction

The calculation of the exact settlement of the subsoil is a challenging task due to the heterogeneous nature of soils. Standards work mainly as a reference in magnitude for the expected settlements. The standard used in this work for the settlement prediction is the DIN 4019 [15]. This standard uses the assumption of elastic half-space to determine the deformation behaviour of the soil ([16], [17], [18]). Equation 1 is used to calculate the stress σ_z at depth z caused by a stress σ_0 at the soil surface

$$\sigma_z = \sigma_0 \cdot i_{S,J} \left(\frac{z}{b}, \frac{a}{b} \right) \quad (1)$$

The factor $i_{S,J}$ depends on the shape of the load at the soil's surface. For a rectangular load, the factor i_S is calculated from equation 2.

$$i_S = \frac{1}{2\pi} \left[\arctan \frac{a \cdot b}{z \cdot R_3} + \frac{a \cdot b \cdot z}{R_3} \left(\frac{1}{R_1^2} + \frac{1}{R_2^2} \right) \right] \quad (2)$$

assuming $a > b$ and

$$R_1 = \sqrt{a^2 + z^2}$$

$$R_2 = \sqrt{b^2 + z^2}$$

$$R_3 = \sqrt{a^2 + b^2 + z^2}.$$

where a and b are the length and width of the tracks as shown in Table 1 and Fig 3.

For the case of a triangular load two different factors are calculated. For the case of the smaller side of the triangle i_J is calculated using equation 3 and for the other end, the same expression as in equation 2 is used.

$$i_J = \frac{1}{2\pi} \left[\arctan \left(\frac{b \cdot a}{z \cdot R} \right) + \frac{b \cdot z}{b^2 + z^2} \cdot \frac{R - \sqrt{b^2 + z^2}}{a} \right] \quad (3)$$

where

$$R = \sqrt{a^2 + b^2 + z^2}.$$

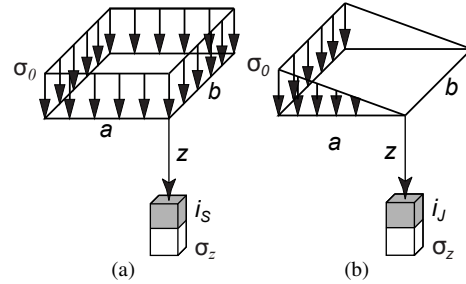


Figure 3: Geometric parameters for a rectangular and triangular load

Finally, the settlement is calculated using equation 4. Superposition is used to calculate the total settlement at the corner points.

$$s = \frac{\sigma_z b}{E_s} \quad (4)$$

Bearing capacity of soil under tracks

The bearing capacity refers to the soil's capability to withstand loads applied on them before developing a shear failure mechanism. This work uses standard DIN 4017 [19] for calculating the bearing capacity of the subsoil underneath the tracks of the drilling rig. Soil failure develops if the vertical load applied to the foundation is larger than the calculated resistance. For the calculations, soil properties are required, such as specific weight, cohesion, friction angle, foundation dimensions, and point load location.

Including the bearing capacity calculation in the digital twin integrates constant monitoring of the stability of the subsoil underneath the drilling rig tracks.

3 MBS of a mobile drilling rig

This section provides a detailed description of the mobile drilling rig's MBS. The environment chosen for the MBS is Simscape Multibody, part of Matlab/Simulink. The reason for choosing this software is the extensive libraries for simulating mechanical characteristics and control, allowing the extension of the proposed work to include more complex behaviour of drilling rigs in the simulation. Furthermore, Simulink provides a simple interface for implementing mathematical functions. Therefore, the calculations for settlement and bearing capacity were implemented using the block language of Simulink in the same model.

Simscape provides a simple solution to deploy stand-alone solutions from the developed model. A stand-alone application can be implemented on different hardware and run parallel to the machines as a digital twin.

3.1 Creation of the parts and assembly

The parts and assemblies in the MBS have intuitive interfaces to define their geometry and inertial properties. The geometry of all parts is parameterized, and the points of their cross-sectional areas are generated using trigonometric functions. Finally, the parts are extruded.

Joints impose primary kinematic constraints and depict each part's interaction with its neighbouring parts. These joints define the degrees of freedom -rotational and translational- between the connected bodies. Additionally, Simscape allows integrating internal mechanics and position or force control.

Figure 4 shows the final model of the drilling rig with three degrees of freedom consisting of the rotation of the uppercarriage (ϕ), the radial movement of the mast (d_M), and the inclination of the mast (α). The overall model's centre of mass (COM) is indicated by point C , and its position is monitored using an inertia sensor provided as a block in Simscape. The COM monitoring allows the calculations of pressure transmitted to the soil and, thus, the inclination and bearing capacity.

Furthermore, an external force (F_M) can be applied at the head of the mast to simulate lifting a load or using an external tool.

3.2 Results and analysis

This section summarizes the results of the MBS and the monitor of different machines' operating conditions. The presented plots are generated from collected data after several simulations. Therefore, they are not generated

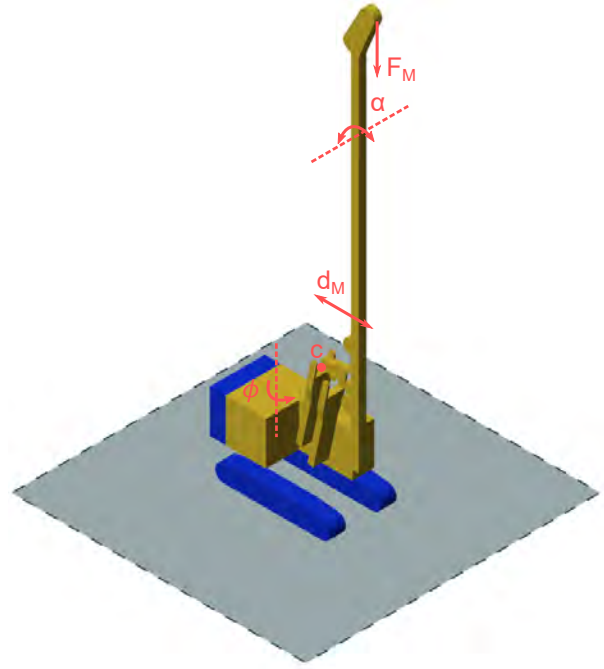


Figure 4: MBS of a mobile drilling rig

online. Nonetheless, when exporting the model as a stand-alone application, similar graphs can be generated using, for example, LabView software as an external interface to implement the real-time application.

The machine dimensions and inertial properties used to obtain the results in this section are summarized in Table 2. The missing dimensions are approximated to a mobile drilling rig of the company Bauer model BG 23 H.

Table 2: Machine's dimensions and inertial properties

Parameter	Value	Units
Total machine's mass	58.97	tons
Tracks' length	5.00	m
Tracks' width	0.80	m
Undercarriage's width	2.58	m

The MBS includes several assumptions and simplifications necessary to keep the model simple enough to deploy a stand-alone online solution and use the methodologies included in the standards. These are summarized as follows.

- The load on each track acts at the centre of the track's width.
- The drilling rig tracks are rigid bodies and, thus, treated as a rigid foundation.
- The soil elastic modulus has a restrained lateral strain.
- The inclination of the drilling rig is neglected, and thus, the predicted settlement is not feedback to the

model.

- e) The subsoil is idealized as horizontal.
- f) The machine's force vector does not incline.

Assumptions a), b), and c) are necessary for all standards considered in this work. Assumption d) simplifies the created model since otherwise, the model would need an iterative calculation, and thus the online monitoring would be compromised. Assumptions e) and f) are considered to simplify the estimation of the allowable bearing capacity.

For the scope of this work, an operating condition of a drilling rig is considered stable if the two following conditions are simultaneously avoided:

1. The force vector of the COM points in a direction towards a point outside the machine's drilling rig's body.
2. The allowable bearing capacity is exceeded for the current location of the COM.

Two approaches are presented as an analysis of the results of the simulation. The first approach includes the creation of radial plots where stability areas are drawn for different combinations of soil parameters and the location of the COM. The second approach provides an insight into online monitoring of the estimated transmitted stress, settlement and bearing capacity.

Stability areas depending on COM location and soil characteristics

The MBS model is used to determine stable operating areas. The COM of the drilling rig is shifted by simulating the lifting of a load. The total weight of the load is increased stepwise until one of the two mentioned conditions is reached, and thus, the machine's stability is lost. The weight acts at the top of the mast as the force F_M in Figure 4. This process is repeated for a 360° uppercarriage rotation (ϕ) with a step size of 10° while keeping the mast's position constant. The load is increased in intervals of 500 kg.

Following this methodology, diagrams for different clays and sands are generated. Tables 3 and 4 summarize the tested soil conditions.

Table 3: Characteristics of different clays

Parameter	Values	Units
Friction angle [φ]	[12, 20, 28, 30]	$^\circ$
Specific weight [γ']	16.5	kN/m ³
Cohesion [c]	20	kPa
Elasticity [E_s]	40	MPa

Figure 5 includes the stable areas for the clays mentioned in Table 3. The closed trajectories show the stable areas for

Table 4: Characteristics of different sands

Parameter	Values	Units
Friction angle [φ]	[25, 32, 41]	$^\circ$
Specific weight [γ']	18	kN/m ³
Cohesion [c]	0	kPa
Elasticity [E_s]	30	MPa

the COM location and the position of the upper carriage. $\phi = 0$ is assumed as the position shown in 4. The two stability conditions are simultaneously checked for each configuration. The points marked in red indicate the loss of the machine's stability due to the appearance of condition 1, the force vector points outside the body. The failing due to condition one only happens at angles close to 90° (or 270°) for small friction angles, and it is completely avoided for the smallest friction angle $\varphi = 12^\circ$. Keep in mind that a failure of the subsoil does not necessarily mean that the drilling rig will overturn. However, it is assumed that this situation is dangerous enough to be considered critical.

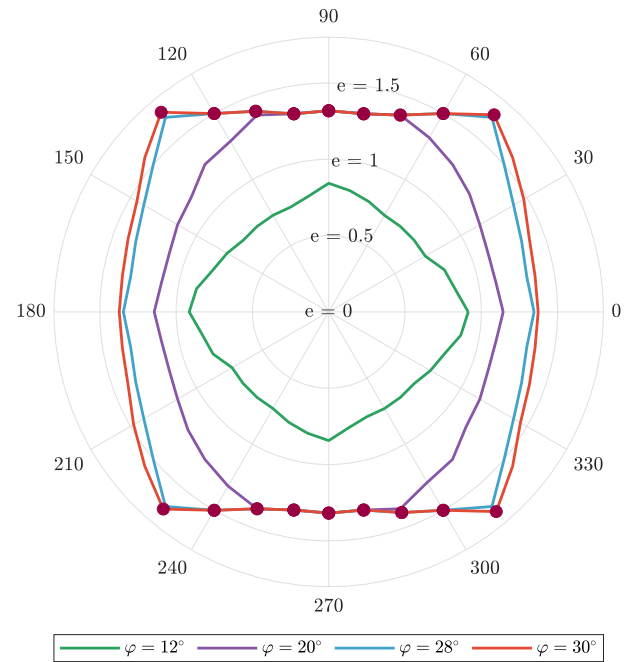
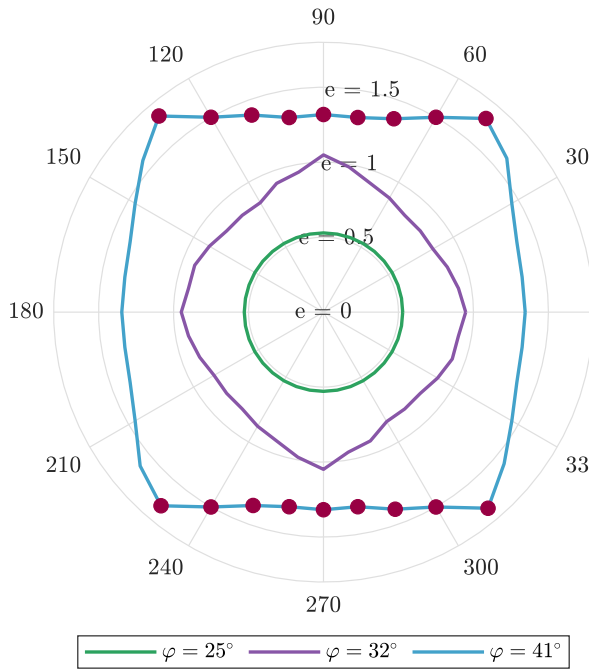


Figure 5: Stable areas for clays with $\gamma' = 16.5$ kN/m³

Figure 6 includes the stable areas for the sands mentioned in Table 4. Notice that for $\varphi = 25^\circ$, the stable area draws a perfect circle. This behaviour occurs because, for that specific friction angle, the soil fails for all possible positions of the uppercarriage; therefore, the second condition for unstable behaviour is fulfilled without the need of adding an extra load. The red dots indicate failure due to condition 1 for that combination of parameters.

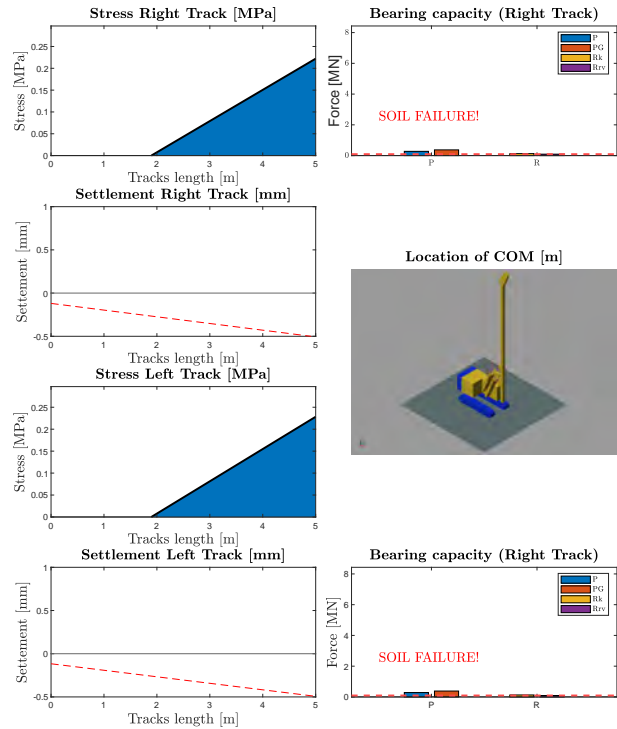
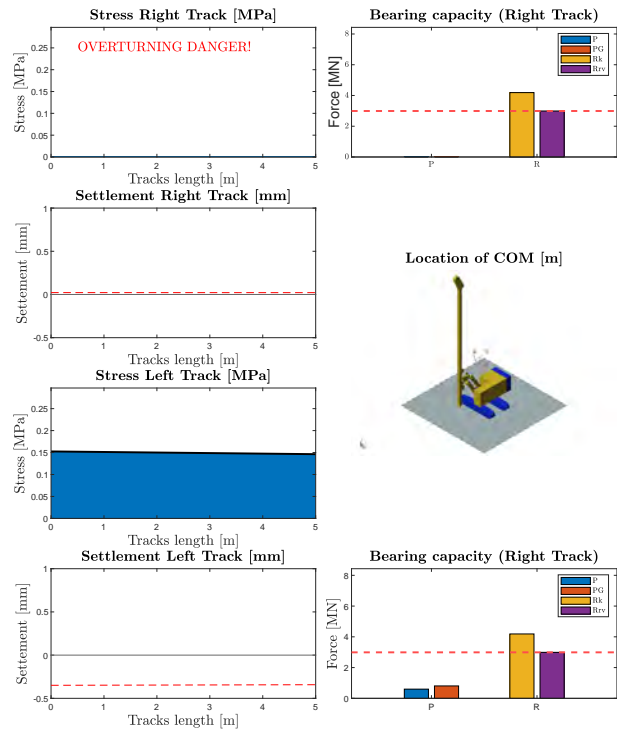
Figure 6: stable areas for sands with $\gamma' = 18 \text{ kN/m}^3$

Towards an online machine's stability monitoring

Figure 7 and Figure 8 show examples of what a dashboard for the monitoring of the stability of the mob drilling rig could look. These plots are again generated offline; however, similar plots will be created after developing the complete stand-alone application. Plots in the dashboard include stress distribution underneath the tracks, the track's settlement and the bearing capacity of the subsoil. If the allowable bearing capacity is surpassed at any point, an alarm is displayed.

Figure 7 displays the warning that the current working condition could cause a soil failure. The warning occurs since the stress distribution is triangular, as seen in the dashboard's left-hand side. Therefore, reducing the contact area between the drilling rig's tracks and soil, concentrating the machine's weight into a smaller contact area and thus, drastically reducing the bearing capacity of the subsoil. In this scenario, stability condition 2 is violated while condition 1 remains unchanged.

Figure 8 shows another dangerous working condition displayed in the dashboard. In this case, the total stress in the right track is zero, meaning that the force vector acting at the COM is located outside the drilling rig body, thus causing a rotational moment. This situation could cause the drilling rig's tip over. The pressure on the subsoil underneath the tracks, for this case, remains far from approaching the bearing capacity. This situation violates stability condition 1 while stability criteria 2 remain unchanged.

Figure 7: Dashboard for the case of $F_M = 90 \text{ kN}$, $\phi = 0$, and $d_M = 2.4$ Figure 8: Dashboard for the case of $F_M = 90 \text{ kN}$, $\phi = 90$, and $d_M = 2.4$

4 Conclusions and outlook

This work introduced the first steps for developing a digital twin to monitor the stability of drilling rigs. The stability monitoring combines existing industry standards and focuses on redefining the concept of stability for drilling rigs to include soil conditions. This work presents an insight into the results of this new stability concept and the perspective of developing a digital twin to online stability monitoring, providing examples of how this system could look for limiting cases. The results of this work seek to include basic information about the soil in the stability assessment of drilling rigs since current standards are based solely on the assumption of a rigid subsoil for their calculations.

Extensive work elaborating on the presented results must still be done to achieve full online stability monitoring and a complete digital twin. Experiments should be conducted to establish communication with the internal sensors on existing drilling rigs since this is crucial for the real-time use of the developed system. Suitable hardware should be selected for this purpose, considering the possible training required for the machine operators to use it and the involved costs.

It should be clear that the stand-alone application works in real-time with current simplifications and assumptions; afterwards, the simplifications should be gradually eliminated to include a more realistic digital twin. Furthermore, numerical calculations using the finite element method (FEM) with state of the art constitutive soil models and field measurements should be carried out to improve the stress estimation underneath the tracks. Preliminary results show that the assumption of the trapezoidal distribution results in an oversimplification. Then, a correction function for the stress should be defined, and a new soil stability criteria should be created.

5 Acknowledgments

This research is sponsored by the German Federation of Industrial Research Associations (AiF).

References

- [1] David Edwards and Gary Holt. Case study analysis of construction excavator H&S overturn incidents. *Engineering, Construction and Architectural Management*, 17:493–511, 2010. doi:10.1108/09699981011074583.
- [2] Stephanie G. Pratt, Suzanne M. Kisner, and Paul H. Moore. Machinery-related fatalities in the construction industry. *American Journal of Industrial Medicine*, 32(1):42–50, 1997. doi:https://doi.org/10.1002/(SICI)1097-0274(199707)32:1<42::AID-AJIM6>3.0.CO;2-T.
- [3] H. Beutinger, Peter and P.A. Vermeer. Ein geotechnischer Beitrag zur Standsicherheit mobiler Baumaschinen, 2005.
- [4] Satoshi Tamate and Tomohito Hori. A case study on the overturning of drill rigs on construction sites. In *Contemporary Topics in Deep Foundations*, pages 135–142. 2009.
- [5] Liebherr. Ground pressure visualization. Online: <https://www.liebherr.com/en/aut/products/construction-machines/deep-foundation/assistent-systems/ground-pressure-indication/ground-pressure-indication.html>, Accessed: 11/01/2022.
- [6] Fei Tao, He Zhang, Ang Liu, and Andrew YC Nee. Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, 15(4):2405–2415, 2018.
- [7] Stefan Boschert and Roland Rosen. *Digital Twin—The Simulation Aspect*, pages 59–74. Springer International Publishing, 2016. doi:10.1007/978-3-319-32156-1_5.
- [8] Qinglin Qi and Fei Tao. Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison. *IEEE Access*, 6:3585–3593, 2018. doi:10.1109/ACCESS.2018.2793265.
- [9] Stephan Weyer, Torben Meyer, Moritz Ohmer, Dominic Gorecky, and Detlef Zühlke. Future modeling and simulation of cps-based factories: an example from the automotive industry. *IFAC-PapersOnLine*, 49(31):97–102, 2016. doi:https://doi.org/10.1016/j.ifacol.2016.12.168.
- [10] Manuel FO Seabra Pereira and Jorge AC Ambrósio. *Computer-aided analysis of rigid and flexible mechanical systems*, volume 268. Springer Science & Business Media, 2012.
- [11] Mathworks. Simscape multibody, getting started guide (2021a). Online: https://www.mathworks.com/help/pdf_doc/physmod/sm/sm_gs.pdf, Accessed: 13/01/2022.
- [12] DIN EN 16228-1. Drilling and foundation equipment – Safety – Part 1: Common requirements. Standard, DIN Deutsches Institut für Normung, Berlin, Germany, 2014.

- [13] BS EN 996. Piling equipment - Safety requirements. Standard, British Standards Institution (BSI), 1996.
- [14] BS EN 791. Drill rigs - Safety. Standard, British Standards Institution (BSI), 1996.
- [15] DIN 4019. Soil - Analysis of settlement. Standard, DIN Deutsches Institut für Normung, Berlin, Germany, 2015.
- [16] Joseph Boussinesq. *Application des potentiels à l'étude de l'équilibre et du mouvement des solides élastiques*, volume 4. Gauthier-Villars, 1885.
- [17] W. Steinbrenner. Tafeln zur setzungsberechnung. *Die Straße*, 1, 1934.
- [18] R. Jelinek. Setzungsberechnung ausmittig belasteter Fundamente. *Bauplanung und Bautechnik*, 1949.
- [19] DIN 4017. Soil – Calculation of design bearing capacity of soil beneath shallow foundations. Standard, DIN Deutsches Institut für Normung, Berlin, Germany, 2006.

Closing the Gap Between Concrete Maturity Monitoring and Nonlinear Time-dependent FEM Analysis through a Digital Twin. Case Study: Post-tensioned Concrete Slab of an Office Building, Barcelona, Spain

H. Posada^a, R. Chacón^a, L.C. Ungureanu^b and D. García^c

^a Department of Civil and Environmental Engineering, Universitat Politècnica de Catalunya, Spain

^b Department of Civil Systems Engineering, Institute of Civil Engineering, TU, Berlin, Germany

^c Department of Architectural Technology, Universitat Politècnica de Catalunya, Spain

E-mail: hector.posada@upc.edu, rolando.chacon@upc.edu, l.ungureanu@tu-berlin.de, david.garcia.carrera@upc.edu

Abstract –

This paper proposes a pipeline for the automation of data between two realms: i) simulation, in a fully nonlinear, and time-dependent structural analysis model and, ii) concrete maturity monitoring data from the construction site. The connection enables an information construct understood for its use within the Digital Twin of the building during construction phases. The pipeline requires a comprehensive coordination between stakeholders at both the site (Construction) and the structural control office, which is challenging. The pipeline consists of a) temperature sensors, b) a mobile app connected to the sensor via Bluetooth with basic instructions for data-gatherers, c) integration and interoperability of BIM, and, d) an advanced Finite Element (FE) model. By measuring the concrete temperature during many days, realistic concrete mechanical properties are inferred and infused into the FE models using adequate calibration. Two applications for the improvement of construction activities are identified. Formwork striking and tendons stressing. The paper describes the testbed of all the connections, for the construction of an in-situ casted concrete building in Barcelona, Spain.

Keywords –

Digital Twin; Concrete Maturity; BIM; Sensor-based Data; Mobile App; Nonlinear Finite Element Analysis

1 Introduction

A Digital Twin is an information construct that may include manifold pipelines from data-gathering to

decision-making. In the specific realm of the built environment, it would be a data representation of the physical infrastructure that takes real-time and other data into the management processes of that real-infrastructure component, as defined by the National Infrastructure Commission in the UK, NIC (2017) [1]. It becomes the living version of a Building Information Model (BIM). Active connections between the virtual and the physical realms enable a data-driven decision-making process [2].

The level of maturity of BIM in the construction sector is significant. Nevertheless, the road to automation to deliver smarter construction services by ensuring information interoperability, a consistent use of information constructs forming the Digital Twin is required [3].

Many technological solutions for automatically monitoring construction works are available and applied commercially, anyway, this data is generally used in disaggregate manners [4]. Providing information pipelines for erecting a richer Digital Twins of an asset is a way of centralizing many potential layers of information. The ability to process and merge data from multiple sources enables a more efficient construction control [4].

This paper aims to present a case study on the digital twinning of buildings during the construction phase. Through the monitoring of concrete maturity, compressive strength evolution is predicted with sensor-based data collected from the construction site. Subsequently, a nonlinear and time-dependent finite element analysis is fed with the adequately calibrated, predicted strength evolution. Data and results are infused and visualized within a Digital Twin platform. This information is thus available for many stakeholders, i.e., construction managers. Decisions on construction tasks such as formwork striking or tendons stressing in post-

tensioned and long-spanned concrete slabs can then be taken based upon a more realistic analysis. More accurate models that account for the rheology of the materials are developed to describe the time-dependent structural behavior of the concrete structure during construction.

The study is developed within the frame of the H2020 European project ASHVIN (Assistants for Healthy, Safe, and Productive Virtual Construction Design, Operation & Maintenance using Digital Twins). The purpose is to pull the digital threads and methodologies into coherent solutions to be implemented over the construction process, as well as wrapped up data systematically for further use during the life-cycle of the asset. Ten democases are under development (design, construction and maintenance phases). The particular demo case regarding this research is the construction of an office building in Barcelona.

This paper proposes a pipeline for the automation of nonlinear and time-dependent structural models with concrete maturity monitoring data from the construction site. A coherent integration of data from the site, sensors, a mobile app, FE-models, and BIM within a Digital Twin platform is presented. An assessment of how this pipeline is enabling efficient decision-making and quality control for formwork striking and tendons stressing activities is the ultimate goal of this case study.

2 The Office Building

The case study is a concrete building under construction located in the Barcelona district of Poble Nou. Currently, there are considerable construction projects under development inside this district under the umbrella of the 22@ innovation district. BIS structures, a Barcelona-based structural engineering office, together with many other stakeholders, agreed upon collaborating with ASHVIN, by facilitating access to MILE – Business Campus project construction site. MILE is an office building project of 38,093m², divided into three complexes: MILE-Badajoz, MILE-Llul, and MILE-Ávila. The access was provided for the specific module of MILE-Ávila, a cast-in-place reinforced concrete building of long-spanned post-tensioned slabs, consisting of eight levels, and a total area of 16,524m².



Figure 1. MILE Ávila office building render (Provided by BIS structures)

This case study represents an ideal scenario to define a framework for building construction digital twinning on several levels:

- The concrete casting of this building is continuous and sequential, monitoring the formwork installation process is vital for project productivity.
- The structure is heavily controlled and dependent on the slab deformation and serviceability limit states, prestressed activities play an important role in securing maximum allowable deflection.
- The construction site is located within an urban area, where space is limited and heavy traffic is constant.

Figure 2 shows the plan view of the post-tensioned slabs under study. These elements span 15.60 m plus a cantilever of 4.40 m. The transverse direction span is 5.40 m. Their cross-section height is 50.0 cm, with hollow core, and top and bottom slabs of thickness 7.5 cm. Ribs of width 38.0 cm are distributed every 1.08 m in the longer span direction. Beams are post-tensioned through single-strand tendons with a diameter of 15.2 mm, an area of 140 mm² and a prestress load of 195kN. The concrete class for the slabs is HP-40/B/20/IIa, defined in the EHE-08 structural concrete Spanish code [5], which is equivalent to C40/50 in the Eurocode classification.

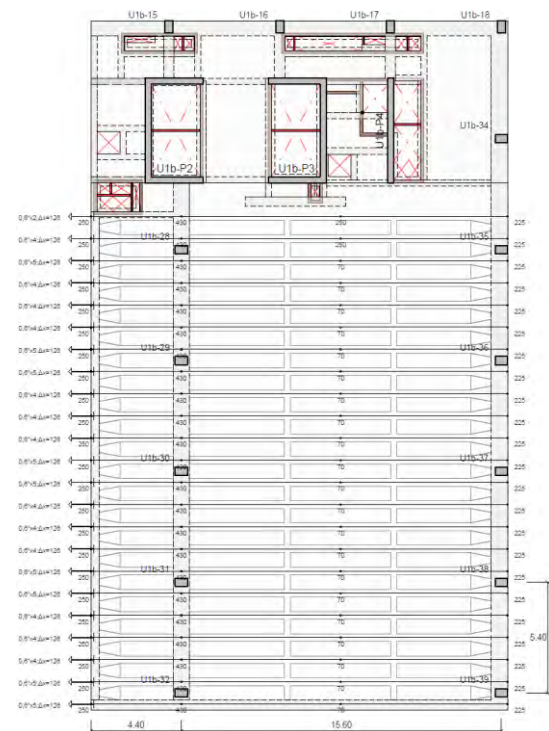


Figure 2. Mile Ávila post-tensioned long-spanned slab plan view (Provided by BIS structure)

3 Concrete Maturity Monitoring

Monitoring the concrete maturity by measuring physical variables such as temperature may provide a good estimate of the compressive strength evolution. Prediction of such properties requires a maturity method. In this section, the theory and attributes of the ASTM C1074 method are first described in 3.1. Subsequently, potential benefits of its use are pointed out in 3.2.

3.1 The Maturity Method

Maturity is the development of the physical properties of the concrete as the hydration process progresses, including the gain of strength [6].

Since the 1950s, a maturity method has been developed to estimate the early-age concrete strengths as a function of temperature and time [6]. The widely-implemented method adopted by ASTM C1074 [7], which assumes a nonlinear relationship between concrete temperature and the rate of development of concrete properties [8], defines an equation to calculate the maturity index:

$$t_e = \sum_0^t e^{\frac{-E}{R}(\frac{1}{T} - \frac{1}{T_r})} \Delta t \quad (1)$$

Where t_e is the equivalent age at the reference temperature, E is the apparent activation energy, R is the universal gas constant, T is the average absolute temperature of the concrete during interval Δt , and T_r is the absolute reference temperature.

According to ASTM C1074, the relationship between concrete strength (S) development vs time (t) can be determined for a particular concrete mix, using Carino's hyperbolic equation, as modified by Knudsen [9]:

$$S = S_u \frac{\sqrt{k(t - t_0)}}{1 + \sqrt{k(t - t_0)}} \quad (2)$$

Where S_u is the ultimate compressive strength, t_0 is the initial time due to the compressive strength development, which corresponds to the time of hydration heat increase, and k is the reaction constant.

Implementing the maturity method provides data nearly immediately and does not involve transporting samples or the management of crushing deadlines [10]. It requires calibration and it is mix-dependent. The model cannot be extrapolated to considerable changes within the mix that may occur in subsequent concrete batches.

Currently, sensor-based techniques are accessible for measuring and monitoring temperature in the interior of the concrete during the curing process. For instance, *SmartRock2*, *TM TEMPCON* or *CONCRETE SENSORS* are some of the brands available in the market [11].

Yikici et al. developed a wireless real-time

monitoring system for concrete maturity calculation, which predicts the compressive strength of steam cured precast concrete, in addition, an IoT technique continuously compares measurements to expected values and sends emails notifications to production engineers [12].

For the prediction of long-term strengths, the existing maturity equations do not provide a good estimation. This model only considers the chemical reaction as the only affecting process [13], however, it is relevant for construction operations.

3.2 Advantages in construction activities

3.2.1 Formworks Striking

The cost of formworks for casted on-site constructions of concrete buildings is significant. For most structures, the time and cost required to erect, and strike the formwork is greater than the time and cost to place the concrete or reinforcing steel [14]. Formworks are often rented, representing the largest single cost component of a structural frame building construction. It varies between 35 and 45% of a reinforced concrete structure's unit cost [15].

Monitoring the temperature of the concrete at early ages allows predicting the compressive strength evolution. This information can be used to reduce the striking time of the formwork [16-17]. Idle time can be minimized and decisions can be taken using real data. For this purpose, an adequate framework that intertwines production, cyber-physical systems, and digital and computing technologies is required.

3.2.2 Tendons Stressing

During the construction of the post-tensioned concrete slabs, the post-tensioning of the tendons must not be initiated before the concrete reaches the transfer strength specified by the structural project engineer [18]. According to ACI 318-19 [19], compressive strength must be at least 17,5 MPa for single-strand or bar tendons, and 28,0 MPa for multistrand tendons. To verify the early age strength of the concrete, samples shall be tested, securing storage and curing conditions similar to those of the in-situ concrete [19].

Adequately calibrated concrete maturity functions prove to give reliable strength development data. This has the potential to simplify the procedure of prediction of the strength and monitor the development on post-tensioned concrete slabs [20].

Luke et al. [21] investigated the ability of the maturity method to predict the prestressing release strength for steam cured precast/prestressed concrete box beams, comparing the compressive strength of several companion cylinders with the one obtained using the maturity method, which led to accurate results. A set of

recommendations was proposed to reduce the number of test cylinders.

4 Methodology

An automated pipeline framework (Fig. 3) is proposed in this study. An information construct that adds a decision-prone layer to the Digital Twin of the building is proposed. The pipeline connects nonlinear time-dependent FE models with concrete monitoring data for cast-in-place reinforced buildings during the construction phase. The rheology of the material is thus considered within the time-dependent structural analysis.

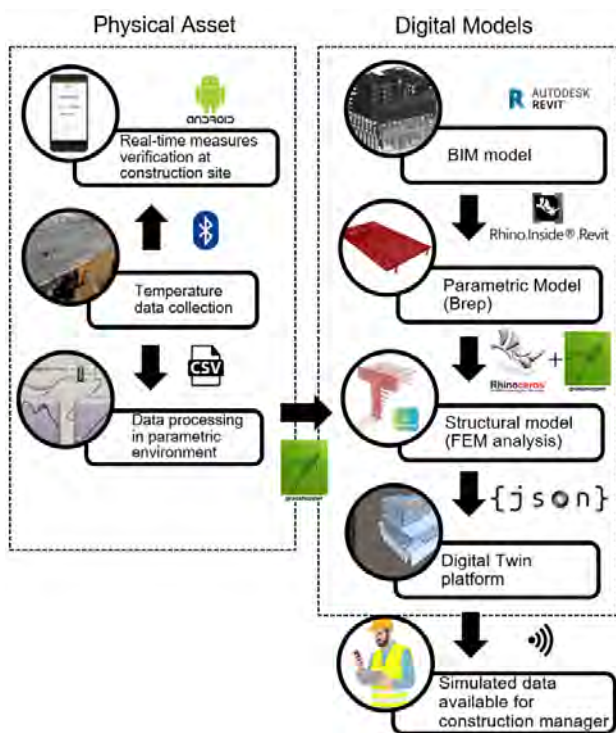


Figure 3. Proposed automated pipeline

4.1 Temperature data collection

For this study, prototypes for data acquisition were developed. Electronic prototyping platforms allowed tailor-making of the data flow. Two different types of temperature measuring sensors were used to collect data from long-spanned slabs. The first sensor type was a thermocouple K-type sensor with a range from 0°C to 400°C and an error of $\pm 2.2^\circ\text{C}$, connected to a MAX6675 module that transforms the analog signals to digital, two units of this sensor were used. The second sensor type was a DS18B20, a waterproof digital sensor that uses a Wire-1 protocol to communicate and measure temperature in a range between -55°C to 125°C with an error of $\pm 0.5^\circ\text{C}$. The energy was supplied by power banks

of 5000mAh. The sensor data was saved on an SD card for further processing. A Bluetooth HC-05 ZS-040 shield sent data to mobile phones for real-time values verification. All these components were plugged and synced to an ESP32 SoC (System on Chip) module through the Arduino IDE (Integrated Development Environment) platform. These hardware modules represent a cost-effective tool with open prototyping capabilities.



Figure 4. DS18B20 hardware module installed during concrete slab pouring

The locations of the sensors are shown in Figure 5. Access during the concrete pouring of the slab was restricted. Sensors were mounted in the allowable areas and fully embedded in the concrete mix to avoid interference with ambient temperature. There is a margin of improvement when it comes to the number of sensors as well as their location for capturing more adequate data. In this particular case, due to the site restrictions, a limited amount of points were set.

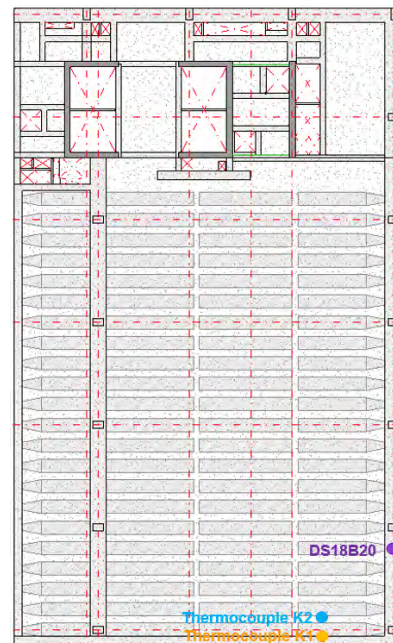


Figure 5. Sensors' locations at building top slab (Image extracted from the BIM provided by BIS structures)

Figure 6 shows data collected from sensors. Temperature evolution was recorded for 34 hours by Thermocouples K and 30 hours by DS18B20 sensor.

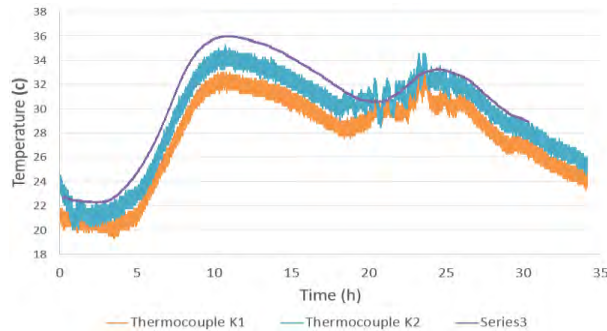


Figure 6. Data collected from temperature sensors

4.2 Real-time temperature verification

A Mobile App was developed to retrieve real-time data from the sensors via Bluetooth (BT) at the construction site. The app was conceived for its use by non-technical staff, with a user-friendly and simple interface. The open prototyping environment also allowed embedding BT protocols. The app was developed using the MIT block-based coding online platform App Inventor, which consists of an intuitive visual programming environment (IDE) for developing Android-based applications [22].



Figure 7. Android-based mobile application developed

The Bluetooth communication between sensors and the mobile app is performed through the HC-05 shield connected to the ESP32 SoC. The data received at the time interval defined by the user can be stored in a CSV file together with a time stamp.

4.3 Data processing

Concrete temperature history is the input data to calculate the maturity index (Eq.1). This leads to

predicting the compressive strength evolution by computing values in the calibrated graph of strength vs. maturity index. This graph is obtained through laboratory testing following the instructions of the ASTM C1074 [7], the results are valid for a particular concrete mix.

The temperature sensors data, stored as CSV files in the SD cards of the hardware modules, is transferred to Grasshopper, in which a developed Python snippet allows estimating the maturity index. Using the temperature data, and predicting the concrete strength evolution within the parametric environment (Fig. 8).

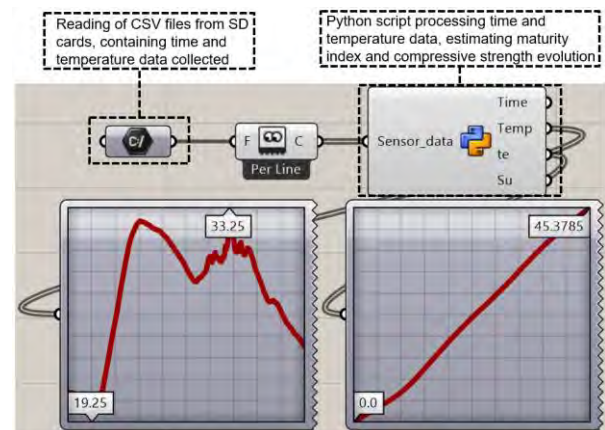


Figure 8. Data processing in a parametric environment (Grasshopper 3D).

4.4 From BIM geometry to FEM structural model

BIM is called to implement a digital track of the building and infrastructure. In this way, BIM significantly improves information flow between stakeholders involved during design and construction phases [23].

The Digital Twin is called to implement a digital track of the building and infrastructure during the entire lifecycle of the asset. This living BIM model becomes useful at all stages (design, construction and operation).

BIS structures provided the BIM structural model for MILE Ávila on Revit 2019. No architectural information was supplied, rebar detailing was retrieved from sheets for retaining walls, slab foundations, floor slabs, and beam active reinforcement. The building was primarily conceived for cast-in-place concrete. The BIM model is classified as LoD (Level of Development) 350.

Using Rhino.Inside.Revit, a Grasshopper plug-in, parametrics models were built. It was thus possible to access the topological components of a three-dimensional Brep (body, face, edge, and vertex). This option enables the model for the further discretization of cross-sections to perform elastic or inelastic analysis (Fig. 9).

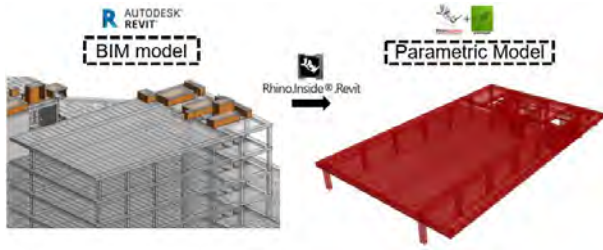


Figure 9. From BIM model to top floor slab parametric model

Having access to the components of Brep (Boundary Representation), the discretization is enabled and calculations are performed using one of the methods developed within MatchFEM. MatchFEM is a tool under development within ASHVIN. Data collected from the sensors, processed in the same parametric environment will feed the FE-models.

4.5 Finite Element Numerical Model

The maturity method estimates with accurate precision the development of the concrete compressive strength as a function of time (see section 3). This information is embedded within the time-dependent analysis. Since one of the purposes of the study is to monitor structural behavior during a construction process, the application of loads and boundary conditions change over time as part of an evolutive construction sequence.

To process accurately the conditions of the construction sequence, structural elements are calculated based on the numerical model of nonlinear and time-dependent analysis for three-dimensional reinforced, prestressed, and composite concrete, developed by Mari [24].

The mathematical model is based on frame finite elements with six degrees of freedom per node. Each element has a length and a prismatic cross-section discretized in filaments, which have a material (concrete or steel) associated and are geometrically defined by their area and position with respect to the sectional local axes.

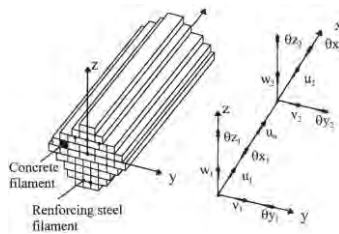


Figure 10. Filament frame element. (Source [24])

For two different materials, strain compatibility exists, therefore, perfect bond is assumed. The plane strain principle is implemented, consequently, is satisfied for mechanical and nonmechanical strains:

$$\Delta \varepsilon(z) = \Delta \varepsilon_m - z * \Delta c \quad (3)$$

Uniaxial stress-strain constitutive models are applied for concrete, active and passive reinforcement, including the rheology of the materials. Loads are imposed to the nodes on specific instants, being able to introduce load increments, allowing a step-by-step analysis over time. Prestressing loads are introduced as an equivalent load vector obtained by balancing the forces of the prestressing tendons [25].

This numerical model was the base for the development of a computer program called CONS, which presents great computational efficiency.

MatchFEM, one of the tools under development for the ASHVIN project digital toolkit, will implement several methods to stochastically adjust input parameters for multi-physics simulations using sensor-based measurements. The aim is to accurately represent many layers of behavior within the Digital Twin of the asset. In this particular case, to perform the structural analysis, MatchFEM benefits from the efficiency and accuracy of CONS and levels up to a powerful graphical environment through parametric design tools such as Rhino and Grasshopper. Notwithstanding, the information construct is conceived for proper interoperable models based on IFC (Industrial Foundation Classes). Sensors, FEM analysis and geometries can be then matched seamlessly.

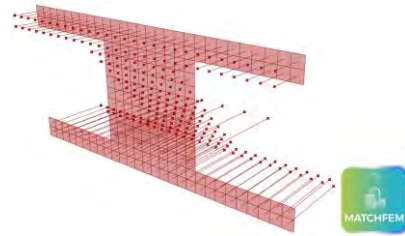


Figure 11. MILE-Ávila top slab cross-section filaments model in MatchFEM tool.

4.6 Digital Twin platform

Within the ASHVIN project, a game engine based Digital Twin platform is under development. It is conceived for enabling connectivity, device management, data acquisition and embeddedness of the different tools. The ASHVIN platform will allow for plug-and-play connectivity with a wide range of devices and real-time data processing functionality. First, the incorporation of as-designed data from BIM models is necessary. An IFC standard format or open file format can be converted to the platform together with all BIM metadata.

Simulated data from the Grasshopper software via MatchFEM ASHVIN tool is converted to JSON string. As a result, visualization of the proper physical behavior of the Digital Twin is possible, including the temperature history, maturity index, concrete strength evolution, and

the nonlinear time-dependent FEM analysis.



Figure 12. Simulated data to Digital twin platform to Construction Manager

Construction managers can track the evolution of realistically fed models of sequential constructions. With such information, it may be possible to take decisions on construction activities, such as formwork striking or stressing of tendons.

5 Discussion

This section summarizes the advantages and challenges of the proposed methodology.

5.1 Automation advantages

The result of the automated pipeline is a digital asset that holds the physics behavior data history of a real building structure at the construction phase. This information is assembled in a Digital Twin environment, where is available in a user-friendly graphical interface. Construction managers can benefit from the Digital Twin for a more data-driven decision-making process.

Applying isolated concrete monitoring systems without an active connection between the real asset and the digital models could lead to deficient data management and increase data loss risk, reducing their impact on construction activities.

In addition, Digital Twins store data for long-term use, therefore, owners or stakeholders have access to the physics history of the structure and could use this information for future interventions or maintenance plans.

Particularly for the case study, prestress activities were planned to start after 5 days of concrete curing time. Maturity index results showed that stressing of tendons could initiate at least one day earlier. Hitherto, data collection time was not enough to evaluate formwork striking.

5.2 Methodology implementation roadmap

Following, limitations, lessons learned, and future work are gathered to establish a roadmap for the implementation of the proposed methodology:

- The hardware modules used for data acquisition had a limited energy supply and a restricted mounting process. Future works based on integrated existing commercial sensor-based

maturity monitoring systems are needed (see section 3.1), capable of sending data wireless being completely embedded in the concrete mix, enabling installation before the slab pouring to avoid interference with the activity.

- As defined by Carino [9], sensors must be mounted at locations with critical exposure or structural requirements. Project engineers must define the numbers and locations of sensors for each slab. For this research, locations were limited to accessible areas. The accuracy of the predictive model requires profuse data collection. However, this may not be possible due to site conditions. In this case, cylinder samples shall be taken.
- The development of procedures for sending data directly from the sensor to an IoT platform and then into Grasshopper, will improve the data flow and avoid human intervention.
- Traditionally, for the quality control of the concrete compressive strength, many cylinder samples are tested. The ASTM C10741 procedure has the potential to reduce cylinder testing, as it only requires crushing samples for a particular concrete mix [7]. Nevertheless, owners and stakeholders have to agree upon accepting concrete maturity as the method for quality control to have a real impact in striking formwork or prestressing activities.

6 Conclusion

A comprehensive pipeline for automation of nonlinear and time-dependent structural models with concrete maturity monitoring data from the construction site was proposed, describing the integration of the BIM model provided by the owner, the hardware modules for sensor-based temperature data, a mobile app for real-time values verification on the construction site, the finite element model analysis fed by the evolution of the concrete compressive strength at early ages, and the incorporation of the results into a Digital Twin platform. The implementation of this framework has the potential to improve the decision-making and quality control for formwork striking and tendons stressing activities.

The developed methodology paves the way for the application of Digital Twins of buildings during the construction phase, with an explicit active connection between digital models and the physics of the real asset.

7 Acknowledgments

All authors acknowledge the funding of ASHVIN, “Assistants for Healthy, Safe, and Productive Virtual Construction Design, Operation & Maintenance using a

Digital Twin” an H2020 project under agreement 958161.

In Addition, the authors are grateful to BIS structures, BCA, DRAGADOS, and FREO for providing access to the construction site for measurements and data collection, as well for their disposition to share project information and follow the development of this research.

References

- [1] Kawalek P and Bayat A. Data as Infrastructure. *National Infrastructure Comision (NIC)*, London, UK, 2017.
- [2] Dávila Delgado J. M and Oyedel L. Digital Twins for the built environment: learning from conceptual and process models in manufacturing. *Advanced Engineering Informatics*, 49, 2021.
- [3] Boje C, Guerreiro A, Kubicki S and Rezgui Y. Towards a semantic Construction Digital Twin: Directions for future research. *Automation in Construction*, 114, 2020.
- [4] Sacks R, Brilakis I, Pikas E, Xie H and Girolami M. Construction with digital twin information systems. *Data-Centric Engineering*, 1(14e), 2020.
- [5] Instrucción de Hormigón Estructural EHE-08 (Spanish Structural Concrete Code). *Spanish Minister of Public Works*, 2008.
- [6] Nixon J, Schindler A, Barnes R and Wade S. Evaluation of the maturity method to estimate concrete strength in field applications. *Research Report, Auburn University, Department of Civil Engineering*, 2008.
- [7] ASTM C1074-19. Standard Practice for Estimating Concrete Strength by the Maturity Method. *ASTM International*, 2019.
- [8] Hoon Lee C and Hover K. Compatible Datum Temperature and Activation Energy for Concrete Maturity. *ACI Materials Journal*, 113(2): 197-206, 2016.
- [9] Carino N and Lew H S. The Maturity Method: From Theory to Application. In *Procedings of the conference Structures Congress and Exposition*, pages: 19, Washington, USA, 2001.
- [10] Benaicha M, Burtschell Y and Alaoui A. Prediction of compressive strength at early age of concrete - Application of maturity. *Journal of Building Engineering*, 6: 119-125, 2016.
- [11] Shima Taheri. A review on five key sensors for monitoring of concrete structures. *Construction and Building Materials*. 204: 492-509, 2019.
- [12] Yikici A, Bekmezci I and Surncu E. Wireless Real-Time Monitoring System for Steam Cured Concrete Maturity Calculation. In *Proceedings of Innovations in Intelligent Systems and Applications Conference (ASYU)*, pages 1-4, Elazığ, Turkey, 2021.
- [13] Kim T and Rens K. Concrete Maturity Method Using Variable Temperature Curing for Normal and High-Strength Concrete I: Experimental Study. *Journal of Materials in Civil Engineering*, 20(12): 727-734, 2008.
- [14] Peurifoy RL and Oberlender GD. *Formwork for concrete structures*, edition 4. McGraw-Hill, 1325 Avenue of the Americas, 10019 New York, 2010.
- [15] Krawczynska-Piechna A. Comprehensive Approach to Efficient Planning of Formwork Utilization on the Construction Site. In *Proceedings of the conference Engineering, Project, and Production Management*, pages 366-372, Bialystok, Poland, 2017.
- [16] Skibicki S. Optimization of Cost of Building with Concrete Slabs Based on the Maturity Method. In *IOP Conference Series: Materials Science and Engineering* 245, 022061.
- [17] Basková R, Struková Z and Kozlovská M. Construction Cost Saving Through Adoption of IoT Applications in Concrete Works. In *Proceedings of the conference Advances in Resource-saving Technologies and Materials in Civil and Environmental Engineering*, pages 452-459, Lviv, Ukraine, 2019.
- [18] BS EN 13670:2009. Execution of Concrete Structures. *British Standards*, 2009.
- [19] ACI 318-19. Building Code Requirements for Structural Concrete. *American Concrete Institue*, 2019.
- [20] Sofi M, Mendis P.A and Baweja D. Estimating early-age in situ strength development of concrete slabs. *Construction and Building Materials*, 29: 659-666, 2012.
- [21] Luke A, Konon W and Punurai S. Maturity Method in Prestressed Concrete Applications. *New Jersey Department of Transportation*, 2004.
- [22] Chacón R, Posada H, Toledo A and Gouveia M. Development of IoT applications in civil engineering classrooms using mobile devices. *Computer Applications in Enginnering*, 26(5): 1769-1781, 2018.
- [23] Borrmann A, König M, Koch C and Beetz J. *Building Information Modeling: Technology Foundations and Industry Practice*, volume 1. Springer, Gewerbestrasse 11, 6330 Cham, Switzerland, 2018.
- [24] Marí A. Numerical simulation of the segmental construction of three dimensional concrete frames. *Engineering Structures*, 22: 585-596, 2000.
- [25] Marí A. Manual Programa CONS para el Análisis No Lineal en el Tiempo de Estructuras de Hormigón Estructural Construidas Evolutivamente. *Departamento de Ingeniería de la Construcción, Universidad Politécnica de Catalunya*, 2005.

ColombiaClass: proposal for a BIM Classification System for public buildings in Colombia

M.P. Revuelta^a, N.P. Garcia-Lopez^b and L.A. Salazar^c

^a Research Assistant, Department of Civil and Environmental Engineering, Universidad de los Andes, Colombia.

^b Adjunct Professor, Department of Civil and Environmental Engineering, Universidad de los Andes, Colombia.

^c Associate Researcher, Young Researchers for Architecture Engineering and Construction Industry, Chile.

E-mail: m.revuelta@uniandes.edu.co, ne-garci@uniandes.edu.co, lasalaza@uc.cl

Abstract –

The use of standardized BIM Classification Systems helps to improve the efficiency of the tendering, planning, and control of construction projects. The Colombian government has identified BIM as a key strategy to help increase the construction industry's productivity and improve value for money in public works projects. At the core of the strategy is the need to implement a standardized Classification System that allows the government to consistently verify project scope, compare bids, and easily control projects. Although numerous Classification Systems already exist, they do not include prevalent construction methods and linguistic variations specific to Colombia, which can lead to misrepresentation issues and lack of standardization across the industry. In this paper, a Colombian BIM Classification System is proposed, called ColombiaClass. ColombiaClass was developed based on the Chilean regulations for the standardization of construction items (NCh 1156), and adjusted to the Colombian context defining the order, classification, and name up to level 3 using current technical regulations, different databases for budgets of official entities and the collaboration with Colombian construction professionals. It was validated in 2 ways: checking ColombiaClass's ability to represent multiple official construction budgets corresponding to five Colombian regions (83% of the country) and using the Delphi Method, conducting two expert panels: one with field engineers and architects, and the other with BIM experts from the Colombian Ministry of Housing, BIM Forum Colombia, and the Colombian National Planning Department (DPN). As future research, the authors call to develop levels 4 and following, using the methodology described in this paper.

Keywords –

Building Information Modelling; BIM Classification Systems; Colombian BIM Classification System; Public works; National BIM strategy; ColombiaClass

1 Introduction

Unlike in other industries, productivity in the construction industry has not increased in the last 50 years [1]. In Colombia, 40% of public work projects during 2019 presented delays [2][3], and one of the main causes was attributed to the uncertainty of project cost and non-compliance with project schedules [4]. Construction plays an essential role in the Colombian economy, representing 6.9% (average between 2015 and 2019) of the Gross Domestic Product (GDP) [5]. For this reason, the Colombian Government has launched a National BIM Strategy 2020-2026 seeking to increase productivity and improve project management practices, especially information management [4].

Public works in Colombia need to have transparent, auditable, non-fragmented processes that articulate the different actors [4]. Since 2019, several government decrees have been published called "standard tenders" that standardize public works bidding processes focusing on legal and professional experience requirements. However, to date requirements for budget structuring and project controls have not been addressed.

Currently, there is no norm that standardizes the construction items or the budget structure for public works in Colombia. Hence, each firm has its own ad-hoc Classification System. This makes it extremely difficult and time consuming for public entities to compare scopes and pricing between bids, and to carry out project controls throughout the project. Moreover, without a standardized BIM Classification System it is impossible to pursue initiatives towards the automatization and robotization of the tendering and control processes.

Since 2020, construction licenses in Colombia can be submitted to public authorities in a digital IFC format [6]. Having a standardized Classification System could enable stakeholders to leverage BIM through the entire building life cycle, taking advantage of IFC's data exchange formats that support interoperability between numerous BIM platforms[7].

Implementing an existing Classification System in Colombia has two main drawbacks. First, most Classification Systems are in English, making them unsuitable for public bids since Colombian contracting law requires that all documents must be in Spanish. On the other hand, Classification Systems that are in Spanish have been adapted to the specific linguistic and construction practices of their home countries (e.g., Chile NCh 1156 standard and Spain GuBIMClass Classification System). Implementing these existing Classification Systems in Colombia would not respect the linguistic diversity of the country and lead to misclassification problems. Moreover, there are several construction methods that are used in Colombia that are not typically used in other parts of the region. For example, Caissons are still a popular deep foundation method used in the country, partition walls are typically in clay masonry, and there are structures in special woods such as Guadua or bamboo.

In this paper, a BIM Classification System is proposed called the ColombiaClass. It is based on the Chilean NCh 1156 (designation of items) standard and was adapted to the Colombian context by comparing, analysing, and classifying the construction items of 33 existing public works budgets from several Colombian Public entities as well as different private databases. The proposed Classification System was validated by using the Delphi method, where two expert panels were held to determine adjustments considering the perspective of builders, designers, academics, and BIM professionals from construction companies and the Ministry of Housing, the National Planning Department (DNP) and the BIM FORUM Colombia. In the short term, ColombiaClass will help the Colombian government's goal to tender all public works projects using BIM methodology by 2026 and push towards standardized tenders.

2 Methodology

Chile is a pioneer in the region in BIM implementation for public works, with a 10-year National BIM Strategy from 2016 to 2025. As a result, by 2020, 53% of the projects of the Ministry of Public Works have been tendered under the BIM methodology described in its PlanBIM [8]. Hence, the Chilean standard NCh 1156 of 2018: construction - technical specifications - management and designation of items [9] was used as a

basis for the ColombiaClass. The NCh 1156 was analysed and fully understood by reviewing documents, databases and consulting with a Chilean engineer who has been working in field for more than 10 years.

Using the NCh 1156 as a base, the ColombiaClass scope and construction item names was reviewed by following three steps: 1) reviewing applicable Colombian technical regulations 2) analysing publicly available budgets for different types of buildings 3) checking the scopes of other Classification Systems. During this process the construction items in the proposed Classification System were named, hierarchized, and designated.

2.1 Review of technical regulations

An exhaustive review was made for the main technical regulations related to public works in Colombia checking the naming and completeness against the NCh 1156:

- RAS2000 [10] is the technical regulation for hydraulic infrastructure networks and NTC1500 [11] is the Colombian Pipework code. Based on these two documents, the correct names for networks, accessories and structures were extracted according to whether the systems were for buildings or infrastructure.
- RETIE [12] and Retilap, [13] both contain the technical specifications for electrical and lighting networks. From these, construction items were extracted related to electrical networks, telecommunication, public lighting, normal and emergency lighting.
- IDU Sidewalks Booklet 2018 [14], Bogotá Planning Office. This technical specification was used to extract the names of the pavement layers and the complementary elements of the sidewalks.
- The book Architects' Data (in Spanish "El arte de proyectar en Arquitectura") [15], was used to understand, classify and name the different architectural elements. In general, Colombian architectural projects use the nomenclature that is used in this book.
- INVIAS technical documents for roads [16]. The authors used the technical specifications published by INVIAS to extract names, components and elements that make up the roads and their complements.
- Health infrastructure regulations of the Ministry of Health[17]. This ministry has published several decrees establishing the guidelines and requirements for health centres. From these documents, the authors extracted mainly line items related to the management of clinical gases and extraction of contaminated water and clinical waste.

2.2 Analysis of official budgets

A Classification System for BIM is not limited to coding the construction elements (3D), it is also used for scheduling (4D) and cost estimating (5D) [1]. Therefore, it is necessary that the construction items in the ColombiaClass allow field managers to link BIM elements with budget elements. In this step, the authors adjusted the construction items that were identified in the previous step according to the results obtained when analysing the 33 official budgets described in Table 1.

The choice of budgets was developed using an information-oriented selection with maximum variation cases strategy [18]. The selection criteria were the type of public building, the contracting entity, and the geographical location of the project. Additional budgets were added to the analysis if they provided added information along a certain dimension. The analysis was closed once no new budgets could be found that added information dimensions to the case studies that had already been analysed. The selected budgets correspond to 25 state entities and different types of public buildings that have been tendered in recent years.

Table 1 Official budgets analysed.

Entity	Qty	Building type
Aerocivil	4	Airport
Ministry of Culture	2	Library
Mayor's Office of Espinal	1	Cafeteria
USPEC	1	Jail
Mayor's Office of Remedios	1	Social centre
Gov. of Cundinamarca and N.S	2	Clinic
Mayor's Office of...	3	School
UAESP	1	School
Pascual Bravo Tech institute	1	School
Gov. of...	2	School
Mayor's Office of...	3	Sports centre
IMDRI	1	Sports centre
Mayor's Office of Cucuta	1	Office
Findeter	2	Park
IDRD	1	Park
IDU	1	Skating rink
College of Cauca	1	University
IDU	2	Transport Station
Metrocali	2	Transport Station
Mayor's Office of San Juan	1	Houses

The analysed budgets correspond to projects located in five of the six regions of the country as shown in Figure 2. Only the island region of San Andres and Providencia is missing.



Figure 1. Budgets analysed by region (adapted from ICA Map)[20]

One of the objectives of the project was to cover the entire range of public building types tendered in Colombia. The authors found and analysed budgets for 14 types of public buildings, which to the author's knowledge, covers the entirety of tendered public buildings in Colombia published in SECOP. ColombiaClass was able to represent the scopes of all the different types of building projects. On the other hand, the variety of budgets selected by location and contracting entity made it possible to compare and adjust the linguistic diversity, construction processes, elements and materials used in different regions.

2.3 Analysis of existing Classification Systems

After adjusting the ColombiaClass line items with the technical regulations and official budgets, a review and comparison was made between the scopes of the Uniformat 1998 [21], UniClass [22], and GuBIMClass [23] classification against the ColombiaClass.

The Uniformat was analysed because it is an American Classification System and most of Colombian design standards have been based on American standards. Additionally, the Uniformat has the same hierarchical criteria as ColombiaClass. The EF table of UniClass was also analysed because it is an extremely complete Classification System that is also structured using a hierarchical organization. Finally, the GuBIMClass was analysed because it is one of the few Classification Systems that were created specifically to be used in conjunction with the BIM methodology [24]. Also, it has an important focus on the design and construction phases, that are the phases in which public projects are tendered in Colombia. Overall, one of the most significant adjustments that were made to the ColombiaClass was

incorporating the health and safety section of GuBIMClass, which is a critical element to ensure project success especially during the past few years due to COVID-19 epidemic.

2.4 Validation using the Delphi Method

The Colombia Class was validated by using the Delphi method, which is a common research method employed in numerous scientific disciplines. Only in 2008, 105 of 15,000 articles published by 4,000 publishers in Scopus employed the Delphi Method [25]. This method consists of a panel of several experts, preferably between five and ten, who are presented with a complex topic or problem and after discussion agree on a solution [26]. The value of the Delphi method lies in the ideas that come out of the expert panel, since the communication in this activity is effective, allowing the consensus of the participants [26].

For this paper, the topic to be addressed by the panel of experts was the ColombiaClass up to level 2: chapters and subchapters through a top-down approach. Two expert panels were held, the first one focused on the validation of the names of the first two levels of the ColombiaClass. The experts were composed of engineers and architects BIM coordinators, designers of hydraulic, gas, and electrical networks, contractors with more than 25 years of experience in public works, project engineers, and a member of the Department of Public Works. It should be noted that these panellists are from different regions of the country, to achieve a correct validity of the terminology.

The second panel focused on the validation of the scope, hierarchy, and order of the ColombiaClass. The experts were composed of engineers and architects with extensive BIM knowledge and experience in BIM implementation in Colombia. These panellists work with the Ministry of Housing, the BIM Forum Colombia, and advising public and private companies in BIM implementation and the creation of Classification Systems for private construction companies.

The Delphi Method provided useful feedback that enriched the research and completing the final adjustments to the ColombiaClass. Some of the most significant changes were a) scope of wet and dry networks, b) inclusion of industrial wastewater according to applicable regulations, c) more intuitive and user-friendly nomenclature.

3 Proposed classification system for BIM: ColombiaClass.

The ColombiaClass has the following coding:

AAA. ##. ##. ##

Where the first 3 letters correspond to the chapter, the first 2 digits that follow correspond to the subchapters, the next 2 digits to the section, and the last 2 digits to the elements. In addition, the numbers increase by 10, giving the user the flexibility to enter information that was not classified while preserving a logical hierarchical representation. It is worth mentioning that the last 2 digits of the coding corresponding to the elements were not developed since they are not part of the scope of this research.

ColombiaClass level one is composed of 21 chapters show in table 2. The letters that name each of these chapters correspond to an abbreviation of the chapter name in Spanish.

Table 2 Chapter letters and Names

Chapter letters	Chapter Name
TRA	Procedures and preliminary studies
CAL	Quality control
PRO	Provisionals
PRE	Preliminaries
MOV	Earthworks and soil stabilization
CIM	Containment and foundations
EST	Structure
MAM	Masonry and partition walls
CUB	Roof
MET	Metal, aluminium, and glass carpentry
MAD	Carpentry wood and other materials
PIN	Painting and waterproofing
ACA	Finishes
HID	Hydraulic and gas networks and installations
CLI	HVAC
COM	Telecommunication networks and installations
ELE	Electrical and lighting networks and installations
ESP	Special networks and installations
MEC	Mechanical conveying systems
URB	Urban planning and exterior works
CIV	Civil works

Table 3 contains the subchapters (level 2) of each one chapter described in table 2.

Table 3 Subchapters of chapter CAL

Coding to level 2	Subchapter Name
TRA	(Procedures and preliminary studies)

TRA.10	Procedures
TRA.20	Study of biological agent
TRA.30	Environmental, hydraulic, and hydrological study
TRA.40	Archaeological survey and paleontological study
TRA.50	Vulnerability study
CAL (Quality control)	
CAL.10	Testing of internal networks and installations
CAL.20	Soil testing
CAL.30	Provisionals
CAL.40	Civil works testing
PRO (Provisional works)	
PRO.10	Provisional construction
PRO.20	Provisional networks
PRO.30	Temporary signs
PRE(Preliminaries)	
PRE.10	Demolition clean-up
PRE.20	Topography
PRE.30	Canalization of water courses
PRE.40	River defence and water table
MOV(Earthworks and soil stabilization)	
MOV.10	Earthworks
MOV.20	Slope stabilization
CIM(Containment and foundations)	
CIM.10	Retaining wall
CIM.20	Shallow foundation
CIM.30	Deep foundation
EST(Structure)	
EST.10	Vertical elements
EST.20	Horizontal elements
EST.30	Reinforcement of existing building
EST.40	Seismic response control system
EST.50	Roof structure
EST.60	Reinforcement of roof
EST.70	Stair
MAM(Masonry and partition walls)	
MAM.10	Masonry wall
MAM.20	Lightweight walls
MAM.30	Concrete wall
MAM.40	Mud wall
CUB(Roof)	
CUB.10	Support and structure
CUB.20	Cover roof
CUB.30	Eaves
CUB.40	Skylight
CUB.50	Pediment
CUB.60	Complementary and water protection elements
CUB.70	Bargeboard
CUB.80	Special details
CUB.90	Green roof
MET(Metal, aluminium and glass carpentry)	
MET.10	Exterior door

MET.20	Gate
MET.30	Interior door
MET.40	Closet door
MET.50	Window
MET.60	Locksmith and railing
MAD(Carpentry wood and other materials)	
MAD.10	Exterior door
MAD.20	Gate
MAD.30	Interior door
MAD.40	Closet door
MAD.50	Window
PIN(Painting and waterproofing)	
PIN.10	Painting
PIN.20	Varnish and sealant
PIN.30	Waterproofing, sealing and water repellent.
ACA(Finishes)	
ACA.10	Ceiling
ACA.20	Façade and exterior wall
ACA.30	Interior wall
ACA.40	Floor
ACA.50	Thermal and acoustic insulation
ACA.60	Safety and security
ACA.70	Overlapping mouldings
ACA.80	Finishing stair
ACA.90	Ventilation grids
ACA.100	Fireplaces
ACA.110	Fixed furniture and accessories
ACA.120	Mobile furniture
ACA.130	Signage
HID(Hydraulic and gas networks and installations)	
HID.10	Plumbing fixture
HID.20	Water system
HID.30	Irrigation system
HID.40	Rainwater system
HID.50	Sanitary system
HID.60	Industrial water
HID.70	Water treatment system
HID.80	Fire protection system
HID.90	Gas system
CLI(HVAC)	
CLI.10	Air-conditioning system
CLI.20	Heating system
CLI.30	Ventilation system
COM(Telecommunication networks and installations)	
COM.10	Radio and TV
COM.20	Access control
COM.30	Voice and data
ELE(Electrical and lighting networks and installations)	
ELE.10	Lighting system
ELE.20	Electrical system
ELE.30	Audio & projection system

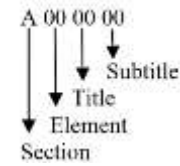
ESP(Special networks and installations)	
ELE.10	Waste extraction
ELE.20	Fuel and specialty gases
ELE.30	Scenotechnics
ELE.40	Facade cleaning system
ELE.50	Security
MEC(Mechanical conveying systems)	
MEC.10	Elevator
MEC.20	Stretcher elevator
MEC.30	Forklift
MEC.40	Car lift
MEC.50	Overhead crane
MEC.60	Escalator
MEC.70	Conveyor belt
MEC.80	Funicular
URB(Urban planning and exterior works)	
URB.10	Enclosures
URB.20	External roads and sidewalks
URB.30	Bicycle lane
URB.40	Landscaping
URB.50	Parks, squares and kiosks
URB.60	Swimming pool
URB.70	Sport and urban furniture
URB.80	Parking
CIV(Civil works)	
CIV.10	Water system
CIV.20	Sanity system
CIV.30	High and very high voltage electricity system
CIV.40	High pressure natural gas system
CIV.50	Earthwork
CIV.60	Pavement structure
CIV.70	Highways complement
CIV.80	Signalling and road safety elements
CIV.90	Green areas

As an example, up to level 3 (section), the shotcrete coding is MOV.20.40, where:

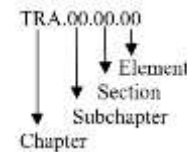
- The chapter (level 1) is MOV which corresponds to the Earthworks and soil stabilization.
- The subchapter (level 2) is .20 corresponding to the Slope stabilization
- The Section (level 3) es .40 which corresponds to the Shotcrete.

4 Analysis.

NCh 1156 of 2018 [9] was the main input to create the ColombiaClass, however, as the research progressed, much of the structure of NCh 1156 was transformed. The coding used by NCh 1156 is shown below.



The section is represented with a letter ranging from A to I, the elements, titles, and subtitles have a number that increases 1 by 1. In addition, the number 80 is used in the section, element, title, and subtitle to include everything that has not been standardized and the letter G to describe generalities. These last two representations were removed in the ColombiaClass and the coding is as follows:



NCh 1156 in its level 3 (title) has 980 items, 653 of which were modified to get the ColombiaClass, this means that more than 66% of the level was modified. Also, level 1, went from 9 sections (NCh1556) to 21 chapters (ColombiaClass), changing all the names and scopes of each item of this level, except for mechanical conveying systems.

Not only does the ColombiaClass have numerous differences from Nch 1156, but it also differs from GuBIMClass, Uniclass and Uniformat. Some examples of scope gaps that were resolved using ColombiaClass are shown below:

1. Clinical gases:
 - ColombiaClass: ESP.20.10
 - UniClass: No classification for specific gases, the nearest is EF_55_05 corresponding to Gas Extraction and Treatment [22]
 - Uniformat: No classification for gases, all types of gases are implicitly included in D2090 corresponding to Other Plumbing Systems [21]
 - GuBIMClass: No classification
2. Escalator:
 - ColombiaClass: MEC.60.10
 - UniClass: No classification for specific type of stair, the nearest is EF_35_10_30 and EF_35_10_40, corresponding to Internal Stairs and External Stairs respectively. [22]
 - Uniformat: No classification for specific type of stair, the most similar is C2010 corresponding to Stair Construction [21]
 - GuBIMClass: 60.30.10.30 [23]

3. Green Roof:

- ColombiaClass: green roof is the section CUB.100 and the elements that make it up are CUB.100.10 to CUB.100.90 where they include the growth medium, drainage layer, etc.
- UniClass: No classification
- Uniformat: No classification
- GuBIMClass: No classification

4. Measuring and protection equipment- biosafety and health:

- ColombiaClass: ACA.10.40
- UniClass: No classification
- Uniformat: No classification
- GuBIMClass: 80.40.20.10 corresponding to Health and safety measurement and detection equipment [23].

ColombiaClass was created with the goal of going beyond 3D BIM modelling, helping field managers to use it for scheduling (4D), cost estimation (5D), and beyond. For this reason, the Procedures chapter was created, which includes codes that cannot be assigned to a BIM element of the 3D model but can be assigned to cost estimation and scheduling.

Table 4 reproduces the comparison between Classification Systems published by Eastman [18], adding the ColombiaClass for comparison.

Table 4. Comparison between existing Classification Systems (adapted from Eastman [18])

Classif. system/ criteria	Scope	Principle Categori- zation	Taxonomy
OmniClass	Organization and sorting of product information for all elements in the project life cycle	Faceted	15 tables
UniFormat	Organization of physical elements of a construction. Used for cost estimation	Hierarchical	1 table with 5 levels
Uniclass	All aspects of the design and construction process.	Faceted	11 tables
Colombia Class	All elements (physical and non-physical) of design and construction process	Hierarchical	1 Table

5 Conclusions and future work

ColombiaClass is a BIM Classification System designed for building projects and in turn, serves as a standard for the definition and hierarchization of building construction items for public buildings in Colombia. This system was conceived for the public sector and is aimed at standardizing the bidding documents that have been developed in recent years, thus allowing bidding processes and works to be more transparent, efficient, and with better information management.

In the discussions generated in the expert panels, it was corroborated that Colombia needs to create a Classification System for BIM that fits the Colombian construction context.

To ensure that the ColombiaClass articulates correctly with BIM models and supports interoperability, data exchange, and cost and time analysis, future research should be done carrying out case studies using a bottom-up approach. This research should be done once the level 4 of the ColombiaClass is completed.

Additionally, as future research, the authors propose that the ColombiaClass be reviewed with ISO 12006 and adjusted if necessary. Also, the ColombiaClass should be expanded to cover the asset management phase and infrastructure projects. We invite readers to contact the corresponding authors to access the ColombiaClass and to collaborate in developing it further.

6 Acknowledgments

The authors thank Jose Luis Ponz for his continued support in this research. We also thank Fabio and Nelson Mendoza, Valentina Sarmiento, Alan Delgadillo from 7D Smart Building, Gustavo Turriago, Juanita Botero, and Andres Gonzales for being part of the Panel of Experts.

References

- [1] T. Chipman, C. Eastman, and T. Liebich, "Bridge Information Model Standardization Volume II: Schema Analysis," *BIM-Infra*, vol. II, no. April, 2016.
- [2] V. Palomino and J. Diaz, "MEJORA DEL CONTROL DE OBRA EN PROYECTOS DE EDIFICACIONES, MEDIANTE LA EXPERIENCIA EN DIRECCIÓN DE OBRA Y LA HERRAMIENTA LEAN CONSTRUCTION," Universidad Católica de Colombia, 2020.
- [3] Portafolio, "El 41% de las obras públicas del país presentan atrasos," 2019. <https://www.portafolio.co/economia/infraestructura/el-41-de-las-obras-publicas-del-pais-presentan-atrasos-524977>.

- [4] Gobierno de Colombia, CAMACOL, and B. FORUM, "Estrategia Nacional BIM 2020-2026," 2020, [Online]. Available: <https://documentos.mideplan.go.cr/share/s/MOQwz7ifQI6vwczIxnFldw>.
- [5] Secretaría de Infraestructura Física, "Análisis del sector de la construcción Modalidad de Selección – Licitación Pública," 2020.
- [6] "SKMBT_C45218061810400.pdf."
- [7] M. Sadeghi, J. W. Elliott, and M. H. Mehany, "Information-augmented exchange objects to inform facilities management BIM guidelines: introducing the level of semantics schema," *J. Facil. Manag.*, 2022, doi: 10.1108/JFM-04-2021-0044.
- [8] PlanBim, "Planbim una iniciativa a 10 años," 2020. <https://planbim.cl/>.
- [9] Instituto Nacional de Normalización, "NCh 1156 Construcción - especificaciones técnicas - Ordenación y designación de partidas." Santiago, Chile, 2018.
- [10] Ministerio de Desarrollo económico, "REGLAMENTO TÉCNICO DEL SECTOR DE AGUA POTABLE Y SANEAMIENTO BASICO República de Colombia Dirección de Agua Potable y Saneamiento Básico," *Minist. Desarro. Económico, Dir. Agua Potable y Saneam. Básico*, vol. 22 De Novi, 2000, [Online]. Available: http://cra.gov.co/apc-aa-files/37383832666265633962316339623934/7._Tratamiento_de_aguas_residuales.pdf.
- [11] ICONTEC, "NTC 1500 - Código de Colombiano de fontanería," 2003.
- [12] Ministerio de Minas y Energía, "RETIE resolución 9 0708 de agosto 30 de 2013," *Resoluc. 90708*, 2013.
- [13] Ministerio de Minas y Energía, "RETILAP resolución 18 0540 de marzo 30 de 2010," 2010.
- [14] Alcaldía Mayor de Bogotá, "Cartilla de Andenes de Bogotá D.C.," *Secr. Dist. planeación*, 2018.
- [15] E. Neufert, *EL ARTE DE PROYECTAR EN ARQUITECTURA*, 1st ed. Barcelona: Gustavo Gili, 1994.
- [16] INVIAS, "Documentos Técnicos INVIAS." <https://www.invias.gov.co/index.php/documentos-tecnicos>.
- [17] Ministerio de Salud, "Infraestructura en salud." <https://www.minsalud.gov.co/salud/PServicios/Paginas/infraestructura-en-salud.aspx>.
- [18] B. Flyvbjerg, "Five misunderstandings about case-study research," *Qual. Inq.*, vol. 12, no. 2, pp. 219–245, 2006, doi: 10.1177/1077800405284363.
- [19] ICA, "Regiones Naturales de Colombia: Dirección Técnica de Epidemiología y Vigilancia Fitosanitaria," 2018. https://www.researchgate.net/figure/Map-of-Colombia-showing-the-natural-regions-and-departments_fig1_323609009.
- [20] ICA, "Instituto Colombiano Agropecuario." <https://www.ica.gov.co/>.
- [21] CSI, "UNIFORMAT." <https://www.csiresources.org/standards/uniformat>.
- [22] NBS, "Uniclass 2015." <https://www.thenbs.com/our-tools/uniclass-2015>.
- [23] GuBIMCat, "GUBIMCLASS v.1.2 Sistema de clasificación BIM de elementos por función," 2017.
- [24] M. J. Converse and J. L. P. Tienda, "Análisis de sistemas de clasificación para entorno BIM," Universidad de los Andes, 2020.
- [25] T. J. Gordon, "The Delphi Method," *Millenn. Proj.*, 2009.
- [26] P. Galanis, "The Delphi method," *Arch. Hell. Med.*, vol. 35, no. 4, pp. 564–570, 2018, doi: 10.4324/9781315728513-10.
- [27] M. Alkasasbeh, O. Abudayyeh, and H. Liu, "A unified work breakdown structure-based framework for building asset management," *J. Facil. Manag.*, vol. ahead-of-print, Aug. 2020, doi: 10.1108/JFM-06-2020-0035.
- [28] A. Cerezo-Narváez, A. Pastor-Fernández, M. Otero-Mateo, and P. Ballesteros-Pérez, "Integration of Cost and Work Breakdown Structures in the Management of Construction Projects," *Appl. Sci.*, vol. 10, p. 1386, 2020.
- [29] A. Ekholm, "A critical analysis of international standards for construction classification - results from the development of a new Swedish construction classification system," *CIB W78 Conf.*, pp. 1–6, 2016.
- [30] J. Gelder, *The design and development of a classification system for BIM*. 2015.
- [31] K. Afsari and C. Eastman, *Categorization of building product models in BIM Content Library portals*. 2014.
- [32] R. Howard, "Classification of building information – European and IT systems," in *Proceedings of the CIB-W78 International Conference IT in Construction in Africa*, 2001, pp. 9–14.
- [33] Y. Jung and S. Kang, "Knowledge-Based Standard Progress Measurement for Integrated Cost and Schedule Performance Control," *J. Constr. Eng. Manag.*, vol. 133, pp. 10–21, Jan. 2007, doi: 10.1061/(ASCE)0733-9364(2007)133:1(10).

Towards Automated Pipelines for Processing Load Test Data on a HS Railway Bridge in Spain using a Digital Twin

C. Ramonell^a and R. Chacón^a

^aDepartment of Civil and Environmental Engineering, Universitat Politècnica de Catalunya, Spain
E-mail: carlos.ramonell@upc.edu, rolando.chacon@upc.edu

Abstract –

This document presents an automated pipeline to process sensor-based data produced during load tests on digitally twinned HS railway bridges. The research is developed within the frame of the H2020 European project ASHVIN, related to Assistants for Healthy, Safe, and Productive Virtual Construction, Design, Operation & Maintenance using Digital Twins. The pipeline is developed within a digital twin application based on event-driven microservices, which integrates the ASHVIN IoT platform, the IFC building information model and an array of services dedicated to automating processes performed during the operation stage of structural assets. A load test carried out on a bridge located on a HS railway in Spain is used as a demonstrator.

Keywords –

Digital Twin; IoT; Load Tests; BIM; Automated Pipeline; Data processing; Event-based Microservice Architecture

1 Introduction

Digital twins are cyber-physical systems able to accurately track the behavior of built assets. In addition, these information constructs are called to accurately represent all relevant processes during their design, construction, and operation stages. A seamless two-way information flow is established between the real and the virtual realms. This is enabled by developing workflows based on novel technologies that are nowadays flourishing within the AECO industry (Architecture, Engineering, Construction, and Operation, namely, Building Information Modelling (BIM), the Internet of Things (IoT), Simulation and Decision-Making services.

BIM models offer high fidelity representations of all components of a project, containing both, geometrical and semantic information, that allow to manage costs, schedules, and processes from a single source of truth,

enhancing collaboration among stakeholders [1]. Extensive research has been dedicated to BIM for the design and construction stages of built assets, however, BIM for operation and facility management is still in its early stage. The value of BIM during this stage of the asset's lifetime is highly dependent on real-time updates of its status. Presently, AEC industry digitalization is conceiving BIM-enabled cyber-physical systems [2]: physical assets are continuously monitored, fusing information from different sources and in multiple formats with virtual models into digital twin environments. Therefore, interoperability among information sources is a key challenge if comprehensive digital twins are expected as a solution. IoT provides the technology stack that enables interoperability, interconnecting models, stakeholders, physical assets, external applications, and other virtual representations, to facilitate the development of intelligent services with self-configuring capabilities [3]. However, there is a lack of studies describing the development of applications based on the digital twin concept for practical implementation in tasks related to civil and structural engineering. Thus, digital twin architectures and development methodologies need to be established.

In this paper, a case study is presented, where a digital twin application is developed to process data gathered during load tests on HS railway bridges in an automated fashion. Load tests on bridges are needed, repeatable and prone to standardize operations that represent the pivots on which the structure turns from construction to the operation stage. Therefore, automation could largely benefit stakeholders who perform this type of assessment, reducing considerably labor times and, thus, increasing productivity while reducing costs. To this end, an automated pipeline is assembled into a digital twin framework, involving users, BIM, sensors, databases, and data processing modules.

The study is developed within the framework of ASHVIN project, which is related to Assistants for Healthy, Safe, and Productive Virtual Construction, Design, Operation & Maintenance, using Digital Twins. The

project provides a series of demonstrators for each scenario, including bridges, buildings, and industrial facilities. This case study is developed on a high-speed railway bridge, located in Spain. The pipeline will perform according to National standards, and results will be stored into an interoperable environment, facilitating data access to users and external applications. This will set off the development towards the full digitalization of load tests in high-speed railway bridges, providing the initial status of the structure for a digitally twinned operation and maintenance of this type of asset.

2 Description of the case study

In this chapter, the context for the development of the pipeline is described, that is, the bridge and the load test.

2.1 The bridge

The viaduct, called La Plata, is part of the Spanish national high-speed railway network. The network is owned and managed by ADIF (Administración De Infraestructuras de Ferrocarril) and is operated by RENFE (Red Nacional de Ferrocarriles Españoles), a railway transport company that offers a high-speed service called AVE (Alta Velocidad Española) in which trains run over the network overpassing 300 km/h. The viaduct is in a rural environment between Plasencia and Cáceres on a 437 km branch connecting Madrid and the Portuguese border, which has been under construction recently.

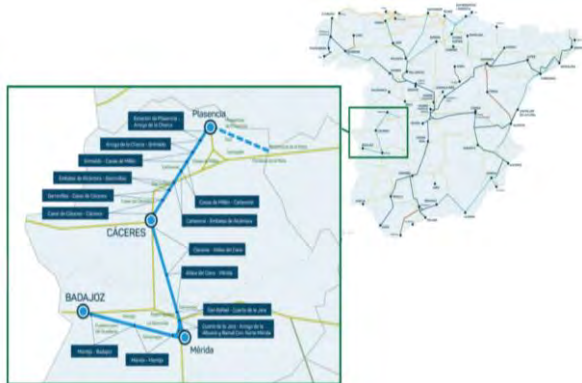


Figure 1. Representation of the high-speed railway network in Spain. The railway branch in which the bridge is located is highlighted.

The structure of the viaduct is arranged in 4 spans with a total length of 114m. The deck is 14m wide and consists of a continuous beam of prestressed concrete casted in situ that is supported by double cylindrical piles. Figure 2 shows a drone view of the Viaduct.



Figure 2. Drone view of the viaduct.

2.2 The Load Test

Load tests on the viaduct were performed by Drace (formerly Geocisa). The procedure followed the National norm provided by ADIF (NAP 2-4-2.0) regarding load tests on railway bridges under construction. According to this norm, static and dynamic actions are systematically reproduced to evaluate the structural performance of the bridge, and subsequently compared with theoretical calculations obtained from physical models.

During the static tests, the structural response is measured in a 5-step loading sequence: initial position, loading, stabilization, unloading and recovery. Dynamic tests are performed as a series of passes where the train circulates at different speeds.

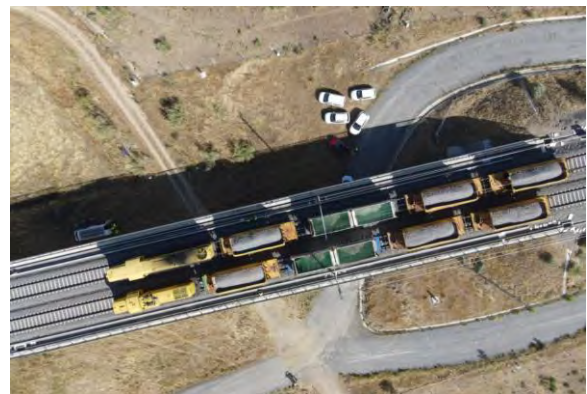


Figure 3. La Plata viaduct during the execution of a static load test.

Accelerometers, strain gauges, displacement transducers and environmental sensors are used to measure vertical displacements of beams, vertical displacements at the supports, accelerations, deformations, temperature, and humidity during the realization of the tests.

After the test is completed, data is processed to obtain the parameters that describe the structural response of the bridge:

- Maximum vertical displacement

- Relative vertical displacements
- Maximum deformation
- Displacement after recovery
- Impact coefficient
- Natural frequencies and vibration modes
- Structural damping

The structure is validated if the results are in correspondence with theoretical design of the structure.

3 Digital twin enablers

BIM tools provide semantic and contextual information while the IoT paradigm enables interconnection among devices such as sensors, actuators, and personal computers, enabling digital twin capabilities. Thus, the development of the pipeline for the digital twin of the target bridge is based on its BIM model, developed according to the IFC standard, and the ASHVIN IoT platform.

3.1 The BIM Model

Industry foundation classes (IFC) is a comprehensive and structured data schema that accurately describes the built environment [4]. Recognized as an open international standard (ISO 16739-1:2018) it is meant to be usable across a wide range of hardware devices, software platforms, and interfaces.

Up to its version IFC4, the schema mainly focused on buildings, however, new extensions are arising that allow the development of infrastructure projects, including bridges, railway, roads, and tunnels [5]. The semantics, relations and properties built in the schema are the core reference for the automatic processes developed in this study.

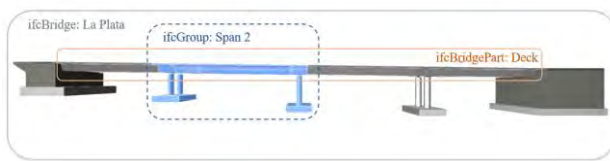


Figure 4. IFC model of La Plata viaduct where examples fort spatial structures (*IfcBridge* and *IfcBridgePart*) and semantic groups (*IfcGroup*) are highlighted.

Load tests are specific activities included within the facility management planning of the bridge. The IFC schema also allows to accurately monitor all the activities within construction projects (*IfcProject*) through a series of controls (*IfcControls*). A Work plan (*IfcWorkPlan*) is

a sub-class of *IfcControl* typically used to group activities within the same life-cycle stage of the asset, i.e., activities for construction management purposes and activities for facility management purposes. Each work plan contains a set of work schedules (*IfcWorkSchedule*), which encapsulate multiple tasks with referenced resources and actors. In this study, the load test is modeled as a work schedule within the facility management work plan. The load test includes a series of tasks (*IfcTask*) representing each individual static and dynamic measurement performed. *IfcTasks* provide a time reference which can be semantically enriched by user-defined parameters. Tasks can also be related to specific built elements of the project on which they are operating. Load test tasks are representations of single assessments that are conducted on specific bridge beams (*IfcBeam*), which have a set of sensors (*IfcSensor*) attached that monitor their behavior. Figure 5 shows a simplified schema of the load test information model for tests performed on a beam with a single sensor attached.

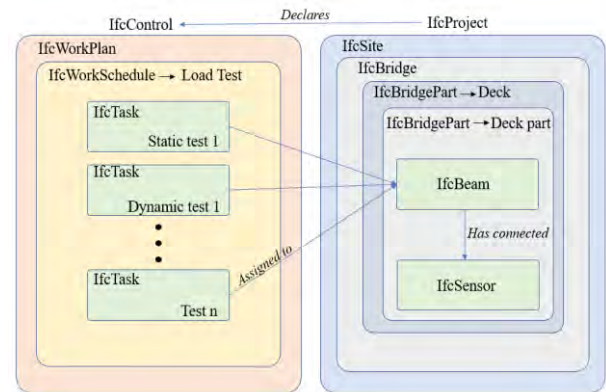


Figure 5. Relational schema of a load test for one beam with one sensor attached.

3.2 ASHVIN IoT platform

ASHVIN IoT platform is based on Mainflux, a high performant, scalable, low foot-print and open-source IoT solution which can be deployed both on the cloud and at the edge. The platform accepts connections over various protocols (HTTP, MQTT, WebSocket, CoAP and LoRaWan) enabling the two-way connection of all sorts of IoT devices.

The platform features three basic entities to perform communication between information producers and consumers: *things*, *channels*, and *users*. A *thing* represents any data source or producer. *Channels* are communication pathways through which *things* send and receive messages. Messages can be addressed to specific *topics*, providing extra semantics to the communication process, enhancing data querying and filtering. *Channels* abstract

away complexities of low-level communication protocols offering a unified and easy-to-use interface for messaging. *Users* represent physical persons and organizations which own channels and things.

Data sent over the platform can be consumed as a stream via MQTT and WebSocket or can be retrieved from a structured time-series databased via a REST API.

4 The BIM-IoT integration

The integration of BIM and IoT into engineering applications remains a challenge for the industry and is positioned as one of the key milestones to generate digital twins for facility management. In this chapter, the BIM-IoT integration is addressed. A brief review of studies regarding this connection is presented. Additionally, the ASHVIN IoT platform is introduced and the integration between the platform and IFC models is proposed.

4.1 Previous works

BIM-IoT integration relays on methods for relating virtual asset contextual data and time-series data coming from their physical counterpart. Some studies linked time-series data stored in SQL (Structured Query Language) databases with private-vendor BIM tools APIs (Application Programming Interface). Desogus et al [6] integrated low-cost humidity and temperature IoT sensors and a Revit model using Dynamo visual programming tool for monitoring the indoor conditions of buildings. Time-series data was called into Dynamo through queries to an external IoT platform. Data stored in the IoT platform was mapped to elements in the BIM matching GUIDs (Global Unique Identifier) through which queries could reach specific sensor data. Quinn et al [7] achieved integration of sensor database and Dynamo Revit using a custom .csv file and naming convention that allowed to match data with BIM model element's ID. These methods are effective; however, its development is limited by private-vendor software characteristics.

Moretti et al [8] proposed an openBIM approach to integrate IoT data and as-built information for digital twins. BIM, and IoT data were integrated in an external digital twin application. IFC data schema and IfcOpenShell were selected to store and query BIM data in JSON (JavaScript Object Notation) format, linking sensor readings to IFC objects GUIDs. The method allowed to maintain separation between dynamic and static data storage while providing a scalable approach using open-source software.

Other approaches transform BIM data into a relational database using new data schemas or create a new integrated query language using semantic web technologies [9]. These approaches are effective, however, may

end up using convoluted methods where data needs to be transformed in advance.

4.2 Integrating IFC and ASHVIN IoT platform

Integration between IFCs and Mainflux communication entities is done by directly matching their GUIDs. Additionally, things and channels metadata are added with relevant contextual information included in the IFC model.

In the bridge information model of La Plata, IfcSensors, IfcTasks and IfcWorkSchedules are mapped to the Ashvin IoT platform. IfcSensors are virtual representation of devices that provide raw physical information; thus, they are represented in the IoT platform as a *thing*. IfcTasks are processes from which a result or report is expected, therefore they can also be considered as information sources, and are also assigned to a *thing*. On the other hand, IfcWorkSchedules represent a context, in which processes are organized and to which information can be referenced. Therefore, they are mapped as a *channel* in the platform. Figure 6 shows an example of this of how information of the IFC model is referenced in the ASHVIN IoT platform.

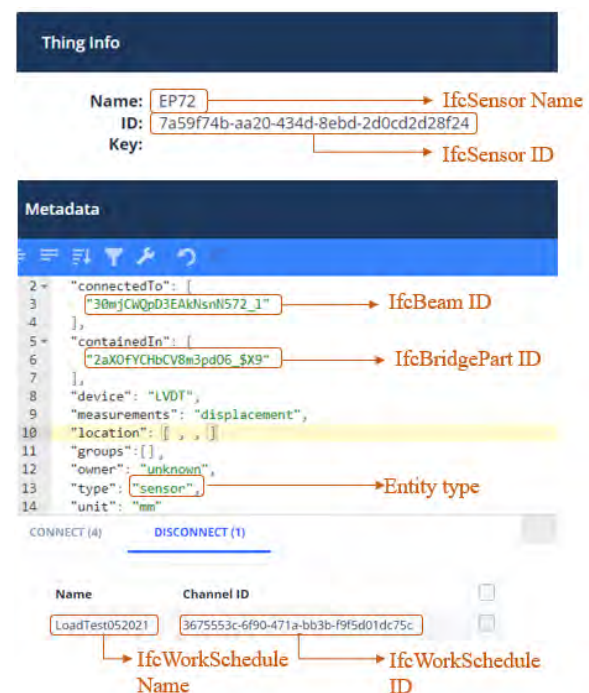


Figure 6. Screenshot of the Ashvin IoT platform interface. Mapping between a thing and an IfcSensor.

5 Contribution to the digital twin application

5.1 Architecture

The integration of IFC models and the Ashvin IoT platform enables the development of applications around asset information models that add advanced functionalities, such as data processing and multi-physics simulation, fostering the transition from BIM to digital twins [10].

In this study, an application based on event-driven microservices is developed to process the information produced during load tests on HS railway bridges. The event-driven architecture (EDA) is a software architecture in which decoupled microservices communicate by producing, detecting, and consuming events. Events are snippets of information triggered by any state change or any update that may provoke some reaction in the system, allowing to generate a chain of processes to accomplish some purpose. Event-driven architectures allow to create agile and scalable IoT applications, that can effectively consume data in real-time [11].

To process the load test data, two services are connected to the Ashvin IoT platform: a *BIM manager* service and a *load test processing* service.

The *BIM manager* parses, distributes, and updates the information contained in IFC file, acting as an integration tool. At the same time, listens to external events that request for information delivery for any type of BIM-based process. It has been developed using Python [13] and IfcOpenShell [12], an open-source library that allows manipulating files in IFC format. The *load test processing* service is a data processing service configured to retrieve load test data from the Mainflux database and return a series of parameters that define the structural performance of the bridge. Analyzing load test data can become computationally intensive due to large data packets which result from recording data at high rates for long periods of time. Therefore, the service is developed with python and Julia [14] to increase performance. Application events are transmitted using MQTT protocol through the Mainflux MQTT broker, thus integrating the information processing services into the IoT platform. Figure 7 shows a schema of the application.

This application will contribute to the development of MatchFEM, one of the applications being developed in the ASHVIN digital toolkit, aiming at integrating IoT with numerical simulations to closely calibrate digital twin models with measured behavior in the real world. The flexibility and decoupled nature of the proposed architecture allow creating simulation tools as an additional service, which will be seamlessly communicated with sensors, IFC models and data

processing modules. Load tests are very comprehensive use cases for demonstrating the tool, since they require data processing, simulation, verification, and calibration, to establish the initial conditions of digitally twinned bridges for further assessments.

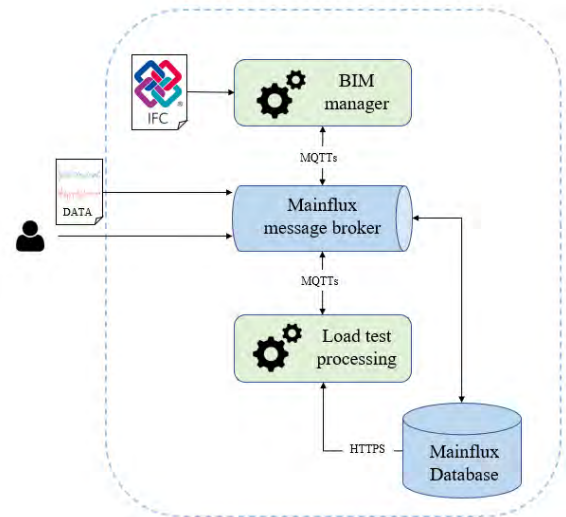


Figure 7. Schema of the digital twin application architecture.

5.2 Proposed pipeline

The asset owner assigns the execution of the load test and provides the structural Model View Definition (MVD) of the bridge, onto which the load test information model is generated. The model is uploaded to the *BIM manager* service, which checks that *IfcSensors*, *IfcTasks* and *IfcWorkSchedules* are correctly mapped in the ASHVIN IoT platform. Then, the load test is carried out, measurements for each test are taken, and data is uploaded to Mainflux database according to the bridge information model. Subsequently, the load test processing pipeline is triggered by the user.

The *BIM manager* service raises an event through the platform MQTT broker. The event message provides the following information in JSON format:

- Event type: load test processing.
- Test: *IfcTask* ID and time schedule information.
- Load test ID: *IfcWorkSchedule* ID.
- Test type: static/dynamic.
- Sensors: information including sensor ID, element to which they are connected and associated bridge span.

The *load test processing* service is subscribed to all events emitted by the *BIM manager* service, receives the

event, and checks if it can be consumed. The *load test processing* service accepts the event as one of its dedicated tasks and retrieves the referenced load test data from Mainflux database. Data is subsequently processed to obtain structural behavior parameters specified in Chapter 3: sensors are grouped according to the span to which they are assigned in the IFC model. Therefore, data is analyzed span-wise. Figure 8 shows an example of Linear Variable Differential Transformer (LVDT) data of a static tests that have produced a structural response in the first span.

Subsequently, time-series are associated to support or midspan sensors. Then, the state of the structure is retrieved at three different instants: at the start of the test, at the time of maximum displacement and after unloading, when the shape of the structure is recovered. Thus, the maximum vertical displacement, the relative vertical displacement and the displacement after recovery can be obtained. A similar process is carried out to obtain the maximum deformation from strain gauges data.

Impact coefficients are obtained for each span calculating the ratio between the maximum displacement obtained in pseudo-static and dynamic tests (see Figure 9). The dynamic excitation of the structure can be allocated in time using time-frequency decompositions [15] (Figure 10) to subsequently extract the natural frequencies, vibration modes and damping factor using current Operational Modal Analysis (OMA) techniques [16]

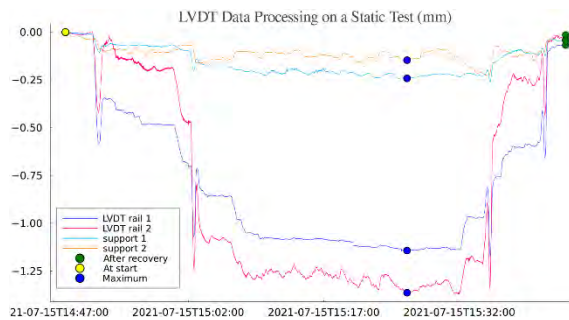


Figure 8. Data retrieved from an LVDT sensors located in the first span during a static load test.

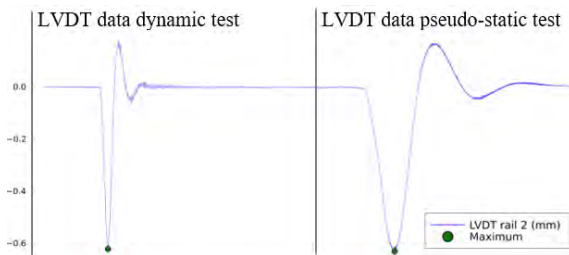


Figure 9. Comparison of LVDT data from pseudo-static and dynamic tests to retrieve impact

coefficient.

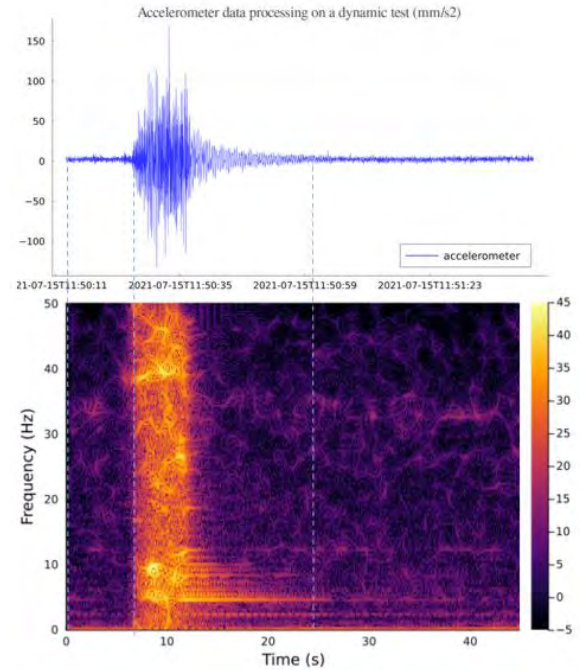


Figure 10. Accelerometer data and Wigner-Ville time-frequency transform.

The *load test processing* service publishes results to the Ashvin IoT platform using each IfcTask ID, and its corresponding key, to channels/IfcWorkSchedule_ID/Span_ID/Result_Name/subName/... Finally, it raises an event to inform the *BIM manager service* that it has accomplished its purpose. In response, the *BIM manager service* updates the status of the task to “completed” in the IFC file. Therefore, results can now be queried to the ASHVIN IoT platform by any user or third-party application using the information contained in the IFC model.

6 Conclusion

In this paper, an automated pipeline for processing data collected during load tests carried out on high-speed railway bridges is proposed. The pipeline encompasses the integration of the IoT platform of the European ASHVIN project and the IFC BIM schema. The integration is achieved within a digital twin application based on event-driven microservices which is flexible and scalable. Two services are proposed: (1) a BIM manager service that matches unique identifiers and contextual data from the IFC file into the IoT platform; (2) and a load test processing service in charge of transforming raw data into parameters that describe the structural behavior of the bridge structure. The pipeline

architecture proposed can be extended with simulation services to generate full digitally twinned load tests that would generate impact at economy and productivity levels.

7 Acknowledgements

All authors acknowledge the funding of ASHVIN, “Assistants for Healthy, Safe, and Productive Virtual Construction Design, Operation & Maintenance using a Digital Twin” an H2020 project under agreement 958161. DEBES INCLUIR A LA AGAUR

References

- [1] Gha A., et al. Building Information Modelling (BIM) uptake: Clear benefits, understanding its implementation, risks, and challenges. *Renewable and Sustainable Energy Reviews*, 75: 1046–1053, 2017.
- [2] Gao X. and Pishdad-Bozorgi P. BIM-enabled facilities operation and maintenance: A review. *Advanced engineering informatics*, 39: 227-247, 2019.
- [3] Colakovi, A. Internet of Things (IoT): A review of enabling technologies, challenges, and open research issues. *Computer Networks*, 144: 17–39, 2018.
- [4] buildingSMART International, buildingSMART, <http://www.buildingsmart.org/>, Accessed: 7/02/2022
- [5] Borrmann A. et al. The IFC-Bridge project—Extending the IFC standard to enable high-quality exchange of bridge information models. *En Proceedings of the 2019 European Conference on Computing in Construction*, pages 377-386 Chania, Greece. 2019
- [6] Desogus G. et al. Bim and iot sensors integration: A framework for consumption and indoor conditions data monitoring of existing buildings. *Sustainability*, 13(8): 4496, 2021.
- [7] Quinn C. et al. Building automation system-BIM integration using a linked data structure. *Automation in Construction*, 118: 103257, 2020.
- [8] Moretti N. et al. An openbim approach to iot integration with incomplete as-built data. *Applied Sciences*, 10(22): 8287, 2020.
- [9] Tang S. et al. A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present status and future trends. *Automation in Construction*, 101: 127-139, 2019.
- [10] Deng M. et al. From BIM to digital twins: A systematic review of the evolution of intelligent building representations in the AEC-FM industry. *Journal of Information Technology in Construction (ITcon)*, 26(5): 58-83, 2021.
- [11] Khriji S. et al. Design and implementation of a cloud-based event-driven architecture for real-time data processing in wireless sensor networks. *The Journal of Supercomputing*, 1-28, 2021.
- [12] IfcOpenshell. On-line: <http://ifcopenshell.org/>, Accessed: 22/02/2022
- [13] Van Rossum G., and Drake Jr F. L. Python reference manual. *Centrum voor Wiskunde en Informatica Amsterdam*. 1995.
- [14] Bezanson J et al. Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1): 65–98, 2017.
- [15] Scholl S. Fourier, Gabor, Morlet or Wigner: Comparison of Time-Frequency Transforms. Signal processing Cornell University
- [16] Ghalishooyan M. and Shooshtari A. Operational modal analysis techniques and their theoretical and practical aspects: A comprehensive review and introduction. *6th International Operational Modal Analysis Conference IOMAC 2015*. 2015.

Potentials of 5D BIM with Blockchain-enabled Smart Contracts for Expediting Cash Flow in Construction Projects

J. H. Yoon and P. Pishdad-Bozorgi

School of Building Construction, College of Design, Georgia Institute of Technology, Georgia, US

E-mail: jyoons37@gatech.edu, pardis.pishdad@design.gatech.edu

Abstract

The low level of digitization in sharing information and the fragmented processes in the construction supply chain negatively affects cash flow across the supply chain, resulting in non- and late-payment issues. Even though a cloud-based 5D BIM platform and Blockchain-enabled smart contracts have the potential to address the issues, their applications for expediting the cash flow are still in their exploration stages in the construction industry. The main objective of this study is to contribute to this transformation by examining one of the most advanced cloud-based 5D BIM platforms, MTWO, to analyze its potential in expediting cash flow when used in conjunction with the blockchain-enabled smart contract. This study contributes to the body of knowledge by proposing an improved way of processing pay applications that leverage blockchain-enabled smart contracts with the 5D BIM platform to expedite the construction projects' cash flow through a semi-automatic payment process.

Keywords –

5D BIM; Automation; Blockchain; Smart Contract; Construction Payment

1 Introduction

According to the McKinsey Global Institute report [1], the construction industry is one of the largest industries in the world economy, with about \$10 trillion in annual expenditures. However, the industry's productivity has trailed that of other sectors with about \$1.6 trillion in gaps of opportunity to close. This low productivity results from the low level of digitization in sharing information and the fragmented processes in the construction supply chain (CSC) involving various stakeholders who represent different interests and requirements [2]. These issues negatively affect cash flow across the CSC, resulting in non- and late-payment issues, and thus negatively affecting the productivity of

construction projects. According to the 2019 Construction Payment Report [3], the cost of slow payments is staggering and involves: (1) higher construction costs, (2) inability to get bids from the best subcontractors (Subs), (3) project delays due to stopped work, (4) discount losses of a minimum 3.7% for not paying faster.

To address the issues of low digitization and fragmented processes and to expedite the cash flow, a cloud-based 5D Building Information Modeling (BIM) platform can be leveraged. A cloud-based BIM platform enables the various stakeholders in construction projects to have higher levels of cooperation and collaboration by providing an effective real-time communication platform [4]. 5D BIM is the most recent BIM technology in which the construction project stakeholders can link a cost database to the construction activities and schedules (4th D BIM data) combined with a 3D model [5]. It enables the stakeholders to conduct model-based cost estimating, tendering, procurement, and cost control [2]. Despite these valuable capabilities of a cloud-based 5D BIM platform, its utilization for expediting the cash flow in construction projects is still rare in the construction industry.

In addition to a cloud-based 5D BIM, applications of Blockchain-enabled smart contracts can expedite the cash flow in construction projects [2, 6-11]. Blockchain is a technology that enables the stored data to be immutable and traceable. A smart contract is a digital contract using blockchain for data immutability and traceability. Using the data stored in blockchain, the smart contract makes the contract execution process automatic and enforces the fulfillment of obligations. Using these capabilities allows an automatic payment process to expedite the cash flow in construction by combining BIM with Blockchain-enabled smart contracts [2, 7, 9].

However, the integration of 5D BIM with blockchain-enabled smart contracts for expediting cash flow in construction projects is still in its early research stages. The primary purpose of this study is to investigate the

potentials of a cloud-based 5D BIM by analysing one of the most advanced cloud-based 5D BIM platforms, MTWO. In addition, this study also investigates how blockchain and smart contracts can improve and enhance the capabilities of 5D BIM to facilitate and expedite the payment process in construction projects. Finally, based on the analysis and investigations, this study proposes a system architecture of a 5D BIM integration with blockchain-enabled smart contracts for expediting cash flow in construction projects. The proposed system enables all the stakeholders including the owner, architects, GC, and Subs to have complete access to the construction progress data along with 5D BIM models. It also enables a semi-automatic interim payment process using 5D BIM data from the models and a blockchain-enabled smart contract.

2 Literature Review

2.1 Late- and Non-payment issues in Construction Projects

The late- and non-payment issues are prevalent and critical issues in the construction industry throughout the world [12-15]. Peters et al. [16] investigated the causes, effects, and solutions for late- and non-payment issues in small- and medium-sized construction companies in developing countries. The study found that the critical factors leading to the payment issues involve the client's poor financial management and delay in payment certification. Consequently, the issues result in cash flow problems, difficulties in procuring materials and equipment, problems acquiring funds from financial institutions, inability to pay wages, and damage to the reputation of contractors [16]. They found that training in cash flow management and speedy dispute resolution are effective solutions to mitigate and prevent the payment issues in construction projects. Xie et al. [17] studied the impact of payment delays on the progress of a construction project by investigating the two links in payment chains (i.e., from owner to general contractor (GC) and from GC to subcontractor (Sub)). In their study, they regulated four payment policies including 1) Shortening the payment period at both links in the payment chain, 2) Shortening only the payment period from the GC to Sub, 3) Increasing the advance fund provided to the subcontractor by the GC and 4) Increasing the percentage of interim payments to be paid to the Sub by the GC. By quantitatively simulating the above four policies, the study found that shortening the payment period at the two links will expedite the flow of funds and mitigate the pressure on contractors in providing advance funds in terms of amount and duration, making them powerful measures to ensure smooth progress in a construction project. Wu et al. [18] also pointed out that non-payment is a common complaint

from contractors in the construction industry. As a solution to the problem, they developed a framework to improve regulative measures that address payment problems. Through an analysis of late- and non-payment dispute cases, Ramachandra and Rotimi [19] found placing charging orders, caveat registration over built properties, and issuance of bankruptcy and liquidation notices have been effective methods in mitigating payment disputes. Nevertheless, the authors admitted that the payment problem is still prevalent within the industry and thus suggested that the rational starting point for real solutions to the payment problem is changing the attitude of upstream construction parties (i.e., Owner and General Contractor), followed by adherence to provisions within the payment-related legislation and contract forms.

As evidenced by the literature review, most of the studies are limited to focusing on investigating the causes of late- and non-payment issues and the negative impacts on construction projects. Even though several studies proposed methods of mitigating and preventing the issues, the approaches have been limited to a regulatory or legislative framework. Additional studies focusing on practical process improvement to facilitate and expedite cash flow is also necessary.

2.2 Potential of Blockchain-enabled Smart Contracts in Expediting Cash Flow

2.2.1 What is Blockchain and Smart Contract?

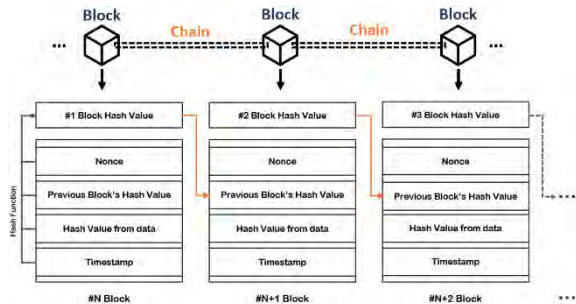


Figure 1. Linkage between the data blocks in Blockchain

Blockchain is a technology that provides an immutable and decentralized digital ledger consisting of the linked blocks with transaction data [20-22]. The data in each block are encrypted via a hash function and transformed to a hash value [21, 22]. Each block also has its own hash value made by the nonce, previous block's hash value, hash value from the transaction data, and timestamp, which creates a chain between two different blocks (Figure 1). Consequently, if someone wants to modify the data in an existing block, they should modify the data along the entire chain. They also must create a corresponding change in the next block, which includes the changed hash value of the modified block. However,

it is technically prevented from “consensus protocol” such as Proof-of-Work or Proof-of-Stake [20]. In addition, the data transaction record is replicated and distributed to every participant node, thus creating a decentralized ledger [23]. As a result, the data stored into blockchain is immutable and cannot be falsified or tampered with. In this environment, the users are able to trust the immutability and traceability of the data.

Given the immutability and traceability of data in Blockchain, blockchain enables the users to trust the data stored in Blockchain and use it as a basis to execute automated smart contracts [24, 25]. Smart contract is a computer protocol intended to digitally facilitate, verify, and enforce the negotiation or performance of a contract [26]. Smart contracts provide the contract terms and conditions with an automatic process, facilitating the fulfilment of obligations without human intervention [26]. The immutability and traceability of Blockchain, and smart contract system have the potentials to promote trustworthiness in information shared in the project and expedite the payment process and the cash flow in construction projects.

2.2.2 Blockchain-enabled Smart Contract for Expediting Cash Flow

The potential of blockchain-enabled smart contract in improving the payment issues in construction projects has been examined by a couple of studies [6, 10]. Nanayakkara et al. [6] conducted discussions with construction professionals and validated that blockchain-powered smart contract solutions can significantly mitigate the payment and related financial issues in the construction industry, including partial payments, non-payments, cost of finance, long payment cycle, retention, and security of payments. Hamledari and Fischer [10] validated why blockchain-based and decentralized smart contracts can provide reliable automation of progress payments by conducting a case study.

Leveraging the advantages of blockchain and smart contracts, other studies have designed a new framework or developed a new system for timely and transparent payment of construction projects. Luo et al. [8] proposed a framework for a semi-automatic construction payment system using smart contracts and blockchain technologies. The framework includes the decentralized environment-based sequential approval process by stakeholders such as an engineer, architect, and owner in processing construction interim payments, which can streamline the cash flow in construction projects. Ahmadisheykhsarmast and Sonmez [27] also proposed a smart contract payment security system (SMTSEC) for expediting the payment process. By using a smart contract in processing the interim payment in construction projects, the system provides a secure, efficient, and trustworthy platform for security of

payments of construction contracts, without requiring a trusted intermediary such as lawyers or banks. Both studies [8, 27] suggested that future research can focus on integration of BIM technology with smart contracts to enable completely automated payments based on the construction progress. Sigalov et al. [2] and Ye et al. [9] presented the concept of implementing smart contracts for automated, transparent, and traceable payment processing for construction projects. They combined BIM technology with blockchain-enabled smart contracts to enable the automatic payment processes in construction projects. Hamledari and Fischer [7] also proposed an autonomous payment administration solution utilizing blockchain-enabled smart contracts and robotic reality capture technologies. In the solution, the construction progress is captured, analyzed, and documented respectively using sensing, machine intelligence, and as-built BIM models. Based on the captured progress, the payment is automatically processed following the smart contract terms and conditions.

While these studies highlight the overall value inherent in using a blockchain-enabled smart contract systems combined with BIM technology for expediting cash flow in construction projects, they have not fully addressed the changes need to be made to the workflow processes to realize this value. First, all the stakeholders including the owner, architects, GC, and Subs should have complete access to the construction progress data along with 5D BIM models. Second, the legal payment documents developed based on the BIM data should be confirmed by each responsible stakeholder. Third, the automatic payment should be based on the construction progress data from the BIM technology linked with the payment documents. Accordingly, a new system architecture and workflow needs to be developed to facilitate a semi-automatic payment process, by enabling all the users to access the construction performance data, confirm the responsibility fulfilled, and provide access to official payment documents.

3 Methodology: Case Analysis of Payment Process Workflow and 5D BIM Technology

This new system architecture is developed by case analysis of current workflow and by leveraging the advantages of the cloud-based 5D BIM technology in the payment process of construction projects.

3.1 Current Cash Flow using AIA Contract Documents and Prompt Payment Codes

Current cash flow in construction projects has a linear flow from owner to suppliers. In the projects using American Institute of Architects (AIA) contract

documents, the suppliers, subcontractors, a general contractor transfers the information related to their payment by using documents including an AIA payment application (G-702) and an AIA continuation sheet (G-703). Based on the transferred documents, the architect reviews the relevant information, and the owner approves the payment. The documents can expedite payment and reduce the possibility of error [28]. Especially on the AIA payment application, the contractor can show the status of the contract sum for the portion that is completed, including the total dollar amount of the work completed to date, the amount of retainage (if any), the total of previous payments, a summary of change orders, and the amount of current payment requested [28]. Along with the payment application, the AIA continuation sheet divides the contract sum into portions of the work in accordance with the predefined construction schedule [29].

In addition to the payment documents, there exist the Prompt Payment Code (PPC) for construction projects in each state in the United States. In this research, the code for Georgia state (GA) is used. The GA code § 13-11-4 (a) describes the time limit for payment to contractors is 15 days, and § 13-11-4 (b) describes the time limit for payment to subcontractors is 10 days.

The payment process workflow using AIA contract documents and abiding by the GA PPCs is illustrated in Figure 2. This current cash flow heavily relies on paper documents and has fragmented stages across the CSC. In this setting, it is difficult for the architect and owner to check the construction performance in detail, thereby delaying the process. In addition, all the stakeholders do not have a single source of truth regarding the construction data, thus leading to misunderstandings or miscommunication regarding the construction schedule and cost estimates. Accordingly, even though prompt payment codes exist, they cannot eliminate the non- and late-payment issues.

3.2 Cloud-based 5D BIM Technology

3.2.1 5D BIM Platform Example: MTWO

5D BIM platform enables all the construction stakeholders to access construction progress data, thus minimizing misunderstandings or miscommunication regarding the construction schedule and cost estimates. The MTWO, which is investigated in this study, is one of the most advanced cloud-based 5D BIM software and provides the ability to use scheduling analysis tools. It provides several modules for scheduling construction works and estimating costs, which can be leveraged to facilitate and expedite cash flow in construction projects. The scheduling modules provide effective tools for the planning of activities such as the Gantt Chart and the Line of Balance. The Gantt Chart shows each construction job or activity with multiple version combinations: planned, current, and the combination of planned and current versions. In addition, it provides relationships, constraints, and events for each construction activity, enabling a critical path analysis for managing the variability of the plan. The Gantt Chart analysis provided by MTWO is illustrated in Figure 3. The Line of Balance shows the activities in different locations along the project time. It facilitates the rescheduling of the planned activities and the control of the necessary resources by allowing the users to detect the collisions of activities and time. The Line of Balance analysis provided by MTWO is illustrated in Figure 4. Both analyses are created based on the 3D BIM model of the project.

In addition to the schedule analysis tools, the scheduling module includes the construction performance measurements indicating the degree of completion for each activity in percentages or quantities in a tabular view. Through this capability, the users can conduct forecasting and plan-actual comparisons. The cost estimation modules enable the users to generate the cost estimate including quantity determination from the 3D component. The users can filter or organize the cost estimate under their own cost management structures and according to each field like quantity-wise, cost-wise, or

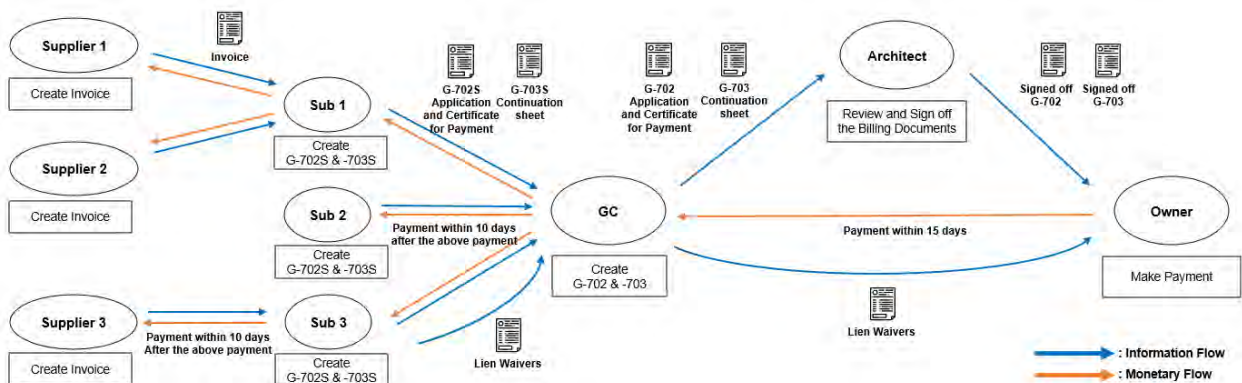


Figure 2. Payment Process using AIA Contract Documents and GA Prompt Payment Codes (PPCs)

assigned parameters-wise. When these cost estimate data are combined with the scheduling data from the scheduling modules, the 5D BIM platform enables the users to conduct a 5D simulation showing the construction progress simulation in a 3D model based on the planned and actual schedule in the Gantt Chart and the cost and budget simulation (Figure 5).



Figure 3. MTWO Gantt Chart Analysis [30]

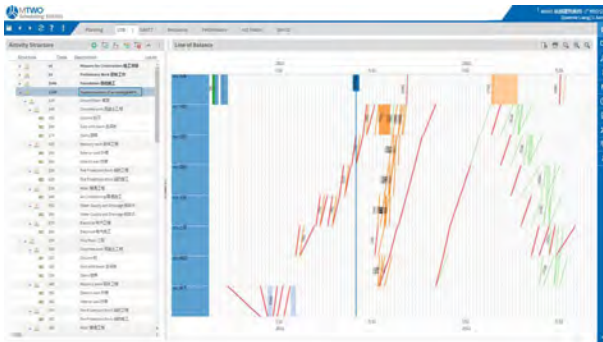


Figure 4. MTWO Line of Balance Analysis [30]



Figure 5. MTWO 5D BIM Simulation [30]

3.2.2 Potential and Limitation of 5D BIM

By leveraging the scheduling and cost estimate modules, the GC can process the raw data from the architects into the information for managing construction schedules and progress payment. First, the raw data including drawings and 3D BIM models from the architect are transferred to the GC. Next, the GC input the raw data into the 5D BIM software. From the estimate module, the GC can have the information including quantity take-off, bill of quantities, and cost estimates. From the scheduling module from 5D BIM software, the

GC can have the information including the planned and current schedules and the overall construction progress and progress lines. The information associated with the cost estimates and scheduling can be used by the subcontractors to develop the invoice and the documents for the progress payment as well as by the GC to develop the progress payment application. In addition, based on the digital documents, the architect can confirm the application and create certificate for the payment.

In summary, the 5D BIM modules facilitate developing the progress payment documents with digital data based on the 3D Model, which can expedite the cash flow by reducing paperwork overload. In addition, the stakeholders including the owner, architect, GC, Subs, and even suppliers can have access to the single source of truth in 5D Digital Twin, representing both the initial plan and in-progress schedule and cost estimate performance. These advantages are the main strengths of the 5D BIM software.

However, 5D BIM technology also has its own limitations. The GC, Subs, and architect need to process the information provided by the 5D BIM software to develop the progress payment invoice, payment application, and certificate for payment, which means the administrative procedures to develop the documents of the payment application and certificate remain. Even though the GC and architect develops the documents of application and certificate for payment by using the 5D BIM data provided by the 5D BIM platform, the owner cannot see the connection between the executed performance data along with the 3D model and the documents of payment application and certificate, which means the knowledge inequality among the stakeholders has not been eliminated. Another challenge is that 5D BIM cannot enforce the payment due. Accordingly, just-in-time payment is not ensured even if the project is using the 5D BIM technology. These limitations can be a new opportunity for expediting cash flow in construction projects. When the digital data related to construction schedule and cost estimates can be used as a resource for automatic payment using smart contract system, just-in-time payment can be ensured.

As introduced in the previous section, blockchain is a technology that provides data traceability and immutability [20, 21], which enables the users to trust and utilize digital data in business life [31, 32]. A smart contract is a digital contract using blockchain to automatically facilitate, execute, and enforce a contract [25, 26, 33]. If the 5D BIM data associated with the cost estimates and schedules can be verified and stored into blockchain at certain key milestones for data traceability and immutability, it can then be leveraged as a single source of truth for an automatic payment system using smart contracts. This is an opportunity for the 5D BIM technology to be used for managing the interim (progress)

payment and expediting cash flow in construction projects.

4 Expediting Cash Flow: Utilizing 5D BIM in conjunction with Blockchain-enabled Smart Contracts

4.1 System Architecture Framework

Upon analysis of the current workflow in the payment process and realizing the potential and limitations of 5D BIM platform in facilitating the creation and approval of pay applications, this study proposes a futuristic system architecture framework that utilizes 5D BIM and blockchain-enabled smart contracts.

The proposed system facilitates and expedites the cash flow in construction projects. First, the system enables all the stakeholders to access the construction progress data and 5D BIM models in real-time using a cloud-based web application. This advantage expedites the payment process by removing misunderstanding and miscommunication issues in confirming the construction progress. Second, the system automatically develops payment documents based on the cost estimates and scheduling modules of the 5D BIM platform. This advantage removes manually and linearly processing payment documents. In addition, the digital payment documents will be linked with the 5D BIM model, thus facilitating the performance examination.

After the examination, each stakeholder creates a digital signature on its responsible document, and the signed documents are uploaded to the web-based decentralized application (Dapp) using blockchain-enabled smart contracts, making the documents

immutable and traceable. In this stage, the owner buys the cryptocurrency to ensure the operation of the smart contract. Next, the smart contract in the system examines whether all the documents are signed and whether or not the predefined stage of work is completed. To examine the construction performance, the system leverages the progress data extracted from the 5D BIM software, which uses an XLSX format. If all the conditions in the smart contract are met, then the payment is processed automatically. This contract execution based on digital data removes the financial institutions like a bank in processing the payment, thus expediting the payment process. The smart contract conditions, XLSX data used in the smart contract, and the signed payment documents are stored in the blockchain to make them immutable and traceable. In this process, the size of the XLSX data can be reduced by encoding it with a hash function (e.g., SHA-256) and changing it into the hash value. The entire process is illustrated in the system architecture, Figure 6.

4.2 Limitations and Future Studies

This study is focused on developing a system architecture for a framework on smart contract-enabled 5DBIM for expediting cash flow. Even though the development is based on the research gaps identified in the literature review and the analysis of the current workflow and the potential of 5D BIM and blockchain-enabled smart contracts, empirical validation of the system is also needed. Accordingly, future studies will develop a pilot system based on the system architecture and validate the effectiveness by conducting a case study using the system in real-world construction projects.

Future studies should also consider the operating costs of the smart contract. When the smart contract is

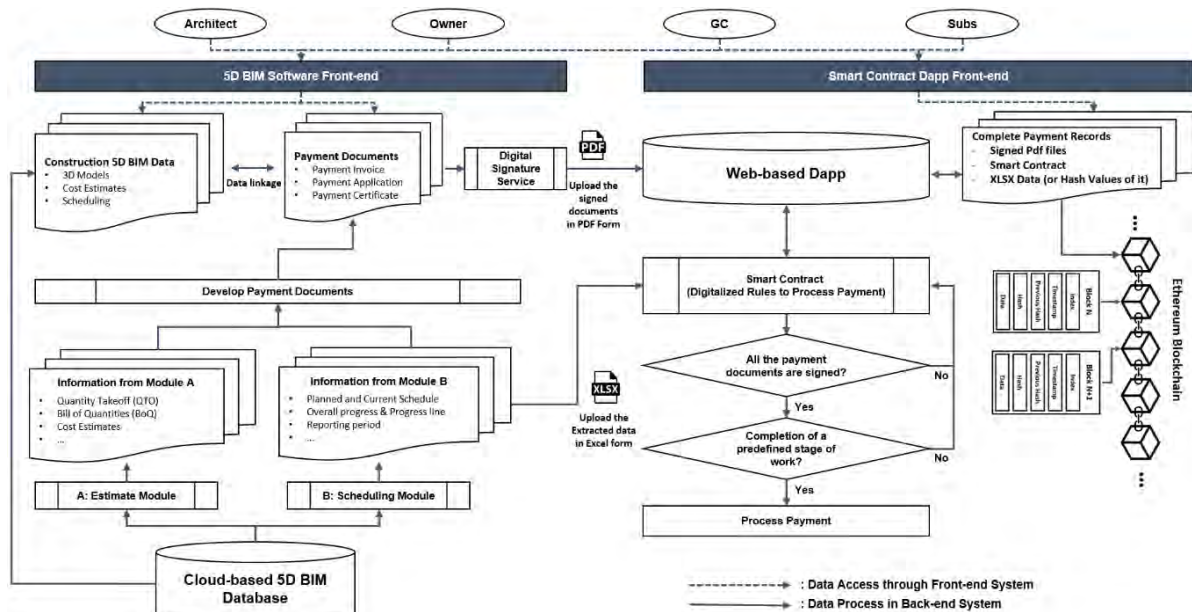


Figure 6. System Architecture of Smart Contract-enabled 5DBIM for Expediting Cash Flow

operated, it uses Gas, which is computing power to create blocks for a smart contract in Ethereum. The used Gas should be calculated because the operating costs can be estimated based on the used Gas. In addition, the pilot system should address the price volatility issue of cryptocurrency. For example, the price of Ether, the cryptocurrency used in the proposed system, varies based on the time in which it is used. To address this issue, future studies can refer to the solutions provided by Hamledari and Fischer [34]. The solutions include employing stable cryptocurrencies introduced by [35, 36], cryptocurrencies pegged against fiat currencies/commodities, asset-backed crypto tokens defined on public or private chains, etc [34].

5 Conclusion

This study contributes to the body of knowledge by proposing an improved workflow for digitalizing and automating the payment process using emerging technologies. The proposed workflow enables the construction stakeholders to access a cloud-based 5D BIM platform to review real-time construction progress data. By having access to both as-planned, as-built 5D BIM progress data, and by linking pay application documents to the supporting 5D BIM data, the stakeholders could visually and easily understand and verify the scope of the work completed, and compare the invoices with the initial estimate, and total contract sum charged so far. The payment documents confirmed by each responsible stakeholder are automatically processed with a blockchain-enabled smart contract, thus expediting the payment process by reducing the paper works and eliminating the intervention of financial institutions. As a result, the proposed workflow enables the integration of construction professionals' insight on construction progress with the automatic payment process using blockchain and smart contract. It also facilitates the industry's adoption of 5D BIM and blockchain technologies by providing a practical and specific application framework in solving a critical issue, late- and non-payment in the CSC.

Acknowledgement

This study is a part of research project, "A Transformative and Integrated Approach to Expediting Cash Flow Across the Construction Supply Chain Using Emerging Technologies", funded by RIB Americas. The authors would like to express their gratitude to RIB Americas for their support and funding of the project.

References

- [1] Filipe Barbosa, J.W., Jan Mischke, Maria João Ribeirinho, Mukund Sridhar, Matthew Parsons, Nick Bertram, Stephanie Brown *Reinventing construction: A route to higher productivity*. McKinsey Global Institute (MGI), 2017.
- [2] Sigalov, K., et al., *Automated payment and contract Management in the Construction Industry by integrating building information modeling and Blockchain-based smart contracts*. Applied Sciences, 2021. **11**(16): p. 7653.
- [3] Rabbet, *Construction Payment Report*. 2019.
- [4] Wong, J., et al., *A review of cloud-based BIM technology in the construction sector*. Journal of information technology in construction, 2014. **19**: p. 281-291.
- [5] Sattineni, A. and J.A. Macdonald. *5D-BIM: A case study of an implementation strategy in the construction industry*. in *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*. 2014. IAARC Publications.
- [6] Nanayakkara, S., et al. *Blockchain and smart contracts: A solution for payment issues in construction supply chains*. in *Informatics*. 2021. Multidisciplinary Digital Publishing Institute.
- [7] Hamledari, H. and M. Fischer, *Construction payment automation using blockchain-enabled smart contracts and robotic reality capture technologies*. Automation in Construction, 2021. **132**: p. 103926.
- [8] Luo, H., et al. *Construction payment automation through smart contract-based blockchain framework*. in *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*. 2019. IAARC Publications.
- [9] Ye, X., K. Sigalov, and M. König. *Integrating BIM-and cost-included information container with Blockchain for construction automated payment using billing model and smart contracts*. in *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*. 2020. IAARC Publications.
- [10] Hamledari, H. and M. Fischer, *Role of blockchain-enabled smart contracts in automating construction progress payments*. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 2021. **13**(1): p. 04520038.
- [11] Das, M., H. Luo, and J.C. Cheng, *Securing interim payments in construction projects through a blockchain-based framework*. Automation in construction, 2020. **118**: p. 103284.
- [12] Ramachandra, T. and J.O. BamideleRotimi, *Causes of payment problems in the New*

- Zealand construction industry*. Construction Economics and Building, 2015. **15**(1): p. 43-55.
- [13] Haron, R.C. and A.L. Arazmi, *Late payment issues of subcontractors in Malaysian construction industry*. Planning Malaysia, 2020. **18**.
- [14] Wu, J., M. Kumaraswamy, and G. Soo, *Payment problems and regulatory responses in the construction industry: Mainland China perspective*. Journal of Professional Issues in Engineering Education and Practice, 2008. **134**(4): p. 399-407.
- [15] Enshassi, A. and L. Abuhamra, *Delayed payment problems in public construction projects: Subcontractors' perspectives*, in *ICCREM 2015*. 2015. p. 567-575.
- [16] Peters, E., K. Subar, and H. Martin, *Late payment and nonpayment within the construction industry: Causes, effects, and solutions*. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 2019. **11**(3): p. 04519013.
- [17] Xie, H., et al., *Effects of payment delays at two links in payment chains on the progress of construction projects: system dynamic modeling and simulation*. Sustainability, 2019. **11**(15): p. 4115.
- [18] Wu, J., M.M. Kumaraswamy, and G. Soo, *Regulative measures addressing payment problems in the construction industry: A calculative understanding of their potential outcomes based on gametric models*. Journal of Construction Engineering and Management, 2011. **137**(8): p. 566-573.
- [19] Ramachandra, T. and J.O.B. Rotimi, *Mitigating payment problems in the construction industry through analysis of construction payment disputes*. Journal of legal affairs and dispute resolution in engineering and construction, 2015. **7**(1): p. A4514005.
- [20] Natoli, C., et al., *Deconstructing blockchains: A comprehensive survey on consensus, membership and structure*. arXiv preprint arXiv:1908.08316, 2019.
- [21] Nofer, M., et al., *Blockchain*. Business & Information Systems Engineering, 2017. **59**(3): p. 183-187.
- [22] Zheng, Z., et al. *An overview of blockchain technology: Architecture, consensus, and future trends*. in *2017 IEEE international congress on big data (BigData congress)*. 2017. IEEE.
- [23] Lisk, *What is Blockchain?* 2019.
- [24] Yoon, J.H. and P. Pishdad-Bozorgi, *State-of-the-Art Review of Blockchain-Enabled Construction Supply Chain*. Journal of Construction Engineering and Management, 2022. **148**(2): p. 03121008.
- [25] Zheng, Z., et al., *An overview on smart contracts: Challenges, advances and platforms*. Future Generation Computer Systems, 2020. **105**: p. 475-491.
- [26] Rosic, A. *Smart Contracts: The Blockchain Technology That Will Replace Lawyers*. 2020 November 25th, 2020; Available from: <https://blockgeeks.com/guides/smart-contracts/>.
- [27] Ahmadiheykhsarmast, S. and R. Sonmez, *A smart contract system for security of payment of construction contracts*. Automation in construction, 2020. **120**: p. 103401.
- [28] AIA. *G702-1992 Application and Certificate for Payment*. 1992 [cited 2022 February 6th]; Available from: <https://www.aiacontracts.org/contract-documents/19661-application-and-certificate-for-payment>.
- [29] AIA. *G703-1992 Continuation Sheet*. 1992 [cited 2022 February 6th]; Available from: <https://www.aiacontracts.org/contract-documents/20631-continuation-sheet>.
- [30] RIB *MTWO Modules Executive Overview*. 2021.
- [31] Samaniego, M. and R. Deters. *Blockchain as a Service for IoT*. in *2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*. 2016. IEEE.
- [32] Abeyratne, S.A. and R.P. Monfared, *Blockchain ready manufacturing supply chain using distributed ledger*. International Journal of Research in Engineering and Technology, 2016. **5**(9): p. 1-10.
- [33] Alharby, M. and A. Van Moorsel, *Blockchain-based smart contracts: A systematic mapping study*. arXiv preprint arXiv:1710.06372, 2017.
- [34] Hamledari, H. and M. Fischer, *The application of blockchain-based crypto assets for integrating the physical and financial supply chains in the construction & engineering industry*. Automation in construction, 2021. **127**: p. 103711.
- [35] Bullmann, D., J. Klemm, and A. Pinna, *In search for stability in crypto-assets: are stablecoins the solution?* ECB Occasional Paper, 2019(230).
- [36] Calcaterra, C., W.A. Kaal, and V. Rao, *Stable cryptocurrencies: First order principles*. Stan. J. Blockchain L. & Pol'y, 2020. **3**: p. 62.

Integrating Digital Twins in Construction Education Through Hands-on Experiential Learning

S. Kangisser^a, J. Irizarry^b, K. Watt^c, R. Borger^d, and A. Burger^e

^{a,b} School of Building Construction, Georgia Institute of Technology, USA

^c Visual Plan Inc. & Digital Twin Instructor, Mohawk College, Canada

^d School of Building Construction Sciences, Mohawk College, Canada

^e Consultant Digital Plant Technology, USA

Email: Stevek@gatech.edu, Javier.irizarry@gatech.edu, Kelly@visualplan.net, Richard.borger@mohawkcollege.ca, Burger.amadeus@gmail.com

Abstract

This paper examines the value of including digital twin technology as a hands-on learning activity in a graduate-level building construction course. The methodology of teaching digital twins as a unit is presented, and the benefits of introducing this topic are examined within the framework of several learning objectives. The campus football stadium provided an opportunity for students to apply the digital twin skill to a real-life case. Feedback from the students was collected and is presented within the context of the intended goals.

Evaluation of the use of digital twins in this course found that the technique was highly valuable in providing a framework for students to understand the potential offered by various technologies in visualizing a facility throughout its lifecycle. Students more easily understood how each technology best fits and how several technologies could be used in concert with one another.

Keywords- Digital Twins; BIM; Asset Management; Facility Management; Preventative Maintenance; Digital Transformation

1 Background

This examination of the benefits that digital twin learning can offer in construction and facilities management education took place within a graduate level course, BC6005, which is offered each semester at The Georgia Institute of Technology. This course is intended to explore the Construction 4.0 framework and the various present and future technologies that contribute to this framework and can be applied in all stages of a facilities' life cycle. Construction 4.0 can be defined as the "organization of production processes based on

technology and devices autonomously communicating with each other along the value chain" [1].

Construction 4.0 relies heavily upon digital technologies and cyber-physical systems. Digital technologies include Building Information Modelling (BIM), Common Data Environment (CDE), unmanned aerial systems, cloud-based project management, Augmented Reality/Virtual Reality (AR/VR), artificial intelligence, cybersecurity, big data, and analytics, blockchain, and laser scanner.

Cyber-physical systems include robotics and automation, sensors, the Internet of Things, workers with wearable sensors, actuators, additive manufacturing, off site prefabricated construction, on-site construction, and equipment with sensors [2]. The course referred to in this report examines software and hardware tools and technologies such as virtual and augmented reality, laser scanning, drones, additive manufacturing, robotics, IoT, and others. The course introduces students to technologies they will use in various classes within the Building Construction curriculum at the Georgia Institute of Technology and industry. During the course, students work on hands-on tasks that provide the opportunity to develop practical skills with many of the technologies covered in the course.

The goal of the course is to provide a broad background and general knowledge through the following objectives, which are to be achieved by all students for successful completion of the course:

- Be able to describe the Construction 4.0 Framework and its importance for the construction industry.
- Understand and explain the various technologies introduced in the course and identify their use cases, including the concept, value, and application of creating digital twins and what technologies can be introduced and deployed on an ongoing basis, based on the use case
- Be able to identify the correct technology for

application to common construction issues.

- Functionally apply some of the technologies introduced in the class to practical situations found in construction environments.
- Work with stakeholders to identify specific needs in a real-life on-campus project.
- Confidently present the results of the application of a technology in solving a project need.

Digital twins are a digital replica of an actual physical asset. They integrate artificial intelligence, machine learning, and data analytics to produce simulation models that can be easily updated from multiple sources. In this manner, a digital twin represents the current condition of the asset [3]. The digital twin may be a physical instrument, social construct, biological system, or composite system. Building construction projects fall within this last category as a composite system, with both physical and social products [4]. Digital twins are comprised of a physical artifact, a digital counterpart, and that which connects these two [5]. This connection is made possible by the development of advanced sensing computer vision, the internet of things, and advanced analytics [6].

This paper explores the use of instruction of digital twin technology to enable students to better understand construction 4.0 and the technologies which enable it. This relates to both facility construction and facilities management. Using a digital twin to monitor a physical asset and analyze real-time parameters makes it possible to improve its operational efficiency. The Digital twin of a building can be used to improve its operation and maintenance efficiency by permitting facility managers to perform what-if analysis. This can be useful to enhance energy efficiency and improve users' comfort [7]. They can also be employed for such uses as life cycle and security planning. Such systems can enable facility operations and maintenance to improve operational efficiency from remote management and pre-planning asset maintenance and repair, such as illustrated in Figure 1. In this case an mechanical, electrical, and plumbing (MEP) room is shown with annotated physical assets connect to computerized maintenance management system (CMMS) and Internet of Things (IoT) sensors.



Figure 1. MEP room with annotated physical assets

Digital twin systems transform business by accelerating holistic understanding, optimal decision making, and effective action. They use real-time and historical data to represent past and predicted future states. Digital twins are motivated by outcomes, tailored to use cases, and powered by integration. They are built on data guided by domain knowledge and are implemented in IT/IoT systems [8].

A valuable collection of wide-angle visual data can be accomplished using Virtual Reality (VR) and Augmented Reality (AR) technology and 360 imaging. VR/AR is generally limited to an asset's design phase and special purpose build stages. This is attributable to the incapability of these technologies to feasibly document on-site progress in real-time [9]. For this reason, 360 imaging is a particularly valuable tool for developing digital twin models. This 360 imaging technology employs basic 360 cameras available at the consumer level. They permit rapid, low-cost visual data collection for current site conditions. Therefore, knowledge of 360 imaging techniques is critical to the understanding of digital twins for this industry sector.

2 Instructional Methodology

The digital twin assignment incorporated in the course included three parts. The first part included a series of three lectures delivered by guest speakers possessing extensive expertise in the digital twin domain. This part of the instructional assignment was conducted as an intensive one-week-long series of lectures and fieldwork. The lectures were conducted during regular class time, which is twice a week for 75 minutes each. The second part was performed by all students, working in groups of 2-3 students each. Here, the students focused only on 360 imaging as applied to digital twins. This promoted an understanding of what a digital twin is, and how they are created and employed on a job site. Upon completing this assignment, students undertook the third part which was a term project that included several technologies brought together by the digital twin framework.

2.1 Lectures

The initial instructional method included a series of 3 lectures. The purpose of the first lecture presented by guest Speakers Kelly Watt, CEO Visual Plan Inc., and Mark Schreiber, President of Safeguard Consulting, was to help students consider security design and safety implications through the building process. Most students had little to no experience with security considerations or what standards to consult. Design Twins were used as a tool to help engage subject matter experts early in the planning, design, and decision phases and throughout a project to provide oversight and ensure best outcomes. The example illustrated in Figure 2 illustrates the use of

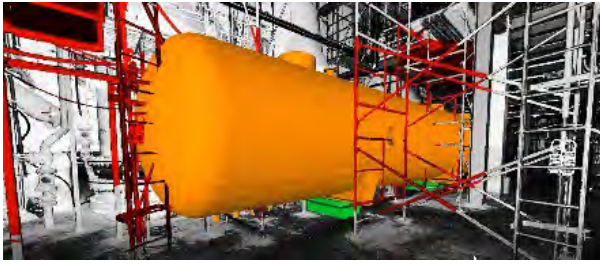


Figure 2. Point cloud and model of US nuclear plant

digital twins in design, planning and simulation to engineer a complex installation where room to navigate was less than one inch. Security was used as an example of an often-neglected process where early collaboration can drive value. Students were taught to think more broadly about facility lifecycle and all aspects and stakeholders pertaining to the site or facility rather than what is narrowly discussed in a project.

Key case examples from brownfield construction to existing facility improvement projects were presented. Comparing existing 2D Computer Aided Design (CAD) and Building Information Modeling (BIM) files to 360-degree panoramas through various project stages provided an effective means to communicate project information, and to identify issues, clashes, and change orders early. Case examples were used to illustrate the benefits of virtual collaboration, pre-planning, and involving subject matter experts early in projects and show where companies are increasing productivity, reducing re-work, and improving their bottom line.

The second lecture in the series was presented by guest speakers Kelly Watt and Amadeus Burger of Digital Plant Technology. Here students learned how digital twins are used post-commissioning for operational processes maintaining assets and facilities. Many of these processes are used together in a digital twin. Amadeus Burger spoke on specific projects creating and working with complex digital twins for nuclear energy and petrochemical plant maintenance. Students gained a clear understanding of the precision required for large equipment installs where lidar was instrumental. These applications drive clear differences between photogrammetry and lidar so students can better understand when to apply one technology over another.

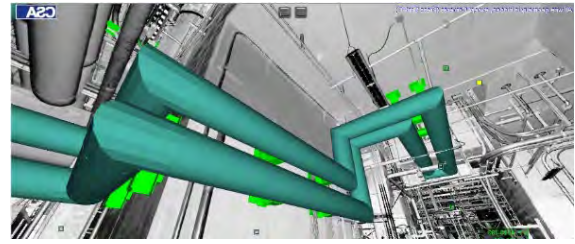


Figure 3. Point cloud/ piping model of a nuclear plant

Kelly Watt discussed methods that tie digital twins to computerized maintenance management systems, asset management, and real-time sensor and IoT devices. These integrations facilitate the ability of the digital twin asset to help produce precise operational efficiency. He presented case studies in nuclear energy and oil and gas production, such as is shown in Figure 3 where a digital model is used to facilitate quick identification of systems through a large and complex plant without the need to visit the site location. These studies illustrated how digital twins had been employed to reduce facility downtime and improve project turnover rates, thereby dramatically impacting project success and profitability. These examples taught students the strength and scenarios of selecting the appropriate technology based on the use case. The expected outcomes showed how digital twins gained unintended value when implemented.

Issues and scenarios impacting organizations from knowledge management, digital knowledge transfer, and employee turnover were presented as well as the challenges of managing operational change. This is where photogrammetry was presented as a game-changer in the speed of data collection, ease of use, and scalability. The benefits of digital representations and breaking data silos were demonstrated through case studies to offer significant gains for entire organizations.

A third lecture, presented by Kelly Watt, took place in small student teams conducted on-site at Bobby Dodd Stadium. Bobby Dodd is the football stadium and is located on the Georgia Institute of Technology campus. This session was given while seated in the stadium stands as shown in Figure 4. This was staged to give students a

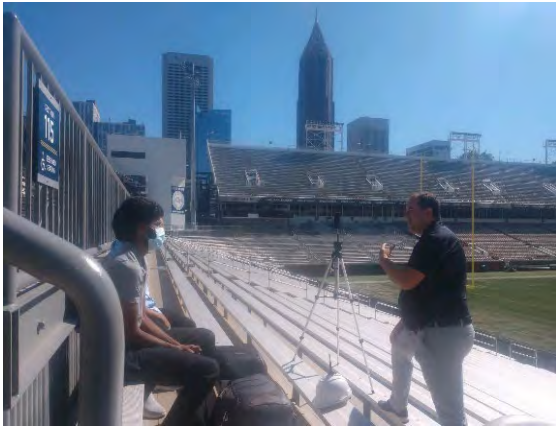


Figure 4. On-site lecture delivery method

feeling of being on a project site. The starting point for this lecture was a team discussion around the site capture approach, techniques, capture method (single panorama, cluster capture, video capture), and deployment methods (tripod, hardhat, pole, drone, 360Rover). Discussions included site preparation, staging, lighting, outdoor conditions like weather, prioritizing critical areas, and discussing what type of documentation is required for the scope of work, expected deliverables, and outcomes.

Also discussed were details concerning data collection using 360 imaging. This included level of data collection required, capture density, imaging within areas of repetitive architectural features, and capturing in difficult locations like ceiling spaces or small, constrained areas. Students were taught the differences in capturing data using lidar compared to photogrammetry, both from a technical approach and project deliverable perspective. The selection of appropriate technology considering specific applications was discussed at length.

The lecture also addressed lighting and capturing text in the field. Students learned how to approach data capture related to general site conditions, detailed site conditions, and asset documentation. Also discussed was scaling a project, taking a known measurement, placing an artifact in the space, or leveraging known object dimensions if accurate CAD/BIM are not available.

The lecture introduced how to use the native capture App and project settings. Students learned how to find their location in the 2D CAD and drop a marker pin at the start and completion of captured areas. Students were able to take this knowledge into their field capture and owner and stakeholder conversations in the term project phase.

2.2 Hands-on Field Data Collection

Following the lecture series, students worked in pairs and were assigned an area in the Bobby Dodd Stadium to capture. Each team applied what they learned on 360 imaging and judged how many panoramas to take and

where to capture them. They employed 360 cameras in conjunction with a field capture phone app to collect data (Figure 5). The camera was either mounted on a hardhat as depicted in Figure 6, or on a mobile robotic device. Each team was instructed on exact procedures to follow in data collection and monitored by the instructor. After completing the assigned capture area, each team reviewed their work and uploaded it for cloud processing. Students also learned how to identify assets, current nomenclature, tagging, and classification, as shown in Figure 7. Additional photography was taken to gather asset attribute details for the asset tagging project.



Figure 5. Digital Twin field capture app

The instructor then reviewed and provided feedback and points for improvement on the data capture exercise through a short video review and introduced the asset tagging assignment through a workflow document and video tutorial for the assignment.



Figure 6. Student collecting data



Figure 7. Asset tagging

Each student was next expected to tag ten assets, add attribute information and upload a photo of the asset nameplate with identification information. Figure 6 shows a stadium mechanical room 360 image on the right, floor plan in the centre and left display shows asset identified with attribute information collected and tagged by students on the left

The project delivery was an asset report from the Visual Plan software in PDF format used for review and grading. Some assets required online search for identification and to find operator manuals or other asset information. Students learned from this practical exercise the importance of organization asset information effectively. The instructor reviewed the tagging exercise and gauged student performance based on a designated rubric that determined the quality and completeness of the work done.

2.3 Term Project

Having gained an understanding of digital twins and essential field tools in creating these twins, students next had an opportunity to select from several of the other technologies introduced in the course to create a digital twin. These technologies included laser scanning and drone photogrammetry. Students also had the option of continuing with 360 imaging. Three teams selected 360 imaging, one team selected laser scanning, and one team selected drone photogrammetry as their technology tool.

The plan was to apply any one of these technology tools in a real-world setting. Bobby Dodd Stadium was used as the project setting. Students were requested to match the technology tool of their preference to a location in the stadium that was appropriate for the selected technology, based on their client's needs, in this case, the Georgia Institute of Technology facilities management department and campus police department. These departments are actual beneficiaries of this project and restricting the number of stakeholders that the students needed to consider helped facilitate this short project. The term project was divided into five parts;

1. Students selected team membership based on technology interest and available time for data collection and team meetings. Each team had from 4 to 6 members. The deliverable was notification to the instructor of team

membership and the technology tool selected.

2. Teams met during class with key stakeholders of the football stadium facility who included the campus Manager of Building Information Modeling/Virtual Design and Construction, Facilities Management, and the Lieutenant of the campus police department. Each student team obtained guidance from these stakeholders in developing a brief project scope that matched their technology tool to an appropriate physical setting which is of interest to the client. Students were instructed to keep their scope small enough to collect data in 1-2 site visits. The deliverable for this part of the project was the team minutes from this meeting and the brief project scope.

3. Data collection was conducted in 1-2 team visits to Bobby Dodd Stadium. Each visit was to be from 2-4 hours. The deliverable was a progress report of 1 page or less.

4. Data compilation and evaluation were then conducted appropriately for the technology tool employed. The deliverable was a progress report of one page or less.

5. Teams then presented their findings during the final class meeting of the semester. The stakeholder attended this presentation in the role of client beneficiary. A final team report was prepared and presented to the instructor at this time. These reports are to be combined with additional data collected on the stadium and presented to the University.

3 Student Feedback

Student feedback was solicited upon completion of the assignment to assess the effectiveness of meeting the learning objectives. This feedback was obtained by including a lesson learned question in the assignment and by posting a discussion board for voluntary feedback, including observations from the assignment. Seventeen comments were received on the discussion board, and 24 student comments received in response to the assignment lessons learned question. These comments were evaluated by the course instructor.

Consistent themes found in student feedback concerned the impact of digital twins on security and pre-incident planning for buildings, facilities, and operations, highlighting further opportunities for in-depth learning and application of effective building and construction pre-planning, documentation, and visualization to safety, security, and risk assessment. It was evident that each student experienced the curriculum differently. The learning outcomes inspired critical thinking on applying digital twin technology as a valuable methodology across various applications within their professional fields.

4 Findings

The use of digital twins provided a framework for understanding a range of technologies used in building construction. This framework proved valuable in a broad scope of learning objectives related to new technologies and relevant facility lifecycle applications.

Discussions with subject matter experts and guest lectures created a foundation for students to understand various visualization technologies and how they integrate to build digital twins to create long-term value during the construction process as well as throughout the building lifecycle post-commissioning: how the technology can be utilized beyond a single application across the organization, and how to set up facilities for success for operations, maintenance, security, and sustainability.

The introduction of 360 image capture was compared to 3D Lidar, SLAM, drone capture, and traditional survey to broaden the students' understanding of each technology and use case's differences. A key takeaway for students was to critically think through the scope of work, project requirements, site conditions, and existing project workflows to determine the best technology tool for the job and how various technologies work together to serve different purposes.

Knowing the stakeholders and their needs is key to approaching the project efficiently to meet the expected outcomes. The capture pre-planning helped students understand how the environment affects the data collection process, time to capture, and overall quality of data collected. Learning stakeholder requirements also helps the best approach to the project to meet desired business outcomes, whether the objective is project capture for construction, project oversight, asset management, or safety/security planning.

Inspecting existing asset tags on a campus facility, Bobby Dodd Stadium, allowed students the opportunity to apply earlier learning to an actual life project. This project incorporated asset characterization, attribution, and taxonomy. Students were presented with a real-world example of the challenges facilities managers face when asset information is incomplete or poorly organized by completing the asset tagging exercise. The term project allowed students to work together to combine technologies to solve a real-world problem. They gained exposure to a project cycle by working with a client to develop a project scope, complete it in a constrained period, and report results to their client. The University's knowledge that data collected would be used provided an opportunity to see firsthand why the technologies introduced in the course are essential within the industry. Finally, students were exposed to the importance of critical thinking to the effective implementation of the technologies introduced in the course.

The BC6005 course is comprised of a mix of students, many with several years of experience in

construction or facilities management and others entirely new to the field. This presented a challenge when covering digital twins within a short time period. Also, the class had an enrollment of 24 students. This made it impossible to provide access to camera and other equipment during the assigned class instructional period. Both students and instructors had to attend sessions outside of the class period. These additional sessions were offered throughout certain days.

5 Conclusions

The students' varied real-world experience, ranging from academic applications through to established professional careers, with specialties including civil engineering, construction, construction management, facilities management, architecture, and transportation, encouraged multiple perspectives and critical thinking on how digital twin technology represents a potentially valuable methodology across various applications within their professional fields.

The combination of lectures, discussions with subject matter experts, guided demonstration, and self-directed real-world application of various scanning techniques and 3D digital twin technology enhanced both the students' understanding of what constitutes digital twins and their ability to identify and assess which of the different visualization technologies can be optimized for each application workflow.

Through the guided hands-on capture project at the stadium, the students were able to engage with critical facilities and security stakeholders, learn how to identify overlapping project goals and outcomes, gain a crucial understanding of when and how to implement various imaging techniques, and how to process and present that data in an accessible, comprehensive deliverable to meet stakeholder needs.

Students effectively and functionally applied this learning through the term paper assignments, working directly with university personnel to identify specific needs in real-world applications on campus.

The real-world application of the course concepts underscored the accessibility and potential for adopting the technology into the construction industry and the process the students need to apply to integrate 3d digital twins into workflows effectively: identifying and understanding the scope of work, how to identify and mitigate potential issues, which technology is appropriate each application, and understanding project goals and outcomes across various stakeholders and their needs.

The benefits demonstrated by this instructional experience can be applied in the curriculum of building construction and facilities management programs at the university, and even community college level. It is expected that the incorporation of digital twin instruction into an institution's curriculum can be highly beneficial

for both graduate and undergraduate students. Students with knowledge of digital twin theory and methodology will have an enhanced ability to perform in the age of Construction 4.0. Such graduates will likely be in high demand as the construction industry further adapts to the digital environment. Further testing of the effectiveness of employing digital twin learning would be beneficial in an undergraduate course and at community colleges, as well as additional testing for graduate level education.

References

- [1] Smit, J., Kreutzer, S., Moeller, C., & Carlberg, M. (2016, February). *Industry 4.0*. Retrieved from [http://www.europarl.europa.eu/RegData/etudes/STUD/2016/570007/IPOL_STU\(2016\)570007_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2016/570007/IPOL_STU(2016)570007_EN.pdf)
- [2] Forcael, E., Ferrari, I., Opazo-Vega, A., & Pulido-Arcas, J. A. Construction 4.0: A literature review. *Sustainability*, 12(22), 9755, 2020.
- [3] Lu, Q., Parlikad, A. K., Woodall, P., Don Ranasinghe, G., Xie, X., Liang, Z., ... & Schooling, J. Developing a digital twin at building and city levels: Case study of West Cambridge campus. *Journal of Management in Engineering*, 36(3), 05020004, 2020.
- [4] Sacks, R., Brilakis, I., Pikas, E., Xie, H., & Girolami, M. Construction with digital twin information systems. *Data-Centric Engineering*. E14. doi:10.1017/dce , 2020.
- [5] Tao, F., Sui, F., Liu, A., Qi, Q., Zhang, M., Song, B., ... & Nee, A. Y. (2019). Digital twin-driven product design framework. *International Journal of Production Research*, 57(12), 3935-3953.
- [6] Rosen, Roland, et al. About the importance of autonomy and digital twins for the future of manufacturing. *IFAC-Papers Online* 48.3 (2015): 567-572, 2015.
- [7] Khajavi, S.H., Motlagh, N.H., Jaribion, A., Werner, L.C., Holmstrom, J. Digital Twin: Vision, Benefits, Boundaries, and Creation for Buildings. *IEEE Access* 7, 147406–147419, 2019.
- [8] Olcott, S., & Mullen, C. (2020). Digital twin consortium defines digital twin. Available at: blog.digital-twin-consortium.org/2020/12/digital-twin-consortium-defines-digital-twin.html.
- [9] Wang, J., Wang, X., Shou, W. and Xu, B. Integrating BIM and augmented reality for interactive architectural visualisation', *Construction Innovation*, 14 (4), 453-476, 2014.

Benefits of Open InfraBIM – Finland Experience

R. Heikkilä^a & T. Kolli^a & T. Rauhala^b

^aCivil Engineering Research Unit, University of Oulu, Finland

^bNovatron Oy, Finland

E-mail: rauno.heikkila@oulu.fi, tanja.kolli@oulu.fi, teppo.rauhala@novatron.fi

Abstract

The achieved benefits of Open Infrastructure BIM in Finland have been studied. The newest concept of Open InfraBIM in Finland is introduced. Industrial experiences and observed remarkable benefits in three large-size infrastructure design and construction projects carried out in 2012-2022 are presented and evaluated. Based on the observations and achieved results, the implementation of the Finnish Open InfraBIM concept has been successful in Finland. Clear benefits were observed and measured in three considered major design and construction projects. Most benefits were found from the latest project Kirri-Tikkakoski with remarkable cost and time savings. The benefits of BIM based machine control and real-time quality control were emphasized. This project has also used the most advanced and widest InfraBIM concept with early integration with all relevant stakeholders. Further development, information sharing and wider utilization of the Open InfraBIM in Finland and globally are suggested.

Keywords – Benefits; Model Sharing; Open InfraBIM; InfraBIM Requirements

1 Introduction

1.1 Background

BIM (Building Information Modeling) for transportation infrastructure has earlier been studied by Costin et al. (2018) [1]. The results showed the use of BIM for transportation infrastructure has been increasing with especially focusing on roads, highways, and bridges. Also they reveal a major need for a standard neutral exchange format and schema to promote interoperability. Benefit assessments were also found: the use of BIM instead of traditional documents was estimated to lead to economic and technical benefits. Also it was stated that application of BIM can help general contractors to reduce their risks and diminish the associated costs.

Dodge Data & Analytics (2017) [2] has provided a report of the Business Value of BIM for mainly infrastructure design phase. According to the study, the

BIM implementation growth in especially US, UK, France and Germany has continued 2015-2019. Most BIM users (87%) have reported positive value from their use of BIM. Top business benefits of BIM related to transportation infrastructure projects have been such as (1) improving ability to show younger staff how projects go together (58%), offering services (56%), establishing consistent and repeatable project delivery process (54%), maintaining business with past clients (52%), and, less time documenting, more time designing (50%). As top ways how BIM has improved project processes and outcomes were found: (1) fewer errors, (2) greater costs predictability and better understanding of the project, and (3) improved schedule performance and design optimization. Based on the results, nearly two thirds (65%) believed that they can get a positive ROI (Return of Investment) from their BIM investment. In the previous publications, the detailed content of InfraBIM, i.e. modeling requirements, nomenclatures, open data transfer formats, as well as methods and applications utilizing these are not presented or described in more detail. [2]

Finland is not the only country in where infraBIM is utilized. Norway has a long experience of infraBIM, and this has been compared with projects in Vietnam [3]. Based on the literature, the business value of BIM for infrastructure sector has been evaluated in France, US, UK, and Germany [1]. In addition, the experiences of OpenBIM in infrastructure projects have been studied in Netherlands by Bergs et al. 2016 [4], Russia by Boykov et al. 2020) [5] and Italy by Giovine 2019 [6]. In those publications, the concept and content of Infra BIM has not described at the level that Infra BIM is nowadays used in Finland.

Halttula (2020) has studied the benefits of early integration of BIM to infrastructure design and construction project. Halttula states that early integration of key stakeholders can lead to prompt and effective optimization of core competencies and knowledge as well as improves communication and decrease fragmentation. The product data model for the project should be designed in the early phase in collaboration with all relevant stakeholders, and also with client. [7]

Related to InfraBIM based machine automation and

the benefits, Caterpillar (2016) has experimentally studied the benefits of utilizing 3-D machine control technology compared to traditional road construction methods. Infrastructure modeling was not considered. The benefits of the machine control technology were clear when considering the reduction in project duration, equipment hours, fuel consumption, total machine cost, operating hours, man-hours and total man-hours. Fewer machine hours protect machines from excessive wear, which reduces fuel consumption and leads to an environmentally conscious solution and reduced greenhouse gas emissions. [8]

Second Caterpillar workstudy was made 2006 in Malaga. The study measured productivity growth with AccuGrade 3-D machine control systems. In the experiments, two identical roads were constructed, one using traditional control methods and the second using 3D machine control systems. The lead times for all the different work steps were measured, the number of passes, buckets or truck loads, fuel consumption and also accuracy with these two different methods. According to the results of the study, the total working time required to build the road was 3½ days with traditional and 1½ days with 3D machine control. Fuel consumption was 43% lower with 3D machine control. [9]

In Finland, the development of Open Infrastructure Building Information Modelling (Open InfraBIM) concept has been ongoing since 2010 (Halttula 2020, Kivimäki et al. 2015), [7, 10] The large research work program RYM Process Re-Engineering (PRE) with Infra FINBIM work package was executed 2010-2014. Open Infra BIM means the concept that enables wide industrial utilization of model-based information at all stages of the construction process. [7,8]

The Finnish Open InfraBIM concept (Figure 1) includes three main parts: 1) modelling guidelines, 2) information classification system, and 3) open information transfer format Inframodel. The open concept has been specified down to detail and published by BuildingSMART Finland Infrastructure [11]. Building Smart Finland Infrastructure Business Group aims to develop and advance the use of infrastructure information modeling in Finland. The goal is to have fully digitized infrastructure design and production processes by 2025. [12]

The Common InfraBIM Requirements (the newest version YIV2019) cover the entire life cycle of an infrastructure project: currently initial material, different phases of design and construction including also as-built documentation. In the future, the aim is to add operation and maintenance to the requirements as well. The modelling guidelines aim to guide, harmonize and improve the modelling practices in the entire infrastructure sector. The guidelines are based on the current best practices, and they will be updated regularly

as the knowledge and tools will be developed.

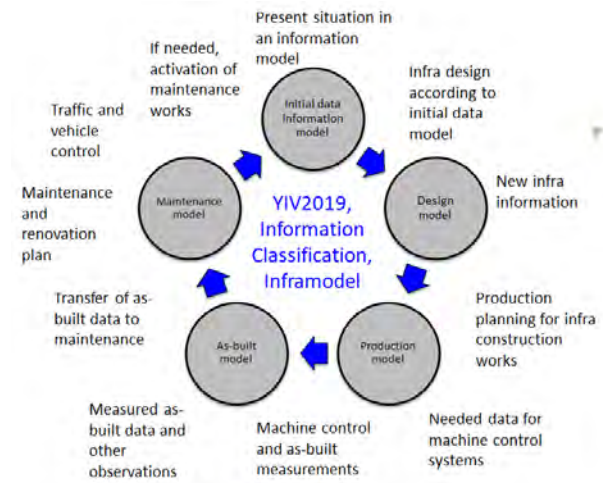


Figure 1. The Concept of Open InfraBIM in Finland (University of Oulu).

Inframodel is an open method for the exchange of infrastructure information. It is based on the LandXML standard. The Inframodel includes parts for terrain models, subsoil surfaces, road and rail geometries and construction layers. It also covers water sewage and supply and some facilities. The Inframodel documentation explains the Finnish method for using LandXML. Inframodel does not gather all the elements in LandXML – Inframodel is a subset of LandXML. On the other hand, some Inframodel specific extensions have been added to support machine readability needed. The most important of these is the Infra classification system. Inframodel is the exchange format required by the Finnish Traffic Agency and major cities since May 1st, 2014. Today, most of the main important client organizations, design offices and contractors use open and real-time cloud services Infrakit to share and utilize of open Infra BIM in different work phases of infra projects.

The InfraBIM Classification specifies the numbering and designation of infrastructures and infrastructure information models covering their entire life cycle (see Fig 2.). The purpose of the classification is to obtain a unified numbering and designation practice that serves infrastructures and infra-structure information models in every phase throughout their lifecycle: in obtaining initial data, design, execution, as-built surveys and maintenance. [13]

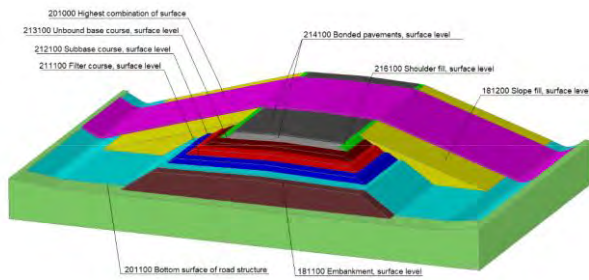


Figure 2. An Example of the InfraBIM Classification System (BuildingSMART Finland).

1.2 Aim

The aim of the research was to study the benefits of Open InfraBIM concept achieved among the selected highway and railway design and construction projects executed in Finland in 2012-2022. The sub-objectives were:

- to identify the BIM utilization level and main achieved benefits of Open InfraBIM in a few selected major infrastructure projects in Finland and assess their economic significance
- to find out whether the continuous development of Open InfraBIM is reflected in the increase of benefits in infrastructure projects
- to evaluate the usefulness and competitiveness of the Finnish Open InfraBIM concept compared to the concepts and measured benefits used in other countries.

2 Materials and Methods

2.1 Study about the use of Open InfraBIM in Finland

The following three major design and construction projects, (1) Kokkola-Ylivieska Dual Rail Design and Construction Project (schedule 2012-2018), (2) Railway Design and Construction Project Pasila-Riihimäki, 1st phase (2016-2022), and (3) Highway Design and Construction project Vt4 Kirri-Tikkakoski (2019-2022), were selected for closer interviews and studies about possible benefits by the use of Open Infra BIM. In this study, some selected key people in the projects were interviewed. In addition, sub-reports and evaluations of the projects were analyzed. In the (1) project the YIV 2015 modelling requirements were first implemented at large scale. There were widely utilized machine control systems (guidance-based control) for all the excavators, some implementations were also made and tested in road graders and compaction machines. Novatron was one example of company which provides machine control system that can open way be connected to cloud service

such as Infrakit. The first version of Infrakit cloud service was used. During the (2) project the InfraBIM concept development was continued and further later utilized for the concept updating. In the (3) project YIV2019 was already utilized with also wider implementation of the real-time Infrakit cloud services for all the model sharing needed.

2.2 Kokkola-Ylivieska Dual Rail Design and Construction Project (2012-2018)

The Kokkola-Ylivieska dual rail design and construction project was implemented in 2012-2017 (Fig 3 and 4). The estimated cost of the project was € 330 million. The project built 80 km of new railways, repaired 80 km of existing railways, built 80 km of maintenance roads, built 20 km of new streets and roads, and built 80 new bridges and 70 culverts. The construction works of the intervals between Kokkola-Riippa and Riippa-Eskola were carried out as two separate sub-projects using the Design-Implement model. The last section, Eskola-Ylivieska, was split into smaller contracts: one electric track contract, which will be implemented with the Design Build model, one design contract, three substructure contracts, three superstructure contracts and the construction of the Vääräjoki bridge. The entire Kokkola-Ylivieska interval will be managed with one safety equipment contract. CC Infra Oy was the builder consultant for the project. The builder consultant was assisted by WSP Finland Oy under a separate agreement.



Figure 3. Kokkola-Ylivieska Dual Rail Design and Construction Project in Northern Finland – General Map.



Figure 4. Kokkola-Ylivieska Dual Rail Construction – The earthworks of the bridge sites were built using BIM model-based machine automation.

2.3 Railway Design and Construction Project Pasila-Riihimäki, 1th phase (2016-2022)

The Pasila–Riihimäki design and construction project (150 MEUR) improved the functionality of Finland's busiest railway section. The aim was to separate freight from passenger traffic and to make the connections at the main transport points faster. Several design companies and contractors participated in the project. All of them utilized Open InfraBIM as well as automated control systems in machinery. Most of them also used Infrakit cloud service.

The first phase of the Helsinki – Riihimäki project focused on improving traffic locations. In addition, an additional freight track was built to Kerava and an additional track to Järvenpää between Ainola and Purola. In the summer of 2021, a new interlocking device was introduced in Kerava to monitor the safe operation of trains, which is a prerequisite for the additional tracks to be built in the second phase.

The Open Infra BIM operation model in Pasila-Riihimäki project was as follows:

- the initial data model was provided by the owner/client
- InfraBIM requirements were written for the procurement
- continuous InfraBIM modeling and the utilization of models
- BIM model coordinators together, several consultants together with BIM coordinator
- InfraBIM modelling plan, operation planning utilizing BIM, BIM management
- continuous development and implementation of the BIM in several workshops with project staff
- implementation of the BIM plan with everyone in the project.

The evaluation with interviews was made by Netlipse organization (Network for the dissemination of knowledge on the management and organization of large infrastructure projects in Europe) on behalf of the Finnish Transport Infrastructure Agency (FTIA). [14]

2.4 Highway Design and Construction project Vt4 Kirri-Tikkakoski (2019-2022)

The Vt4 highway is one of Finland's most important heavy transport routes (see Fig 5). According to the FTIA, the traffic volume at Kirri in Jyväskylä was about 20,000 vehicles a day. Heavy traffic accounts for about 10 percent of this. The aim of building the motorway has been to improve traffic safety as well as uniform driving conditions, predictability of travel times and operational reliability.

The project Vt4 Kirri-Tikkakoski built a new motorway a total of 16 kilometers from Kirri in Jyväskylä to Tikkakoski and further to Laukaa Vehnä. In addition to the motorway, more than 20 kilometers of new roads and about 30 kilometers of new light traffic routes have been built during the contract. The design and development phase of the Kirri-Tikkakoski area began in the autumn of 2018. Construction work began in the spring of 2019. Destia was the main contractor in the project.



Figure 5. Highway Design and Construction project Vt4 Kirri-Tikkakoski (Finnish Transport Infrastructure Agency, FTIA).

3 Observations and Results

3.1 Kokkola-Ylivieska Dual Rail Design and Construction Project (2012-2018)

During the project, an interview survey was conducted for all different parts of the project, involving individuals from design, construction, and supervision. According to the survey results:

- 97% of respondents had been in contact with BIM
- 68% thought the data models had made the job of

design easier

- 69% said the turnaround time was faster in construction
- 55% answered that some or significant additional costs were incurred in the contracts due to BIM
- the savings from the use of BIM were achieved to some extent or by a significant 64%
- the quality of the design was estimated to have improved by 80%
- estimation of the quality of construction 93% of the respondents have improved data modeling



Figure 6. A Combination BIM Model Example of Kokkola-Ylivieska project.

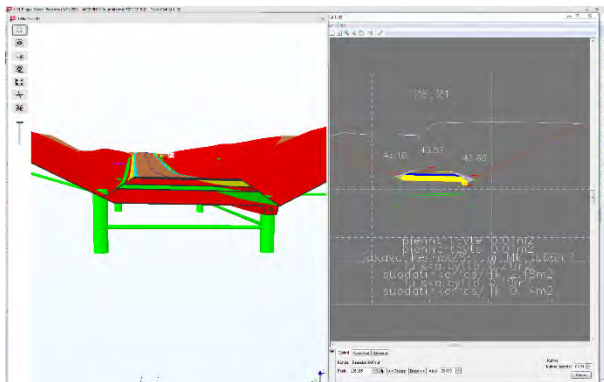


Figure 7. An Example of BIM model in Tekla Civil Design Software.

In the design, the greatest benefits were felt from more illustrative design, coordination of different types of technology and reduction of errors (see examples in Fig 6 and 7). In construction, there has been an improvement in quality in terms of both design and construction, as well as faster work. The cost has been somewhat increased by learning and developing new technologies and practices. The savings have been achieved through a reduction in errors, a reduction in waste and faster work. Software and the continuous development of operating methods were seen as important, as can see in Figures 8 and 9.

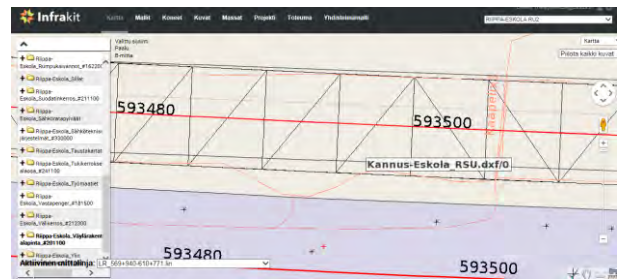


Figure 8. Model of track protection gauge in the Infrakit user interface. This real-time function was developed in a machine control system (Novatron Oy) that warns the operator when the machine approaches the area. The implementation and final tests were carried out in the Kokkola-Ylivieska project.

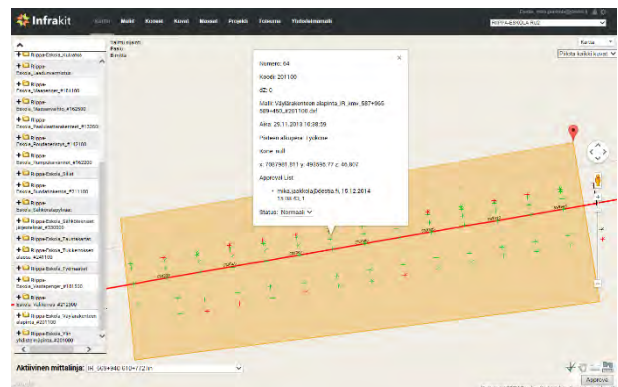


Figure 9. An Example - Supervisor's tool for validating point quality in Infrakit cloud service. The inspected area can be delimited and accepted, causing the color of the actual points to turn green.

3.2 Railway Design and Construction Project Pasila Riihimäki, 1th phase (2016-2022)

The Pasila–Riihimäki design and construction project was the second major infrastructure design and construction project in Finland, in which the wider implementation and utilization of the Open InfraBIM was developed (Figs 10, 11 and 12). The success of the implementation was evaluated by an external consultant Netlipse.

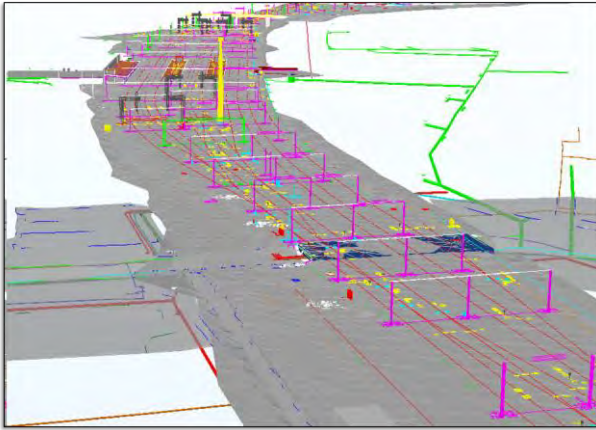


Figure 10. Initial data model of Pasila-Riihimäki project (Trimble Connect cloud service). The railway yard included lots of different structures. BIM models were seen necessary to handle the information.

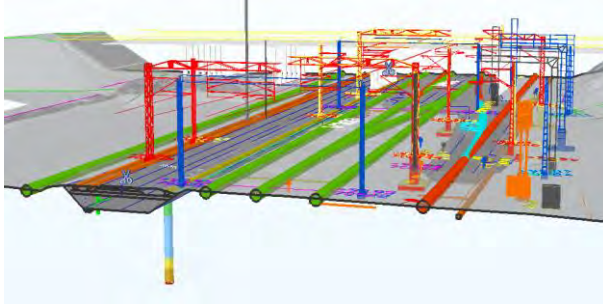


Figure 11. An Infra BIM example (Trimble Connect) – A Combination model of Pasila-Riihimäki project.

According to the external evaluation report of the construction project (Netlipse 2018) [14], the team has been in the project relatively thin and the members have felt themselves to have meaningful roles and responsibility for making decisions within their areas of competence. Issues referred to the Project Director (PD) for decision making have been dealt with quickly. The team appeared very confident in the PD's leadership and experience to ensure success. Motivation to deliver a quality product has been high.

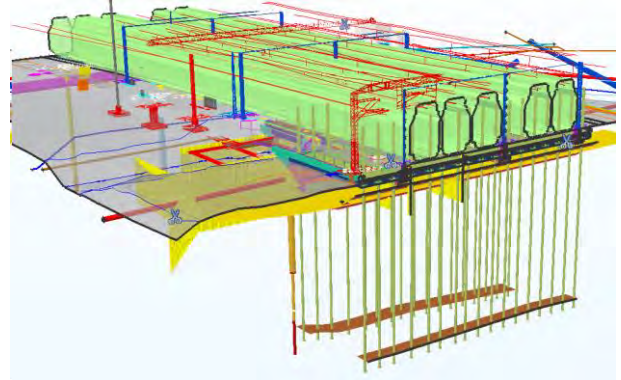


Figure 12. An Example - Utilization of InfraBIM in design (visualization, clash detection) and construction (Trimble Connect).

The decision to fully utilize the opportunities presented by BIM 3D modelling for design, construction and maintenance by the PD and team has been highly successful. The project has achieved significant cost savings (estimated at approximately Euro 10 million), efficiencies and quality improvement through this approach.

Cost management and control came out very strongly in most interviews with the PD proactively seeking to reduce anticipated final costs and deliver additional value through underspend. Procurement of contracts also very strongly focused on price.

The PD was an experienced leader and was involved in almost all aspects of the project delivery function. He has significant knowledge of the key issues and risks but has not documented much of this knowledge. In order for his successor and other team members to manage this transition, some form of lessons learned / knowledge capture was essential:

- a knowledge management plan (or lessons learned register) is strongly recommended due to the limited nature of formal project documentation and the strong reliance upon the PD and the team's personal knowledge and intervention.
- it is strongly recommended that the lessons learned through the adoption of BIM modelling should be captured and shared for the benefit of future schemes.

In the evaluation report it was also suggested that to ensure that original benefits are realized the FTIA could act in a more proactive client role to continue to measure benefits case during construction and then test those benefits are realized post completion. One of the key suggestions to FTIA was to continue BIM utilization development in future construction and maintenance project i.e. to exploit BIM opportunities nationally.

3.3 Highway Design and Construction project Vt4 Kirri-Tikkakoski (2019-2022)

In December 2021, the construction contract was up to eight months ahead of schedule [15]. According to the fairway agency, the rapid progress of the work was due to new ways of working in design, construction and quality control.

The design has been done entirely on a model basis, and no paper drawings have just been commissioned from the project. The time saving was about 20-22 weeks, i.e. almost half a year has been saved in design time. This has also been reflected in the construction schedule. A lot of new construction technology, quality management technology and design have been introduced and utilized in the project, which has enabled significant time savings. Automatic 3-D machine control systems were used for all the machinery, thus measuring sticks have not been installed on the site at all.

Project members can use combination model from Infrakit to monitor the current progress (see Fig 13). Infrakit can also use to real-time Model-based quality control using color to show tolerances (Fig 14) has been done as well as by using for example with drones (Fig 15). The quality information was produced in real time. The bridges were modeled, and all approvals were model-based using Trimble Connect, no papers were delivered to the client or for construction. Reinforcements were ordered from the factory according to IFC models, not with paper reinforcement lists. Bridge measurements were performed entirely according to IFC models, and no separate measurement data was created for bridge measurements.



Figure 13. Highway Design and Construction project Vt4 Kirri-Tikkakoski – a BIM model View in Infrakit Cloud Service (Destia Oy).

Their large-scale road construction project has remained very well within its budget. According to FTIA,

the initial cost estimate for the Kirri-Vehniä subscription was EUR 156 million. The actual construction costs were EUR 139 million. In addition to the Kirri-Tikkakoski project, 139 million were completed on an additional Vehniä section, i.e. one interchange and six to seven kilometers of parallel road longer than was included in the original project.

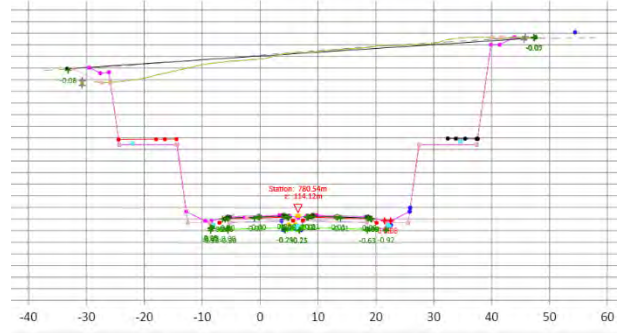


Figure 14. A sample of Infrakit Cross-section showing real-time quality deviations (green points inside tolerances, red points outside tolerances).



Figure 15. Example – a drone image used for model-based quality control (Destia Oy).

4 Conclusion

This study was carried out at the request of the industry in order to find out the benefits of Open InfraBIM in design and construction projects in Finland. It was especially hoped that the research would provide support for the introduction of the Finnish Open InfraBIM concept internationally.

In the earlier first project (2012-2018) already most of the employees interviewed had been in contact with the Finnish Open InfraBIM. In design, the greatest benefits were achieved from more illustrative design, coordination of different types of technology and reduction of errors. In construction, faster works, material savings and quality improvements were reported. Turnover time was evaluated to be clearly faster in construction. Significant savings from the use of BIM were achieved.

In the second project (2016-2020) the report stated that the project has achieved significant cost savings of Euro 10 million, i.e. about 7% of the total project cost, efficiencies and quality improvements.

In the later third project (2019-2022), the realized reduction in construction time was about eight months in the end of 2021. The rapid progress of the work was due to new ways of working in design, construction and quality control. The time saving in design was about 20-22 weeks. All the machines were equipped with automatic machine control systems. Real-time quality control was seen to be preventing potential faults at an early stage during construction.

The implementation of the Open InfraBIM concept has been successful in Finland. Clear economic and technical benefits were observed and measured in all of the considered major design and construction projects carried out in 2012-2022. The most remarkable benefits were faster construction with reduction of construction time, cost savings up to 7% of total design and construction costs, and better completed quality of infrastructures. Most benefits were found from the latest project Kirri-Tikkakoski with remarkable time savings. The benefits of BIM based machine control and real-time quality control were emphasized in that project. This project has also used the most advanced and widest InfraBIM concept with early integration with all relevant stakeholders. The technology had already evolved considerably further, and people (an experienced team) had learned to build model-based.

The continuous development of Open InfraBIM seems evidently to be reflected in the increase of benefits in design and construction operations and processes. The progressiveness and competitiveness of the Finnish BIM concept in similar systems in other countries is difficult to assess on the basis of this study. Further development, wider information sharing and utilization of the Open InfraBIM in Finland and globally are suggested.

References

- [1] Costin, A. & Alireza, A. & Hanhin, H. & Stuart, S. C. (2018) Building Information Modeling (BIM) for Transportation Infrastructure – Literature Review, Applications, Challenges, and Recommendations. *Automation in Construction* 94 (2018), pp. 257-281.
- [2] Dodge Data & analytics (2017) Smart Market Report Top Business value of BIM for infrastructure.
- [3] Nhat Nam Bui (2021) Implementation of Building Information Modeling in Infrastructure Construction. Doctoral Dissertations at the University of Agder 329.
- [4] Bergs, R. A. & Winkels, H. J. (2016) *InfraBIM in den Niederlanden*. https://henry.baw.de/bitstream/20.500.11970/102344/1/04_Bergs-Winkels_InfraBIM.pdf
- [5] Boykov, V. N. & Skvorsov, A. V. & Gurev, V.A. (2020) InfraBIM Open paradigm as the driver of information of the road sector in Russia. *IOP Conf. Ser.: Mater. Sci. Eng.* **832** 012045.
- [6] Giovine, A. (2019) Thesis: InfraBIM e construction management: valutazione di interoperabilit  OPEN BIM. <https://webthesis.biblio.polito.it/11128/1/tesi.pdf>
- [7] Halttula, H. (2020) Enhancing Data Utilization in the Construction Project Lifecycle through early Involvement and Integration. University of Oulu, Graduate School, C Technica, C750, 112 p.
- [8] Caterpillar (2016) Comparison report, Caterpillar jobsite information study, 22 p.
- [9] Caterpillar (2006) Road construction – production study. 26 p.
- [10] Kivim ki, T. & Heikkil , R. (2015) Infra BIM based Real-time Quality Control of Infrastructure Construction Projects. ISARC'2015, The 32nd International Symposium on Automation and Robotics in Construction and Mining, 15-18 June, 2015, Oulu, Finland.
- [11] Final report of Inframodel3 pilot project for data modelling-based planning and construction process. Finnish Transport Agency, Planning Department. Helsinki 2015. Research reports of the Finnish Transport Agency 17/2015. 45 pages and 2 appendices. ISSN-L1798-6656, ISSN 1798-6664, ISBN 978-952-317-077-3.
- [12] Liikennevirasto (2015) Tietomallipohjaisen suunnittelu- ja rakentamisprosessin Inframodel 3 pilotti, loppuraportti.
- [13] BuildingSMART Finland (2022) Infrastructure, YIV2019 guidelines, Inframodel, InfraBIM classification system, InfraBIM Glossary of Terms, <https://buildingsmart.fi/en/infrabim-en/>
- [14] Netlipse (2018) Preliminary Feedback Session. IPAT Pasila Riihim ki 1th phase. Evaluation report.
- [15] YLE (2021) – the News Agency in Finland. Mit  ihmett ? Moottoritie valmistuu kuukausia etujassa ja pysy  budjetissa

The Neural Basis of Risk Attitude in Decision-Making Under Risk: fNIRS Investigation of the Simulated Electrical Construction Task

S. Pooladvand,^a D. Ay,^b S. Hasanzadeh^{c*}

^a PhD Student, Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Dr., Hampton Hall B147, West Lafayette, IN 47907.

^b Undergraduate Research Assistant, Industrial Engineering, Purdue University, 550 Stadium Mall Dr., West Lafayette, IN 47907.

^c Assistant Professor (corresponding author), Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Dr., Hampton Hall 1229, West Lafayette, IN 47907.

E-mail: spooladv@purdue.edu, ayd@purdue.edu, sogandm@purdue.edu

Abstract –

Risk propensity, or individuals' attitude toward risk, can highly impact individuals' decision-making in high-risk environments since those who merely focus on positive consequences associated with high-risk acts are more likely to engage in risk-taking behaviors. Previous studies identified activation in the prefrontal cortex during decision-making under risk to be a sign of an individual's attitude toward risks. To investigate whether such past work—prevalent in behavioral research domains—translates into construction safety, this study conducted an experiment in a mixed-reality environment using functional near-infrared spectroscopy (fNIRS) technology to examine whether positive risk attitudes cause individuals to adopt risky construction behaviors and whether the activation of the prefrontal cortex of the brain can represent such risk attitudes. The results show that participants with a higher risk propensity had a higher brain activation during the risky electrical tasks; these individuals merely focused on gains, which motivated them to increase their risk-taking behavior and consequently experience more electrical accidents. Understanding workers' attitudes toward risk will thus influence future understandings of decision behavior under risk.

Keywords –

Risk attitude; Construction safety; Decision-making; Risk-taking behavior; fNIRS neuroimaging; Mixed-reality (MR)

1 Introduction

Despite various efforts to reduce the number of incidents occurring within the electrical construction industry, this area still experiences a high rate of fatalities, representing a 3.75% increase over recent years [1]. In part, these fatalities may be sourced in construction workers' behaviors, which can be easily influenced by their individual characteristics. Consequently, investigating the individual characteristics that can affect workers' unsafe behaviors may help avoid future accidents.

Risk propensity, or one's attitude toward risks, is one influential characteristic that can affect jobsite safety as high-risk propensity causes individuals to adopt risky behaviors [2]. Previous research highlighted the direct connection between risk attitude and risk decision-making [2,3], the latter of which ties into cost-benefit analysis weighing the costs (risks) against the benefits (gains) delivered by the behavior. Thus, the extent to which one engages in risky behaviors is a function of individuals' positive attitudes (i.e., focused on gains) and/or negative attitudes (i.e., focused on losses) related to risk consequences [4]. Individuals with positive attitudes mostly consider positive consequences over negative ones, which stimulates them to take more risks.

The impacts of risk propensity conceivably manifest profoundly within such competitive and dynamic workplaces as construction jobsites, since the business nature of construction is highly competitive and may turn stakeholders' focus toward gains (e.g., earning more money) [5]. In such situations, managers stimulate workers by offering extra compensation as an incentive to speed up or perform simultaneous tasks in order to

offset delays or increase the company's profits. As a result, risk propensity in terms of expecting pleasurable outcomes and benefits may guide individuals to engage in more risky actions.

While the impacts of risk propensity in gain-loss decision-making under risk have been widely discussed in behavioral research domains, there is a paucity of research within the construction sector despite the fact this industry's high-risk work environment may be considerably impacted by the concept of risk attitude. Therefore, this study examined whether perceiving greater benefits during risky activities causes workers to engage more in risky behaviors on jobsites. To achieve this objective, this study asked subjects to perform a simulated high-risk electrical activity under conditions with varying benefits. The research team then used traditional (questionnaires) and emerging neuroimaging (functional near-infrared spectroscopy) techniques to document subjects' risk propensity; the latter method monitored subjects' cortical hemodynamic responses (i.e., brain activation) during the high-risk situations to quantify cognitive appraisals associated with risky decisions. Combined, this methodology enabled the team to both better understand subjects' attitudes towards risk and discern whether functional near-infrared spectroscopy (fNIRS) signals could be considered a useful method for studying individual risk attitudes in dynamic risky decision-making. The outcomes of this research, therefore, deliver both an innovative methodology for monitoring construction workers' real-time risk propensity and a deeper understanding of workers' attitudes towards risk to enhance evaluations of decision-making behaviors under risk.

2 Background

2.1 Risk Propensity and Expected Consequences

Generally, risk propensity is defined as individuals' attitudes toward risk and reflects their orientation toward taking or avoiding risks [2]. Therefore, risk attitude includes both risk-seeking and risk-aversion and signifies "the degree to which a person has a favorable or unfavorable evaluation or appraisal of a behavior" [6, p. 188].

Risk attitude can be quite influential in explaining individuals' risk-taking behaviors and risk decision-making. As with the cost-benefit analysis discussed above, one's behavior evaluation will include gains or losses, depending squarely on individuals' attitude toward the risk (i.e., risk-seeking or risk-aversion). Workers who are risk-takers primarily look forward to gaining potential benefits from the risky activity, which they perceive as worth any associated potential negative consequences [2].

Previous literature showed that individuals may adopt risky behaviors when the balance between the perceived losses of a situation and the perceived gains of that situation is considered favorable [4]. In a related study, Slovic and his colleagues observed that individuals who were more engaged in risky activities perceived greater associated benefits and also greater control over potential losses than those who did not engage in risky activities [4]. In one of the recent studies, Hasanzadeh and her colleagues examined risk propensity as a factor of individuals' risk-taking behavior in a simulated mixed-reality environment. They observed that risk propensity moderated the relationship between safety protection and risk-taking behaviors since individuals with higher risk propensity took more risks when protections were in place [2,7]. Therefore, it is crucial to investigate the substantial differences at play in individuals' risk attitude and how these gain expectancies in non-targeted risky events are linked to individuals' at-risk decisions on construction jobsites.

2.2 Cortical Brain Activation and Decision-making Correlates

Previous studies showed that neuroimaging provides an excellent understanding of the underlying cognitive processes involved in considering trade-offs between costs (loss) and benefits (gains) under risky conditions [8,9]. The prefrontal cortex (PFC) plays a substantial role in these decision-making processes [10]. Particularly, the increase of cerebral oxyhemoglobin and blood flow within the PFC reflects processing variances, uncertainties, risks, expected values, and probabilities [9]. Furthermore, previous studies showed that increased expected benefits of risky actions (greater gain) will increase the prefrontal area's brain activation, specifically among risk-seeking individuals (i.e., those with higher risk propensity) [11,12]. As such, cortical brain activity can functionally reveal the correlation between risk and associated benefits and can signify how individuals perceive the consequences of risky decisions (i.e., whether they focus on gains or losses). Several neuroimaging and human behavior studies have investigated risk attitude and decision-making under risk using cortical brain activation (e.g., [13,14,15]), but there is limited research in this field within the construction safety domain.

3 Methodology

This study examined risk perception and risk decision-making during a risky construction activity to identify the correlation between risk-taking behaviors, expected benefits, and brain cortical responses. To accomplish this objective, this study used the

transmission and distribution of energized powerlines as a high-risk task since linemen are required to work in close proximity to high-voltage powerlines while they are also at the height [16,17]. We hypothesize that risk attitude (i.e., concentrating on expected gains or expected losses) serves as a key contributor to stimulating risk-taking behaviors and exacerbates the likelihood of incidents (e.g., experiencing arc flash, which is an electrical discharge that includes burns, blasts, and electrocution hazards) within high-risk tasks among those with high-risk propensity.

3.1 Experimental Design

A mixed-reality environment consisting of virtual and physical models was developed to simulate an electrical task in a U.S. suburban area (Figure 1). The physical model included passive haptics (i.e., bucket, hot-stick, fall-arrest system, and insulating gloves). The virtual model entailed the simulated setting as well as five virtual reality trackers attached to the subject's body to capture individuals' postures and adjust the virtual avatar accordingly; these trackers also registered interactive behaviors—e.g., simulated electrical arc flash—and the virtual reality system included any corresponding visual and audio representations to enhance participants' sense of presence within the mixed-reality simulation. Environmental modalities, including wind and sound effects, were also added to increase realism and subjects' sense of presence. Most of the participants reported a high presence score (Mean = 4, SD = 0.5), given a 5-point Likert scale post-trial presence questionnaire (with 1 = low and 5 = high). This conveys that the developed MR environment offered a valid and appropriate framework to trigger the naturalistic behaviors of line workers. All subjects wore a wireless functional near-infrared spectroscopy (fNIRS, Brite) neuroimaging cap so the research team could monitor subjects' decision-making and risk attitude while the subjects completed the electrical tasks.

3.2 Data Collection

Thirty-three healthy subjects—11 females and 22 males aged 21.3 ± 2 years, with at least 1.5 years of work experience in the construction industry—were recruited to participate in this study. All procedures were approved by Purdue University's Institutional Review Board (IRB).

After a 30-minute comprehensive training regarding the experimental process and electrical tasks, each participant filled out several questionnaires, including the cognitive appraisal of risky events (CARE) questionnaire. The CARE questionnaire, developed by Fromme et al. (1997), evaluates individuals' expectations of gains (i.e., positive outcomes known as PCARE) and losses (i.e., negative outcomes known as NCARE) as consequences of risky behaviors [18]. Subjects responded to this questionnaire based on a 7-point Likert-scale that ranged from 1 (not at all likely) to 7 (extremely likely). Their responses evaluated the expected positive and negative outcomes of six various types of risky behaviors, including (I) Illicit drug use, (II) Aggressive and illegal behaviors, (III) Risky sexual activities, (IV) Heavy drinking, (V) High-risk sports, and (VI) Academic/work behaviors. This study considered the positive outcome expectancies of subjects (i.e., PCARE) for further analysis.

After completing the questionnaires, participants were equipped with the fNIRS cap, and their brain cortical responses were captured for 120 seconds as the baseline. Thereafter, they were required to complete the line replacing task, which included two sub-tasks: (1) move the energized powerlines from an old pole to a new pole, (2) remove conductor hoods from energized lines. The participants were equipped with complete safety interventions and performed the task under two experimental conditions: (A) normal condition, and (B) high-risk with incentive. For this latter, high-risk with incentive, condition B, the research team added productivity demand, time pressure, and cognitive demand to the expectations: subjects were given 10 fewer

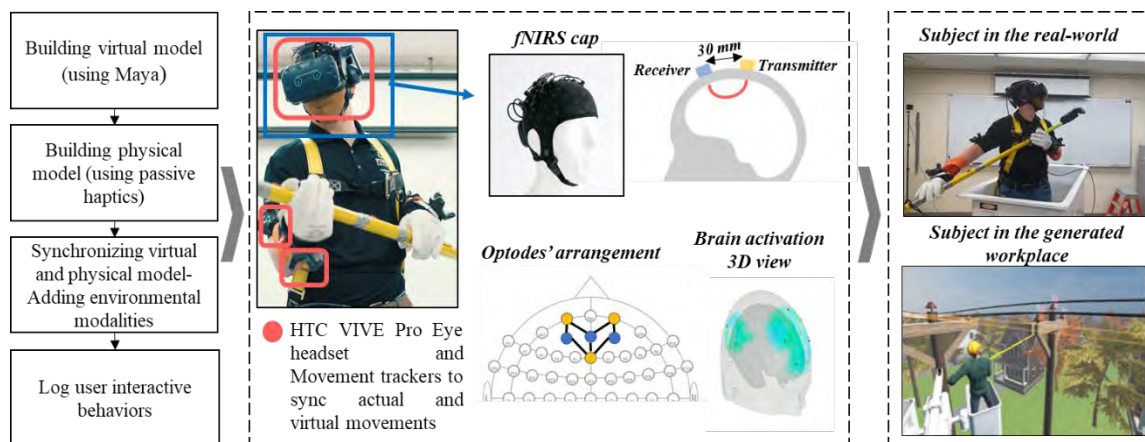


Figure 1. Research framework for real-time mixed-reality environment synchronized with fNIRS

seconds than they needed to complete the task under the normal condition, and they were asked to complete a 2-back working memory task simultaneously while they performed the main task. Critically, under Condition B, participants were told that if they could complete the task in a timely manner while completing the secondary task accurately, they would receive \$10 additional compensation. At the end of the experiment, the research team conducted a semi-structured interview to assess participants' risk perception within each condition.

As explained, brain activation manifests as increases in both cerebral oxyhemoglobin and blood flow throughout the brain, which appears to serve as a proxy for risk-seekers concentrating on gains during risky decision-making [10]. The arrangement of the fNIRS optodes' locations along with the PFC is demonstrated in Figure 1, which covers both right and left hemispheres. Specifically, a trajectory of 7 optode channels was implemented, which covered mostly the dorsolateral prefrontal cortex (DLPFC). The neural activity from the hemodynamic response function (HRF) that specifies BOLD signals overtime was used for the analysis.

4 Results and Findings

This study investigated subjects' behavioral responses and safety-related decisions under risk when gains and losses were in place. To do so, the research team began by differentiating the 33 participants' responses to the CARE questionnaire across the six categories of risk activities, discussed above. Then, correlations between the PCARE six categories and the PFC activations under Condition B, as well as the correlations between the PCARE categories and the subjects' Δ brain activation (i.e., changes in brain

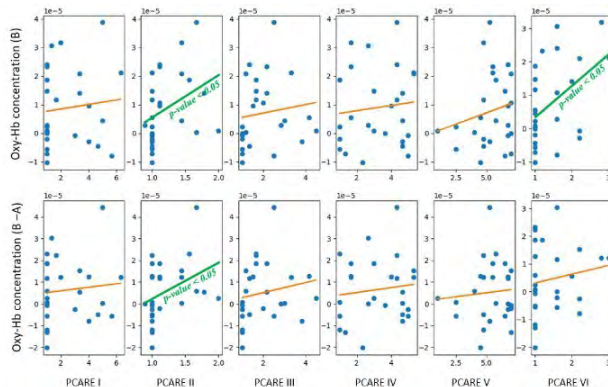


Figure 2. Correlation among brain activation and different categories of PCARE

activation from the normal Condition A to the risky Condition B), were identified (Figure 2).

While only the correlation between subjects' PCARE score and brain activation under Condition B in

categories *II* and *VI*, and their PCARE score and changes in brain activation (B - A) in category *II* are significant, all scatterplots demonstrate positive correlations between these two factors. Such insight reveals that as participants perceived more benefits than harm from being involved in irrelevant risky events, they perceived more gains in a risky construction task when there is an incentive in place.

For further analysis, participants were divided into two groups based on their average score in each category of the PCARE: (1) those more likely to focus on gains (high PCARE) versus (2) those less likely to focus on gains and more likely to focus on losses (low PCARE). Then, the changes in hemodynamic responses (oxy-Hb) in Condition B compared to Condition A (i.e., B-A) were compared between high-PCARE and low-PCARE groups across the six categories. Figure 3 demonstrates that, on average, there is more brain activation (changes in oxy-Hb) among participants in the high-PCARE groups than those in the low-PCARE groups in all six risky-activities categories. Moreover, there is a significant difference between the average brain activation among the high-PCARE group compared to the low-PCARE group within category *II* (i.e., aggressive and illegal behaviors).

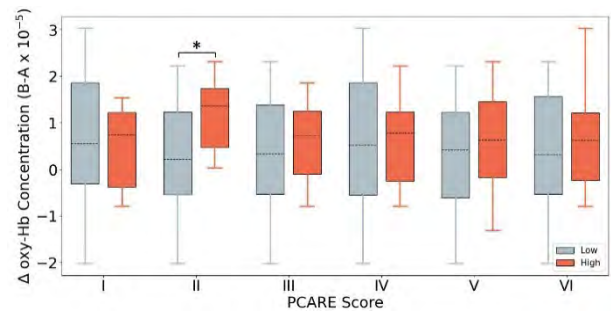


Figure 3. Box plot representing the distribution of brain activation (oxy-Hb concentration) in high-PCARE and low-PCARE groups

To examine whether there are significant differences in decision dynamics and associated PFC brain activation among individuals with positive risk attitudes (high-risk propensity group) when they need to complete a risky electrical task with (Condition B) and without incentive (Condition A), the Oxy-Hb changes were compared between the two conditions across all six risky event categories. As Table 1 shows, there are statistically significant differences in oxy-Hb concentrations in PCARE categories *II*, *III*, *IV*, *V*, *VI* between the normal condition (A) and the risky condition (B) (p -value *II* = 0.016, p -value *III* = 0.048, p -value *IV* = 0.040, p -value *V* = 0.025, p -value *VI* = 0.048). Partially significant changes in oxy-Hb were identified in PCARE category *I* (p -value *I* = 0.070). Further investigation of their safety performance indicated that those with higher risk

propensity took additional risks under situations with heightened risk-benefit dynamics and ended up experiencing more electrical arc flashes during the experiment.

Table 1. Statistical results comparing changes in brain activation of high-risk propensity group between Conditions A and B across six PCARE categories

Cond.	P CARE	MEAN	STD	Test Statistics (t)	p-value
A	I	0.471	0.712	-1.579	0.070**
B		1.128	1.473		
A	II	0.257	0.875	-2.672	0.016*
B		1.622	1.421		
A	III	0.173	0.562	-1.725	0.048*
B		1.000	1.423		
A	IV	0.280	0.674	-1.899	0.040*
B		0.995	1.311		
A	V	0.329	0.560	-2.154	0.025*
B		0.959	1.120		
A	VI	0.687	0.640	-1.832	0.048*
B		1.314	1.380		

**p-value < 0.1, *p-value < 0.05

Further, while the mean brain activation values across all categories were higher in Condition B than A for individuals with lower risk propensity (low-PCARE groups), there was no significant difference in oxy-Hb between Conditions A and B (p -value > 0.05) (Table 2), suggesting those with lower risk propensity will likely not take additional risks under situations with heightened risk-benefit dynamics.

Table 2. Statistical results comparing brain activation of low-PCARE groups within conditions A and B across six categories

Cond.	P CARE	MEAN	STD	Test Statistics (t)	p-value
A	I	0.144	0.634	-1.428 ^a	0.088
B		0.694	1.266		
A	II	0.180	0.609	0.003 ^a	0.499
B		0.192	0.886		
A	III	0.240	0.747	-1.598 ^a	0.089
B		0.790	1.349		
A	IV	0.226	0.732	-1.337 ^a	0.102
B		0.755	1.385		
A	V	0.164	0.915	-0.343 ^a	0.370
B		0.301	1.179		
A	VI	0.058	0.583	-1.151 ^a	0.134
B		0.386	0.954		

**p-value < 0.1, *p-value < 0.05

5 Discussion

Construction activities are known as high-risk activities, so proper perception of risks inherent to the surrounding environment is crucial for worker safety [19]. However, there are substantial differences among individuals in how risk is perceived. Such differences especially manifest in how individuals exhibit different sensitivities to losses and gains when making decisions under risk, a factor rooted in individuals' various risk attitudes.

This study examined the neural correlates and safety performance measures (i.e., number of arc flashes they have experienced while completing the task) among individuals with different attitudes toward gain and loss to assess risk-taking behaviors under varying risk-benefit conditions. The findings indicate that there is a positive correlation between each category of PCARE and brain activation during risky-with incentives tasks; thereby, risk attitude modulated brain activation in the prefrontal cortices more in participants who perceived greater positive consequences from risky actions than in participants who perceived more losses from risky actions. This finding is well-aligned with other neuroimaging studies that reported the involvement of PFC in risk decision-making behaviors [20,21]. As an example, a related study empirically showed a higher brain activity for subjects with more consideration of gains than losses [20]. In addition, previous studies observed decreased and increased hemodynamic responses in individuals who were focused on losses (i.e., having negative attitudes toward risks) versus those who mostly considered gains (i.e., having positive attitudes toward risks), respectively [9]. Therefore, activation of the PFC can serve as a proxy of individuals' risk attitude: Those more focused on gains (i.e., incentives in Condition B) will have higher brain activation.

These correlation results also indicate that subjects who are often highly focused on positive outcomes in other risky activities (e.g., the thrill of driving while intoxicated outweighs the perceived risk of arrest) are those who are highly concentrated on gains rather than losses in the simulated electrical construction task. Here, response generalization theory may come into play, as this theory explains that individuals who tend to be involved in a targeted risky behavior can also be involved in non-targeted risky behaviors [22].

Problematically, underestimating the risk of a hazardous situation increases the likelihood of taking more risks [2,7], especially as individuals who have positive attitudes toward risks tend to adopt risky behaviors by assigning higher expected values to the outcome. In contrast, people with negative attitudes perceive lower benefits and higher negative outcomes when involved in risky actions [2,7]. Well-aligned with this discussion is our observed changes in brain activation—i.e., the

differences in oxy-Hb concentration from the normal condition (A) to the risky condition (B, with incentives). These values showed higher average values for the high-PCARE groups versus low-PCARE groups within each category.

Although the risk level was higher in Condition B due to the time pressure and productivity demand—which may increase the risk of potential losses (experiencing arc flash)—the presence of incentives (i.e., gains) caused subjects to concentrate merely on achieving the gains and correspondingly increase their risk-taking behavior by speeding up to complete the defined task faster to obtain the incentive. These participants, who were also grouped high in various CARE categories, experienced more arc flashes in Condition B. So, the associated gains (i.e., additional compensation as an incentive) increased their perceived benefits and caused them to overlook losses as they found more value in taking risks. Collectively, these findings regarding subjects' assessment of expected value (i.e., gain) and harm (i.e., electrical accident) provided empirical evidence regarding the contributing role of risk attitude in workers' unsafe behaviors and at-risk decisions.

6 Conclusion

By employing a mixed-reality environment and neuroimaging technology, this study empirically investigated participants' risk propensity in a simulated high-risk construction scenario when gains and losses were in place. Results indicate that oxy-Hb concentration captured by fNIRS sensors may serve as a proxy of participants' risk attitudes since this value positively correlates with the evaluated PCARE scores. The present study also shows that expected value signals (gain) in the prefrontal cortex are considerably increased among risk-seeking individuals, which indicates that the more participants focus on gains during a risky situation, the greater their brain activation will be. Further, as subjects perceived more benefits associated with a situation, they valued positive consequences (i.e., gains) over negative ones (i.e., losses), which stimulated them to engage in greater risk-taking behaviors.

Together, the findings argue that fNIRS signals are reliable to provide behavioral information regarding individuals' risk propensity, decision-making, and risk-taking behaviors. This study provides insights into identifying at-risk workers whose positive attitudes toward risky situations may put them at high risk of engaging in potentially dangerous activities on jobsites. Future studies may incorporate different physiological sensors (e.g., Electrodermal activity (EDA)) to see the correlation between physiological responses and fNIRS signals, in investigating individuals' risk attitude and risk-taking behaviors. Using such knowledge can help in

suggesting behavioral interventions that incorporate educational information regarding risk perception, as a modifiable construct, to counteract excessive risk-taking.

Acknowledgment

The authors would like to thank National Science Foundation (Award Number 2049711) and Electri International for supporting the research presented in this paper. We would like to extend our appreciation to Aditya Mane, Beyza Kiper, Makayla Simpson, Kaylee Dillard, and Doug Hermo who assisted in data collection. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

References

- [1] ESFI: Workplace Fatalities and Injuries 2003 - 2019, (n.d.). <https://www.esfi.org/resource/workplace-fatalities-and-injuries-2003-2019-791> (accessed June 29, 2021).
- [2] Hasanzadeh S., la Garza J.M.D, Geller E.S., How Sensation-Seeking Propensity Determines Individuals' Risk-Taking Behaviors: Implication of Risk Compensation in a Simulated Roofing Task, *Journal of Management in Engineering*. 36: 4020047, 2020. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000813](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000813).
- [3] Morshedi M.A., Kashani H., Assessment of vulnerability reduction policies: Integration of economics and cognitive models of decision-making, *Journal of Reliability Engineering and System Safety*. 217: 108057, 2022. <https://doi.org/10.1016/J.RESS.2021.108057>
- [4] Benthin A., Slovic P., Severson H., A psychometric study of adolescent risk perception, *Journal of Adolescence*. 16: 153–168, 1993. <https://doi.org/10.1006/JADO.1993.1014>.
- [5] Park H.-S., Thomas S.R., Tucker R.L., Benchmarking of Construction Productivity, *Journal of Construction Engineering and Management*. 131: 772–778, 2005. [https://doi.org/10.1061/\(asce\)0733-9364\(2005\)131:7\(772\)](https://doi.org/10.1061/(asce)0733-9364(2005)131:7(772)).
- [6] Ajzen I., The theory of planned behavior, *Organizational Behavior and Human Decision Processes*. 50: 179–211, 1991. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T).
- [7] Hasanzadeh S., la Garza J.M.D., How May Risk Tolerance, Cognitive Appraisal, and Outcome Expectancy Motivate Risk-Taking Behavior? The

- Implication of Risk Compensation through Multi-Sensor Mixed-Reality System, Construction Research Congress 2020: Safety, Workforce, and Education - Selected Papers from the Construction Research Congress 2020. 424–433, 2020. <https://doi.org/10.1061/9780784482872.046>.
- [8] Hu M., Shealy T., Application of Functional Near-Infrared Spectroscopy to Measure Engineering Decision-Making and Design Cognition: Literature Review and Synthesis of Methods, *Journal of Computing in Civil Engineering*. 33: 04019034, 2019. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000848](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000848).
- [9] Holper L., ten Brincke R.H.W., Wolf M., Murphy R.O., fNIRS derived hemodynamic signals and electrodermal responses in a sequential risk-taking task, *Brain Research*. 1557: 141–154, 2014. <https://doi.org/10.1016/J.BRAINRES.2014.02.013>.
- [10] Rao H., Korczykowski M., Pluta J., Hoang A., Detre J.A., Neural correlates of voluntary and involuntary risk-taking in the human brain: An fMRI Study of the Balloon Analog Risk Task (BART), *NeuroImage*. 42: 902–910, 2008. <https://doi.org/10.1016/j.neuroimage.2008.05.046>.
- [11] Rogers R.D., Ramnani N., Mackay C., Wilson J.L., Jezard P., Carter C.S., Smith S.M., Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition, *Biological Psychiatry*. 55: 594–602, 2004. <https://doi.org/10.1016/J.BIOPSYCH.2003.11.012>.
- [12] Breiter H.C., Aharon I., Kahneman D., Dale A., Shizgal P., Functional Imaging of Neural Responses to Expectancy and Experience of Monetary Gains and Losses, *Neuron*. 30: 619–639, 2001. [https://doi.org/10.1016/S0896-6273\(01\)00303-8](https://doi.org/10.1016/S0896-6273(01)00303-8).
- [13] Tyagi O., Hopko S., Kang J., Shi Y., Du J., Mehta R.K., Modeling Brain Dynamics During Virtual Reality-Based Emergency Response Learning Under Stress, *Human Factors*, 2021. 001872082110548. <https://doi.org/10.1177/00187208211054894>.
- [14] Shi Y., Zhu Y., Mehta R.K., Du J., A neurophysiological approach to assess training outcome under stress: A virtual reality experiment of industrial shutdown maintenance using Functional Near-Infrared Spectroscopy (fNIRS), *Advanced Engineering Informatics*. 46: 101153, 2020. <https://doi.org/10.1016/J.AEI.2020.101153>.
- [15] Pooladvand S., Kiper B., Mane A., Hasanzadeh S., Effect of time pressure and cognitive demand on line workers' risk-taking behaviors: Assessment of neuro-psychophysiological responses in a mixed-reality environment, *Construction Research Congress 2022*, American Society of Civil Engineers, Reston, VA, 2022: pp. 759–769. <https://doi.org/10.1061/9780784483985.077>.
- [16] Gholizadeh P., Onuchukwu I.S., Esmaeili B., Trends in Catastrophic Occupational Incidents among Electrical Contractors, 2007–2013, *International Journal of Environmental Research and Public Health* 2021, Vol. 18, Page 5126. 18: 5126, 2021. <https://doi.org/10.3390/IJERPH18105126>.
- [17] Gholizadeh P., Esmaeili B., Cost of occupational incidents for Electrical Contractors: Comparison using robust-factorial analysis of variance. *Journal of Construction Engineering and Management*. 146(7): 04020073, 2020. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001861](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001861).
- [18] Fromme K., Katz E.C., Rivet K., Outcome Expectancies and Risk-Taking Behavior, *Cognitive Therapy and Research* 1997 21:4. 21, 421–442, 1997. <https://doi.org/10.1023/A:1021932326716>.
- [19] Pooladvand S., Taghaddos H., Eslami A., Tak A.N., (Rick) Hermann U., Evaluating Mobile Crane Lift Operations Using an Interactive Virtual Reality System, *Journal of Construction Engineering and Management*. 147 (2021) pp. 04021154. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0002177](https://doi.org/10.1061/(ASCE)CO.1943-7862.0002177).
- [20] Tobler P.N., Christopoulos G.I., O'Doherty J.P., Dolan R.J., Schultz W., Risk-dependent reward value signal in human prefrontal cortex, *Proceedings of the National Academy of Sciences*. 106: 7185–7190, 2009. <https://doi.org/10.1073/PNAS.0809599106>.
- [21] Schonberg T., Fox C.R., Mumford J.A., Congdon E., Trepel C., Poldrack R.A., Decreasing ventromedial prefrontal cortex activity during sequential risk-taking: an fMRI investigation of the balloon analog risk task, *Frontiers in Neuroscience*. 6: 1–11, 2012. <https://doi.org/10.3389/FNINS.2012.00080>.
- [22] Ludwig T.D., Geller E.S., Improving the Driving Practices of Pizza Deliverers: Potential Moderating Effects of Age and Driving Record, *Journal of Applied Behavior Analysis*. 24: 31–44, 1991. <https://doi.org/10.1901/JABA.1991.24-31>.

Toward Personalized Safety Training: Automating the Classification of Construction Workers' Cognitive Failures

K. Lee^a, Y. Shinde^a, S. Hasanzadeh^{a*}, and B. Esmacili^b

^aLyles School of Civil Engineering and Construction Engineering & Management, Purdue University, USA

^bDept. of Civil, Environmental and Infrastructure Engineering, George Mason University, USA

*Corresponding author

E-mail: lee2490@purdue.edu, yshinde@purdue.edu, sogandm@purdue.edu, besmaeil@gmu.edu

Abstract –

Safety training has long been considered a promising method to enhance workers' hazard identification skills within construction sites. To improve the effectiveness of safety training, such varied features as a training environment, individuals' learning ability, and lesson personalization have been investigated. However, as records show workers still miss hazards even after receiving safety training, understanding the fundamental cognitive reasons for unrecognized hazards becomes a crucial step toward developing effective personalized safety training. This study used various 360° panoramas of construction scenarios to empirically examine 30 workers' visual search strategies and assess workers' hazard identification skills. Results suggest several cognitive limitations caused failures in hazard recognition, including attentional failure, inattention blindness, and low perceived risk. Based on these findings, this study proposes a personalized safety training framework to address such cognitive limitations to improve occupational safety in the construction industry.

Keywords –

Hazard identification; Construction safety; Cognitive failures; Personalized safety training; Attentional failure; Inattention blindness; Risk Perception

1 Introduction

Given over 1,000 recent fatal injuries in the construction industry in the U.S. [1], researchers have been trying to improve workers' hazard identification abilities to avoid injuries [2,3]. A promising approach to counteracting injuries is to provide effective safety training to enhance workers' hazard recognition performance [4]. Previous studies revealed that workers often missed hazards in their surrounding environment due to different cognitive limitations (e.g., failed

attention [5], flawed risk perception [6,7]). Thus, to properly identify hazardous conditions within a dynamic construction environment, workers need to appropriately detect hazards and perceive them as risks [3], and training programs should address failures affecting this skill set. However, safety training has neither comprehensively covered these various cognitive limitations nor proven capable of customizing training per the cognitive failures of individual workers.

This study uses eye-tracking technologies to identify the types of cognitive failures affecting construction workers' safety and thereby recommend opportunities for automating personalized safety training. This study contributes to the body of knowledge and practice by proposing an advanced personalized safety training framework that can automatically translate workers' subjective test results and objective psychophysiological responses into personalized training recommendations. The outcomes of this paper will lay the necessary foundations required to build tailored training regimens to improve construction worker safety.

2 Background

2.1 Assessing Construction Workers' Cognitive Limitations via Eye-tracking Technology

Identifying hazardous situations in dynamic construction environments is a complex cognitive process. Advanced sensing technologies (e.g., electroencephalograms, eye-tracking) have been actively utilized in several studies to evaluate human cognitive processes and safety-related behaviors under hazardous conditions [3,8]. Among these sensors, eye trackers have been widely used to assess workers' cognitive failures and low-hazard identification skills because eye-movement data represent the most direct manifestation of visual attention [9,10].

In a study conducted by Hasanzadeh and her colleagues [3], three fixation-related metrics (i.e.,

fixation count, dwell-time percentage, and run count) were utilized to predict workers' hazard identification skills. The results indicated that hazard recognition skills remarkably affect workers' visual scanning patterns. For instance, workers with higher hazard identification skills showed higher fixation counts and run counts, and lower dwell-time percentages on various hazard types. Accordingly, eye-tracking technology provides considerable opportunities for assessing workers' different attentional distributions and for predicting cognitive failures.

2.2 Personalizing Safety Training

Researchers have investigated various aspects of safety training, such as training format [4,11,12] and workers' learning ability [13]. Among these efforts, personalized training recently received attention as the next generation of safety training [13]. Compared to traditional safety training, personalized training aims to include the assessment of individuals' differences and their resulting decisions when exposed to assorted risks on a jobsite [14]. For example, Xu et al.'s study argued that workers' learning abilities during the safety training varied, which led the study to develop a learner model that could capture and evaluate individual workers' cognitive capabilities and learning abilities [13]. Further, some studies showed the feasibility of automatically capturing and analyzing workers' visual search patterns [15]. While many studies theoretically discussed the importance of developing personalized safety training, no studies empirically develop a training framework to

address these cognitive limitations.

To develop personalized safety training, it is essential to understand the reason for cognitive failures and select an appropriate training approach to counteract the problem. For instance, if someone has a poor visual search strategy, showing an expert's visual search path—which has been used among marines and radiologists to enhance visual search strategies—could function as a suitable training approach [16]. Incorporating such a design can address unrecognized hazards and promote the development of more effective personalized training.

3 Research Method

To identify the types of cognitive failures impacting construction workers' safety and thereby recommend opportunities for automating personalized safety training, this study conducted a hazard identification experiment presenting videos of a realistic construction environment to monitor subjects' visual behaviors. To create realistic scenarios able to capture the complexity and dynamics within a construction experiment, the design used several 360° video panoramas captured using an Insta360 OneX camera. The scenarios covered various construction activities (e.g., painting, erecting the structure, installing HVAC, and welding) and were recorded at commercial construction sites in Washington D.C. and northern Virginia to include different static and dynamic hazards of varying risk. Professional safety managers carefully reviewed all video scenarios in advance and identified hazards within each scenario.

For this study, the research team recruited thirty

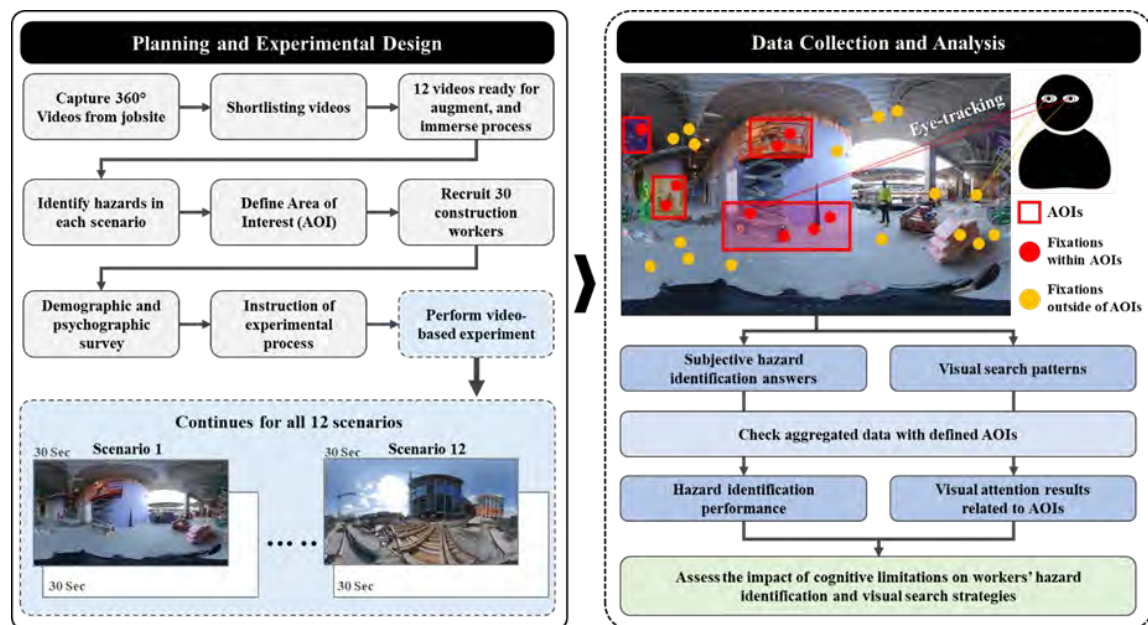


Figure 1. Research framework

experienced construction workers (29 males, 1 female; aged 34.5 ± 10.6 years) from jobsites to collect realistic behavioral data. The participants had, on average, 8.5 years of experience, and all had received multiple safety trainings before. Each participant performed a single 60-minute session, which was delivered via the HTC VIVE Pro Eye head-mounted display. While the participants were searching for hazards in each scenario, their visual scanning patterns were captured using eye-tracking sensors embedded in HMD. In total, workers were asked to view twelve scenarios for thirty seconds, and then report the types of hazards they recognized in each scenario (Figure 1).

To analyze the subjects' resulting eye-tracking data, the research team marked multiple areas of interest (AOIs) which were predefined by safety professionals. AOIs are the boundary range of the hazardous regions in the scenarios; in this paper, analysis focuses on two major hazard categories (fall, and struck-by), though the results reflect additional findings not detailed at this time. The research team mapped subjects' fixation points on static and dynamic AOIs using image processing algorithms. Those AOIs that did not receive fixations were deemed "attentional distribution failures" whereas those AOIs that had fixations were deemed either "recognized" or "risk-perception failures" based on whether subjects self-reported identifying the hazard. Additionally, spatial attention proportion is calculated when fixation points are within the AOIs boundaries to the total number of fixation points in the entire scene. Then, by coupling the eye-tracking data with workers' self-reported hazard identification results, the research team classified the cognitive reasons behind the unrecognized hazards. The contrasts between the empirical (eye-tracking) results and subjective (self-reporting) results were then analyzed to identify training opportunities.

4 Results and Findings

Figure 2 indicates the average cause-specific rate of hazard identification failures—e.g., those caused by failed attentional distribution or those caused by failed risk perception—for all hazards and the two main hazard categories detailed in this paper. Generally, more than half of hazards remained *unrecognized* even when the worker allocated considerable attentional resources to those hazardous areas, which indicates that the worker either experienced *inattentional blindness* or *did not perceive the risk* of hazards within the scene due to high-risk tolerance or lack of knowledge. On average, 43% of workers who failed to identify fall hazards illustrated inefficient visual search strategies and improper attentional allocation for fall hazards. Additionally, on average, 67% of struck-by hazards remained unrecognized because subjects failed to identify the

hazardous conditions as risks, even if they allocated sufficient attentional resources to those hazards. The remaining 33% of failed struck-by identifications were hazards completely missed by workers who did not properly distribute their visual attention to struck-by hazards. These results clearly illustrate that workers missed identifying hazards due to various cognitive failures and raise the necessity of various training approaches that rely on targeted problems.

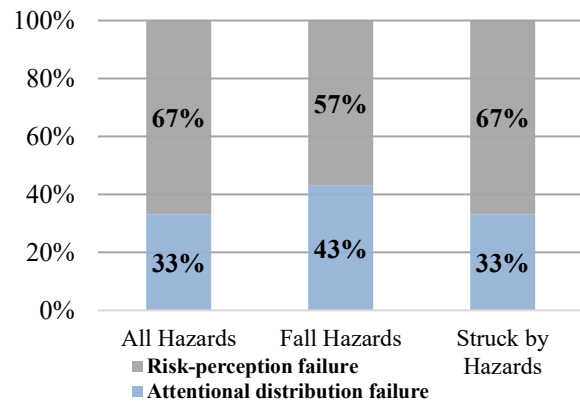


Figure 2. Average cause-specific rate of hazard identification failures

Figure 3 depicts a scenario that was selected for further investigations. In this scenario, two scissor lifts provide work platforms for workers installing wall panels and an HVAC system. In addition to the operators being at fall risk while working at height (Figure 3, *d* and *e*), workers on the ground were also at risk due to struck-by hazards from the elevated work platforms; workers on the ground would need to be aware of their surroundings and avoid working in close proximity to the lift or passing underneath it when it is raised (marked area, Figure 3, *c* and *f*) because they might get injured or killed by objects falling from the lifts or may be struck by the lift itself. At one point during the video, a worker entered the work zone (see AOI *a* in Figure 3) and passed underneath (marked areas) without checking the status of the lifts' position and without attending to the workers operating the lifts to avoid any potential struck-by hazards. In this scenario, there were also few workers performing a welding task without fire protection (Figure 3, *b*). Due to the spatial arrangement of the camera, fixations on this fire hazard (AOI *b*) overlapped with AOI *a* for a short period, a point we discuss below.

The subjective, self-reported hazard identification results show that 87% of participants (26 out of 30) failed to identify the dynamic hazard (Figure 3, AOI *a*). In these cases, the research team investigated how the subject's spatial attention was distributed over the scene—a factor in situational awareness theory [17]—to explore which



Figure 3. A representative example of a 360° video construction scenario with associated AOIs (Dynamic hazard: *a*, and Static hazard: *b*, *c*, *d*, *e*, *f*)

of three causal factors accounted for the failed identification (i.e., inappropriate attentional distribution, inattention blindness, and lack of safety knowledge/low perceived risk). Figure 4 shows the cumulated attentional allocation (the dots represent fixation points) for all 26 participants who missed the hazard and grouped based on the cognitive challenges observed. Note: these coordinate data were extracted for the period where AOI *a* was activated.

The results indicate that 52% of subjects (Figure 4a) did not appropriately allocate their attention across the scene to recognize hazards, and therefore they failed to identify hazards. Interestingly, only 2% of their spatial attentional resources were allocated to the dynamic AOI *a*, whereas most of their fixations were on other static hazards or environmental objects. Such *inappropriately distributed attentional resources* may be counteracted through training in situational awareness, and therefore represent a way the observed eye-tracking data could provide an opportunity for improved personalized training.

Furthermore, the results show that 20% of subjects failed to identify the potential struck-by hazard (AOI *a*) due to inattention blindness (Figure 4b). Overall, these participants allocated 32% of their attentional resources toward the area related to AOI *a*, where the worker passed underneath the two active lifts (Figure 4b), but the subjects failed to name this hazard, an indicator of *inattention blindness*. Such blindness may manifest when the cumulative attentional distribution map demonstrates that although a subject pays close attention to an AOI—and even brings attention back to it several times—the subject never “sees” the risky behavior and

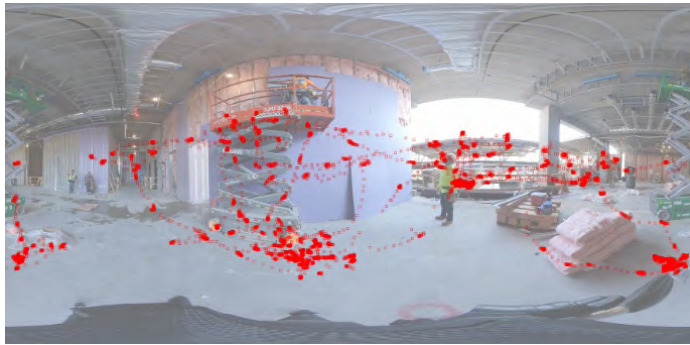
does not report the hazard in the follow-up oral report. Alternatively, workers may not “see” what they are directly looking at because they are attending to something else within the same environment (e.g., welding without fire protection). In either case, inattention blindness appears in the data when a subject fails to perceive a clearly visible stimulus (AOI *a*) located exactly where she/he is looking (fixating), and thereby represent an opportunity for automating personalized training.

Lastly, 28% of workers failed to identify AOI *a* despite a relatively efficient visual search strategy compared to the other two groups. Specifically, these subjects distributed 60% of their spatial attentional resources within the target boundary (Figure 4c), but they still failed to identify the hazards involved in this space. Such results indicate that this group of subjects may need a different training approach to target their *knowledge level* or *risk perception skills*.

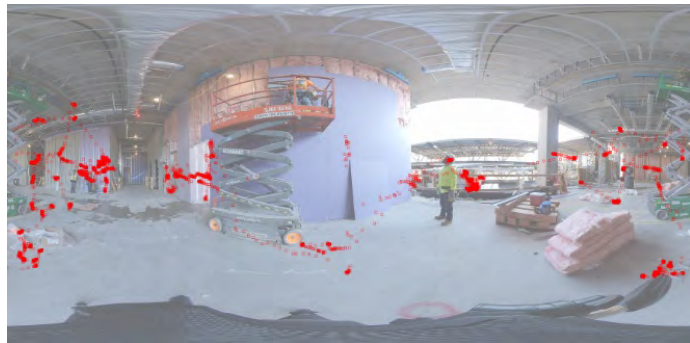
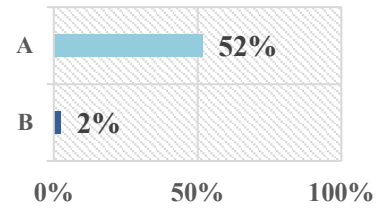
These three cognitive-failure based causes for failed hazard identification represent opportunities for personalizing safety training based on workers’ true limitations. Figure 4 contrasts the differences in eye-tracking data between the different groups’ behaviors, revealing an inroad for automating this personalization process to improve the safety levels at jobsite. In the next section, based on these findings, this study proposes the framework of personalized safety training.

4.1 Personalized Safety Training Framework

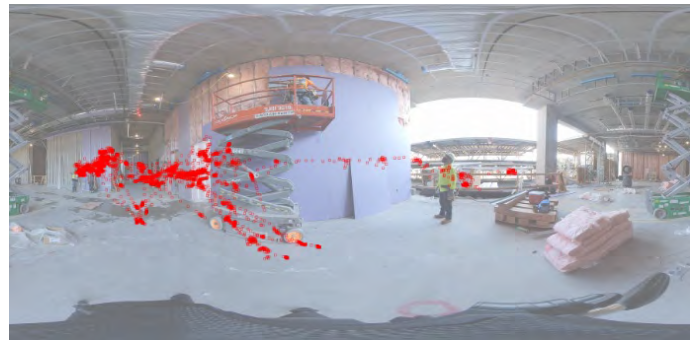
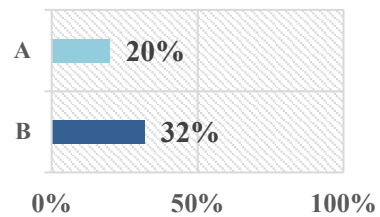
As illustrated in Figure 5, this study proposes an advanced personalized safety training framework



(a) Inappropriate attentional distribution



(b) Inattentive blindness



(c) Lack of safety knowledge or low perceived risk

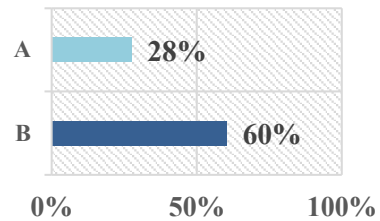


Figure 4. Cumulative attentional distribution for each discussed cognitive limitations A: Percent of workers who failed to identify hazard *a* due to associated causal reason, and B: Spatial attention proportion distributed to the related hazard “*a*” over total fixation counts across the scene

consisting of three main stages: primary setup, assessment, and customized training. The proposed training adopts various advanced technologies (e.g., 360° panoramas, eye-tracking, wearable sensors, and artificial intelligence) to present realistic hazardous scenarios, assess workers’ visual search strategies, identify their risk-perception state, and classify individual’s true cognitive challenges necessitating improvement. Such a training platform could be designed in two versions: (1) desktop and (2) virtual reality delivered via VR headset to provide an immersive education experience.

During the assessment step, workers will be immersed with 360° video and images and be asked to

scan the scene while their neuro-psychophysiological responses are being continuously collected to obtain information regarding the worker’s attentional distribution, risk perception, and decision dynamics. Then, the workers will be presented with a quiz to assess their hazard identification performance. Like eye-tracking data directly links to workers’ visual attention, physiological responses data (e.g., EDA, EEG, fNIRS) are highly connected to risk perception. Thus, by synchronizing their hazard identification performance with multiple data aggregated from the above-mentioned sensors, an automatic classification model will determine which training regimens the workers need to receive. For

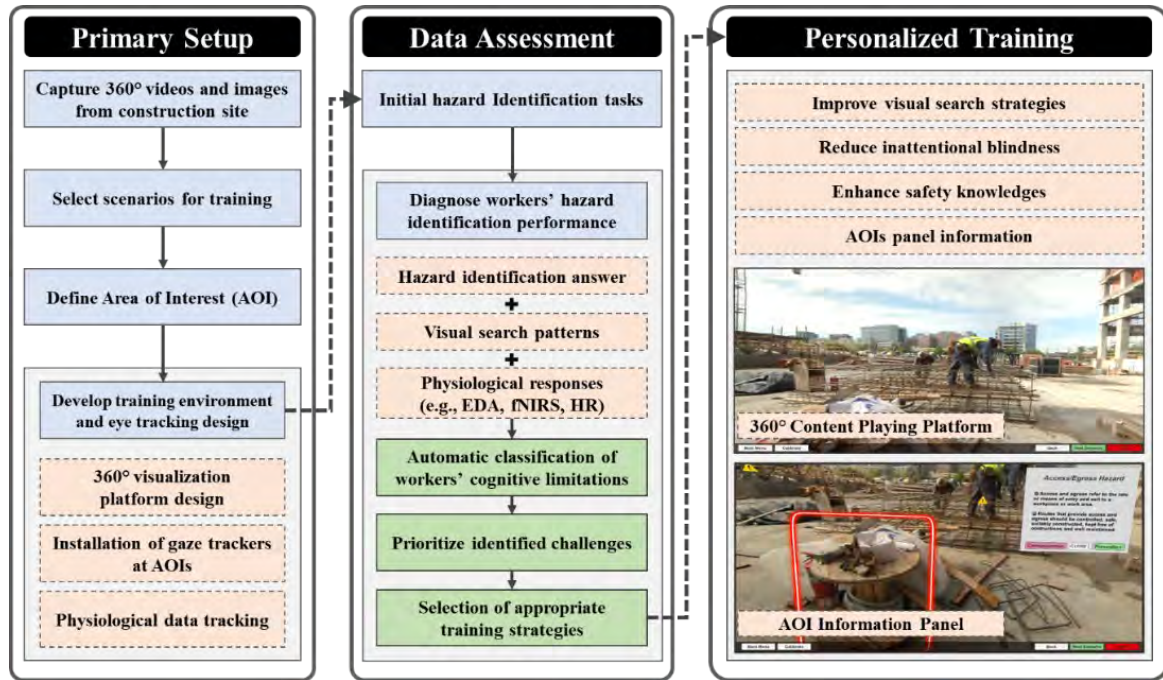


Figure 5. Proposed personalized training framework

example, the group with poor visual search abilities (those who present visual search strategies similar to Figure 4a) would be assigned training in which they would learn more about how to effectively scan the scene and allocate their limited attentional resources to hazardous areas. Another example is at-risk workers who need to receive more training regarding specific hazards they have missed. In that case, additional, user-friendly components (e.g., AOI info panels) would be embedded in the platform to highlight the overlooked hazards and provide auditory and visual information for workers regarding the description of the hazard, its consequences, and how the hazard could be prevented. Our research team is developing such a training program, and the results of this proposed framework are forthcoming.

5 Discussion

Construction environments are complex, dynamic, and rich in detail, whereas human perceptual and cognitive resources are limited. Therefore, workers may fall prey to various failures of awareness, leading to injury. The findings of this study indicate that this failure of hazard identification is caused by different cognitive limitations, each of which essentially requires divergent training strategies. While the importance of tailoring training to an individual's cognitive limitations is pivotal, no studies to date have empirically explored these cognitive challenges to propose personalized safety training.

In this study, the research team investigated workers' cognitive failures based on subjective hazard identification performance and objective physiological data, which we combine to propose a personalized training framework. Due to the demanding and dynamic nature of jobsites, some workers may not be able to remain situationally aware of their surroundings to identify hazards. These workers may have low modal hazard anticipation skills to predict whether and/or how a specific hazard might materialize at a particular time in the near future—as particularly evidenced in the discussed case of subjects missing a worker passing underneath two raised lifts. These subjects—all experienced construction workers—primarily need to be trained to improve their visual search strategy, distribute their visual attention properly across the surrounding environment, and make the best use of their limited attentional resources to identify hazards.

The improvement of visual search strategies is crucial in various industries (e.g., military, driving, and lifeguarding), and several training approaches have been utilized [16,18]. For example, a driving-related study suggests training regimens that show an expert's visual search pattern, including more consistent and systematic scan paths [18]. In addition, the marines have recommended providing expert feedback about an individual's search path to provide another effective training method [19]. Such training systems, if combined with this study's approach to diagnosing cognitive limitations, would feasibly provide excellent inroads to improved construction safety.

Although the inattention blindness concept is now well-established in cognitive psychology and can be prevented through education and training, it has rarely been discussed in the construction safety setting. Studies addressing inattention blindness [6] showed that workers may allocate their attentional resources to some areas within the scene without perceiving the scene, a factor that may put these workers at a very high risk of being involved in an accident. This phenomenon has its roots in a selective looking paradigm presented by Neisser (2019) [20] and may also have roots in a tendency among individuals in high-risk environments to miss a second target after detecting the first target, a factor known as subsequent search misses [21]. Previous literature also indicated that certified training and frequent exposure to accidents showed positive impacts on reducing inattention blindness [6]. Therefore, although inattention blindness is a natural human cognitive limitation, an educational training method that allows workers to recognize their cognitive limitations and try to control them must be developed.

In our proposed training, workers will be shown the hazards they have missed despite looking at them, which will provide a first step toward enhancing these workers' awareness about this phenomenon. Then, they will receive training on how to allocate their attention throughout the scene, remain mindful, and avoid premature search termination.

Lastly, our results show workers may have appropriately distributed their visual attention when the dynamic hazard was activated, but they did not perceive the situation as a risky condition due to their limited safety knowledge or inordinately higher risk tolerance. Therefore, this group may require more safety knowledge-based training to understand why the condition is considered hazardous as well as the condition's risk level and consequences.

While our findings provide a unique perspective on workers' cognitive limitations and require personalized interventions, several limitations need to be noted. Due to the page and space limit, the current paper classified the different cognitive limitations by only analyzing eye-tracking data and hazard identification results. Therefore, future research needs to explore other psychophysiological responses (HR, EDA, and brain activity) to have better classification results. Second, future studies may conduct a pre-post experiment to examine the effectiveness of proposed personalized training.

6 Conclusion

Construction jobsites are complex and dynamic environments requiring constant attention, so the ability to recognize static, dynamic and emerging hazards in a

surrounding environment is highly associated with worker safety. The results of this study suggest workers' hazard identification failures were predominantly affected by workers' various cognitive limitations (e.g., attentional failure, inattention blindness)—a factor discernable in the subjects' empirically identified eye-movement behaviors. Aligned with this finding, the study proposes a framework for advanced personalized training. Such recommended training will adopt multiple sensing and visualization technologies to automate the individualized assessment of workers' true cognitive limitations and thereby select optimal training methods. The results of this paper are expected to motivate more efforts into creating a highly effective personalized training platform and ultimately improve workers' hazard recognition abilities, thereby decreasing the number of injuries and fatalities in construction.

Acknowledgments

The National Science Foundation is thanked for supporting the research reported in this paper (1824238 and 2049711). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation and the supporting construction company. The authors also would like to thank the workers and professional safety managers who participated in and supported this study.

References

- [1] U.S. Bureau of Labor Statics (BLS). "Census of Fatal Occupational Injuries (CFOI)." On-line: <https://www.bls.gov/iif/oshcfoi1.htm#2019>, Accessed: 03/06/2021.
- [2] Choudhry, R. M., and Fang, D. "Why operatives engage in unsafe work behavior: Investigating factors on construction sites." *Safety science*, 46(4):566–584, 2008.
- [3] Hasanzadeh, S., Esmacili, B., and Dodd, M. D. "Impact of Construction Workers' Hazard Identification Skills on Their Visual Attention." *Journal of Construction Engineering and Management*, 143(10):04017070, 2017.
- [4] Eiris, R., Gheisari, M., and Esmacili, B. "Pars: Using augmented 360-degree panoramas of reality for construction safety training." *International Journal of Environmental Research and Public Health*, 15(11), 2018.
- [5] Lee, K., Hasanzadeh, S., and Esmacili, B. "Spatial Exposure to Dynamic Safety

- Hazards in Construction Sites Through 360-Degree Augmented Panoramas: Ecological Validity in Safety” In *Processing of Construction Research Congress*, Washington, USA, 2022.
- [6] Park, S. J., Park, C. Y., Lee, C., Han, S. H., Yun, S., and Lee, D. E. “Exploring inattention blindness in failure of safety risk perception: Focusing on safety knowledge in construction industry.” *Safety Science*, 145(July 2021), 105518, 2022.
- [7] Pooladvand, S., Kiper, B., Mane, A., and Hasanzadeh, S. “Effect of Time Pressure and Cognitive Demand on Line Workers’ Risk-Taking Behaviors: Assessment of Neuro-Psychophysiological Responses in a Mixed-Reality Environment” In *Processing of Construction Research Congress*, Washington, USA, 2022.
- [8] Jeon, J., and Cai, H. “Classification of construction hazard-related perceptions using: Wearable electroencephalogram and virtual reality.” *Automation in Construction*, 132, 103975, 2021.
- [9] Hasanzadeh, S., Esmaeili, B., and Dodd, M. D. “Measuring the Impacts of Safety Knowledge on Construction Workers’ Attentional Allocation and Hazard Detection Using Remote Eye-Tracking Technology.” *Journal of Management in Engineering*, 33(5), 04017024, 2017.
- [10] Solomon, T., Hasanzadeh, S., Esmaeili, B., and Dodd, M. D. “Impact of Change Blindness on Worker Hazard Identification at Jobsites.” *Journal of Management in Engineering*, 37(4):04021021, 2021.
- [11] Bükür, S., Wolf, M., Golovina, O., and Teizer, J. “Using Field of View and Eye Tracking for Feedback Generation in an Augmented Virtuality Safety Training.” In *Proceedings of Construction Research Congress*, Reston, USA, 2020.
- [12] Eiris, R., Jain, E., Gheisari, M., and Wehle, A. “Online Hazard Recognition Training: Comparative Case Study of Static Images, Cinemagraphs, and Videos.” *Journal of Construction Engineering and Management*, 147(8):04021082, 2021.
- [13] Xu, S., Zhang, M., and Hou, L. “Formulating a learner model for evaluating construction workers’ learning ability during safety training.” *Safety Science*, 116(August 2018), 97–107, 2019.
- [14] Lee, K., Hasanzadeh, S., and Esmaeili, B. “Assessing Hazard Anticipation in Dynamic Construction Environments Using Multimodal 360-Degree Panorama Videos.” *Journal of Management in Engineering*, 2022.
- [15] Jeelani, I., Han, K., and Albert, A. “Automating and scaling personalized safety training using eye-tracking data.” *Automation in Construction*, 93(October 2017), 63–77, 2018.
- [16] Kramer, M. R., Porfido, C. L., and Mitroff, S. R. “Evaluation of strategies to train visual search performance in professional populations.” *Current Opinion in Psychology*, 29, 113–118, 2019.
- [17] Hasanzadeh, S., Esmaeili, B., and Dodd, M. D. “Examining the Relationship between Construction Workers’ Visual Attention and Situation Awareness under Fall and Tripping Hazard Conditions: Using Mobile Eye Tracking.” *Journal of Construction Engineering and Management*, 144(7):04018060, 2018.
- [18] Leff, D. R., James, D. R. C., Orihuela-Espina, F., Kwok, K.-W., Sun, L. W., Mylonas, G., Athanasiou, T., Darzi, A. W., and Yang, G.-Z. “The impact of expert visual guidance on trainee visual search strategy, visual attention and motor skills.” *Frontiers in human neuroscience*, 9, 526, 2015.
- [19] Carroll, M., Kokini, C., and Moss, J. “Training effectiveness of eye tracking-based feedback at improving visual search skills.” *International Journal of Learning Technology*, 8(2):147–168, 2013.
- [20] Neisser, U. “The control of information pickup in selective looking.” *Perception and its development*, Psychology Press, 201–219, 2019.
- [21] Cheng, P. X., and Rich, A. N. “More is better: Relative prevalence of multiple targets affects search accuracy.” *Journal of vision*, 18(4), 2, 2018.

iVisit-Collaborate: Online Multiuser Virtual Site Visits Using 360-Degree Panoramas and Virtual Humans

R. Eiris^a and M. Gheisari^b

^aDepartment of Civil, Environmental, and Geospatial Engineering, Michigan Technological University, USA

^bRinker School of Construction Management, University of Florida, USA

E-mail: reiris@mtu.edu, masoud@ufl.edu

Abstract –

Construction professionals need to understand collaboration within the context of jobsites. Consequently, many construction programs utilize learning activities to enable student learning of collaboration. One of those key learning activities used by construction programs is site visits. Although site visits are greatly beneficial for students, these present several challenges to be completed. This study centers on developing and testing a multiuser online system to support student learning of collaboration within the context of site visits. A case study virtual site visit was completed with 14 student dyads to understand collaboration in terms of teamwork perception, spatial visualization, and knowledge retention. It was found students perceived iVisit-Collaborate as a medium that supported effective teamwork with scores greater than five on a 7-point Likert scale. Furthermore, it was noticed the average student score 42% in spatial visualization and 64% in knowledge retention. However, it was observed that spatial visualization was not correlated to knowledge retention.

Keywords –

Site Visits; Collaboration; 360-degree Panoramas; Virtual Reality; Construction Education

1 Introduction

Construction professionals are required to collaborate with diverse stakeholders to complete any project [1]. Typically, construction educational programs utilize classroom exercises to enable students to practice collaboration [e.g., 2,3]. Although classroom exercises have been shown to be effective for teaching collaboration in-class, there are existing limitations linked to the lack of exposure to the spatiotemporal contexts where construction operations occur [4,5,6]. Learners require the use of spatial, temporal, or social contextual information (e.g., a certain space, changes with time, and interactions with peers) associated with a

specific site to learn collaboration in the context of the construction disciplines. Consequently, site visits are a common method used within construction curricula to offer opportunities for learning collaboration within the context of construction jobsites [7]. However, site visits have severe limitations to be completed within the traditional construction curricula. Standardized curricula, limited financial resources, and strict class schedules introduce challenges for an educational organization, program managers, and instructors to integrate this type of learning methodologies [5]. These challenges that exist in face-to-face classroom instruction become intensified when remote online learning delivery methods. Learners that are geographically dispersed pose blockades for learning collaboration – impossible to reach locations disconnect the student from the context and cause reductions in satisfaction, motivation, and overall learning [8]. Remote learning challenges have become prevalent during the COVID-19, where students were not allowed to perform face-to-face activities for safety reasons. These deficiencies in online learning result in student difficulties to learn collaboration common in today's construction professional workforce.

To address the issues associated with real-world site visits and remote instruction, virtual environments have been employed to create educational spaces for learning collaboration. Virtual environments are defined as the 3D rendered representation of construction jobsites that enables students to explore the virtual locations and control digital objects [9]. As part of these simulated construction sites, often construction professionals are represented as virtual humans (embodied depictions of experts that contain digital knowledge) [10]. These digital environments that contain expert virtual humans have been utilized in educational construction applications such as work coordination [11], building design [12], safety awareness [13], and scheduling and planning [14]. Within these educational applications, researchers have found that virtual environments offer opportunities for students to learn collaboration similarly to real-world settings. Virtual environments that utilize virtual humans create efficient communications channels

between students [14], enhance student feeling of presence [15], and student knowledge exchange [16]. However, effective collaboration within these digital settings is constrained by the realism offered to portray construction jobsites. Researchers point to the lack of realism as a major limitation for current virtual environments, as students experience a reduced feeling of being present at the construction site [16,17]. Existing 3D modeled digital environments require high computational power and large time commitments to create close-to-reality simulations of the construction site [18]. Furthermore, it has been found that students that learn in unrealistic simulations do not perform with the same proficiency as they do in the real world [19]. An increasingly utilized reality-capturing technique to address the realism constraints of virtual environments is 360-degree panoramas. These easy-to-capture illustrations of reality offer an unobstructed view of a location with a high sense of “being at the present” produced by the high sense of realism [20].

This study aims to address the current limitations for virtual environments by creating a multiuser virtual site visit system – iVisit-Collaborate – using highly realistic 360-degree panoramas and interactive virtual humans. iVisit-Collaborate offers opportunities for instructors to perform virtual site visits to enable students’ learning of collaboration within the spatiotemporal contexts of construction jobsites while removing the inherent challenges in these types of location-dependent interventions. A case study virtual site visit was completed by student’ dyads using iVisit-Collaborate within a developed online learning system. Student teamwork perceptions, spatial visualization, and knowledge retention were collected and analyzed to understand how iVisit-Collaborate supported the learning of collaboration within online site visits. The contribution of this study centers on better understanding how multiuser systems can be used to learn contextualized collaboration within the educational constraints of classroom instruction.

2 Background

Researchers have leveraged the realism embedded into 360-degree panoramas to represent real-world jobsites across multiple educational applications. For example, Gheisari et al. [21] employed 360-degree panorama technologies to demonstrate free-body diagrams of structural elements within building structures. Students using the 360-degree panoramas could understand the building structure without having to visit the construction jobsite. In another study, Eiris et al. [22] developed a system for students to learn hazard identification using 360-degree panoramas. Within this system, students used augmented 360-degree images

with text, signifiers (e.g., arrows, circles, lines), and audio to learn safety contents. In a follow-up study, Eiris et al. [23] found that the hazard identification index for trainees that learned about multiple fall hazards (e.g., unprotected edges, floor openings, ladders) was 52% using the system. More recently, 360-degree panoramas have been used to offer virtual site visits to students. Eiris et al. [24] created a platform to offer students replicable, interactive site visit experiences guided by a virtual human. The platform was further developed to enable problem-solving opportunities within the context of the construction jobsite for students in-class [24]. The researchers found that using 360-degree panoramas for site visits significantly increased the student development of problem-solving skills by allowing direct observation of the spatiotemporal context of the jobsite. In another recent study, Eiris et al. [25] evaluated a multiuser 360-degree panorama-based environment to understand collaborative problem-solving behaviors in students. It was found that collaborative and problem-solving behavior occur sequentially and that discussion engagement was critical for collaboration. Although some studies have explored the use of 360-degree panoramas as a tool to offer realistic site visits for problem-solving and collaboration [e.g., 7, 24, 25], many aspects of collaboration within these digital experiences remain largely unexplored.

3 Research Goals and Point of Departure

The goal of this study is to understand how iVisit-Collaborate supported the learning of student collaboration. This study builds on the system development and the outcomes of prior studies [7,24,25] that have explored virtual site visits for student learning of problem-solving and collaboration. The point of departure of this study is to study how students perceived their teamwork performance while using such systems. Moreover, the relationship between student knowledge retention and spatial visualization during collaborative activities was also explored. The following subsections of this study describe the development of iVisit-Collaborate, a case study where the system was tested, and the results as the outcome of the student collaborative learning activities.

4 iVisit-Collaborate System Development

The iVisit-Collaborate system was developed using the Unity® game engine and the Photon® networking engine. These game engine software support the creation of multiuser site visits for students to collaborate online. Although the Unity® game engine supports multiple platforms for implementing software (e.g., head-mounted displays, mobile devices, desktop computers),

this study selected desktop-based computers to enable access to students to the systems during the COVID-19 pandemic. Figure 1 shows the three main components included in iVisit-Collaborate. The following subsections contain a description of those components.



Figure 1. iVisit-Collaborate Platform Components

1. **Spatiotemporal Contexts:** Construction spatiotemporal contexts were used to demonstrate ideas and concepts to students in the virtual environment. The creation process of these spatiotemporal contexts required the use of 360-degree panoramas. The 360-degree panoramas were captured from real-world construction sites using commercially available 360-degree cameras (e.g., Insta 360 One X2, Ricoh Theta V) and their associated software. Using these 360-degree panoramas from real-world construction sites as background, text, objects, sounds, or signifiers were used as augmentations to clearly demonstrate information to students within the Unity® game engine.
2. **Virtual Humans:** Virtual humans were used to provide students an expert guide within the augmented 360-degree panorama spatiotemporal contexts through the use of audio narrative descriptions. There were three major components of the virtual human: a 3D digital representation, animations, audio voicings. The 3D digital representation was created using software such as Adobe's Fuse CC® or Unreal's Metahuman Creator. These software packages allowed to create a geometrical representation that embodied real humans in the digital space and animations to match the human reaction to speech or communication (e.g., facial expressions, body movements). Voice recordings were created using IBM Watson®'s text-to-speech tools for the audio voicings. This process produced audio files that contained the narrations required from the expert construction professionals. These three components were combined into the Unity® game engine to make a cohesive representation of humans as a construction professional tour leader within the virtual site visit.

3. **Online Multiuser Interaction Affordances:** Interaction affordances were developed for the iVisit-Collaborate desktop-based online application. Within the online multiuser platform, students interacted using a keyboard-and-mouse interface with each other, the spatiotemporal contexts, and the virtual humans. Color-coded laser pointers affordances were used for student communication, enabling each student to point at locations, objects, or augmentations within the 360-degree images. In addition to the pointers, two interactable objects were offered to the students – hotspots and clipboards. Hotspots served as a hub for activating augmentations and directing the attention of the students. Upon clicking a hotspot, the virtual human walked towards it and provided descriptions that matched the augmentations superimposed to the 360-degree panoramas. Clipboards served as a support affordance for student information management, keeping track of the learning objectives for the specific location, showing a transcript of the audios narrated by the virtual human, and displaying a map of the site with the locations of the 360-degree panoramas.

5 Case Study – Masonry and Wood Multiuser Construction Virtual Site Visits

To evaluate how the created iVisit-Collaborate system supported the learning of student collaboration, a case study was created for a Construction Techniques (BCN3224) class at the University of Florida. This class teaches students the principles of construction methods and techniques. The focus of this class is the basic understanding of vertical construction processes that includes topics such as masonry construction, wooden platform frame construction, cast-in-place and pre-cast concrete construction, and steel erection. Site visits to active jobsites are often used within this class to deliver active hands-on activities, facilitating the learning of spatiotemporal-dependent topics. The visits for the class are planned on a semester-by-semester basis, depending on location availability. However, due to the COVID-19 pandemic, these site visits were canceled. iVisit-Collaborate enabled the delivery of these site visits using an online multiuser format.

From the topics covered in this class, two site visits were selected by balancing topic simplicity, ease of visualization, and recommendations from the class instructor. The two topics selected were masonry and wood construction. For the masonry site visit, a masonry fire station construction site was used (Figure 2-a). Within this masonry site visit, students obtained an understanding of what are clay and concrete masonry

units and the techniques required to build and reinforce masonry walls. The virtual human within the masonry site visit offered expert knowledge regarding the construction materials (e.g., nominal/actual sizing, grout composition) and the means and methods of construction (e.g., joint reinforcement, expansion joints). For the wood site visit, a commercial multifamily residential platform frame building was used. Within this wood site visit, students observed the materials installed on-site (e.g., lumber, plywood, OBS panels) and the techniques to construct the platform frame structure. Similarly, the virtual human served as an expert construction professional that described the process of building floors, walls, and roofs structures within the site.

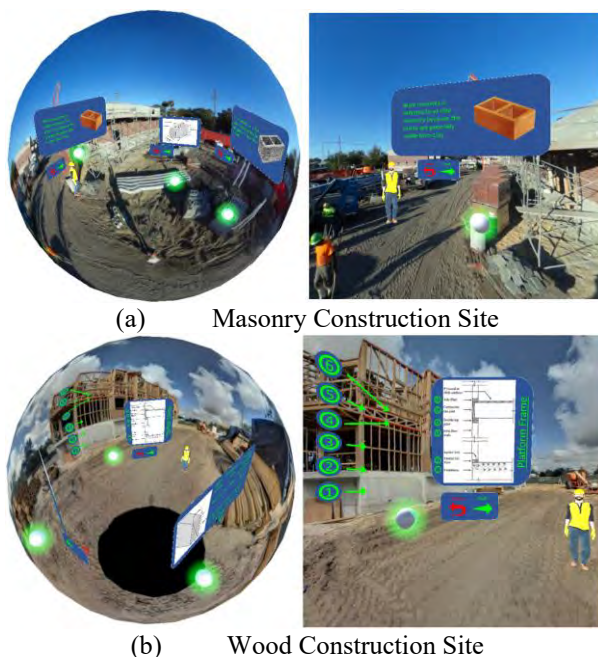


Figure 2. Case Study Construction Virtual Site Visits

Data collection: This study was conducted remotely, using a set of online Zoom® videoconference meetings throughout three weeks (under IRB201903043). Students completed the virtual site visits in pairs or dyads. Upon completing the iVisit-Collaborate platform experience, the students individually completed a set of four surveys using Qualtrics®. The following paragraphs provide the details for each of these surveys.

Demographics: This metric aims to understand the student background information, including age, sex, years in the construction program, experience in construction, and familiarity with virtual reality, 360-degree panoramas, construction materials, and construction techniques. A 7-question survey was used to collect this metric.

Teamwork Experience: This metric aims to understand the students' perspective regarding team effectiveness during the iVisit-Collaborate virtual site

visit. A 40-question survey was used to collect this metric. The survey was adapted from the "Team Learning Beliefs and Behaviors" validated survey, developed by Van de Bossche et al. [26] to understand team learning for successful collaborations. This survey contains 28-question items to explore the social process of knowledge building within the iVisit-Collaborate. The Likert-scale questions use 7-points granularity (1-Strongly Disagree to 7-Strongly Agree) to evaluate the students' level of agreement with statements that relate collaborative learning, individual beliefs, and interpersonal contexts as part of the fundamental knowledge building process for effective teamwork.

Spatial Visualization: This metric aims to measure the student's ability to visualize objects from different perspectives in three-dimensional space using their imagination [27]. Within the context of iVisit-Collaborate, this metric measures the ability of students to envision objects on the 360-degree panorama virtual environments. For this study, Vandenberg & Kuse's [28] Mental Rotation Test was used. The Mental Rotation Test requires students to select the two correct rotations of a depicted object from four possible responses. Positive credits are assigned to correct answers, and negative credits are assigned to incorrect answers, reducing the guessing effect from the participants. A total of twenty mental rotations were done by the students. As a result of this metric, a spatial visualization score was obtained that represented the ability of students to perform mental rotations. The mental rotation score was computed individually for each student, ranging from 100% (no errors and all twenty rotations done correctly) to 0% (all twenty rotations done incorrectly or multiple incorrect rotation done that reduced the score to zero).

Knowledge Retention: This metric aims to measure the student's obtained knowledge from the iVisit-Collaborate virtual site visit. A 10-question test was used to collect this metric. The test mimics the type and scope of the questions contained in a typical class quiz for "BCN3224 – Construction Techniques". Previous quizzes utilized in this class were reviewed and adapted to develop this data collection tool. The scope of test contents was adjusted to match the information delivered through the iVisit-Collaborate site visit. From the ten questions contained in the knowledge retention test, five related to the masonry module and five related to the wood module. For each topic, one concept recall question, fill-in-the-blank question, one multiple section question, and two open-ended questions were completed by the students. The questions were graded, and partial fractional credits were awarded for the fill-in-the-blank and open-ended questions when partially correct answers were provided by the students. As a result of this metric, a total knowledge retention score across the masonry and wood was obtained for each individual

student. The knowledge retention score ranged from 100% (all ten question correctly responded) to 0% (zero questions correctly responded).

5.1 Results and Discussion

In this study, a total of 14 dyads of students completed the iVisit-Collaborate virtual site visits. Students had an average age of 21 years (standard deviation: 0.7), with a high percentage of being male (82%). Over 50% of the students had low or no experience in the construction industry. Overall, a large proportion of students had no to low experience with virtual reality (no experience: 7%; low experience: 50%) and 360-degree panoramas (no experience: 4%; low experience: 60%). Contrary, most students reported having medium to high experience with construction materials (medium experience: 29%; high experience: 50%) and construction techniques (medium experience: 33%; high experience: 46%).

5.1.1 Teamwork Perceptions

The teamwork survey was analysed to understand how students perceived team effectiveness during the iVisit-Collaborate virtual site visit. The scores obtained for each dimension of this survey are illustrated in Table 1. It can be observed that all scores were very positive, with overall medians around six (Agree in the Liker-scale response). In the teamwork survey, task cohesion statements were negative, which the students responded with consistently low scores. These overall patterns indicate that the students were pleased with their teamwork performance within the iVisit-Collaborate platform. Some students supported these scores in the open-ended questions by stating comments such as “iVisit-Collaborate was overall a good learning tool and I enjoyed using it” and “Communication was great. No one had any problems and we worked well as a team together”.

Figure 1. Teamwork Student Perceptions

Collaboration Variable	Mean	SD
Interdependence	5.8	0.47
Social Cohesion	6.1	0.68
Task Cohesion	2.0	0.65
Psychological Safety	5.6	0.54
Group Potency	5.8	0.53
Construction	6.0	0.45
Co-Construction	6.0	0.41
Constructive Conflict	5.6	0.79
Shared Cognition	6.0	0.62
Team Effectiveness	6.2	0.55

5.1.2 Spatial Visualization and Knowledge Retention

The spatial visualization and knowledge retention

metrics were analysed to investigate student performance during the iVisit-Collaborate virtual site visit. Table 2 illustrates the descriptive statistics for these tests. Spatial visualization scores indicate the student ability for performing a mental rotation, which has been associated to a better understanding of serial operations in three-dimensional spaces. It was found that students had an average spatial visualization score of 41.8% (Standard Deviation = 59.3%). These results indicate that student scores had large variances how they would be able to understand the three-dimensional space of iVisit, potentially causing differences in the student dyads' understanding of the contents during the site visit. Furthermore, the knowledge retention score averaged 63.5% (SD = 13.4%). Knowledge retention scores describe student recall of concepts from the exposure to the site conditions using the iVisit platform. These obtained results indicate that students answered correctly over half of the ten test questions for masonry and wood after experiencing the site visits. Furthermore, a correlation analysis using a Spearman Rank Correlation Coefficient as described by Spearman [29] was performed between spatial visualization score and knowledge retention score. However, no significant correlation was found between spatial visualization and knowledge relation scores. The lack of correlation found suggests that ability to visually understand the three-dimensional space is independent from the knowledge that can be recalled from from the collaborative virtual site visits.

Table 2. Spatial Visualization and Knowledge Retention Scores

Variable	Mean	SD
Spatial Visualization Score	41.8%	59.3%
Knowledge Retention Score	63.5%	13.4%

6 Conclusion, Limitations, and Future Research

This study aimed to understand the use of virtual site visits such as iVisit-Collaborate to support the learning of student collaboration. A multiuser online system was developed and a set of virtual site visits were created (masonry construction and wood construction) to test the system. A total of 14 student dyads participated in the virtual site visit, where data was collected in terms of teamwork perception, spatial visualization, and knowledge retention. It was found students perceived iVisit-Collaborate as a medium that supports effective teamwork for virtual site visits, as all scores provided were greater than five on a 7-point Likert scale. Furthermore, it was found that spatial visualization scores greatly varied across the students and that the average student scored 63% in knowledge retention.

However, it was found that spatial visualization was not correlated to knowledge retention.

This study had three major limitations, the sample size, the limited topics covered in the system, and the remote nature of the data collection. Only 14 student dyads were collected, which limits the generalizability of the results shown in this study. Additionally, only masonry and wood construction sites were virtually visited in this study. Other types of construction (e.g., cast-in-place concrete, steel) might produce different results from the ones reported here. Moreover, the relationship between the level of realism and the variables explored in this study (teamwork attitudes, spatial visualization, knowledge retention) should be further investigated. Finally, data collection was conducted online, which limited the experimental control that the research team had on the students during the virtual site visit. Potential distractions during the remote sessions might have increased the variance in the results reported. Future studies should increase the sample size and cover a broader range of topics to better understand the usage of iVisit-Collaborate. Furthermore, an experimental setting with a control or baseline condition should be completed to understand how iVisit-Collaborate compares to traditional interventions.

References

- [1] Harty, C. (2005). Innovation in construction: A sociology of technology approach. *Building Research & Information*, 33(6), 512–522.
- [2] Boeykens, S., Somer, P. De, Klein, R., & Saey, R. (2013). Experiencing BIM collaboration in education. In *Proceedings of the 31st international conference on education and research in computer aided architectural design in europe (eCAADe 2013: Computation and performance)*, delft, Netherlands, 18-20th september (pp. 505–513).
- [3] Korkmaz, S. (2012). Case-based and collaborative-learning techniques to teach delivery of sustainable buildings. *Journal of Professional Issues in Engineering Education and Practice*, 138(2), 139–144.
- [4] Forsythe, P. (2009). The construction game – using physical model making to simulate realism in construction education. *Journal for Education in the Built Environment*, 4(1), 57–74.
- [5] Eiris, R., & Gheisari, M. (2017-a). Site visit application in construction education: A descriptive study of faculty members. *International Journal of Construction Education and Research*, 15, 2019(2), 83–99.
- [6] Mutis, I. (2018). Spatial-temporal cognitive ability: Coupling representations to situations and contexts for coordinating activities in the construction project environment. In *Transforming engineering education: Innovative computer-mediated learning technologies* (pp. 5–25). American Society of Civil Engineers (Chapter 2).
- [7] Eiris, R., Wen, J., & Gheisari, M. (2021-a). iVisit-Collaborate: Collaborative problem-solving in multiuser 360-degree panoramic site visits. *Computers & Education*, 104365.
- [8] Richardson, J. C. (2001). Examining social presence in online courses in relation to students' perceived learning and satisfaction. State University of New York at Albany.
- [9] Wen, J., & Gheisari, M. (2020-a). "Using virtual reality to facilitate communication in the AEC domain: A systematic review." *construction Innovation*. Emerald Publishing Limited.
- [10] Eiris, R., & Gheisari, M. (2017-b). Research trends of virtual human applications in architecture, engineering and construction. *Journal of Information Technology in Construction*, 22(9), 168–184.
- [11] Woksepp, S., & Olofsson, T. (2006). Using virtual reality in a large-scale industry project. *Journal of Information Technology in Construction*, 11, 627–640.
- [12] Van Nederveen, S. (2007). Collaborative design in second life. In *Second international conference world of construction project management 2007*. The Netherlands: TU Delft.
- [13] Guo, H., Li, H., Chan, G., & Skitmore, M. (2012). Using game technologies to improve the safety of construction plant operations. *Journal of Accident Analysis and Prevention*, 48, 204–213. <https://doi.org/10.1016/j.aap.2011.06.002>
- [14] Anderson, A., & Dossick, C. S. (2014). Avatar-model interaction in virtual worlds improves distributed team collaboration through issue discovery. In *Computing in Civil and Building Engineering (2014)* (pp. 793-800).
- [15] Wu, T. H., Wu, F., Liang, C. J., Li, Y. F., Tseng, C. M., & Kang, S. C. (2019). A virtual reality tool for training in global engineering collaboration. *Universal Access in the Information Society*, 18(2), 243-255.
- [16] Taylor, J. E., Alin, P., Comu, S., Dossick, C. S., Hartmann, T., Mahalingam, A., & Mohammadi, N. (2018). CyberGRID: A virtual workspace for architecture, engineering, and construction. In *Transforming engineering education: Innovative computer-mediated learning technologies* (pp. 291–321). American Society of Civil Engineers (Chapter 10).
- [17] Du, J., Shi, Y., Mei, C., Quarles, J., & Yan, W. (2016). Communication by interaction: A multiplayer vr environment for building

- walkthroughs. Construction Research Congress, 2281–2290, 2016. doi:10.2307/1412159
- [18] Wang, X., & Dunston, P. S. (2007). Design, strategies, and issues towards an augmented reality-based construction training platform. *ITcon*, (12), 363–380.
 - [19] Eiris, R., Gheisari, M., & Esmacili, B. (2020-a). Desktop-based safety training using 360-degree panorama and static virtual reality techniques: A comparative experimental study. *Automation in Construction*, 109, 102969.
 - [20] Lee, J., Kim, B., Kim, K., Kim, Y., & Noh, J. (2016). Rich360: optimized spherical representation from structured panoramic camera arrays. *ACM Transactions on Graphics (TOG)*, 35(4), 1-11.
 - [21] Gheisari, M., Sehat, N., and Williams, G. (2015) Using Augmented Panoramic Views as an Online Course Delivery Mechanism in MOOCs. 51st ASC Annual International Conference Proceedings, Washington DC.
 - [22] Eiris, R., Moore, H.F., Gheisari, M., and Esmacili, B. (2018-a). Using Panoramic Augmented Reality to Develop a Virtual Safety Training Environment. In *Construction Research Congress 2018* (pp. 29-39).
 - [23] Eiris, R., Gheisari, M., and Esmacili, B. (2018-b). PARS: Using augmented 360-degree panoramas of reality for construction safety training. *International journal of environmental research and public health*, 15(11), 2452.
 - [24] Eiris, R., Wen, J., & Gheisari, M. (2020-b). iVisit: digital interactive construction site visits using 360-degree panoramas and virtual humans. In *Construction Research Congress 2020: Computer Applications* (pp. 1106-1116). Reston, VA: American Society of Civil Engineers.
 - [25] Eiris, R., Wen, J., & Gheisari, M. (2021-b). iVisit-Collaborate: Collaborative problem-solving in multiuser 360-degree panoramic site visits. *Computers & Education*, 104365.
 - [26] Van den Bossche, P., Segers, M., and Kirschner, P.A. (2006) Social and Cognitive Factors Driving Teamwork in Collaborative Learning Environments – Team Learning Beliefs and Behaviors. *Small Group Research*, Volume 37, Number 5, pp. 490-521.
 - [27] Lieu, D. K., and Sorby, S. A. (2009). Visualization, modeling, and graphics for engineering design. Clifton Park, NY: Delmar, Cengage Learning.
 - [28] Vandenberg, S. G., and Kuse, A. R. (1978). Mental rotations; A group test of three-dimensional objects. *Perceptual and Motor Skills*, 47(2), 599–604.
 - [29] Spearman C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*. 15 (1): 72–101.

Vulnerability Assessment of Construction Equipment: An Example for an Autonomous Site Monitoring System

M. S. Sonkor ^a, X. Xu ^a, S. A. Prieto ^a, B. García de Soto ^a

^a S.M.A.R.T. Construction Research Group, Division of Engineering, New York University Abu Dhabi (NYUAD), Experimental Research Building, Saadiyat Island, P.O. Box 129188, Abu Dhabi, United Arab Emirates
E-mail: semih.sonkor@nyu.edu, xx927@nyu.edu, samuel.prieto@nyu.edu, garcia.de.soto@nyu.edu

Abstract –

The digital transformation of the construction industry is accelerating with the advances in information technology (IT) and operational technology (OT) and their convergence. While the benefits of such transformation in construction are apparent, cybersecurity aspects are usually overlooked. Cyber-attacks against project information can cause the exposure of confidential project data, intellectual property, and personal information and interruption of project tasks. On the other hand, cybersecurity incidents affecting the OT utilized in the project can lead to misinformation, harm people nearby, or even cause loss of life.

Given the criticality of providing robust cybersecurity in construction projects, this study aims to (1) point out the cybersecurity considerations to be taken into account before utilizing autonomous equipment on-site, (2) raise awareness about cybersecurity in construction, and (3) present an example of using vulnerability assessment systems in construction projects. This paper utilizes the Common Vulnerability Scoring System (CVSS) to assess the vulnerabilities and risks of different levels of an autonomous site monitoring system. Four assessors with different backgrounds in cybersecurity and robotic systems performed the assessment. The results revealed the most vulnerable levels of the assessed robotic system, which can be considered as a warning. The assessment suggested in this study can help construction decision-makers identify the levels they need to pay extra attention to before employing a cyber-physical system (CPS) on-site. Utilizing CVSS to conduct a vulnerability assessment for CPSs during the construction phase has not been proposed by any previous study, making this paper novel.

Keywords – Construction 4.0; Cybersecurity; Cyber-Physical Systems; Vulnerability Assessment; Construction Robots

1 Introduction

Automation emerges in various forms (e.g., process automation, information technology (IT) automation) and disrupts many industries, including construction. Automation is defined as the use of technology to minimize human input [1]. It is a considerable part of the digital transformation in the construction industry, also referred to as Construction 4.0 [2]. The automation in construction affects both IT and operational technology (OT)-related processes. As an example of the IT-related ones, generative design using visual programming tools (e.g., Dynamo) automates design optimization and helps reach the desired design outcomes much faster than conventional methods. OT-related automation utilizes cyber-physical systems (CPSs) such as robotic systems and can be seen in construction and operation and maintenance (O&M) phases. For example, repetitive construction tasks such as excavation and digging can be automated using autonomous construction equipment. Some construction machinery manufacturers such as Caterpillar and Komatsu and start-ups such as Built Robotics have been developing such equipment. Another example is autonomous data acquisition systems for progress monitoring on construction sites proposed by several studies such as [3] and [4]. The common purpose of all given automation examples is to improve efficiency, accuracy, quality of work, and the safety and well-being of workers by minimizing human involvement.

While the benefits of automation in construction are apparent and well presented in previous studies, the concerns related to cybersecurity did not receive the same degree of attention from the industry and academia [5]. Potential cyberattacks against digital platforms and tools or CPSs utilized in construction projects can lead to the disclosure of sensitive information and cause physical damage to the surrounding environment, including humans and equipment [6]. The volume of sensitive information grows depending on the criticality of the building involved in a project. For example, in 2013, hackers gained access to the blueprints of the Australian Intelligence Service headquarters when it was still under construction [7]. It shows that design documents can also

become a target for hackers when the constructed building has a critical function. A famous example of the cyberattacks against CPSs is the Stuxnet attack that targeted the Natanz nuclear plant in Iran in 2010 and damaged nearly one-fifth of the centrifuges [8]. It can be considered a cyberattack that occurred during the O&M phase.

Given the high impact of potential cyberattacks during different phases of construction projects, maintaining a robust cybersecurity level through the entire life-cycle of the project is crucial. Therefore, this study focuses on construction cybersecurity and proposes using an existing vulnerability assessment system (VAS) to evaluate the vulnerabilities of autonomous robotic systems used in construction projects. This evaluation targets identifying the most vulnerable levels of such systems so that greater attention can be paid to keeping them secure. Several VASs were analyzed to choose the most suitable one for this study. In order to demonstrate the implementation of the selected VAS, an autonomous site monitoring system (ASMS) was used. Different levels of the system were scored according to their vulnerabilities.

2 Research Methodology

The research methodology followed in this paper is divided into three main sections, described as follows.

- **Providing background information on cybersecurity:** Sections 3 and 4 provide background information by presenting prominent studies on cybersecurity efforts in the construction industry and OT cybersecurity, since the scope of the paper and the assessment conducted as a part of the study are primarily related to these two topics.
- **Overview of the different VASs:** An overview of the prominent VASs is provided in Section 5, starting from the first examples and including the most recent and widely used ones. This overview aims to give an understanding of the commonly used VASs and show their main characteristics. This overview helps decide the VAS to be utilized in the following section.
- **Vulnerability Assessment of an ASMS:** The assessment shown in this study includes an ASMS that has a prototype at the S.M.A.R.T. Construction Research Group's lab at NYUAD. A high-level overview of the system is provided in Section 6.1. For the assessment, CVSS was chosen. Its main characteristics are presented in Section 6.2 before demonstrating the proposed implementation. Finally, in Section 6.3, the proposed assessment is demonstrated. The assessment was conducted by four assessors with different backgrounds. The authors of this paper performed the assessor roles in this study for the sake of simplicity and demonstration purposes.

3 Cybersecurity Efforts in the Construction Industry

Even though cybersecurity does not stand as one of the popular topics in construction research, some studies have been conducted to point out the necessity of strong cybersecurity levels in projects considering different phases. One of the examples is the study by Zheng et al. [9]. Their study targeted to prevent BIM data leakage by proposing a context-aware access control for BIM systems instead of the conventional role-based ones. They provided examples of two different possible contexts that can be used for access control: location and time. Mantha et al. [10] developed a construction-specific cybersecurity threat model for different phases of projects. They demonstrated the use of the proposed model at the commissioning phase. Possible intrusions into the data collection process by malicious actors and a countermeasure to prevent such actions were presented in their study. Alshammari et al. [11] investigated the cybersecurity aspects of digital twins, which are envisaged to be commonly used to monitor and simulate built environments in the near future. Grundy [12] discussed cybersecurity during the O&M phase by underlining the increasing utilization of interconnected sensors/devices in smart buildings. He suggested using generic cybersecurity frameworks published by internationally recognized institutions (e.g., ISO, National Institute of Standards and Technology (NIST)) to address the raised concerns that stem from such interconnectivity. Finally, Sonkor and García de Soto [5] focused on the automated and remote-controlled equipment starting to be a bigger part of construction processes. They reviewed the literature and accentuated the lack of studies on the cybersecurity aspects of such equipment utilized on construction sites.

Some scholars suggested blockchain-based solutions to different cybersecurity problems in construction. For example, Turk and Kline [13] did one of the first studies that analyzed the potentials of blockchain in the construction industry. They considered blockchain to provide a trustworthy environment for managing information exchange during different phases of projects. Pärn and Edwards [13] scrutinized the cyber threats affecting the built environment, mainly focusing on critical infrastructures (CIs). They proposed using blockchain technology to improve the confidentiality of sensitive information in BIM common data environments (CDEs). Lee et al. [14] stressed the diversity of stakeholders involved in construction projects and proposed an integrated framework consisting of blockchain and digital twin technologies for enhanced data traceability. Last but not least, Sonkor and García de Soto [15] proposed a data-sharing architecture that utilizes a decentralized storage approach and blockchain technology to improve cybersecurity in construction

networks. They suggested using blockchain to store the fingerprints of BIM files for validation purposes while storing the actual files in a decentralized manner over the construction network.

Besides the academic studies, several reports published by private institutions indicated the increasing cyber threat surface in the construction sector. One of them is AECOM's report [16] showing the results of a survey that involved 509 civil infrastructure professionals from Europe, North America, and Asia-Pacific. Their survey aimed to understand the cybersecurity awareness and preparedness of the civil infrastructure sector. The report concluded that "To support economic growth and social prosperity, future-proofing and protection against cyber and physical attack are essential". Moreover, a recent study by NordLocker [17] showed that the construction industry had been the primary target of ransomware attacks in 2021. This concerning finding should be considered a wake-up call for the industry to start taking immediate actions.

4 Cybersecurity of Operational Technology

OT in construction is new and still not at a mature level—especially during the pre-occupational phases—unlike in some other sectors that use industrial control systems (ICSs) and are ahead of construction in terms of digitalization. These sectors include but are not limited to water treatment, energy, oil & gas, electric power distribution, and manufacturing. Since most of these sectors are considered a part of CI, they have been the primary target of hackers [18] (e.g., patriot hackers, organized cybercriminals). The importance of availability in such environments further increases the outcomes of potential attacks [19].

Large-scale OT attacks in history have repeatedly proven the criticality of robust cybersecurity. An example is a recent attack against the largest pipeline in the US, Colonial Pipeline, that caused the operations to stop for two weeks and led to severe outcomes such as fuel price increases and fuel shortages [20]. The operations could go back to normal only when the hackers received a \$4.4 million ransom. Increasing attention of hackers to these sectors caused CI owners and operators to take additional cybersecurity measures and researchers to direct their efforts toward proposing preventive methods against possible attacks.

Some examples from the extensive literature on this topic are [21]–[23], that proposed cyberattack and intrusion detection systems for ICSs. These detection systems aim to gain the cybersecurity teams of the attacked entities valuable time before any unrecoverable damages occur. In 2015, two cybersecurity researchers published a white paper showing that they could remotely

control a passenger vehicle [24]. They gained access to the car's entertainment system by exploiting the software vulnerabilities. After they gained access, they could remotely control different functionalities of the car, such as the dashboard, brake, steering, and air conditioner. Another research to prove the vulnerabilities of commonly used OT was conducted by Trend Micro Research [25]. They tested radio frequency (RF) remote controllers from 17 vendors installed on cranes in industrial environments. Their results showed that millions of cranes using these remote controllers are vulnerable to cyberattacks.

5 Vulnerability Assessment Systems

Cybersecurity researchers and white-hat hackers discover new vulnerabilities every day, and these vulnerabilities have different levels of impact when they are exploited. Since it is almost impossible for organizations to address every discovered vulnerability, it is crucial to know which ones to prioritize. Therefore, government and private institutions have developed various VASs to identify the severities of vulnerabilities over the years. The Escal Institute of Advanced Technologies (SANS) published one of the early vulnerability assessment documents in 2001 [26]. It provided insights into the necessity of vulnerability assessments and gave an overview of the recommended process for conducting one. In 2016, Bugcrowd (a crowdsourced security platform) released the Vulnerability Rating Taxonomy [27], which is simpler and less comprehensive than the widely used Common Vulnerability Scoring System (CVSS). It was developed to be used by the bug bounty community. Therefore, it mainly focuses on the vulnerabilities frequently seen by bug hunters. Microsoft Security Response Center (MSRC) developed the Microsoft Exploitability Index (MEI), which mainly aims to help Microsoft customers assess the exploitability potential of vulnerabilities [28]. MEI has four levels of exploitability: 0 – Exploitation Detected, 1 – Exploitation More Likely, 2 – Exploitation Less Likely, and 3 – Exploitation Unlikely. MSRC developed MEI as a separate vulnerability scoring system that is independent from the CVSS. However, MSRC also contributes to the improvement efforts of CVSS [28].

The mentioned VASs have particular use-cases and scopes. On the other hand, CVSS released by the Forum of Incident Response and Security Teams (FIRST) has been commonly used by internationally recognized cybersecurity organizations, such as NIST, for a wide range of software vulnerabilities. CVSS is an open framework developed to assess the severities of discovered vulnerabilities [29]. It has been used by the National Vulnerability Database (NVD), one of the most extensive vulnerability databases available on the

internet. There are three groups of metrics provided by CVSS: base metric group, temporal metric group, and environmental metric group [29]. Base metrics show the characteristics of a vulnerability that do not change in different environments and over time. Temporal metrics show the vulnerability characteristics that do not change in different environments but might change over time. Finally, environmental metrics reflect the characteristics of vulnerabilities that are specific to the end user's environment. While scoring the base metrics is mandatory, temporal and environmental metrics are optional but recommended for higher precision.

Besides the original purpose of CVSS, which is to describe the severity and characteristics of vulnerabilities, Mantha and García de Soto [30] used it to assess the cybersecurity vulnerabilities of different participants in construction projects, such as the public owner, local contractor, and construction worker. Their study provided an alternative use of CVSS, showed its usefulness in the construction industry, and inspired the implementation demonstrated in the following section.

6 Vulnerability Assessment of an Autonomous Site Monitoring System

Several OT uses on construction sites have attracted attention from construction equipment manufacturers

and academics. One of them is utilizing ASMSs on-site to track construction progress with minimized human intervention. This section demonstrates the implementation of CVSS (version 3.1) to perform vulnerability assessments for different levels of an ASMS. CVSS was chosen due to its well-established assessment structure and suitability to be used for various cases besides software vulnerabilities, as proven in [30]. The following subsections provide an overview of the assessed ASMS and the assessment details.

6.1 Overview of the Assessed Autonomous Site Monitoring System

The different components and levels of the assessed ASMS are depicted in Figure 1. The structure of the communications and integrations within the robotic platform is divided into five different levels based on their level of abstraction.

In Level 0 of the structure, the physical components of the robot responsible for acquiring data and interacting with the environment (i.e., sensors and actuators) can be found. They can be grouped into three major subgroups: the 3D scanner, all the sensors embedded on the robot (i.e., LiDARs, RGBD camera, encoders, and IMU), and the platform itself—housing all the different hardware such as sensors, computers, and locomotion means.

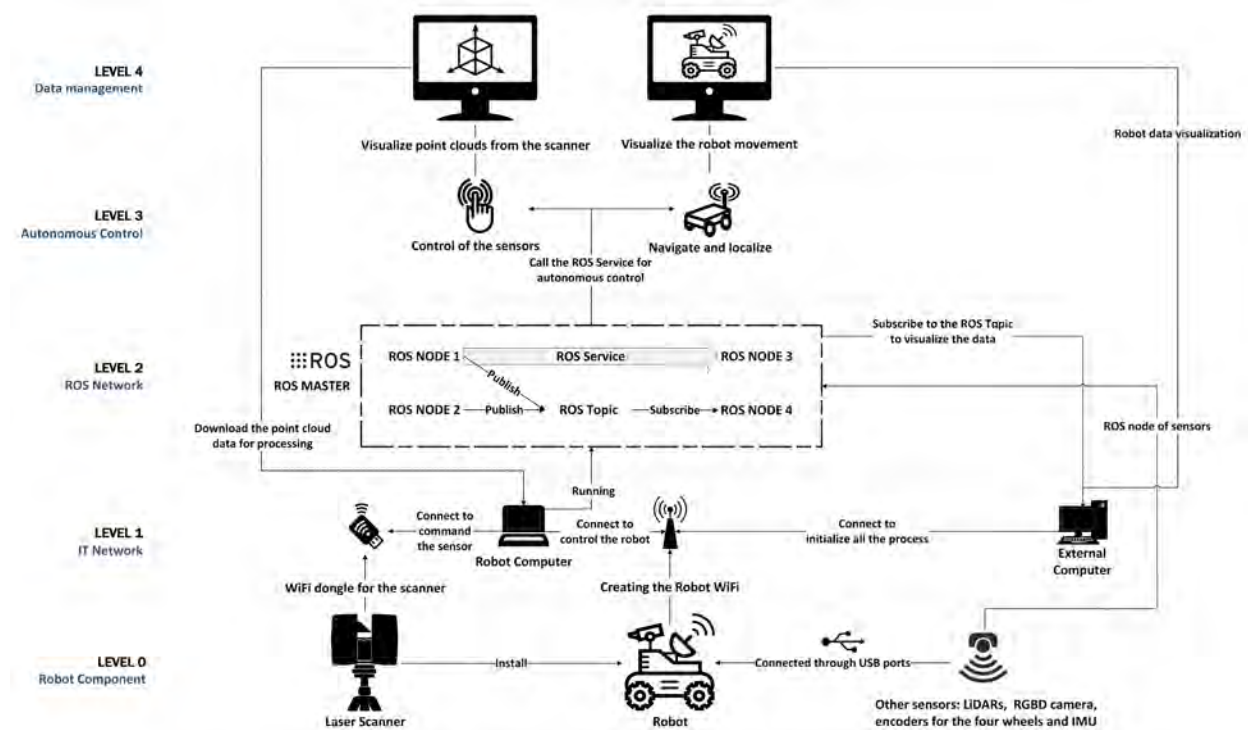


Figure 1. Diagram showing the different levels and components of the ASMS

Level 1 involves the basic connections between the elements mentioned above and the computer embedded in the robot responsible for controlling everything. Most sensors are connected to the Robot Computer by a USB protocol. The platform is equipped with a Wi-Fi router to create a Master Network, where the Robot Computer is connected via Ethernet. The laser scanner is only reachable through Wi-Fi communication and is connected to the Robot Computer through a Wi-Fi dongle. To provide a Human-Machine Interaction (HMI) interface, a second External Computer can be connected to the Master Network.

Level 2 consists of the Robot Operating System (ROS) [31] Network. The ROS Network can involve as many devices and computers as needed, as long as there is a device running the Master. In this case, the Robot Computer acts as a Master, and the External Computer connects to the ROS Network to interact with it. This interaction is bi-directional, allowing the robotic platform to attend to any command given by the External Computer, and the External Computer to visualize any data coming from the robotic platform. Within the ROS Network, multiple nodes are running, each managing all the different functionalities of the robot (e.g., the autonomous navigation, the localization, all the different sensors, control of the motors). The ROS nodes publish/receive information in the way of ROS Topics and can receive/give commands in ROS Services.

Level 3 demonstrates the basic tasks fulfilled by the robotic system, which involves the autonomous control of the robot. By using the ROS Services and publishing into the ROS Topics, the platform can communicate with all the different sensors and actuators autonomously.

Level 4 is the Human-Machine Interaction (HMI) layer. ROS provides multiple graphical user interfaces (GUIs) to interact and visualize the information of the robot, the most important one being the visualization software for the robot sensors, running in the External Computer to achieve autonomous mapping and localization using SLAM algorithms. This level overrides any command issued by Level 3.

6.2 Overview of the Scoring System

As mentioned earlier, CVSS has three groups of metrics (i.e., base, temporal, environmental), and the base metric group is mandatory to be scored. Therefore, to keep the scoring demonstrated in this study brief, only the base score will be calculated for different levels of the ASMS shown in Figure 1. Different metric categories within the base group, different metrics under each category, and possible metric values for each metric are shown in Table 1. Each possible metric value has a different numerical equivalent to be used in the formulas of the base score. Explanations of each metric, the corresponding numerical values, and the formulas to

calculate the base score are presented in detail in the CVSS version 3.1 Specification Document [29].

Table 1. Different metrics and possible metric values to calculate the CVSS base score

Metric Category	Metric Name	Possible Metric Values
Exploitability	Attack Vector	[Network, Adjacent, Local, Physical]
	Attack Complexity	[Low, High]
	Privileges Required	[None, Low, High]
	User Interaction	[None, Required]
Scope	Scope	[Unchanged, Changed]
	Confidentiality	[High, Low, None]
Impact	Integrity	[High, Low, None]
	Availability	[High, Low, None]

The exploitability category covers the characteristics of the component considered in the scoring process and the properties of a successful attack that can exploit this component. For example, the attack vector metric refers to the distance of a successful attacker that exploits the vulnerable component. As the possibility of remote exploitation increases, the number of possible attackers increases, thus the base score in this assessment. Since different levels of the robotic system are scored in this paper rather than software vulnerabilities, the possibility of exploiting the corresponding level was considered while making the assessment presented in the following subsection. For example, if the assessor thinks that the components in Level 0 (shown in Figure 1) can be exploited only with physical interaction, he/she should select the “Physical” option for the attack vector.

The scope metric reflects whether a successful breach of a component can affect another component that is not in the same security authority. The security authority in this definition refers to a mechanism (e.g., an operating system) that controls access to specific components of a system. Therefore, if the same mechanism controls access to two different components, these two components are considered under the same security authority. Considering the ASMS in Figure 1, if the assessor thinks that a successful attack against Level 0 can also affect another level, the “Changed” option should be selected.

The impact category indicates the potential outcomes of a successful attack against the vulnerable components. The most likely outcomes of the potential attacks should be considered while scoring these metrics. For example,

if the breach of Level 4 is likely to cause a complete confidentiality loss, the “High” option should be selected for the confidentiality metric.

6.3 Demonstration of the Assessment and Discussion

For demonstration purposes, four assessors evaluated the different levels of the ASMS (in Figure 1) using CVSS. Descriptions of different levels of the ASMS and different metrics of CVSS were given to the assessors before they provided their input. The base scores for each level of the system were calculated using the inputs from

each assessor and the formulas in [29]. As an example, Table 2 shows the input received from Assessor 1, numerical equivalents, and the calculated total base scores for each level.

Input from the four assessors was considered and combined to determine the vulnerability of the different levels of the system. In order to account for the assessors’ knowledge of cybersecurity and the assessed system, different weights were applied (Table 3). Weighted averages for each level were calculated (as summarized in Table 4). The risk ratings (derived from CVSS [29]) reflect the exploitation risk for each level.

Table 2. CVSS assessment from Assessor 1 and numerical equivalents

Metric Category	Metric Name	Level 0	Level 1	Level 2	Level 3	Level 4
Exploitability	Attack Vector	Physical / 0.20	Adjacent / 0.62	Local / 0.55	Adjacent / 0.62	Adjacent / 0.62
	Attack Complexity	Low / 0.77	High / 0.44	High / 0.44	Low / 0.77	Low / 0.77
	Privileges Required	Low / 0.68	Low / 0.68	Low / 0.68	Low / 0.68	Low / 0.68
	User Interaction	Required / 0.62	None / 0.85	None / 0.85	None / 0.85	Required / 0.62
Scope	Scope	Changed	Changed	Changed	Changed	Changed
Impact	Confidentiality	None / 0	None / 0	High / 0.56	None / 0	High / 0.56
	Integrity	Low / 0.22	Low / 0.22	High / 0.56	Low / 0.22	High / 0.56
	Availability	Low / 0.22	Low / 0.22	High / 0.56	Low / 0.22	High / 0.56
Total Base Score		3.6	4.4	7.8	5.4	8.4

Table 3. Different weights for the assessors according to the level of knowledge

Knowledgeable in cybersecurity?	Knowledgeable about the assessed robotic system?	Weight
Yes	Yes	100%
Yes	No	80%
No	Yes	60%

The weighted base score averages (Table 4) indicate a relatively higher exploitation risk and vulnerability against cyberattacks for Level 2 (CVSS score: 8.6, Risk rating: High). On the other hand, Level 0 (CVSS score: 3.7, Risk rating: Low) can be considered as the lowest risk level of the assessed ASMS. In the context of this study, a higher risk implies a larger threat surface, a greater number of potential attackers, a higher probability of a successful attack, and a higher impact when a vulnerable component is exploited. The higher risk rating

of Level 2 can be due to two main reasons. The first one is the high impact of a potential breach of Level 2, which involves the ROS Network. Since ROS can be used to give commands to the robot and request data, potential exploitations of the ROS Network can have severe impacts on the confidentiality, integrity, and availability of the data that can be accessed through the ROS nodes. A successful attack can affect the robot’s movement—which can cause harm to the surrounding environment and people—, alter the data received from the robot—which can negatively impact the decisions made using this data—, make the robot or data unavailable—which can disrupt the related tasks—, and expose some information that the attackers can use to plan their future attacks against the construction site. The second main reason for the high-risk rating is that the ROS Nodes and Services can be accessed without a need for credentials in the assessed ASMS. Even though the computers used for accessing the ROS Network require credentials, they are in Level 1 and thus not considered while scoring Level 2. Therefore, the privileges required and the attack

complexity to exploit Level 2 are low, making the risk high. Additional security features are employed in ROS 2, such as encrypting the communication traffic and authentication [32]. However, the robotic system in this study uses ROS 1, which does not have the mentioned features. The main reasons for the low-risk rating of Level 0 include the physical interaction required to launch an attack against this level's components. Not

being able to perform an attack remotely significantly reduces the number of potential attackers for this level. Moreover, potential exploitations of the components in this level do not have a high impact since the data exchange and all the communication between the robot, operating system, sensors, and utilized software happens at the other levels.

Table 4. Weighted averages of the CVSS scores from all assessors for each level

	Level 0	Level 1	Level 2	Level 3	Level 4
Total base score (weighted average)	3.7	6.6	8.6	7.5	6.5
Risk rating	Low	Medium	High	High	Medium

7 Conclusions, Limitations, and Future Work

Automation and digitalization are increasingly affecting the construction processes in different project phases. The complexity of maintaining robust cybersecurity is elevated due to the fragmented nature of construction projects, unstructured and unstable environments, and the variety of the utilized equipment in terms of purpose and security levels. CPSs, such as autonomous earthmovers and site monitoring equipment, are particularly open to cyberattacks since they are relatively new in the construction industry. Therefore, construction decision-makers should employ specific methods to assess the risk and vulnerability levels of the CPSs that they are planning to use in their projects. This study addresses this need by proposing CVSS to assess the vulnerabilities of different levels in an ASMS. Four assessors conducted the assessment, and the weighted averages according to their knowledge levels of cybersecurity and the assessed system were calculated for each level. The results provided the levels with the highest cybersecurity risks, which can be considered as a warning to pay additional attention, particularly to these levels.

One of the limitations of the study is the number of assessors. The involvement of more assessors knowledgeable about cybersecurity and robotic systems can improve the accuracy and reliability of the results. Moreover, the components in the system were not scored individually, which can be considered another limitation. Instead, the scoring was performed level by level (i.e., for each level consisting of different components). Different characteristics of the components in some levels made it more challenging to make a representative assessment. For example, in Level 1, the attack complexity required to exploit the Wi-Fi dongle and the External Computer are inherently different due to their varying software and hardware characteristics. In Level 4, two different

visualization software developed by different software companies were considered together, which also caused an inaccurate assessment due to these companies' different cyber defense mechanisms. The authors are extending this study to address these limitations. The future assessment will include a larger pool of assessors and more detailed information about each component of the ASMS. It will be conducted by examining each component instead of each level. Suggestions to mitigate the potential cybersecurity risks related to the most vulnerable components will be provided. Moreover, the usefulness of the proposed assessment will be tested on other construction robotic applications.

Acknowledgment

The authors thank the Center for Cyber Security at New York University Abu Dhabi (CCS-AD) for their support.

References

- [1] IBM. What is automation? *IBM*. Online: <https://www.ibm.com/topics/automation>, Accessed: 24/01/2022.
- [2] Kline R. and Turk Ž. Construction 4.0 - digital transformation of one of the oldest industries. *Economic and Business Review*, 21(3):393–410, 2019.
- [3] Moselhi O., Bardareh H., and Zhu Z. Automated Data Acquisition in Construction with Remote Sensing Technologies, *Applied Sciences*, 10(8). 2020.
- [4] Prieto S., García de Soto B., and Adan A. A Methodology to Monitor Construction Progress Using Autonomous Robots. Oct. 2020.
- [5] Sonkor M. S. and García de Soto B. Operational Technology on Construction Sites: A Review from the Cybersecurity Perspective. *Journal of Construction Engineering and Management*,

- 147(12):4021172, 2021.
- [6] García de Soto B., Georgescu A., Mantha B. R. K., Turk Ž., and Maciel A. Construction Cybersecurity and Critical Infrastructure Protection: Significance, Overlaps, and Proposed Action Plan. *Preprints 2020*, 2020.
- [7] Watson S. Cyber-security: What will it take for construction to act? *Construction News*. Online: <https://www.constructionnews.co.uk/tech/cyber-security-what-will-it-take-for-construction-to-act-22-01-2018/>,
- [8] Hemsley K. E. and Fisher R. E. History of Industrial Control System Cyber Incidents. 2018. Online: <https://www.osti.gov/servlets/purl/1505628>,
- [9] Zheng R., Jiang J., Hao X., Ren W., Xiong F., and Zhu T. CaACBIM: A context-aware access control model for BIM. *Information*, 10(2):47, 2019.
- [10] Mantha B. R. K., García de Soto B., and Karri R. Cyber security threat modeling in the AEC industry: An example for the commissioning of the built environment. *Sustainable Cities and Society*, 66:102682, 2021.
- [11] Alshammari K., Beach T., and Rezgui Y. Cybersecurity for digital twins in the built environment: Current research and future directions. *Journal of Information Technology in Construction*, 26(March):159–173, 2021.
- [12] Grundy C. Cybersecurity in the built environment: Can your building be hacked? *Corporate Real Estate Journal*, 7(1):39–50, 2017.
- [13] Pärn E. and Edwards D. Cyber threats confronting the digital built environment: Common data environment vulnerabilities and block chain deterrence. *Engineering Construction & Architectural Management*, 26(2):245–266, 2019.
- [14] Lee D., Lee S. H., Masoud N., Krishnan M. S., and Li V. C. Integrated digital twin and blockchain framework to support accountable information sharing in construction projects. *Automation in Construction*, 127, 2021.
- [15] Sonkor M. S. and García de Soto B. Towards Secure Construction Networks: A Data-Sharing Architecture Utilizing Blockchain Technology and Decentralized Storage. 2021.
- [16] AECOM. The Future of Infrastructure. 2018. Online: <https://tinyurl.com/mrx3rj2c>, Accessed: 20/06/2022.
- [17] Nordlocker. Top industries hit by ransomware. *Nordlocker*. Online: <https://nordlocker.com/recent-ransomware-attacks/>, Accessed: 20/01/2022.
- [18] McLaughlin S. *et al.* The Cybersecurity Landscape in Industrial Control Systems. *Proceedings of the IEEE*, 104(5):1039–1057, 2016.
- [19] Drias Z., Serhrouchni A., and Vogel O. Analysis of cyber security for industrial control systems. In *2015 International Conference on Cyber Security of Smart Cities, Industrial Control System and Communications (SSIC)*, pages 1–8, 2015.
- [20] Turton W. and Mehrotra K. Hackers Breached Colonial Pipeline Using Compromised Password. *Bloomberg*. Online: <https://www.bloomberg.com/news/articles/2021-06-04/hackers-breached-colonial-pipeline-using-compromised-password>, Accessed: 10/08/2021.
- [21] Zhang F., Kodituwakku H. A. D. E., Hines J. W., and Coble J. Multilayer Data-Driven Cyber-Attack Detection System for Industrial Control Systems Based on Network, System, and Process Data. *IEEE Transactions on Industrial Informatics*, 15(7):4362–4369, 2019.
- [22] Adepu S. and Mathur A. Distributed Attack Detection in a Water Treatment Plant: Method and Case Study. *IEEE Transactions on Dependable and Secure Computing*, 18(1):86–99, 2018.
- [23] Sugumar G. and Mathur A. Testing the Effectiveness of Attack Detection Mechanisms in Industrial Control Systems. In *2017 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, pages 138–145, 2017.
- [24] Valasek C. and Miller C. Remote Exploitation of an Unaltered Passenger Vehicle. 2015. Online: https://ioactive.com/pdfs/IOActive_Remote_Car_Hacking.pdf,
- [25] Andersson J. *et al.* A Security Analysis of Radio Remote Controllers for Industrial Applications. 2019. Online: https://documents.trendmicro.com/assets/white_papers/wp-a-security-analysis-of-radio-remote-controllers.pdf,
- [26] Cima S. SANS Institute - Vulnerability Assessment. 2001. Online: <https://www.sans.org/white-papers/421/>,
- [27] Bugcrowd. Bugcrowd's Vulnerability Rating Taxonomy. Online: <https://bugcrowd.com/vulnerability-rating-taxonomy>, Accessed: 04/02/2022.
- [28] Microsoft. Microsoft Exploitability Index. *Microsoft*. Online: <https://www.microsoft.com/en-us/msrc/exploitability-index>, Accessed: 06/02/2022.
- [29] FIRST. Common Vulnerability Scoring System version 3.1. 2019. Online: https://www.first.org/cvss/v3-1/cvss-v31-specification_r1.pdf,
- [30] Mantha B. R. K. and García de Soto B. Assessment of the Cybersecurity Vulnerability of Construction Networks. *Engineering, Construction and Architectural Management*, 2020.
- [31] Quigley M. *et al.* ROS: an open-source Robot Operating System. In *ICRA workshop on open source software*, 3(3.2), 2009.
- [32] The Construct. How to Enable and Use Security in ROS 2 | Sid Faber | ROSDevDay 2021. *Youtube*. Online: <https://youtu.be/UJa4XWRA6EY>, Accessed: 20/06/2022.

Fire Evacuation and Management Model Based on Building Information Modeling and Virtual Reality

R.-J. Gao^a, K.-C. Wang^{a*}, X.-H. Lai^a and W.-H. Hung^a

^a Department of Civil and Construction Engineering, Chaoyang University of Technology, Taiwan
E-mail: s10911615@gm.cyut.edu.tw, wkc@gm.cyut.edu.tw, lonermiker@gmail.com,
william20008917@gmail.com

Abstract –

Fire management is a key aspect of building management. During a fire, the primary reason why evacuees cannot evacuate in time and casualties occur is their lack of familiarity with the evacuation route. In practice, two dimensional images or videos are used as aids for firefighting educational training. However, they can only display a partial view of an evacuation route (not a 360° panorama) and thus have limited effectiveness in helping evacuees to understand the evacuation route. We proposed a fire evacuation management model that combines building information modeling (BIM) and virtual reality (VR) to support fire evacuation. The model is divided into two modules, namely (1) the evacuation route, and (2) educational training modules. The evacuation route module is used to analyze the shortest evacuation route from a room (where evacuees are located) to an exit and to estimate the shortest evacuation time. To verify whether the simulated route generated by the VR fire management system is applicable to evacuations, we further calculated the allowable evacuation time with the empirical equation provided in the Building Fire Refuge Safety Verification Technique Manual. The VR-based educational training module helps evacuees to evacuate from a building by providing a first-person perspective simulation. The immersive characteristic of a VR environment allows people to view an evacuation route from a 360° perspective, which deepened their impression of an environment. In this study, the proposed model was introduced and tested in a case study, which verified its ability to assist management units with fire management assessments and support fire evacuation training for users of buildings.

Keywords –

Fire evacuation; Building information modeling (BIM); Virtual reality (VR)

1 Introduction

In recent years, the frequency of fires in Taiwan has increased. According to the statistics released by the National Fire Agency, Ministry of the Interior, 17,319 fires occurred from January to September, 2021, causing 110 deaths and NT\$252,881,000 in property damage [1]. Therefore, tragedies can be reduced by implementing disaster prevention measures.

Studies on fire escapes have indicated that in fire education and training for the public, fire evacuation knowledge can only be conveyed through text, posters, or videos. However, to improve the efficiency of fire education as well as equipment maintenance and management, virtual reality (VR) technology can be incorporated to support evacuation training for the public, and building information modeling (BIM) can be employed in setting the parameters for firefighting equipment models and achieving the informatization of equipment maintenance and management.

In the present study, we incorporated optimal evacuation route guidance into a VR system to effectively guide personnel who are unfamiliar with a building to avoid dangerous areas safely and easily and reduce accidents caused by prolonged evacuation time. We also used VR to simulate the emergency stress that personnel cannot experience when they are performing an evacuation drill in a safe environment to improve their emergency response to actual disasters and increase the effectiveness of such drills.

2 Literature Review

2.1 Incorporation of BIM into firefighting management

Studies on disaster prevention management have mostly verified whether the structural design of a building is appropriate or whether an evacuation route is safe. Few studies have comprehensively evaluated

personnel safety by examining both indoor structures and evacuation routes. Therefore, Shih established a disaster prevention management model on the basis of BIM to implement disaster prevention measures before disasters occur, increase emergency response efficiency, and reduce casualty rates and property loss, thereby achieving the goals of disaster prevention management [2].

When building fires occur, information about dangerous factors is crucial for emergency responders, equipment management personnel, and rescue teams. Inadequate management limits the accuracy and speed of fire rescue operations. In addition, relying solely on the experience of emergency responders may have a negative effect on fire response operations. BIM can support disaster prevention by providing information on the locations of key elements and enabling effective decision-making. However, the integration of the life cycle information of buildings is a challenging task. In particular, information about building fires should be retrieved through BIM software because it mostly has spatial characteristics. Suhyun et al proposed the Building Fire Information Management System, which provides reliable fire-related information through computerized and systematic methods and BIM tools. This system enables emergency responders to intuitively identify the location of indoor facilities by referring to information that is presented in a three-dimensional format. Through scene-based application, the system was proven to accelerate access to relevant information [3].

2.2 VR evacuation simulation

When a fire occurs in a building, the timely evacuation of residents from the building is the key to protecting the safety of residents and property. In a literature review, Chiu et al. examined the fire-related regulations in Taiwan, process through which fires occur, methods for calculating evacuation time, and fire simulation software that were commonly used. A fire simulation model was built using the Unity software, and variables (e.g., size and location of the rescue opening and number and age of victims) were adjusted to estimate the time required for evacuation under varying conditions. The final simulation result was consistent with the results obtained by other researchers who used other fire simulation software. This indicated that the Unity software can be used in disaster prevention simulation. This software can also output fire simulation models in a mobile platform mode that is compatible with mobile phones or tablets, such that onsite furniture layouts can be considered when conducting an evacuation simulation [4].

Wang et al. employed a BIM-based virtual environment and a game engine to solve the key challenges in building emergency management, which include the provision of timely two-way information

updates and enhanced emergency response awareness training [5]. They reported on how real-time fire evacuation guidance can be provided by adopting BIM to provide comprehensive building information and combining BIM with VR technology to create applicable and immersive game environments.

Because of the high-risk environment and complex superstructures of offshore oil and gas platforms (OOGPs), accidents on OOGPs typically result in severe casualties and economic loss. Accordingly, Cheng et al. developed an OOGP simulation model in which evacuation plans are evaluated through BIM to improve evacuation technology and achieve an enhanced agent-based model. To increase their simulation performance, they developed and input environment perception and dynamic evacuation routing functions into the agent [6].

2.3 Navigation function of Unity

The navigation function of the Unity game engine can be used to create characters that move intelligently within a virtual world. It can create a navigation mesh that serves as the moving area of the agent on the basis of geometry. The agent then simulates movement behavior and controls a character to cause it to move toward a target. [7]

The factors that should be considered in the navigation function are speed, angular speed, acceleration, and stopping distance. The setting for each parameter is determined based on the needs of the navigation process. [8]

3 BIM-VR Based Fire Evacuation and Management Model

In this section, we demonstrate a BIM-VR-based fire evacuation and management model (Fig. 1) in which a BIM model was integrated into a VR environment. To increase drill effectiveness, immersive VR technology was applied to induce in users the emergency stress that they cannot experience through conventional fire drills and educational animations.

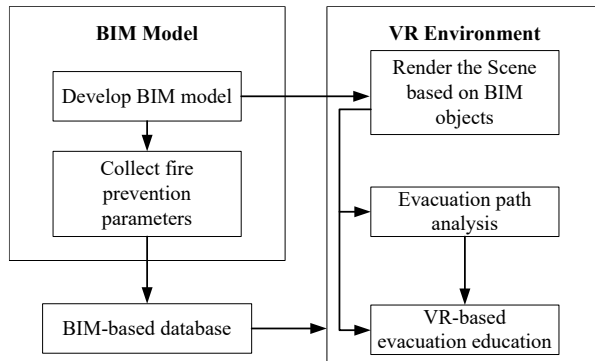


Fig. 1. BIM-VR-based fire evacuation and management model.

3.1 BIM module

Before the integration of VR software, BIM was conducted. The case model contained basic elements, such as columns, beams, slabs, and walls. Because a discussion on the maintenance and management of firefighting equipment was required, the relevant data about the elements of firefighting equipment were input into a database that was subsequently linked to the VR software.

3.2 VR environment

The VR system was divided into three modules, which are as follows.

3.2.1 Evacuation route module

Through the use of the navigation function of the Unity game development engine, we created a bake mesh map in the building information model. Parameters were added in accordance with the settings for the game agent. The script function was used to move the agent to the safe area. The information of each room in the model was substituted into the building refuge safety recommendation equation (provided in the Building Fire Refuge Safety Verification Technique Manual) to calculate the required evacuation time, which was calculated by summing the start time of an evacuation (e.g. the period after the fire occur and before the evacuees start to escape, t_{start}), travel time to the ground (t_{travel}), and the time required to move pass the ground exit (t_{queue}). The floor area in a space was required to estimate the start time of an evacuation (t_{start}), the distance from a room to an exit was required to estimate the travel time to the ground (t_{travel}), and the floor area and exit width were required to estimate the time required to move pass the ground exit (t_{queue}). The aforementioned information could be obtained through BIM. The empirical equation is displayed in Eq. (1).

$$T_{escape} = t_{start} + t_{travel} + t_{queue} = ((\sqrt{\Sigma A})/30 + 3) + \max li/v + (\Sigma \rho A)/\Sigma NeffBd \quad (1)$$

Subsequently, the estimated evacuation time for an entire building was compared with the descent time of a smoke layer to determine whether the residents of the building had enough time to evacuate. The calculation of the required time for smoke in each room to descend to the threshold height is performed using Eq. (2).

$$t_s = \frac{A_{room} \times (H_{room} - H_{lim})}{\max(V_s - V_e, 0.01)} \quad (2)$$

Where, A_{room} = floor area of the room (m^2), H_{room} = height from the floor to the ceiling of the room (m), H_{lim} = height of smoke boundary (m), V_s = smoke volume (m^3/min), and V_e = effective smoke exhaust volume (m^3/min).

Finally, we verified whether the evacuation times calculated using the empirical equation and the evacuation route set using Unity had allowable values to determine whether the residents could evacuate safely in time. In our model, the simulated building was treated as a fire-resistant structure. The fire-resistant properties of the building (e.g., situation in which the building collapses because of a prolonged fire) were not considered. Furthermore, the physical and mental factors associated with dangerous situations (e.g., panic, and coma) were not considered in the parameter setting for the agent.

3.2.2 Educational training module

After comparing the evacuation time that was estimated based on the evacuation route (set using Unity) with that obtained through the empirical equation, we discovered that the evacuation time estimated based on the Unity-derived evacuation route was within limits of the allowable evacuation time. Therefore, VR educational training is a feasible means for guiding users on the use of an evacuation route to reach a safe area. Moreover, we added evacuation route guidance to the educational training module, which enabled users to clearly identify evacuation routes. The particle system of the Unity software was used to simulate the flow of smoke during an actual fire; this increased the feelings of tension and stress experienced by users, giving them an experience that conventional fire drills and educational animations cannot provide.

4 Case Study

The case in this study was a 4-story duplex. We combined the building information model of the duplex with a fire management and maintenance VR system. The test process is described in the following subsections.

4.1 BIM module

The case model had basic elements, such as columns, beams, slabs, walls, a firefighting system, and firefighting equipment. In addition, to import the maintenance and management module into the VR system, firefighting equipment maintenance information was added to the model parameters.

4.2 VR environment

When a building information model is imported into Unity as an FBX file, texture loss tends to occur. Therefore, we re-rendered the model and added light and shadow through Unity (Fig. 2) by making a shader with the more realistic texture maps provided by the Unity developer. We then added the shader to the model elements for reattachment and set the lighting effect to increase shadow detail. With the Unity lightmap function, the realism of scenes can be enhanced. After a scene was rendered, and light and shadow settings were adjusted, the setting of the module function could be conducted.



Fig. 2. Rendered model with light and shadow added through Unity.

4.2.1 Evacuation route module

The navigation function in Unity was applied for mesh map baking (Fig. 3). According to the Building Fire Refuge Safety Verification Technique Manual, the average and stair-descent walking speeds of an evacuee are 60 and 36 m/min, respectively. Hence, the waking speeds of the agent was set accordingly. The agent would move to the safe area according to the script function, and the system would record the overall time they spent to

move to the safe area. Subsequently, the relevant evacuation information (e.g., building sections, building areas, and width measurements; Table 1) was exported by using the schedule export function in the building information model. We estimated the allowable evacuation time by applying the empirical equation provided in the Building Fire Refuge Safety Verification Technique Manual and verified whether the actual evacuation time of the agent was within the limits of the allowable evacuation time calculated using the empirical equation. During the test, the agent was placed in the second-floor bedroom in Flat B, and the ignition was set to occur in the second-floor dressing room. The actual evacuation time of the agent in the VR system was 129 s, and the estimated allowable evacuation time obtained using the empirical equation was 360 s. The actual evacuation time calculated using the VR system was shorter than the estimated allowable evacuation time. Therefore, the optimal evacuation route produced by the VR system (Fig. 4) can assist the agent to evacuate safely.



Fig. 3. Mesh map baking.



Fig. 4. Evacuation route of agent.

Table 1. the evacuation information retrieved from BIM model.

Floor	1	2
Flat	Flat B	Flat B
Section	Living room	Master bedroom
Agent	Agent 01	Agent 01
Minimum average ceiling height	2.8 m	2.1 m
Area	13.6 m ²	13.43 m ²
Effective smoke exhaust level	0 m ³ /min	0 m ³ /min
Exit width	1.2 m	1 m

4.2.2 Educational training module

Conventional fire-related education training is conducted through actual drills or educational videos. However, actual drills tend to require additional cost and time (e.g., creation of props and dispatch of training personnel), and the content in educational videos are easily forgotten. After comparing the actual evacuation time of the agent in the VR system with the allowable evacuation time estimated using the empirical equation, we verified that the evacuation route recommended by the VR system was a feasible route. On the basis of this finding, we created a navigation agent (Fig. 5) that guides users to evacuate to a safe area. In addition, the Unity particle system was employed to simulate an actual fire scene and increase users' feelings of tension and stress.



Fig. 5 Navigation agent.

5 Conclusion

In the present study, we proposed a BIM–VR-based fire evacuation and management model. Through the integration of the building information model and VR technology and the use of the navigation and script functions of Unity, the game agent was able to move to a safe area by taking the shortest evacuation route. We estimated the allowable evacuation time by using the empirical equation provided in the Building Fire Refuge Safety Verification Technique Manual and verified that the agent in the VR system could evacuate to the safe area through the shortest evacuation route within the estimated allowable time. In addition, the navigation agent was set to guide users to evacuate to the safe area through the verified evacuation route and assist users in identifying an evacuation route during a fire. We used the Unity particle system to simulate smoke to increase users' sense of tension and stress during a simulated fire scene; we also employed the Unity script function and an SQL Server to enable access to the firefighting equipment information of the building information model in the VR system, enabling users to understand information about firefighting equipment maintenance.

The case model was a fire-resistant building. The fire-resistant properties of the building (e.g., the situation where the building collapsed due to prolonged fire) were not considered. Furthermore, for the parameter setting of the agent, the physical and mental factors associated with dangerous situations (e.g., panic and coma) were not considered. Future studies should include the two aforementioned factors or other factors to enhance our proposed fire management model.

References

- [1] National Fire Agency, Ministry of the Interior. Statistical table of the number of fires, causes of

- fires and fire losses in Taiwan from January to September 2011. On-line: <https://www.nfa.gov.tw/cht/index.php>, Accessed: 03/11/2021.
- [2] Shih-Yu Shih. *Applying building information modeling in building disaster prevention and management*. master's thesis, Department of Civil Engineering National Chiao Tung University, 2013.
 - [3] June S., Cha H. S., and Jiang S, *Developing a Building Fire Information Management System Based on 3D Object Visualization*, 10(3), 772, Applied Sciences, 2020.
 - [4] Yung-Piao Chiu, Shu-Chen Lai, and Yan-Chyuan Shiau, *Study on Simulating Building Fire Evacuation Using Unity Software*, Taiwan Society Of Construction Engineers 12th, Chung Hua University, 2014.
 - [5] Wang B., Li H., Rezgui Y., Bradley A., and Hoang N. Ong, *BIM Based Virtual Environment for Fire Emergency Evacuation*, 1, 22, The Scientific World Journal, 2014.
 - [6] Cheng J. C. P., Tan Y., Song Y., Mei Z., Gan V. J. L., and Wang X, *Developing an evacuation evaluation model for offshore oil and gas platforms using BIM and agent-based model*, 89, 214-224, Automation in Construction, 2018.
 - [7] Unity Documentation, Navigation and Pathfinding. On-line: <https://docs.unity3d.com/Manual/Navigation.html>. Accessed: 28/09/2021.

Evaluation of Construction Site Layout Using Virtual Reality Linked with 3D CAD and Body Tracking

T. Hidetoshi^a and I. Kosei^b

^aGraduate School of Creative Science and Engineering, Waseda University Japan

^bDepartment of Architecture, Waseda University, Japan

E-mail: lkmk39@akane.waseda.jp

Abstract

Lately, virtual reality (VR) is being investigated for the field of construction. VR enables investigation from a stage where there is no actual object. The most common approach in VR is interaction with a controller. Therefore, the sense of body movements is often far from reality. Reflecting actual body movements is required for a more near-realistic experience in VR. In addition, VR linked to 3D CAD instantly enables the examination of design change. This paper proposes a method and framework for reflecting real body movements by tracking the body and real-time linkage of 3D CAD and VR. Azure Kinect and Leap Motion can track body movements and provide skeletal data. The VR body is linked to a real body by using a skeleton. The VR body enables to recognize how one's body moves in the VR space and its position in relation to other objects in the VR space. The ability to observe the distance between the body and VR space will expand the use of VR for this type of safety and buildability check. In addition, the tracked skeleton can be applied to measure the energy consumption of the body and extend the scope of analysis in VR. A real-time linkage between 3D CAD and VR was achieved by Rhinoceros and RhinoInside. In this study, steel-frame construction is used as an example of construction at heights and in confined spaces. The background behind the creation of this system is the liberalization of architectural shapes through 3D CAD. As the use of free shapes and dimensions in design and construction increases, the number of projects that cannot be handled by existing construction methods also increases. Therefore, a method for studying the implementation of each building should be developed to verify its buildability.

Keywords –

Virtual Reality; Construction Training; Tracking; Simulation; Parametric Modeling; Buildability

1 Introduction

Virtual reality (VR) has recently been investigated for construction purposes. VR enables training and examination in a completely safe environment [1]. Using head-mounted displays (HMDs), the environment created in the VR space can be viewed in three dimensions, which is an effective method for examination in construction where components are intricately linked. Furthermore, VR activities include reflections on real body movements, which are similar to real-life experiences [2]. Owing to these characteristics, VR can be considered as an effective tool for examining buildability. In this study, Unity was used to develop VR applications, and HTC VIVE Pro Eye was used for the HMDs.

This study presents a framework for evaluating parametric modeling using VR that incorporates body tracking. Parametric modeling creates a portion of the construction site and verifies buildability by searching for parameters. A steel-frame construction is used as an example of construction at heights and in confined spaces. 3D models created during design and construction studies are utilized in real time from 3D CAD and reflect real body movements. This enables verification of the relationship between the body and construction components in terms of buildability, for example, whether the envisioned site is sufficiently large for construction activity.

The liberalization of architectural shapes through 3D CAD is the background for the development of this system [3] [4]. It is possible to introduce programming into 3D CAD for parametric modeling. The use of parametric modeling is expanding, for example, in modeling small-scale buildings [5] and the construction of shield tunnels [6]. Therefore, the use of free shapes and dimensions in design and construction is expected to increase. However, some problems cannot be solved using existing construction methods. Therefore, it is important to confirm the buildability and safety of new construction methods.

2 Literature Review

2.1 Virtual Reality for Construction

Various studies have been conducted on VR technology, with VR in safety education being a typical example. Construction sites are considered dangerous environments, and safety education must be promoted; VR education is an alternative to passive learning methods [7]. Studies have been conducted in recent years to use VR not only for safety education but also for technical education. Adami et al. reported that training with VR improves knowledge, operating skills, and safety behaviors compared with face-to-face training [8]. They also cited cost as an advantage of VR, in addition to its superior safety features. Costs can be saved as real objects are no longer required and site constraints are eliminated. Furthermore, efforts to dynamically update VR environments also exist. Harichandran et al. proposed a framework for creating realistic and dynamically evolving VR training scenarios using digital twins [9]. In this study, we describe a method of linking 3D CAD and VR to speed up studies by reflecting changes in design and assumptions in real time, during construction studies.

2.2 Application of tracking technology

The tracking technology tracks the movements of real-world objects. An example of the use of tracking technology is to acquire data from remote local environments to operate robots remotely [10]. Tracking technology acquires the three-dimensional information of real objects. Commands can be transmitted based on three-dimensional data. Subsequently, a depth camera was used to acquire point cloud data as 3D data. Point cloud data can be used to record the shape of a building [11]. One of the advantages of using a depth camera is that there is no need to attach tracking equipment to the target object. Tracking technology can be applied to not only objects but also humans. In this study, tracking was targeted at human movement. When humans are the target of tracking, technology in the form of human skeletons is used to obtain information. Some construction projects have used tracked data of human skeletons, that is, acquiring the location of workers at a site using color-coded hard hats as tracking targets [12], and it has been suggested that the posture of workers can be analyzed. In this study, Azure Kinect [13] was used. Azure Kinect has a built-in depth camera and can acquire information about the human skeleton. There is a debate on whether the Azure Kinect is sufficiently accurate for use in energy consumption measurements. It has been suggested that the Azure Kinect could be used as an alternative to standard accelerometers [14]. In this study,

skeletal data acquired using the Azure Kinect were utilized to reproduce body movements in VR, by enabling the VR body model to follow the acquired skeletal information. Leap motion [15] is another technology for tracking parts of a body with high accuracy. Leap motion can track hand movements with a high accuracy. At construction sites, many operations involve hands. Therefore, tracking the movements of hands with high accuracy is important for reproducing the movements of the real body.

3 Reproduction of body movements in VR

3.1 Research Aim

This study presents a framework for evaluating parametric modeling using VR that incorporates body tracking. By reproducing body movements in VR, it is possible to evaluate the width and height of a construction site, based on one's own body. The model can also be evaluated by performing actions such as moving hands, stretching and contracting the body, and performing construction. This provides an experience similar to that of evaluating the size required for construction in a real mockup. Incorporating human behavior into a VR experience is an effective way to increase the level of detail in the reproduction of construction sites [16]. Furthermore, the parametric modeling transmitted to VR in real time enables the evaluation of dimensions with an efficiency that is not possible in reality because exporting objects from 3D CAD is not required. This system provides a method to examine, in an exercise format, whether construction work can be performed smoothly and safely in a prepared environment on a construction site, by visually checking the body and model of the job site.

3.2 The method of reproduction of body movements in VR

3.2.1 The elements for reproduction of body movements

Three main elements are required to reproduce the real body movements. The first is the use of HMDs, and there are various methods for using stationary monitors and HMDs to handle VR. However, when examining constructions using objects in a VR space, it is necessary to recognize the VR space as 3D. The second is to capture rough movements of reality. There are several techniques for this purpose, such as wearing a tracker on the body or using a camera. Considering the types of structures available in the construction industry and the large number of people involved, easy application is essential. To this end, tracking using a camera is considered an effective method, because people experiencing VR can

easily take turns. The third element is the tracking of individual body parts, which must be reproduced in detail. The operations performed by craftsmen at construction sites involve extensive detailed work. Therefore, when considering work in VR, it is necessary to reproduce movements in detail. For this it is necessary to track it separately from the rough movement tracking. In this study, rough movements were tracked using Azure Kinect, and detailed movements were tracked using Leap Motion.

3.2.2 Reproduction of whole-body movements using Azure Kinect

The Azure Kinect tracks whole-body movement. The information obtained through tracking is provided as a skeleton. The body model in VR moves as real movement using a skeleton. Movement control was performed using inverse kinematics (IK). We use the final IK [17] assets as a control method in Unity.

3.2.3 Reproduction of hands movements using Leap Motion

The Leap Motion tracks hand movements with high accuracy, and Azure Kinect also tracks hand movements. Therefore, when Leap Motion tracks the movements of the hands, the information transmitted to the IK is provided by the Leap Motion, and when it does not, the information transmitted to the IK is provided by the Azure Kinect. Furthermore, Leap Motion sends finger information that is directly reflected in the body model. Figure 1 shows the body movement reproduction system in a VR environment.

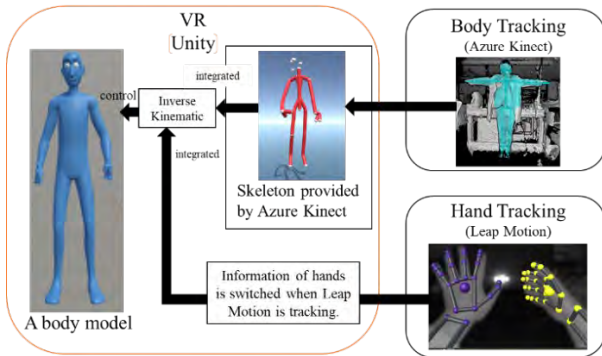


Fig.1 System of body movement reproduction in VR

4 Real-time linkage of 3D CAD and VR

4.1 Advantages of real-time linkage of 3D CAD and VR

By reflecting the 3D CAD model used in the VR design in real time, changes in the design plan can be immediately examined. Additionally, when a 3D model

is created such that the dimensions can be changed according to the parameters, the optimal values for the parameters can be determined without updating the 3D model. When a 3D model is used in VR, it must be exported as an object and then ported. However, by omitting this step, the time required can be shortened, and a speedy study can be achieved. In this study, Rhinoceros was used for 3D CAD, and RhinoInside [18] was used as an add-in to link Rhinoceros with Unity in real time.

4.2 Coordinate setting for 3D CAD

If the coordinates of the 3D model and HMD are incorrect, the positions of the models or HMD must be corrected after starting VR. This problem can be solved by setting the coordinates during the modeling stage. As the origin is shared between Rhinoceros and Unity, it is easy to handle the 3D model from the start of the interaction by setting the HMD as the origin in Unity and the center point of the location to be used in the 3D model as the origin in Rhinoceros. Ideally, the origin should be set at the start of modeling; however, if this is not the case, it is also effective to move the origin at the end of modeling. Furthermore, it is important to note that the scale varied between Rhinoceros and Unity. Unity was measured in meters, whereas rhinoceros was measured in millimeters. To smoothen the model for rhinoceros, we reduced the model transmitted to Unity by 1/1000.

4.3 Modeling of scaffold and steel-frame

Steel-frame 3D models were created using Rhinoceros software to evaluate the buildability of the scaffold. In this study, to examine the real-time reflection of changes in plan or design, 3D models were created using parametric modeling. Figure 2 and 3 show the parameters of the model. The parameters are as follows:

- (a) Width of scaffold
- (b1) Handrail depth (below)
- (b2) Handrail depth (above)
- (c) Width of pillar
- (d) Width of beam
- (e) Depth of beam

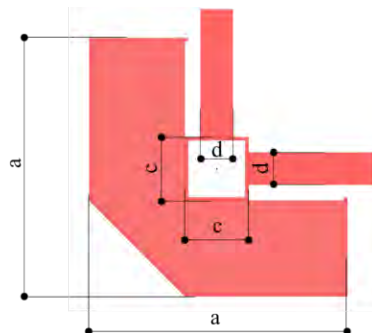


Fig.2 Top view of the steel-frame and scaffold models

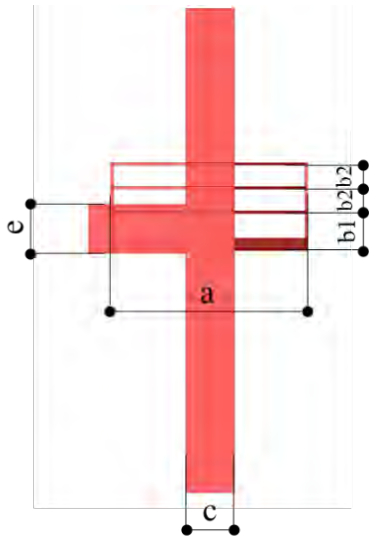


Fig.3 Front view of the steel-frame and scaffold models

4.4 Transmitting 3D CAD data to VR

RhinoInside transmits the Rhinoceros mesh data to Unity through a C# script. In this study, the Rhinoceros mesh was used as the default. Three-dimensional (3D) models were created using Grasshopper, a visual programming feature of Rhinoceros. The mesh list was transmitted to Unity using C# script. The material information was not transmitted but was provided to Unity as soon as it was received. Figure 4 shows a photograph of the grasshopper model. The program consisted of three parts. The first part is modeling, the second is listing, and the third is transmission. Any number of meshes can be supported by inserting the components to be listed and by interspersing the process of counting the list.

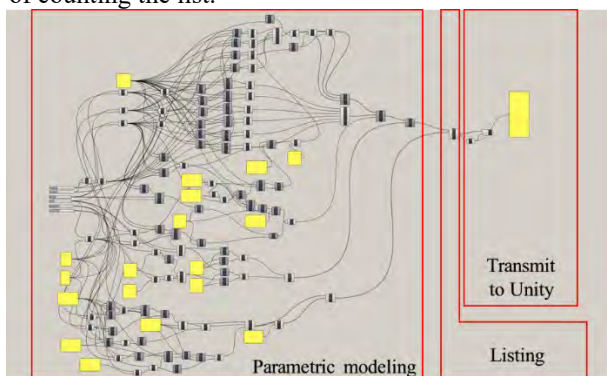


Fig.4 Composition of the grasshopper program

There are two advantages of transmitting each mesh individually. First, because mesh data are divided, hit detection is more likely to be reproduced for less complex shapes when a convex mesh collider is used. The higher the reproducibility of the hit decision, the

more accurate the simulation. The second advantage is that each component can be assigned a function in VR. For example, when adding functions such as gravity judgment and hand tracking, if the meshes are integrated, all of them will be affected. However, if the meshes are transmitted individually, it is possible to add functions only to the desired parts. Figure 5 shows the models transmitted from Grasshopper to Unity.

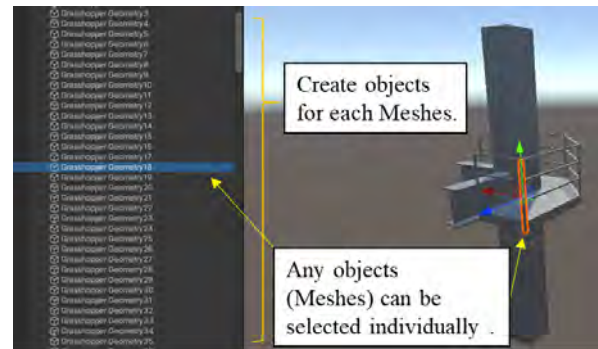


Fig.5 Models transmitted from grasshopper in Unity

4.5 Example of evaluation of construction site layout

First, we launched a VR device and then started linkage with 3D CAD. Once the linkage is initiated, the Rhinoceros model can be verified using VR. At this time, the 3D model is updated in real time, and changes to Rhinoceros can be immediately confirmed in VR. In the 3D model created in this study, the scaffold size, column height, handrail height, beam width, and beam height were set as parameters, and various dimensions were considered individually. In this study, the column height was fixed, and the width of the scaffold was varied to determine the necessary and sufficient width for safe construction. Additionally, when the construction area was fixed, the scaffold size was examined. The height of the scaffold railing can be determined to be safe and the appropriate height for construction can be determined. Figure 6 shows an example of checking the width of the scaffold. The most suitable width was determined by varying the width of the scaffold. Utilizing the reproduction of the body, the user performs movements such as turning the body and turning the hand, and examines whether the set parameters allow construction to be performed safely and constructively. Using various parameters, optimal scaffold dimensions were derived, and the use of the scaffold was envisioned to determine the actual temporary construction materials to be used. Without real-time linkage, the model must be exported every time this parameter is changed. However, this hassle was eliminated by sending mesh data in real time, allowing for a continuous study. The dimensions of the columns and beams can also be varied, which has the

advantage that multiple locations with different dimensions can be consecutively considered. Figure 7 shows examples from these studies. Even when examining multiple parameters, such as these, the 3D CAD model is immediately reflected, allowing for an examination with no time lag. By observing the body, it is possible to examine the width that can be walked, the width needed to move, and to determine the necessary

and sufficient parameters. While performing the construction operation, it is possible to examine whether the legs and body protrude from the scaffold, and whether there is excessive space. For the handrail, a gripping motion can be used to determine whether it is sufficiently or excessively high. Thus, the plan can be examined in VR in real time while planning for changes and decisions.

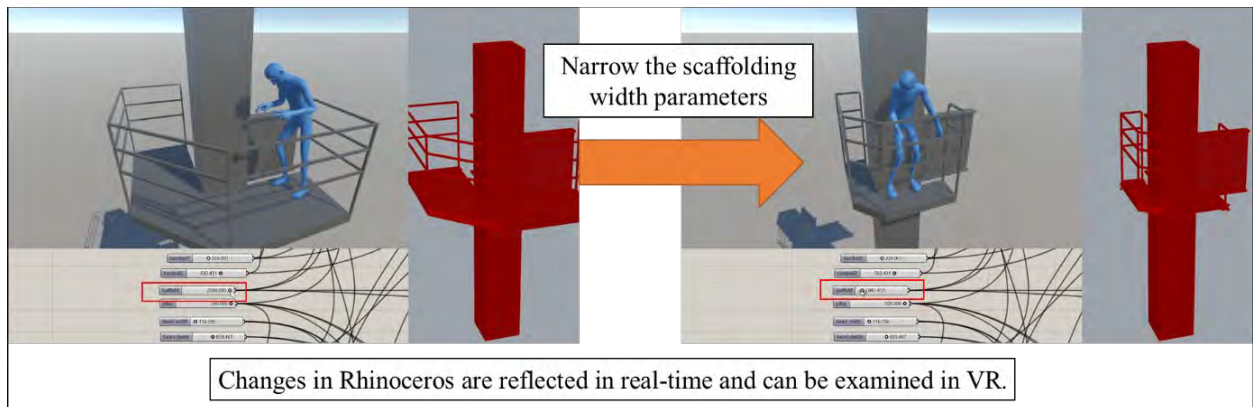


Fig.6 Checking the width of the scaffold

First-person view	Third-person view	Parameters(mm)	First-person view	Third-person view	Parameters(mm)
		(a) = 1600 (b1) = 305 (b2) = 200 (c) = 400 (d) = 200 (e) = 400			(a) = 1600 (b1) = 305 (b2) = 400 (c) = 400 (d) = 200 (e) = 400
First-person view	Third-person view	Parameters(mm)	First-person view	Third-person view	Parameters(mm)
		(a) = 2000 (b1) = 305 (b2) = 40 (c) = 400 (d) = 200 (e) = 400			(a) = 2000 (b1) = 305 (b2) = 400 (c) = 800 (d) = 200 (e) = 400
First-person view	Third-person view	Parameters(mm)	First-person view	Third-person view	Parameters(mm)
		(a) = 1200 (b1) = 305 (b2) = 400 (c) = 800 (d) = 200 (e) = 400			(a) = 1300 (b1) = 305 (b2) = 400 (c) = 800 (d) = 200 (e) = 400

Fig.7 Example of differences in first- and third-person viewpoint because of changes in parameters

5 Discussion

Using this system, we can immediately review the

plans after modeling, which enables the early expression of opinions in the process. Various people, such as construction workers, can easily express opinions, and they will connect to reflect workers' opinions on the

design.

One limitation of this system is that if the body is outside the range of the camera performing the tracking, real-life movements are not reflected in the VR. Therefore, it is difficult to study a wide range of areas. To improve this, it is necessary to set the means of movement in VR or attach a tracker to the body. In addition, this study confirms that only Rhinoceros and Unity can communicate in real-time. Therefore, if other applications are used, new means of communication will need to be developed, which is a limitation of this system. Large amounts of data may overload the computer and cause a lag, or may not be transmitted.

There are several possible developments for this system. Further information can be analyzed using the method of connecting applications. For example, by transmitting the data of skeletons tracked by Azure

Kinect, the energy consumption of humans was analyzed. If Rhinoceros is used, a plugin such as Karamba 3D [19] can also be used. The advantage is that a single application can perform multiple modeling and analysis tasks simultaneously, and is simple to configure. Figure 8 shows the configuration of the system. 3D CAD, tracking systems, and VR were linked based on the VR platform. The system is composed of components that select the information to be used in VR and process it appropriately. In this study, the elements used in these functions were investigated. Therefore, if the necessary elements and data transmission methods are available, any application or platform can perform in a similar manner. In addition, as more functionality is required, similar discussions will enable the integration of additional systems into the development of VR applications.

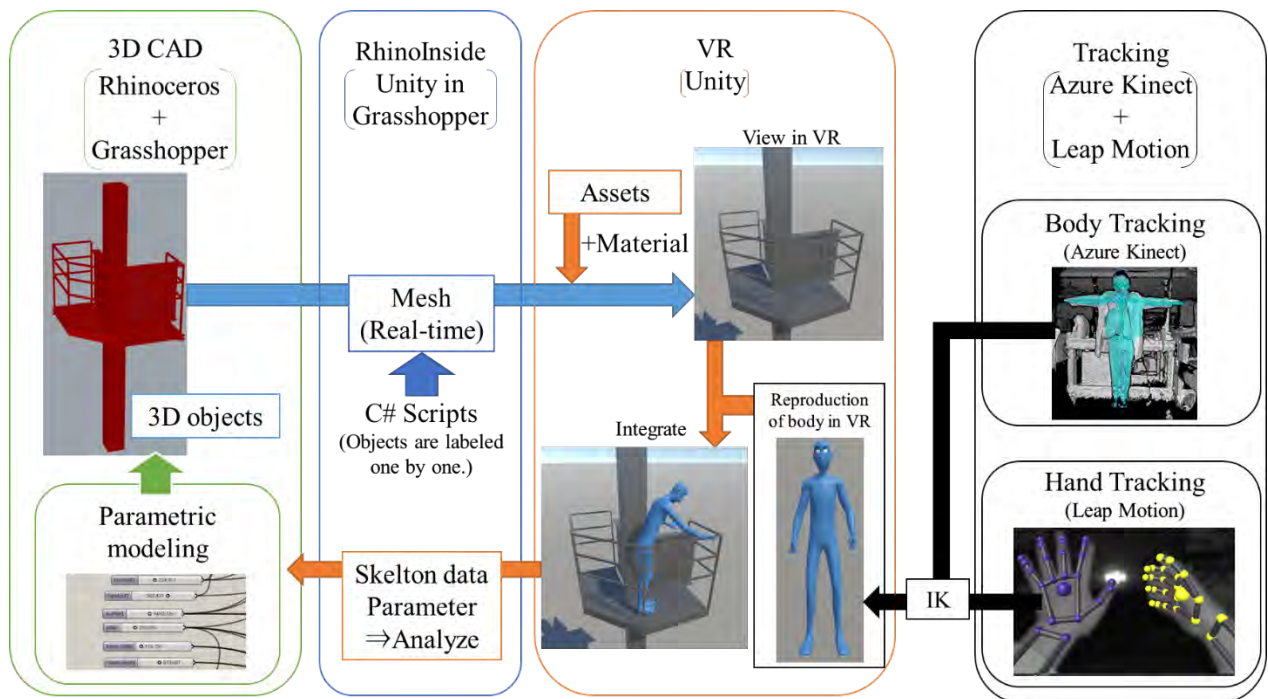


Fig.8 System configuration of VR system using tracking and 3D CAD

6 Conclusion

In this paper, the VR method connected to 3D CAD and the tracking system in real time were explained. The real-time tracking and reproduction of a body in VR can verify buildability by viewing the body and elements of a construction project. The real-time linkage of 3D CAD and VR enables us to view the plan and design changes instantly, as well as reflect the results of the study for planning. These linkages were achieved by examining the elements used in the functions. Rhinoceros was used for 3D CAD, Unity for the VR platform, Azure Kinect

for body tracking, and Leap Motion for hand tracking. If these elements are in place, other 3D CAD, VR platforms, and tracking systems can be used. Furthermore, information from tracking and VR applications can be analyzed. Using this system, we can review the plans immediately after modeling, and express opinions. Various people will easily express opinions, and they will connect to reflect workers' opinions on design.

Acknowledgements

We would like to thank Editage (www.editage.com)

for English language editing.

References

- [1] Alan Ferguson. Virtual reality and safety training. On-line: <https://www.safetyandhealthmagazine.com/articles/19440-virtual-reality-and-safety-training>, Accessed: 8/2/2022.
- [2] Program ace. 7 Ways to benefit from VR in construction. On-line: <https://program-ace.com/blog/vr-in-construction/>, Accessed: 8/2/2022.
- [3] Zaha Hadid Architects. Zaha Hadid architects. On-line: <https://www.zaha-hadid.com>, Accessed: 8/2/2022.
- [4] A.R.U.P. Arup. On-line: <https://www.arup.com/>, Accessed: 8/2/2022.
- [5] Adeline Stals, Sylvie Jancart, and Catherine Elsen. Parametric modeling tools in small architectural offices: Towards an adapted design process model. *Design Studies*, 72, article 100978, 2021, doi: [10.1016/j.destud.2020.100978](https://doi.org/10.1016/j.destud.2020.100978).
- [6] Hanbin Luo, Linhao Li and Ke Chen. Parametric modeling for detailed typesetting and deviation correction in shield tunneling construction. *Automation in Construction*, 134, article 104052, 2022, doi: [10.1016/j.autcon.2021.104052](https://doi.org/10.1016/j.autcon.2021.104052).
- [7] Pooya Adami, Patrick B. Rodrigues, P.J. Woods, Burcin Becerik-Gerber, Lucio Soibelman, Yasemin Copur-Gencturk and Gale Lucas. Effectiveness of VR-based training on improving construction workers' knowledge, skills, and safety behavior in robotic teleoperation. *Advanced Engineering Informatics*, 50, article 101431, 2021, doi: [10.1016/j.aei.2021.101431](https://doi.org/10.1016/j.aei.2021.101431).
- [8] Mikko Nykänen, Vuokko Puro, Maria Tiikkaja, Henriikka Kannisto, Eero Lantto, Frans Simpura, Jose Uusitalo, Kristian Lukander, Tuula Räsänen, Tarja Heikkilä and Anna-Maria Teperi. Implementing and evaluating novel safety training methods for construction sector workers: Results of a randomized controlled trial. *Journal of Safety Research*, 75, pp. 205-221, 2020, doi: [10.1016/j.jsr.2020.09.015](https://doi.org/10.1016/j.jsr.2020.09.015).
- [9] Aparna Harichandran, Karsten W. Johansen, Emil L. Jacobsen and Jochen Teizer. A conceptual framework for construction safety training using dynamic virtual reality games and digital twins. *International Symposium on Automation and Robotics in Construction*, 38th, pp. 621-628, 2021, doi: [10.22260/ISARC2021/0084](https://doi.org/10.22260/ISARC2021/0084).
- [10] Yufeng Li, Jian Gao, Xinxin Wang, Yimin Chen and Yaozhen He. Depth camera based remote three-dimensional reconstruction using point incremental cloud compression. *Computers and Electrical Engineering*, 99, article 107767, 2022, doi: [10.1016/j.compeleceng.2022.107767](https://doi.org/10.1016/j.compeleceng.2022.107767).
- [11] Kosei Ishida. Evaluation of drainage gradient using 3D measurement data and a physics engine. *International Symposium on Automation and Robotics in Construction*, 37th, pp.1037-1044, 2020, doi: [10.22260/ISARC2020/0143](https://doi.org/10.22260/ISARC2020/0143).
- [12] Bereket H. Woldegiorgis, Chiuhsiang J. Lin and Riotaro Sananta, Using Kinect body joint detection system to predict energy expenditures during physical activities, *Applied Ergonomics*, 97, article 103540, 2021, doi: [10.1016/j.apergo.2021.103540](https://doi.org/10.1016/j.apergo.2021.103540).
- [13] Microsoft. Azure Kinect DK documentation. On-line: <https://docs.microsoft.com/en-us/azure/Kinect-dk/>, Accessed: 10/2/2022.
- [14] I.P. Tharindu Weerasinghe, Janaka Y. Ruwanpura, Jeffrey E. Boyd and Ayman F. Habib. Application of Microsoft Kinect sensor for tracking construction workers. *Construction Research Congress 2012*, pp. 858-867, 2012, doi: [10.1061/9780784412329.087](https://doi.org/10.1061/9780784412329.087).
- [15] Ultraleap. Digital worlds that feel human. On-line: <https://www.ultraleap.com/>, Accessed: 12/2/2022.
- [16] Tsuda Hidetoshi and Ishida Kosei. Study on the reproduction of human motion units in construction work in VR space. *Proceeding of Architectural Research Meeting, Kanto Chapter, Architectural Institute of Japan*, I (91), pp.473-476, 2022.
- [17] Unity technologies. Final IK | Animation tools | Unity asset store. On-line: https://assetstore.unity.com/packages/tools/animation/final-ik-14290?aid=1100IGdc&utm_campaign=unity_affiliate&utm_medium=affiliate&utm_source=partnerize-linkmaker, Accessed: 12/2/2022.
- [18] GitHub. Rhino.Inside®. On-line: <https://github.com/mcneel/rhino.inside>, Accessed: 13/2/2022.
- [19] Karamba. Karamba3D – Parametric engineering. On-line: <https://www.karamba3d.com/>, Accessed: 13/2/2022.

Automated Recognition of Hand Gestures for Crane Rigging using Data Gloves in Virtual Reality

A. Harichandran^a and J. Teizer^b

^aDepartment of Civil and Architectural Engineering, Aarhus University, Denmark

^bDepartment of Civil and Mechanical Engineering, Technical University of Denmark, Denmark

E-mail: aparnaharichandran@cae.au.dk, teizerj@byg.dtu.dk

Abstract –

Construction safety training through Virtual Reality (VR) environments offers workers an interactive and immersive experience. Many complex interactions and realistic scenarios are possible in VR using accessories such as data gloves and trackers which allow data recording on incidents. For example, misinterpretation of hand signals construction workers give during operations may result in severe accidents on construction sites. This study proposes automatic gesture recognition to identify hand signals for crane rigging operations in VR training. The developed gesture recognition algorithm tracks information from a data glove and a tracker. Preliminary data on movements and orientation of hands and fingers were recorded. Mathematical models of hand gestures were created based on finger movement data. Gesture rules were created based on the rotation and orientation of the hand. The gesture models were combined with the gesture rules to develop the algorithm for automated gesture recognition. Final experiments estimated the efficacy of the proposed method in automatically recognizing crane rigging signals in real-time. The performance is evaluated by comparing the identified hand gestures with independently created ground truth labels. The proposed method identified static hand gestures with an average accuracy of 96.55 percent. This method recognizes the gestures along with the hand movements and displays the results in real-time equivalent to dynamic gesture recognition. More refined dynamic gesture recognition based on this method is in progress.

Keywords –

Gesture recognition; Virtual reality; Construction safety; Safety training; Data gloves; Hand motion tracking

1 Introduction

Virtual Reality (VR) environments provide new

possibilities for advancing construction safety training. These computer-aided training methods develop an interest in workers and ensure their active participation [1], [2]. The workers can be subjected to complex and dangerous scenarios in the virtual environment without the possibility of injuries [3]. The trainees can receive personalized feedback based on their performance assessed from the data collected during training [4]. Multiplayer serious games create collaborative learning spaces for the participants [5] and dynamically updated VR training exposes the trainees to the latest work environment according to the construction progress [6].

Hand gestures are widely used for communication on construction sites. They enable effective communication irrespective of the construction noises and language barrier between the workers [7]. There are standard hand signals for operations such as crane rigging [8]. However, misinterpretation of these predefined hand signals by the operator of the cranes or any other machines may cause severe accidents [9]. According to the U.S Bureau of Labor Statistics, an average of 42 crane-related deaths occur per year and 43 percent of the fatal crane injuries are from the construction industry [10]. Human error account for 90 per cent of crane accidents and proper training of crane operators and signalers is essential [11]. Automatic recognition and interpretation of the hand gestures may potentially assist in effective communication between the signaler and the operator. It improves communication in VR training scenarios involving multiple workers similar to the actual construction site. Automated gesture recognition in VR coupled with the collection of trainees' behavioral data [6] can quantitatively estimate the effectiveness of VR based training methods. The objective of the current study is to automatically recognize hand signals for crane rigging in virtual reality (VR) through data gloves. The scope of this study is limited to identifying six classes of hand gestures: five hand signals, including stop, raise boom, lower boom, hoist load and lower load, and any other unidentified hand gestures.

This paper is organized as follows. Section 2 provides the background on gesture recognition in VR and related

terminologies. The study's methodology is described in Section 3 and an overview of the experiment and data collection is provided in Section 4. The development of the gesture recognition algorithm is in Section 5. Section 6 presents the gesture recognition method and Section 7 describes its validation. The results and discussion are given in Section 8. Finally, Section 9 concludes the paper with findings and an outlook for future work.

2 Background

2.1 Hand Anatomy and Movements

The human hand is a complex biomechanical device that evolved over millions of years to achieve the current level of motor skills. It is capable of performing numerous tasks that involve a variety of movements. A typical human hand consists of a wrist, a palm and five fingers. The hand is composed of 27 bones as illustrated in Figure 1. The bones are categorized as 1) carpals in the wrist [short bones, 8 no.], 2) metacarpals in the palm [long bones, 5 no.], and 3) phalanges in fingers [long finger bones, 14 no.]. Each of the four fingers except the thumb is composed of proximal phalange, intermediate phalange, and distal phalange. The thumb has only two phalanges, proximal phalange and distal phalange. The placement of the metacarpal bone of the thumb enables its distal phalanges to oppose the distal phalanges of other fingers. This configuration of the thumb allows humans to grab objects in hand.

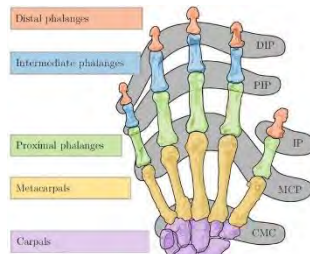


Figure 1. Finger joints of a human hand [12].

The joints in the hand facilitate various movements of fingers, as illustrated in Figure 2. Hinge joints provide 1 DOF (Degree of Freedom), i.e. flexion or extension. Saddle joints provide 2 DOF, i.e., flexion or extension and abduction or adduction. The metacarpophalangeal (MCP) is a saddle joint, whereas proximal interphalangeal (PIP), distal interphalangeal (DIP) and interphalangeal (IP) are hinge joints. In addition to carpometacarpal (CMC) joints, the thumb has two joints (MCP and IP) and all other fingers have three joints (MCP, PIP, and DIP).

2.2 Gesture Interactions and Recognition in Virtual Reality

Gesture interactions in VR are facilitated through different input devices such as wearable sensor-based devices, touch-based devices, and computer vision-based devices [13]. The wearable devices include data gloves, inertial sensors and myoelectricity sensors [14], [15]. The touch-based devices comprise touch screens and stylus pens. Different types of cameras such as monocular, binocular, and depth cameras constitute computer vision-based devices. The data gloves can act as an input device to collect finger postures and movements. Besides, some data gloves have additional features to provide haptic feedback to the users, creating a more realistic experience [16]. Some of the advantages of data gloves include high recognition accuracy, no environmental influence, small data sets, and low computational power requirement. However, it has some shortcomings such as high cost, low flexibility and the need for frequent calibration.

Gesture recognition methods in VR can be categorized as methods based on 1) wearable devices, 2) touch technology, 3) computer vision, and 4) multimodal interaction technology. Data gloves are one of the most commonly used wearable devices in VR. The gesture recognition based on wearable devices involve collecting finger posture data, extracting spatiotemporal parameters, selecting effective parameters, and model training [17] or identification by an intelligent algorithm [18]. The gesture recognition based on touch technology can be further divided into single touch and multitouch recognition. The \$1 algorithm is a simple algorithm for identifying single-stroke gestures [19]. It has been modified to identify multi-stroke gestures with reduced complexity using point clouds [20]. Currently, computer vision-based gesture recognition methods are being widely implemented. This method consists of data collection through cameras, preprocessing, gesture segmentation, gesture analysis and gesture recognition [21]. Multimodal interaction technologies use two or more modalities of communication to recognize the instructions from the user [22]. These methods attempt to incorporate a more natural way of communication by the user for virtual interaction.

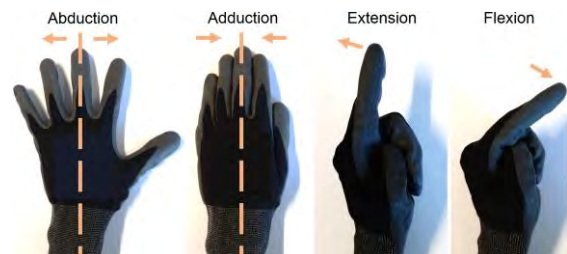


Figure 2. Motions of human fingers of a construction worker wearing protective gloves.

2.3 Significance of the Current Study

Data-driven methods such as deep learning have been widely used for gesture recognition. However, the existing methods have several drawbacks. These methods often require large datasets for training that may not be available for newly created VR training scenarios. The performance of computer vision-based methods depends on environmental factors such as light, skin color, and occlusion. The methods require several image processing techniques which might affect the recognition accuracy. The computer vision-based method may not capture the signals correctly in crane rigging operations unless multiple cameras are installed at different levels. Autonomous cranes might depend on the ground-based human operator who wears intelligent gloves.

The interactions with VR controllers often lack the realistic experience during construction safety training. Besides, the nuances of finger movements and seamless coordination between real and virtual avatars are essential in collaborative training environments[23]. Therefore, data gloves and trackers were introduced to enhance the interactive and realistic experience in VR training. Automatic gesture recognition during the training potentially improves the communication between the trainees. Thus, a gesture recognition method for VR training has been developed in this study. A training scenario involving a signaler and a crane operator is envisioned where the communication has been enhanced by automatic gesture recognition. A gesture recognition algorithm has been developed based on the information streams from data gloves and trackers. The proposed method identifies and displays the gestures in real-time during the VR training.

3 Methodology

The overall methodology for this study consists of experiments and data analytics. First preliminary experiments were conducted to collect various data such as finger movement and hand orientation required for gesture recognition. Then, mathematical models of the hand gestures were created based on finger movement data. These gesture models were combined with the orientation data to develop the algorithm for automated gesture recognition. After that, experiments were designed to capture the efficacy of the proposed method in automatically recognizing crane rigging signals. Next, experiments were conducted to collect the data in run-time and implement the proposed method. The performance of the method is evaluated by comparing the predicted hand gestures with independently created ground truth labels. The subsequent sections will describe more details regarding each of these steps.

4 VR Experiments and Data Collection

The current study conducts virtual reality experiments in two stages. The serious games for the experiments were developed in the game engine software Unity. The preliminary experiments in the first stage are for developing the gesture recognition algorithm (described in Section 6). The VR experiments in the second stage are for validating the gesture recognition method (described in Section 8). The user wears a data glove and tracker while making the specified gestures in both cases. The data corresponding to hand and finger movements were collected in real-time during the experiments. Figure 3 shows the user wearing the data glove and tracker ready for the experiment. The Prime X Haptic VR data glove [24] is used for finger tracking. This data glove contains a flex sensor skeleton to capture the bending of the fingers. It also contains an Inertial Measurement Unit (IMU) with nine DOF (Degree of Freedom) per finger to capture relative movements. Collection and visualization of the data from the data glove are enabled by the data handling software Manus Core. This software also helps to live stream the finger movement data to Unity. The VIVE Tracker (3.0) [25] is mounted on the data glove to track hand movements. The tracker enables seamless coordination between the real hand and its virtual counterpart in the VR environment. The tracking data is collected and streamed to Unity through the SteamVR application [26].



Figure 3. A user wearing the data glove and tracker ready for gesture recognition.

5 Development of Gesture Recognition Algorithm

An overview of developing the gesture recognition algorithm is illustrated in Figure 4. It starts with preliminary experiments containing five selected hand signals for crane rigging. The experiments were conducted to understand the nature of hand gesture data for developing the recognition algorithm. The user makes the selected gestures in the experiment wearing a data glove and a tracker. Each set of the experiment contains one hand gesture. The current study has selected five crane rigging signals as shown in Figure 5: 1) stop, 2)

raise boom, 3) lower boom, 4) hoist load, and 5) lower load. These hand signals are dynamic i.e., they involve movements of the hand along with the hand gestures. The proposed method is designed to recognize hand gestures in a static position (static hand signals). However, continuous real-time identification of gestures enables dynamic gesture recognition.

The finger data is live-streamed to the Manus Core and visualized in the Manus dashboard. The data viewer displays values of flexion, extension, abduction, adduction, thumb rotation and wrist rotation. The flexion and extension are estimated with respect to the finger joints, whereas abduction and adduction are estimated with respect to the midline of the hand and a finger joint. The finger data is simultaneously streamed from the Manus Core to Unity. Currently, the gesture models were created in Unity as described in the next paragraph.

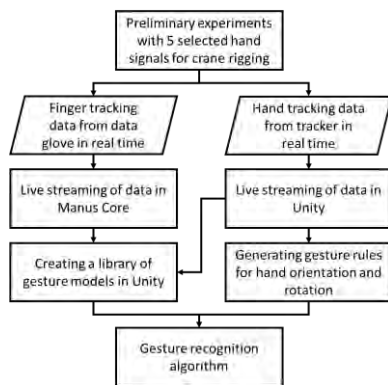


Figure 4. Development of gesture recognition algorithm

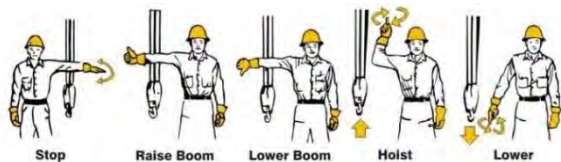


Figure 5. Selected crane rigging signals for this study [8]

Three hand gestures were selected to represent the finger positions in the five crane rigging signals. The selected hand gestures are: 1) 'thumbs up' (only thumb is straight, all other fingers are flexed), 2) 'pointing' (only index finger is straight, all other fingers are flexed) and 'high five' (all fingers are straight). Note that only finger movements can be tracked with the data glove. Additional information about the hand movement is required for identifying the crane rigging signals. Mathematical models were created for the three selected gestures in Unity. The mathematical model of a gesture defines the movement of each finger in a relative scale with respect to the finger joints and/or midline of the hand. Flexion and extension of a thumb are specified

based on CMC, MCP and IP; and that of other fingers based on MCP, PIP and DIP. Abduction and adduction of a thumb are specified based on CMC, whereas that for other fingers based on MCP. Thus, a library of predefined gesture models was created in Unity. Currently, the library contains three gesture models, each of which represents the finger positions of the selected hand signals as shown in Table 1.

After creating the gesture library, gesture rules were generated from the hand tracking data. The movements of the hand are tracked in real-time through the VIVE tracker. The real hand will appear as a game object in Unity and the same object will be seen by the users in the virtual environment. The game object of the hand is hereafter referred to as the hand object. The hand object has an attribute called 'transform' that contains the position, rotation and scale of the object in the virtual environment. The current study uses the transform of the hand object to create rules for recognizing hand gestures.

Table 1. Hand signals and corresponding gesture models in the library

Label	Hand signal	Gesture model	Need more information for recognition?
0	No recognition	-	Yes
1	Stop	HighFive	Yes
2	Raise boom	ThumbsUp	Yes
3	Lower boom	ThumbsUp	Yes
4	Hoist load	Pointing	Yes
5	Lower load	Pointing	Yes

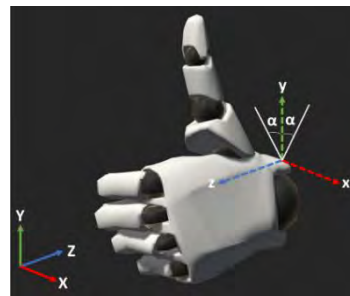


Figure 6. Schematic of the Gesture rule for raise boom hand signal

Consider the example of the hand signals for 'raise boom' and 'lower boom'. Both of these hand signals have the same gesture model (ThumbsUp) to represent their finger positions. Therefore, rotation or orientation of the hand in the virtual environment is essential to distinguish between these hand signals. Thus, the gesture rules for recognizing these hand signals involve a specific range of values for these parameters from the viewer's perspective. Figure 6 shows a schematic representation of one of the gesture rules for the raise boom hand signal.

Here, the solid arrow marks represent the global coordinate system (X, Y, Z) for the virtual environment and dotted arrow marks represent the location coordinate system (x, y, z) for the hand object. The tolerance of rotation of the hand object about the y axis is denoted by α . The gesture rule in this scenario is: if $\alpha \leq 30^\circ$ for the gesture model 'ThumbsUp', the hand gesture is 'raise boom'. Note that this is a simplified illustration of a gesture rule. Orientation and rotation of the hand object with respect to all other axes will be specified in the actual gesture rule. The gesture recognition algorithm is developed by combining the gesture rules and gesture models. A detailed description of the gesture recognition method is given in the next section.

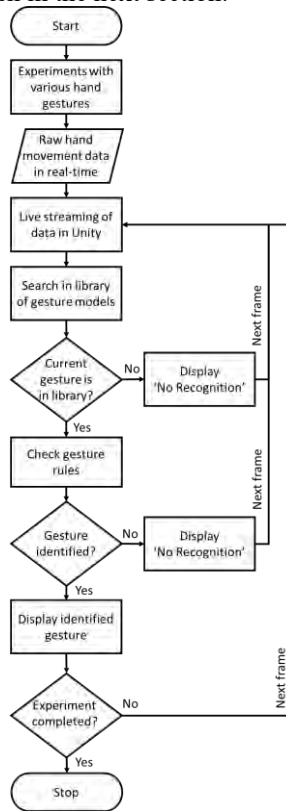


Figure 7. Gesture recognition method.

6 Gesture Recognition Method

The automated gesture recognition method proposed in this study is shown in Figure 7. First, the algorithm is implemented in actual virtual reality experiments containing various hand gestures. The raw hand movement data collected by the tracker and the finger movement data from the data glove are live-streamed into Unity. The gesture recognition algorithm is attached as a script to the hand object in Unity. The recognition algorithm runs in a fixed interval for accurately capturing the physics movements of the hand object. The gesture

data from the user is evaluated in each run. First, the gesture library is searched to see if the current finger movement data match any of the predefined gesture models. If none of the gesture models matches the current gesture data, display 'No recognition' and proceed to the next frame. If any of the gesture models match with the current finger movement data, check the associated gesture rules. The orientation and rotation of the hand object are estimated, and the gesture rules are evaluated. If none of the rules is satisfied, display 'No recognition' and proceed to the next frame. Otherwise, determine the hand signal based on the gesture rules satisfied. Then display the identified hand signal and check whether the experiment is completed. If the experiment is complete stop the iteration. Otherwise, proceed to the next frame.

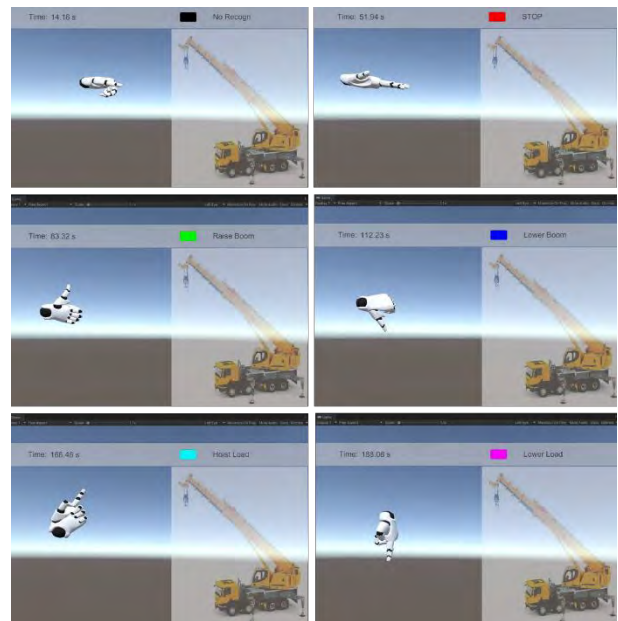


Figure 8. Conveying hand signals to the crane operator in virtual reality. The real-time predictions by the recognition method for each hand signal are displayed with the time stamp.

7 Validating Gesture Recognition Method

The proposed gesture recognition method (Figure 7) is validated by virtual reality experiments where the user acts as a signaler for a crane operator. In the experiments, the user wearing the data glove and the tracker make various hand gestures for a specific interval. The gestures involve hand signals for crane rigging operations and some random gestures. The gesture data are collected from the data glove and trackers in real-time during the experiments. Simultaneously the collected data is analyzed using the proposed method and results were also displayed in real-time. The display contains the predicted hand gesture, an associated color and time

stamps as shown in Figure 8. The predictions are also logged in a text file (.txt) as entries in the format “time (in seconds), predicted gesture, label”; e.g., “72.82 RaiseBoom 2”. The entire experiment is recorded in the game view using the Unity recorder. The recorded videos of the experiments with time stamps were used to create ground truth labels for the gestures. The experiment was repeated five times. The accuracy of the recognition method is estimated as an average of all the repetitions.

8 Results and Discussion

The results of gesture identification and related discussion are presented in this section. The accuracy of the proposed recognition method for the dataset generated from each repetition of the experiment is presented in Table 2. The gesture recognition method has an average accuracy of 96.55 percent. The method delivers an accuracy above 95 % for all datasets except the first one. The proposed method is designed to recognize static hand gestures. However, the experiments for validation used dynamic gestures to understand the potential of the method in identifying actual crane rigging signals. Nevertheless, the gesture recognition method delivered an overall good performance.

Table 2. Performance of the recognition method.

Dataset	Accuracy
1	93.86%
2	96.86%
3	98.58%
4	95.60%
5	97.87%
Average accuracy	96.55%

The recognition results for dataset 4 is illustrated in Figure 9 and the results with highlighted misclassifications are presented in Figure 10. The gesture recognized by the proposed method is plotted in grey and the ground truth values are plotted in blue. The lines overlap whenever the recognized gesture is correct. Thus, the misclassifications can be seen are the misaligned parts of the grey line that was highlighted in black in Figure 10. The hand signal for ‘stop’ (Label 1) and other undefined or random gestures (Label 0) are identified without any mistakes. However, some instances of the hand signals for ‘raise boom’ (Label 2), ‘lower boom’ (Label 3), ‘hoist load’ (Label 4) are misidentified as the undefined class. The number of misidentifications is fewer compared to the frequency of the function call (50 times per second) for gesture recognition. Therefore, these misidentifications may not result in serious

communication problems while displaying the results continuously. Similarly, some instances of the hand signal ‘lower load’ (Label 5) are wrongly identified as ‘hoist load’ (Label 4). These misidentifications need to be addressed carefully since they belong to the opposite category of hand signals. The potential reasons for the misidentifications are the strict bounds for hand orientation and rotation. More flexible and robust gesture rules are being explored with information from additional trackers.

The overall accuracy of the recognition method per hand gesture is given in Table 3 and illustrated in Figure 11. All gestures have been recognized with accuracy above 90 percent, and some of them have close to 100 percent accuracy. Therefore, the proposed method shows the potential for improving the VR safety training scenarios involving construction site communication. The undefined or random gestures and hand signals for ‘stop’, ‘lower boom’ and ‘hoist load’ are identified with high accuracy. However, identifying the hand signals for ‘raise boom’ and ‘lower load’ needs further improvement. Since opposite hand signals of these classes were identified with high accuracy, the gesture models seem robust to represent the finger movement. More attention is required to refine the gesture rules that define hand orientation and rotation during signaling.

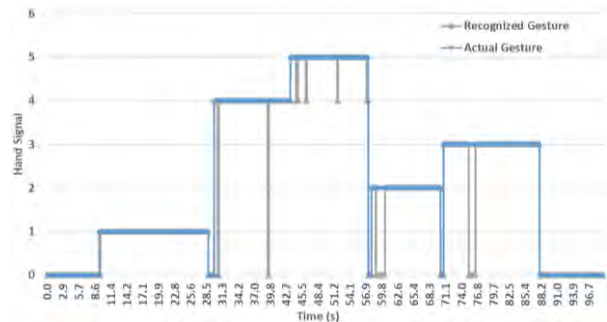


Figure 9. Gesture recognition results for dataset 4. Gesture labels are 0: No recognition, 1: Stop, 2: Raise boom, 3: Lower boom, 4: Hoist load, and 5: Lower load.

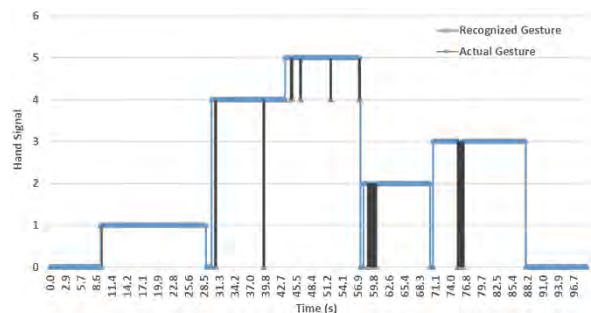


Figure 10. Gesture recognition results for dataset 4 with misidentifications highlighted in black. Gesture labels are 0: No recognition, 1: Stop, 2:

Raise boom, 3: Lower boom, 4: Hoist load, and 5: Lower load.

Table 3. Performance of recognition per hand gesture.

Label	Hand gesture	Overall prediction accuracy
0	No recognition	99.78%
1	Stop	99.29%
2	Raise boom	90.61%
3	Lower boom	98.60%
4	Hoist load	97.64%
5	Lower load	91.13%

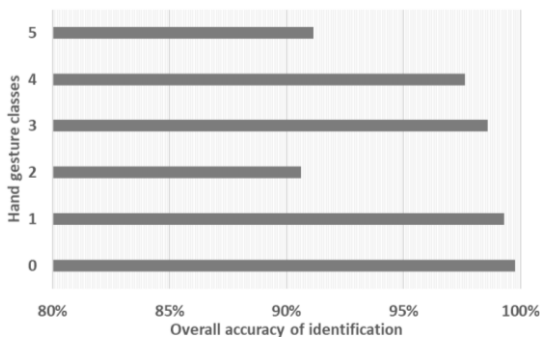


Figure 11. Accuracy of identification per hand gesture classes.

9 Conclusions and future work

An automatic gesture recognition method for identifying hand signals for crane rigging in virtual reality is proposed in this study. The gesture recognition method is developed based on tracking information from a data glove and a tracker. Gesture models were created to represent the finger movement during hand signals. The rotation and orientation of the hand for a signal are defined by gesture rules. A gesture recognition algorithm is developed by combining the gesture models and rules. The proposed method is validated by virtual reality experiments containing various hand gestures.

The gesture recognition method delivered an average accuracy of 96.55 percent. Most of the gesture classes were identified with high accuracy. The misidentifications were mainly attributed to the bounds of the gesture rules. More robust gesture rules are being developed based on independent information from other tracking devices. The good recognition performance shows that this method can improve VR safety training scenarios involving communication between workers. Incorporating data gloves in VR helps articulate the hand signals better than conventional controllers.

This study can be further extended to quantitatively estimate the effectiveness of the current communication

methods that are being utilized in the construction site operations. Although the proposed method is designed to recognize static hand gestures, it can continuously recognize the hand gestures and display the results in real-time. Therefore, it is equivalent to dynamic gesture recognition to a certain extent. The future study involves incorporating additional trackers to capture more complex and dynamic hand gestures. Besides, haptic feedback from the data glove is being implemented to enhance the learning experience of the workers.

References

- [1] Y. Gao, V. A. Gonzalez, and T. W. Yiu, "The effectiveness of traditional tools and computer-aided technologies for health and safety training in the construction sector: A systematic review," *Computers & Education*, vol. 138, pp. 101–115, 2019, doi: 10.1016/J.COMPEDU.2019.05.003.
- [2] L. Chittaro and F. Buttussi, "Assessing knowledge retention of an immersive serious game vs. A traditional education method in aviation safety," *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, no. 4, pp. 529–538, 2015, doi: 10.1109/TVCG.2015.2391853.
- [3] O. Golovina, C. Kazanci, J. Teizer, and M. König, "Using serious games in virtual reality for automated close call and contact collision analysis in construction safety," in *36th International Symposium on Automation and Robotics in Construction (ISARC 2019)*, 2019, pp. 967–974. doi: 10.22260/isarc2019/0129.
- [4] M. Wolf, J. Teizer, B. Wolf, S. Bürkü, and A. Solberg, "Investigating hazard recognition in augmented virtuality for personalized feedback in construction safety education and training," *Advanced Engineering Informatics*, vol. 51, p. 101469, 2022, doi: 10.1016/J.AEI.2021.101469.
- [5] E. L. Jacobsen, N. S. Strange, and J. Teizer, "Lean Construction in a Serious Game Using a Multiplayer Virtual Reality Environment," in *Proc. 29th Annual Conference of the International Group for Lean Construction (IGLC)*, Jul. 2021, pp. 55–64. doi: 10.24928/2021/0160.
- [6] A. Harichandran, K. W. Johansen, E. L. Jacobsen, and J. Teizer, "A Conceptual Framework for Construction Safety Training using Dynamic Virtual Reality Games and Digital Twins," in *38th International Symposium on Automation and Robotics in Construction (ISARC 2021)*, 2021, pp. 621–628. doi: https://doi.org/10.22260/ISARC2021/0084.
- [7] P. D. Bust, A. G. F. Gibb, and S. Pink,

- “Managing construction health and safety: Migrant workers and communicating safety messages,” *Safety Science*, vol. 46, no. 4, pp. 585–602, 2008, doi: 10.1016/J.SSCI.2007.06.026.
- [8] Occupational Safety and Health Administration (OSHA), *29 CFR Part 1926 Cranes and Derricks in Construction; Final Rule*. USA, 2010. Accessed: Feb. 27, 2022. [Online]. Available: www.osha.gov
- [9] X. Wang and Z. Zhu, “Wearable Sensor-based Hand Gesture Recognition of Construction Workers,” in *38th International Symposium on Automation and Robotics in Construction (ISARC 2021)*, 2021, pp. 498–504.
- [10] U.S Bureau of Labor Statistics, “Fatal Occupational Injuries Involving Cranes,” 2019. <https://www.bls.gov/iif/oshwc/cfoi/cranes-2017.htm>, Accessed: 04/20/2022.
- [11] Crane Inspection & Certification Bureau (CICB), “With Crane Related Injuries on the Rise Don’t Become Another Statistic,” 2022. <https://cicb.com/with-crane-related-injuries-on-the-rise-dont-become-another-statistic/>, Accessed: 04/20/2022.
- [12] T. Feix, “Anthropomorphic Hand Optimization based on a Latent Space Analysis,” Technische Universität Wien, Wien, Austria, 2011. Accessed: Feb. 22, 2022. [Online]. Available: <http://me.xief.net/publications/thesis.pdf>
- [13] Y. LI, J. Huang, F. Tian, H. A. Wang, and G. Z. Dai, “Gesture interaction in virtual reality,” *Virtual Reality & Intelligent Hardware*, vol. 1, no. 1, pp. 84–112, 2019, doi: 10.3724/SP.J.2096-5796.2018.0006.
- [14] C. Demolder, A. Molina, F. L. Hammond, and W. H. Yeo, “Recent advances in wearable biosensing gloves and sensory feedback biosystems for enhancing rehabilitation, prostheses, healthcare, and virtual reality,” *Biosensors and Bioelectronics*, vol. 190, p. 113443, 2021, doi: 10.1016/J.BIOS.2021.113443.
- [15] J. Baek, I. J. Jang, K. Park, H. S. Kang, and B. J. Yun, “Human computer interaction for the accelerometer-based mobile game,” in *Lecture Notes in Computer Science*, vol. 4096 LNCS, Springer Verlag, 2006, pp. 509–518. doi: 10.1007/11802167_52.
- [16] A. Ahmad, C. Migniot, and A. Dipanda, “Hand pose estimation and tracking in real and virtual interaction: A review,” *Image and Vision Computing*, vol. 89, pp. 35–49, 2019, doi: 10.1016/J.IMAVIS.2019.06.003.
- [17] J. Maitre, C. Rendu, K. Bouchard, B. Bouchard, and S. Gaboury, “Object recognition in performed basic daily activities with a handcrafted data glove prototype,” *Pattern Recognition Letters*, vol. 147, pp. 181–188, 2021, doi: 10.1016/J.PATREC.2021.04.017.
- [18] N. H. Adnan, K. Wan, A. B. Shahrman, S. K. Zaaba, S. N. Basah, Z. M. Razlan, D. Hazry, M. N. Ayob, R. M. Nor, and A. A. Aziz, “Measurement of the Flexible Bending Force of the Index and Middle Fingers for Virtual Interaction,” *Procedia Engineering*, vol. 41, pp. 388–394, 2012, doi: 10.1016/J.PROENG.2012.07.189.
- [19] J. O. Wobbrock, A. D. Wilson, and Y. Li, “Gestures without libraries, toolkits or training: A \$1 recognizer for user interface prototypes,” in *Annual ACM Symposium on User Interface Software and Technology (UIST 2007)*, 2007, pp. 159–168. doi: 10.1145/1294211.1294238.
- [20] R. D. Vatavu, L. Anthony, and J. O. Wobbrock, “Gestures as point clouds: A \$p recognizer for user interface prototypes,” in *ACM International Conference on Multimodal Interaction (ICMI 2012)*, 2012, pp. 273–280. doi: 10.1145/2388676.2388732.
- [21] L. Huang, B. Zhang, Z. Guo, Y. Xiao, Z. Cao, and J. Yuan, “Survey on depth and RGB image-based 3D hand shape and pose estimation,” *Virtual Reality & Intelligent Hardware*, vol. 3, no. 3, pp. 207–234, 2021, doi: 10.1016/J.VRIH.2021.05.002.
- [22] I. Chatterjee, R. Xiao, and C. Harrison, “Gaze+gesture: Expressive, precise and targeted free-space interactions,” in *International Conference on Multimodal Interaction (ICMI 2015)*, 2015, pp. 131–138. doi: 10.1145/2818346.2820752.
- [23] I. Moelmen, H. L. Grim, E. L. Jacobsen, and J. Teizer, “Asymmetrical Multiplayer Serious Game and Vibrotactile Haptic Feedback for Safety in Virtual Reality to Demonstrate Construction Worker Exposure to Overhead Crane Loads,” in *38th International Symposium on Automation and Robotics in Construction (ISARC 2021)*, 2021, pp. 613–620. doi: <https://doi.org/10.22260/ISARC2021/0083>.
- [24] MANUS, “Prime X Haptic VR.” <https://www.manus-meta.com/haptic-gloves>, Accessed: 02/23/2022.
- [25] VIVE United States, “VIVE Tracker (3.0).” <https://www.vive.com/us/accessory/tracker3/>, Accessed: 02/23/2022.
- [26] V. Corporation, “SteamVR.” <https://www.steamvr.com/en/>, Accessed: 02/25/2022.

A Knowledge Graph for Automated Construction Workers' Safety Violation Identification

Y. Zhu^{a,b} and X. Luo^{a,b}

^aDepartment of Architecture and Civil Engineering, University of Hong Kong, Hong Kong, China

^bResearch Center of Architecture and Civil Engineering, Shenzhen Research Institute, City University of Hong Kong, Shenzhen, China

E-mail: yifanzhu4-c@my.cityu.edu.hk, xiaowluo@cityu.edu.hk

Abstract

Identifying workers' safety violations on construction job sites is critical for improving construction safety performance. The advancement of sensing technologies makes automatic safety violation detection possible by encoding the safety knowledge into computer programs. However, it requires intensive human efforts in turning safety knowledge into computer rules, and the hard-coded rules limit the expandability of the developed applications. This study proposes a condition-based knowledge graph for the safety knowledge representation to support the reasoning on safety violations. The improved knowledge graph's structure solves the limitation by presenting the public knowledge and safety rules for condition structure, respectively. A natural language processing supported automatic knowledge graph development approach is developed in this paper to extract the safety knowledge from safety knowledge texts automatically and to construct the knowledge graph. To validate this construction framework, an initial knowledge graph containing 1,200 rules is developed based on construction safety regulations. The proposed automatic safety knowledge extraction model achieves an F1 value of 67%.

Keywords

Knowledge Graph; Natural Language Processing; Construction Safety; Workers' Safety Violation

1 Introduction

Timely identifying workers' safety violations onsite is essential to construction safety, and safety knowledge provides guidance to such identification tasks. Recently, research efforts have been made to automatically identify the workers' onsite safety violations using sensing technologies and artificial intelligence. Each of those research focuses on specific safety rules and hard-coded them into the computer applications, limiting the

expandability of those developed applications. A generic safety representation that the computer could understand is needed to support the development of automatic worker safety violation identifications.

This research proposes a condition-based knowledge graph to store the safety knowledge, which the computer could understand to support automatic safety violation identification. The proposed knowledge graph consists of two main components: a Rule Knowledge Graph (RKG) used to store the safety rules and an Association Knowledge Graph (AKG) to save corresponding safety knowledge. Furthermore, a knowledge graph development framework based on Natural Language Processing (NLP) is proposed to support the automatic knowledge graph development using safety regulations, reducing the labor cost, and improving the efficiency in constructing such knowledge graph. In this case study, an initial knowledge graph will be established to demonstrate the knowledge graph development and validate the effectiveness of the development framework.

2 Literature Review

2.1 The Application of Artificial Intelligence in Construction Safety

In the construction industry, automatic safety inspection has been applied in various fields related to worker safety, with the most prevalent domain divided into three aspects: Personal Protective Equipment (PPE) detection, exposure to hazardous areas, and unsafe behavior [5, 6]. In the area of PPE detection, Zhang et al. proposed an improved BiFPN-based deep learning method to detect workers and their hardhats [3]. Mneymneh et al. provided an intelligent monitoring framework for hardhats detection by applying motion detection algorithms and object detection tools to capture the required data [4]. In the aspect of identifying worker's exposure to hazardous areas, Fang et al. have tried to detect whether workers are on or across structural support using a convolutional neural network [5]. Konstantinou

et al. designed a vision-based approach to track workers with similar appearances and abrupt changes in a complex environment due to congestion, background clutter, and occlusions [6]. In the field of unsafe behavior, Yan et al. established an ergonomic posture recognition technique to capture injury-prone postures so as to prevent accidents and injuries [7]. Distinguishing workers' dangerous and fatigable postures, Seo and Lee tried to prevent work-related musculoskeletal disorders using an ergonomic assessment system in a computer vision-based assessment method [8].

Despite various studies in automatically detecting workers' unsafe behavior, only specific safety rules are used in those researches, without much consideration given to exhaustive references. The knowledge application lacks versatility since the selection of rules is targeted at fitting only a single construction activity rather than a broader range of areas.

2.2 Knowledge Graph

As a representation technique derived from the semantic web, a knowledge graph is widely recognized as one of the essential technologies for knowledge storage and management. Compared with traditional knowledge structures, a knowledge graph allows complex queries across multiple data sources so as to manage construction information on site [9], thereby realizing its applications in complicated and dynamic construction environments. Besides, a knowledge graph can save more time and labor costs by providing in-depth knowledge management methods [10]. In rule-based methods, constructors set up isolated rules, which do not have connections with each other. For the knowledge graph-based methodology, users can flexibly set their mechanisms to find out the results required. In addition, the graph-based knowledge representation can facilitate the discovery of new knowledge hidden behind. Therefore, the knowledge graph representation is more flexible and economical in management while enabling quick reflection of answers through a comprehensive retrieval.

Some researchers have designed several structures for knowledge graphs. For example, Ding et al. proposed an event logic graph, with its nodes being the event, and the edges representing the sequential, casual, is-a relations [11]. Li et al. suggested an AND/OR graph-based knowledge point organization model to represent the selective knowledge that is hard to describe previously [12]. Yu et al. proposed a tax graph to express the calculation logic about specific tax topics; it includes thousands of interconnected calculation models to indicate the calculation statement and contains calculation function nodes and input/output data nodes [12]. Such structures of knowledge graphs above generally have a specific application domain and cannot

be used in construction safety fields. Hence, a generalized knowledge graph is essential for furnishing a dependable reference in the construction safety domain.

3 Methodology

To apply numerous construction safety-related knowledge to aid artificial intelligence in spotting safety violations in an automatic, swift and correct manner, this study refines new knowledge graph representation and suggests a relevant automatic construction model. The results of the study could be used to create a knowledge graph that can be used for the automatic identification of safety violations on construction sites. The framework consists of two major components: a knowledge graph storing the knowledge and an automatic construction framework used to generate the knowledge graph. Figure 1 depicts the structure of the construction framework. The input is a series of one-sentence rules and public knowledge, and the output is the constructed knowledge graph.

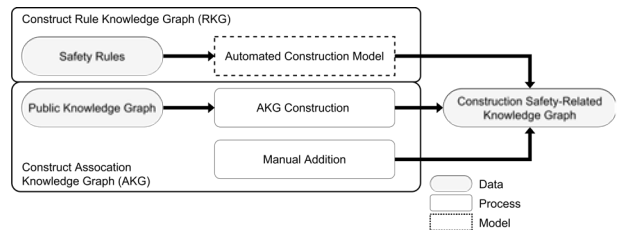


Figure 1. The construction framework structure

3.1 The Structure of the Construction Safety-Related Knowledge Graph

After training the artificial intelligence to automatically identify the safety violation, the artificial intelligence should understand which safety rules the workers should follow. In this process, it is necessary to construct a knowledge graph that can represent rules. Sensing technology, on the other hand, has to learn the public knowledge connected with these rules, meaning that a knowledge graph linked with these rules shall be built up. Figure 2 shows the overall structure of the knowledge graph conceived in this study. It comprises two main knowledge graphs: RKG expresses the construction-related rule knowledge graph, and AKG indicates the public knowledge graph associated with the knowledge on construction safety. The two main knowledge graphs enable a more flexible representation of safety rules and public knowledge and promote knowledge expandability.

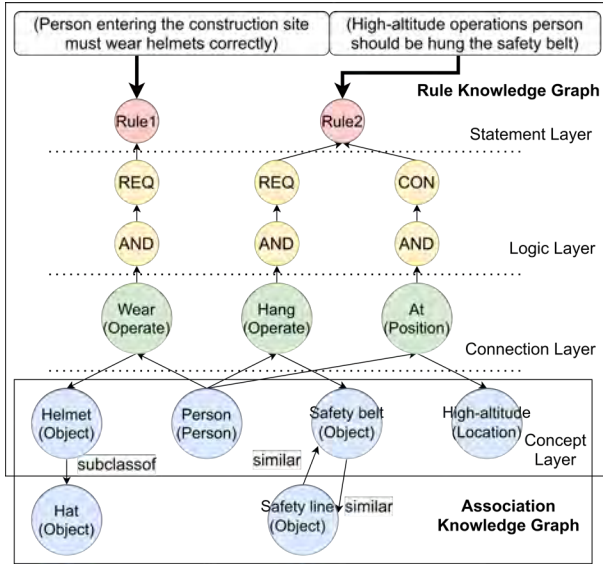


Figure 2. The overall structure of the knowledge graph

3.1.1 The Structure of Rule Knowledge Graph

Because of the dynamic and complex nature of construction environments, the safety criteria and associated circumstances for various types of job contents are diverse, with examples including to use supplementary PPE while conducting different tasks. The following issues are fundamental for safe operation by workers in the context of construction. Which kinds of PPE should be worn by workers? Which workplaces and conditions are permitted for workers to stay? Which process or activity is considered to be safe for workers? [6, 20]. Thus, the primary safety aspects in aiding artificial intelligence to perceive safety events that demand attention are related above. The above-mentioned elements serve as the basis for the RKG's construction. In addition, construction regulations and manuals have codified the safety elements that should be observed into the rules that must be followed. Therefore, artificial intelligence should understand the rules by using a knowledge graph to automatically detect workers' risky activities.

The RKG is made based on the condition-based knowledge graph offered by Jiang et al. [14]. It is composed of a sequence of rules retrieved from safety regulations, with each rule defined by a collection of triplets, shown in equations (1) and (2):

$$\begin{aligned} t_{u_1} &= (\{n_1: a_1\}, n_2, \{n_3: a_3\}) \\ t_{u_2} &= (\{n_1: a_1\}, r, \{n_3: a_3\}) \end{aligned} \quad \begin{matrix} (1) \\ (2) \end{matrix}$$

where $n_1, n_3 \in C$, and C is a set of concept nodes in the triplet; $n_2 \in O$, and O is a set of connection nodes in the triplet; $a_1, a_3 \in A$, and A is a set of attributes. '1' is for

the subject, and '3' is for the object; r is the relation between n_1 and n_3 . Triplet t_{u_1} uses the connection node as a connection entity to represent the relation, and t_{u_2} applies the relations 'Con_Belongto' to show the connection. Furthermore, triplet t_{u_1} represents an event such as a worker standing at a height, a worker moving rebar, or a worker wearing the hardhat. Triplet t_{u_2} denotes the affiliation between entities in a specific rule. Furthermore, in an RKG, the attribute and affiliation of an entity are valid only in the rule it is in.

Jiang et al. have introduced a condition-based knowledge graph [14], but it has no thorough description of the logic execution and sequence of triplets in the requirement and condition components. Furthermore, there are no detailed type definitions for the triplets' subjects, objects, and relation parts in triplets, leading to ambiguous expressions, i.e., some entities owning several meanings. Thus, this study adds a logic layer to the graph structure to showcase the details of the rules' logical execution and the relationship between the triplets in the requirement and condition part. In addition, this study adds the entity type to the subjects, objects, and relation in triplets, which will play an assistance role in the follow-up querying procedure. The proposed structure of the knowledge graph after the modification is shown in Figure 2, which is categorized into four layers: concept layer, connection layer, logic layer, and statement layer.

The concept layer contains concept nodes, which are the subject and object entities in triplets. The concept nodes have five types to display five categories of entities: **Person**, **Work**, **Object**, **Location**, and **Environment**. **Person** indicates the roles and professions on the construction sites, such as the Person concept node in Figure 2; **Work** represents the behavior and action of the people and machinery like Climb; **Object** indicates the objects usually appear on the sites, and this type of entity has its own attributes to additionally define its requirements or circumstances within the rules they are in like Helmet concept node in Figure 2; **Location** indicates a range belonging to the machinery, object, people or region like High-altitude concept node in Figure 2; **Environment** indicates the weather or times on construction sites such as the wind, and it uses three predefined attributes to show the degrees and levels: **Unit**, **Value**, and **Property**. These three attributes are illustrated in the selection range of weather conditions in the current rule. Particularly, **Unit** presents the weather's unit including level, m/s; **Value** determines the value of the weather's level, which is composed of digital form like 6 or six; and **Property** exhibits the weather's value range, such as larger than, not less than. Furthermore, the Object- and Location-typed nodes have 'Con_Belongto' links to describe the attribution relationships.

The connection layer includes connection nodes and

links connected with the concept nodes, and the link direction shows that a concept node is a subject or an object in triplets. Connection nodes have three types applied in three distinct domains: **Operate**, **Position**, and **Predicate**. Generally, **Operate** is applied in a worker's PPE domain, demonstrating the relationship between worker and PPE, such as wear, hang. **Position** regulates the place and weather relation like up, down, or around. **Predicate** represents the working condition of workers and things and how they operate other things including use and operate.

The logic layer denotes the logic execution in the requirement and condition part as well as the logic order between triplets. The logic layer is like an event tree, in which the nodes are similar to tree branches. A larger branch can continue to be extended to smaller branches, with most terminals being the connection nodes. The nodes in the logic layer have two categories: the logic node and the part node. The former one plays a crucial role in the logic execution representation; in safety rules, it is an objective existence, meaning that some requirements and conditions have distinct constraints. The logic node is indispensable to describe this circumstance. The logic node is divided into three categories: AND node, OR node, and NOT node. The AND node means that all child nodes should be followed. The OR node means that at least one of the child nodes should be obeyed. The NOT node means that all children nodes should not be followed. The result of logic nodes will be concluded to part nodes, i.e., the Req and Con nodes. The part node denotes which triplets belong to the requirement and condition part through connection to the Req and Con nodes. For example, in Figure 1, the triplet [Person, Hang, Safety Belt] is the requirement part due to the connection to the Req node, and this triplet should be followed as it is connected with the AND node.

The statement layer uses statement nodes to define each rule, and is the root node in the knowledge graph to represent rules. The statement node indicates which requirement and condition part belongs to a certain rule by connecting the requirement and/or condition-typed nodes.

3.1.2 The Structure of Associated Knowledge Graph

Except for rules, artificial intelligence is also required to learn the knowledge related to construction safety so as to facilitate the safety violation identification ability. For instance, the relationship between the subclasses of hats or objects has a similar meaning. Therefore, this study constructs the knowledge graph associated with the construction safety-related rules (AKG). The entities come from the concept layer of RKG and artificial additions. Triplets in AKG work in any rule. In contrast, some triplets in RKG only work in the rules they belong

to. Compared with RKG, the relation in triplets of AKG exists in the form of edges, not nodes. Formula (2) denotes triplets in AKG.

AKG implements two links: 'Similar' and 'subclassof'. The former one indicates the entities that have a similar meaning; the latter one means that a subject entity is the subclass concept to the object entity. The relation 'subclass' is an edge in AKG to connect the subject and object. Likewise, the relation 'similar' is an edge in a triplet, e.g., [Safety line, similar, Safety belt]

3.2 The Construction Framework of the Knowledge Graph

In the construction framework, in addition to implementing the automatic extraction model as proposed by Wei [15], this study designs a knowledge graph construction procedure. The construction procedure is responsible for transforming the extracted file into a knowledge graph based on the predefined rules and methods.

In practice, some texts describing workers' normative requirements lack the subject, while supplementing the subject words to each text is a time-consuming task. Additionally, in RKG, triplets are composed of three parts: subject entity, connection entity, and object entity. If a connection entity is fixed as several relations to link the subject and object entities, it will not represent more elements flexibly, and the complexity will increase dramatically. Meanwhile, the representation of belonging parts and types for triplets in a tuple label is indispensable. Therefore, the extraction model needs the specific relation label representation to suit the RKG with multiple layer structures. To solve this issue, this study divides the relation label into five types: the relation label for labelling subject and connection entities; the connection entity for object label; the entity for attributes; the subject for objects; and the connection entity for the connection entity.

The information of relation labels includes the triplet type, a requirement or a condition part, and two entities' types (the subject and object). For example, the tuple label 'Operate-Req_AND-Object' means that the subject type is Operate, the object type is Object, and the relation between the two entities is Req_AND, where 'Req' indicates this tuple belongs to the requirement part, and 'AND' indicates the tuple's execution logic. Thus, the relation label can show the variety of data in the relation label.

In addition, some relation labels are disparate from the labels mentioned above. For instance, 'Object-Con_Belongto-Object' means that the object and subject entities' type is Object, and the subject entity belongs to the object. Besides, some triplets do not need a connection entity, and thus their relation label will be like 'Object-Con_AND-Work' to describe this circumstance,

while a connection entity will be added in the triplet in the construction of the knowledge graph. Likewise, the relation type also has the attribute representation like 'Object-Con_AND-Attribute', and the connection entity's sequence representation of 'Predicate-Connect_Dis-Predicate' indicates that the triplet belonging to each connection entity should be judged separately.

3.2.1 The Automated Construction Procedure of Rule Knowledge Graph

After finishing the part of triplet extraction, the next step is to construct the knowledge graph based on the extracted partial triplet. Firstly, for the extraction content of the automated extraction model this study calls tuples, the tuple sets will be inputted for expansion. For instance, the type 'Object-Con_AND-Work' means the condition that the object is doing some work. Nevertheless, this form is not appropriate for the RKG, so a division is necessary. This tuple will be divided into 'Object-Con_AND-Predicate' and 'Predicate-Con_AND-Work'; the Predicate-typed entities in two new tuples through the predefined procedure are the same, both the entity 'conduct'. Secondly, some triples represent the entity's attributes, such as the 'Environment-Con_OR-Property' and 'Object-Con_AND-Attribute'; they will not be constructed as the triples in the knowledge graph. In contrast, the attribute entity will show the attributes to the subject in the corresponding rules. Thirdly, the two tuples with the same connection entities and same relations will be combined with the new triplet in the knowledge graph. In addition, some rules do not own the subject, so the labeled triples in these rules will all be the latter part of the triple in the RKG, and the new first half of the corresponding triple will be added. Fourthly, the corresponding logic layer will be generated, followed by the relation in the triple, and the triple containing two connection entities will also participate in the logic layer's generation. Finally, after finishing the four steps above, each text will create a statement node to associate with the logic node and indicate the belonging of the requirement and condition parts.

3.2.2 The Construction of Association Knowledge Graph

For the generation of AKG, the data usually come from two sources: the public knowledge graph and the knowledge added manually. For the public knowledge graph, this research will search each concept node in the public knowledge graph for similar entities and the subclass or upper-class entities. Furthermore, this study will observe the relations between the same-type entities in the concept layer and add the SubClassof, with similar relations in these entities.

4 Case Study

To demonstrate the construction process of the knowledge graph, a case study is carried out mainly in two steps to verify the feasibility. First, the case study will manually establish the construction safety-related knowledge graph by relying on safety documents and the public knowledge graph. Second, this study will develop an automated extraction model to improve the automatic construction model of the knowledge graph. Figure 3 shows the overall process, with these steps explained in detail below.

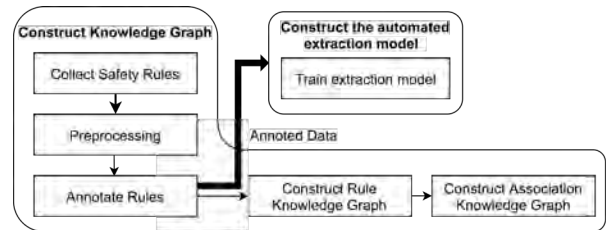


Figure 3. The process of case study

4.1 Construct the Knowledge Graph

4.1.1 Collect Safety Rules

Nearly 97 specifications have been collected, including Chinese National Standards (GB), construction industry standards (JGJ), and safety manuals. These documents were chosen because they all involve construction safety and workers' behavior safety, while showcasing a certain universality.

4.1.2 Preprocessing

This study starts with data preprocessing to handle the construction codes collected. First, this study determines the extraction range of documents, eventually picking out provisions in the three areas mentioned in Section 3.1, and excluding those requirements related to other construction safety conditions. To represent it clearly, this study divides the relatively complex rules into simple contents, and rules will be replenished with the subject and object, if missing, based on the chapter titles. Finally, 1,236 rules are collected as the corpus used for the next step.

4.1.3 Annotate Rules

The rule annotation extracts the imperative data from unstructured text and facilitates artificial intelligence-based safety inspection, and it further marks the entities and the relations between the entities in the rule. On one side, this study adopts the extracted data as the corpus to train the automated extraction model. On the other side, to ensure the accuracy and efficiency of the knowledge graph, this study uses the extracted data as the input to

construct the RKG, without using the output of the automated extraction model. This study extracts information by applying brat rapid annotation tool (BRAT), which is a web-based tool for annotating by adding notes to the existing text documents [16]. Designed for systematic annotation, it has a defined structure that artificial intelligence can process and understand. BRAT involves two types of annotations: text span and relation annotation. The former one marks the entities and their types.

4.1.4 Construct the Rule Knowledge Graph

The annotated corpus will be fed into the automated construction procedure to construct the RKG. As one of the most popular tools for most knowledge graph storage, Neo4j is extensively used in many studies [10]. Therefore, this research uses Neo4j as the storage medium [17]. After annotation, this study outputs the annotated entities, including their name, type, and id number with interrelated relations; then, they are saved in the graph database management system Neo4j, as shown in Figure 4. The example rule shown in figure 4 has a set of triplets: [Person, at, high altitude] (Con_AND), [Person, Hang, safety belt] (Req_AND), indicating that when people are at a high altitude, they should wear the safety belt.



Figure 4. One rule in the Neo4j

4.1.5 Construct the Association Knowledge Graph

The corresponding AKG is constructed to be added to the relevant knowledge graph. OwnThink is a knowledge graph-based public knowledge graph, which incorporates twenty-five million entities with billions of entity-attribute relationships [18]. This study extracts similar word relations from the Ownthink knowledge graph. Searches for the Concept entity both in the knowledge graph and Ownthink find the similar entity through relations 'Also Known As (又名)' and 'Another Name (别名)'. Furthermore, this study manually identifies the 'subclassof' and 'similar' relations between person-type entities for better querying. Finally, The AKG is constructed in this step based on the public knowledge graph resources and manuals. The searched relations and entities are inputted into Neo4j manually.

4.2 Training and Performance of the Automated Extraction Model

The automated extraction model needs to be trained

so as to make the model extract the data more precisely. After annotation of safety rules, the annotated files will be transformed into the json-type files and are regarded as the corpus for the extraction model. The BERT model in the extraction model adopts the Chinese_wwm_ext_L-12_H-768_A-12 pre-training model proposed by Cui et al. [19]. The corpus is divided into three components: train corpus, valid corpus, and test corpus, with a ratio of 7.5:1.5:1.5, respectively.

The results acquired after the training of the extraction model are shown in Table 1, illustrating the performance of the automated extraction model. The title 'Type' means the tuple's type; 'Precision' means the precision in each type; 'Recall' means the recall value in each type. The overall precision, recall, and F-1 value are 78.07%, 58.92%, and 67.15%, respectively. More specifically, this study divides the types into two parts: the first half of the triplet, and the second half of the triplet. The first half of the triplet is the subject, the connection, and the link between these two nodes, the second half is the connection, object, and the links between these two nodes. The performance of the first half is Precision (95.24%), Recall (60.60%), and F1-value (74.07%); the second half of the triplet is Precision (76.42%), Recall (59.54%), and F1-value (66.93%). Observation shows the performance of the first half is better than the second half. In addition, this study divides the tuple type into four parts based on the function: Operate, Position, Predicate, and Attribute. The performance of Operate is: Precision (79.66%), Recall (78.99%), and F1 (79.32%); the performance of Position is: Precision (85.31%), Recall (60.70%), and F1 (70.93%); the performance of Predicate is: Precision (71.11%), Recall (50.96%), and F1 (59.37%); the performance of Attribute is: Precision (86.05%), Recall (55.22%), and F1 (67.27%). These results indicate that the performance of Predicate and Attribute is the weakest, the performance of Position is strong, and the performance of Operate is the best.

Table 1. The performance of the automated extraction model in part type

Type	Precision (%)	Recall (%)	F1 value (%)
Operate-Req_NOT-Object	72.73	66.67	69.57
Position-Con_AND-Object	75.00	40.00	52.17
Person-Req_NOT-Position	92.31	85.71	88.89
Person-Req_AND-Operate	100.00	77.78	87.50
Object-Object	50.00	10.00	16.67
Con_Belongto-Object			
Environment-	100.00	76.92	86.96

Con_OR-Property			
Environment-Con_OR-Unit	92.31	92.31	92.31
Environment-Con_OR-Value	92.31	92.31	92.31
Predicate-Req_NOT-Work	40.00	12.50	19.05
Position-Req_NOT-Location	81.82	47.37	60.00
Object-Con_AND-Attribute	50.00	18.18	26.67
Position-Con_AND-Location	87.50	60.87	71.79
Predicate-Con_AND-Work	63.64	36.84	46.67
Object-Con_AND-Work	78.26	64.29	70.59
Predicate-Con_OR-Object	65.52	65.52	65.52
Position-Con_OR-Environment	96.30	96.30	96.30
Position-Req_NOT-Object	76.32	55.77	64.44
Predicate-Con_AND-Object	71.74	52.38	60.55
Operate-Req_AND-Object	79.00	85.87	82.29
Predicate-Req_NOT-Object	72.48	54.48	62.20
Overall	78.07	58.92	67.15

5 Discussion

For the automated extraction model, this study excludes the tuple types with their number less than 30 to analyze the results since most of them have F1-values of zero. In the results, the tuple type with a relatively fixed structure will have a strong influence. The tuples like 'Position-Con_OR-Environment', 'Environment-Con_OR-Unit', 'Environment-Con_OR-Value' and 'Environment-Con_OR-Property' have a fixed text structure about 'when in some weather'. Thus, the performance of these tuples is better than other tuples. In addition, the tuple types 'Person-Req_AND-Operate' and 'Operate-Req_NOT-Object' have the fixed text structure 'should/shouldn't be equipped with something '. Nevertheless, the number of 'Operate-Req_NOT-Object' is smaller than that of other types, and it may impact the final performance. Furthermore, both types of 'Person-Req_NOT-Position' and 'Person-Req_AND-Operate' have the fixed structures, where 'Person-Req_NOT-Position' has the structure of 'Non-worker cannot go into it ', while 'Person-Req_AND-Operate' has the text structure of 'A certain types of worker needs to be equipped with it'. Therefore, the tuple types with fixed

structures will deliver a good performance. The types of 'predicate-Req_NOT-Work', 'Object-Con_Belongto-Object' and 'Object-Con_AND-Attribute' give the worst performance, with variable text structures; and the extraction model faces some challenges in the annotation. Some potential solutions are proposed. On one side, more corpus could be furnished for training the model. On the other side, using the hidden connections between different concepts in post-processing can also enhance accuracy.

6 Conclusion

This research develops a unique knowledge graph and corresponding construction framework to assist artificial intelligence-based identification of workers' safety violations. The knowledge graph includes two parts: a Rule Knowledge Graph (RKG), including four layers (Statement layer, the Logic layer, the Connection layer, and the Concept layer), and an Association Knowledge Graph (AKG). The F1-value performance of the automated extraction model in the construction framework can reach 67%.

The proposed framework showcases the following advantages. First, it can assist artificial intelligence to automatically identify safety violations by searching for all rules related to the scene description. Second, users can flexibly adjust the identification range in constructing the knowledge graph. Third, the proposed construction framework is beneficial for the automatic construction of knowledge graphs, while reducing the labor cost incurred in such constructing. In practical applications, the acquired information from the variety of sensors will be transformed to a text-based description. The artificial intelligence identifies the worker's violation in description based on the constructed knowledge graph.

Despite the successful advantages above, several limitations with this study have to be overcome. First, the automated extraction model still needs to be enhanced for better performance. Second, the knowledge graph shall collect more related rules, whether on construction codes or even corporate regulations. Third, the proposed framework cannot convert construction-site videos into textual description, and it has not been verified in a practical application.

Future research directions include: First, to further boost the performance of the automated extraction model. Moreover, a more accurately semantic matching method relying on machine learning based on the current mechanism can be implemented in the querying model. Second, by collecting more relevant regulations to be stored in the knowledge graph, a more comprehensive range of construction safety domains will be suitable for this model. Lastly, the proposed framework can be

integrated with some practical applications on construction sites, so as to verify its feasibility. For example, in order to improve detection efficiency, it can be combined with computer vision in camera monitoring.

Acknowledgment

This work was supported by the Shenzhen Science and Technology Innovation Committee Grant (PJ#JCYJ20180507181647320). The conclusions herein are those of the authors and do not necessarily reflect the views of the sponsoring agency.

References

- [1] W. Fang, P. E. D. Love, H. Luo, and L. Ding, Computer vision for behaviour-based safety in construction: A review and future directions, *Advanced Engineering Informatics*, 43, August 2019, page 100980,2020
- [2] W. Fang *et al.*, Computer vision applications in construction safety assurance, *Automation in Construction*, 110, February 2019, page 103013,2020
- [3] C. Zhang, Z. Tian, J. Song, Y. Zheng, and B. Xu, Construction worker hardhat-wearing detection based on an improved BiFPN, page 8600–8607,2021
- [4] B. E. Mneymneh, M. Abbas, and H. Khoury, Vision-Based Framework for Intelligent Monitoring of Hardhat Wearing on Construction Sites, *Journal of Computing in Civil Engineering*, 33, 2, page 04018066,2019
- [5] W. Fang *et al.*, A deep learning-based approach for mitigating falls from height with computer vision: Convolutional neural network, *Advanced Engineering Informatics*, 39, January, page 170–177,2019
- [6] E. Konstantinou, J. Lasenby, and I. Brilakis, Adaptive computer vision-based 2D tracking of workers in complex environments, *Automation in Construction*, 103, November 2018, page 168–184,2019
- [7] X. Yan, H. Li, C. Wang, J. O. Seo, H. Zhang, and H. Wang, Development of ergonomic posture recognition technique based on 2D ordinary camera for construction hazard prevention through view-invariant features in 2D skeleton motion, *Advanced Engineering Informatics*, 34, June, page 152–163,2017
- [8] J. Seo and S. Lee, Automated postural ergonomic risk assessment using vision-based posture classification, *Automation in Construction*, 128, March 2017, page 103725,2021
- [9] P. Pauwels, S. Zhang, and Y. C. Lee, Semantic web technologies in AEC industry: A literature overview, *Automation in Construction*, 73, page 145–165,2017
- [10] X. Li, M. Lyu, Z. Wang, C. H. Chen, and P. Zheng, Exploiting knowledge graphs in industrial products and services: A survey of key aspects, challenges, and future perspectives, *Computers in Industry*, 129, page 103449,2021
- [11] X. Ding, Z. Li, T. Liu, and K. Liao, ELG: An Event Logic Graph, [Online]. Available: <http://arxiv.org/abs/1907.08015>2019
- [12] S. Li, X. Li, and L. Wang, Knowledge points organization model based on AND/OR graph in ICAI, *Proceedings - 2010 6th International Conference on Natural Computation, ICNC 2010*, 4, Icn, page 2121–2124,2010
- [13] W. Fang, P. E. D. Love, H. Luo, and L. Ding, Computer vision for behaviour-based safety in construction: A review and future directions, *Advanced Engineering Informatics*, 43, August 2019,2020
- [14] T. Jiang *et al.*, Biomedical Knowledge Graphs Construction from Conditional Statements, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 18, 3, page 823–835,2021
- [15] Z. Wei, J. Su, Y. Wang, Y. Tian, and Y. Chang, A Novel Cascade Binary Tagging Framework for Relational Triple Extraction, page 1476–1488,2020
- [16] NLPlab. Online:<https://brat.nlplab.org/introduction.html>
- [17] Neo4j. Online:<https://neo4j.com/>
- [18] Ownthink. Online:<https://www.ownthink.com/docs/kg/>
- [19] Y. Cui, W. Che, T. Liu, B. Qin, and Z. Yang, Pre-Training with Whole Word Masking for Chinese BERT, *IEEE/ACM Transactions on Audio Speech and Language Processing*, 29, November, page 3504–3514,2021

Wearable Technology for Highway Maintenance and Operation Safety: A Survey of Workers' Perception and Preferences

S. Sabeti¹, O. Shoghli², N. Morris³ and H. Tabkhiv⁴

¹Department of Engineering Technology and Construction Management, University of North Carolina at Charlotte, USA

²Department of Engineering Technology and Construction Management, University of North Carolina at Charlotte, USA

³Department of Mechanical Engineering, University of Minnesota, USA

⁴Department of Electrical and Computer Engineering, University of North Carolina at Charlotte, USA

ssabeti@uncc.edu, oshoghli@uncc.edu nlmorris@umn.edu htabkhiv@uncc.edu

Abstract -

Securing the safety of highway work zones is one of the most pressing issues in the highway maintenance and operation community. Recent studies have indicated that highway workers keep suffering from fatal injuries and death caused by traffic, long-night shifts, and limited space for maneuvering. In the meantime, recent advances in wearable technology have provided promising potential in the context of safety in different disciplines, especially building construction. However, highway workers have been underrepresented and limited information is available about their perception, preferences and ideas toward wearable technology. With this, in this paper, we document our early results in investigating the perception of highway workers toward wearable technology and their preferences among available devices to be used for safety purposes. Our results highlight a promising potential for the application of wearable technology in highway work zones and an acceptable level of engagement from the body of highway workers. Therefore, we envision this study to energize developers for further research and investment in the application of wearable technology in highway work zone safety.

Keywords -

Wearable Technology; Highway Work Zones; Safety; User Experience; Worker-centered Design

1 Introduction

In 2018, Federal Highway Administration (FHWA) reported that 124 workers lost their lives at highway construction sites. With an average of 135 workers fatality, FHWA also documented that lethal crashes in highway work zones climbed by 3 percent between 2016 and 2017. Fatal injuries on top of death is another substantial threat to the safety of highway workers. A total of 158,000 crashes and 42,000 associated injuries were reported in 2016 alone [1]. Meanwhile, recent investments in highway infrastructure construction would further exacerbate this issue by creating more work zones. Therefore, securing the safety

of highway work zones is one of the most pressing issues that the maintenance and operation face in the years to come [2, 3].

The recent boom in wearable technologies has provided significant potential in addressing some of the challenging problems in the engineering world [4]. Latest research trends have demonstrated that researchers and practitioners from different disciplines have rushed toward wearable-enabled systems for enhancing the status-quo, especially in building construction safety [5, 6, 7]. In the meantime, highway maintenance, operation, and construction community is and has been only relying on fairly reactive systems and have not yet departed toward more modern technologies [8, 9]. Meanwhile, building construction researchers have also attempted to increase the usefulness and usability of the newly developed systems by investigating the perception of construction workers toward wearable technology [10]. However, highway workers have been underrepresented in the body of knowledge. The majority of previous studies have targeted vertical building construction safety, and highway maintenance community's perception has not been investigated in comparable detail [11, 8, 12]. Limited space for maneuvering, often long and night shifts, and dealing with drivers with a broader range of behaviors are some of the reasons that make the needs of highway maintenance workers unique. This makes prior information in building construction directly inapplicable to highway work zones and could potentially hinder future developments for highway workers [13, 14].

In this study, we investigated the perception and preferences of highway workers toward wearable technology to be used in safety-related systems. In specific, the contributions of this paper to the body of knowledge are:

- This article is among the first research studies investigating the perception of highway workers toward wearable technology.
- This article specifically compares the highway worker's perception about some of the most common

wearable options in the market.

- The results of this study paves the path toward future interaction and safety system designs for highway workers.

In the following, we will first lay the foundation by explaining the related works of this study. Then, we will move on to explaining our methodology followed by results and discussion.

2 Related Works

2.1 User Perception and Technology Development

It is commonly believed that achieving success in information technology projects requires careful user research. Previous studies have demonstrated that there is usually a gap between what developers think of the system and what users' actual perceptions are [15]. Therefore, investigating whether the intended users would adopt the developed technology or not is of importance [8]. Recent investments in wearable technology have accelerated modernizing the concept of work [16]. However, several researchers have already discussed that there is a potential for the end-users to resist adopting the developed technologies, regardless of their benefits. This fact majorly highlights the importance of early user studies [17]. Therefore, a few researchers in the past performed some studies investigating the reaction of construction workers to new technologies. For example, [18] developed an extended Technology Acceptance Model (TAM) for including the future alterations in the attitude of workers with respect to time [19]. We also identified such studies in other disciplines [20, 21].

2.2 A Survey of Wearable Devices and Their Applications

Smart wearable devices have a long and rich history [22]. Recent boom in wearable technology has provided industry leaders, researchers and practitioners in different industries with a new source of power for increasing the cognitive capabilities of humans in their daily life and decision making [23, 24]. [25] categorized the state-of-the-art wearable devices under three main categories, accessories, e-textiles, and e-patches. They considered wrist-worn (smart watches and wristbands), head-mounted (smart eye-wears, hard hats, and ear-buds) and others such as vests as the main subgroups of the accessories wearable devices. Moreover, wearable technology has already been proven as an efficient tool in different disciplines, ranging from healthcare to education [26]. In an interesting study, [4] discussed the historical and current trends in wearable technology and concluded that health related issues are still among the most well-received applications of wearable technology. There have been already multiple review

articles that investigated the application of wearable devices in different disciplines [27]. Recent trends show that interest in applying wearables in safety-related contexts are exponentially growing. For example, [28] investigated the application of wearable devices in securing the safety of miners. Another recent trends in wearable technology is using technologies such as Artificial Intelligence (AI) or Internet of Things (IoT) as the backend and wearable technology as the frontend and the means of interaction with users. For instance, the authors in [29] studied how wearable technology can be leveraged in securing the safety of women by sending an emergency notifications to their relatives and adjacent police stations using IoT.

2.3 Wearable Devices and Their Application in Construction

The application of wearable devices in construction industry has been continuously and exponentially growing in recent years for mitigating safety issues. Researchers and practitioners have used wearable technology for monitoring and collecting different information, including kinematic movement, cardiac activity, skin response, and muscle engagement [30, 11]. For this purpose, they have deployed different devices, ranging from smart wristbands, Electroencephalogram (EEG) headsets, and Augmented Reality glasses [31, 32]. These devices were used in order to study safety risks such as falling, engagement to hazardous behaviors, preventing extreme fatigue or other within-site accidents [33, 34]. Some studies also suggested the use of wearable technology for designing active safety systems in construction sites. For instance, [35] proposed a novel system architecture for a safety system that warns construction workers and prevents them from accidents using wearable technology and IoT.

3 Methodology

Given the limited information available about highway workers and wearable technology, our main goal in this study is to investigate the perception of highway workers toward wearable technology to be used in safety-related systems. In specific, we want to investigate:

1. Perception of highway workers toward the practicality of wearable technology in highway work zones.
2. Likelihood of highway workers using wearable devices in highway work zones.
3. Preferences of highway workers among the currently available wearable options.
4. Major concerns of highway workers toward using wearable technology.

For this purpose, we designed a two-step methodology, including a semi-structured interview followed by an extensive survey of the body of highway operation and maintenance community. Below, we will explain our methodology in detail.

3.1 Semi-structured Interview

Before designing our questionnaire, we first conducted an in-depth interview with an experienced highway maintenance crew member. The reason behind this was twofold. Firstly, given the unique needs of highway maintenance workers, we wanted to gain an initial impression of their thoughts and beliefs to be reflected in the design of the questionnaire. Secondly, we wanted to run the survey questions by him and ensure the language and general structure of the survey is suitable for highway workers. Our interviewee was a senior and former member of the highway maintenance and operation division of a state Department of Transportation (DOT) with more than 20 years of experience.

Our interview was semi-structured. We first started off with basic demographic questions, including age, experience, and familiarity with this field. Then, we moved on to open discussions about the wearable technology and its application in highway work zones. Our interviewee mentioned that he had no prior experience with wearable technology as a safety mean in highway work zones. Our interviewee believed that wearable technology could be useful in highway work zones. Ultimately, in this step, we selected smart glasses, smart wristbands, smart hard hats, and smart clothes such as smart vests among different available wearable options in the market. We chose these devices after consultation with our interviewee, availability in the market, our literature review, competitive cost, and compatibility with outdoor environment. However, our interviewee at first was concerned about the use of smart glasses for the users who wear prescribed glasses. While this concern is completely valid, some of the currently available options in the market are compatible with prescribed glasses and offer a solution for this problem (see Vuzix Blade [36]).

3.2 The Structure of the Questionnaire

After our initial interview and with the suggestion of our interviewee, we decided to separate our participants into two personas: highway maintenance crew and affiliated participants. The former only includes workers while the latter consists of state DOT members, private consultants, managers, researchers, and other individual who are acquainted with maintenance and operation community and are not a worker. The main reason behind this personification was twofold. Firstly, our target population

is all individuals that are physically present in highway work zones and will interact with wearable technology as a safety measurement. However, maintenance crew are more frequently present in work zones than other members. This could lead to these groups having developed a different set of ideas and beliefs about wearable technology to be used as a safety measure. Secondly, we believed that crew members might have different expectations from wearable technology than managers given the differences in the nature of their jobs. For instance, it is logical to assume that workers should not care about the cost of the technology when it could be an important contributor to what managers and supervisors think of the technologies. It should be noted that hereinafter, we will be calling the highway maintenance crew “maintenance crew” and the affiliated participants with highway maintenance and operation community “affiliated members”.

In the next step and based on our feedback from the interview, we designed our questionnaire to survey the body of highway work zone maintenance and operation community. We started off the survey with demographic questions, where we asked participants their age, role, and the frequency of their presence in the highway work zones in general. Next, we asked them about their previous experience with wearable devices, if any. Then, we asked them our major questions, which were:

1. Practicality of the provided wearable devices (Question 1).
2. Likelihood of users utilizing the provided devices (Question 2).

Finally, we asked participants to share with us their concerns in specific about the application of wearable technology in highway work zones (Question 3). We asked them to select from "impractical for operation in highway work zones, unpleasant experience with devices", "influence on the performance such as vision obstruction", "slow and painful adaptation to the technology as a routine", "the unreliability of the devices in identifying potential dangers", "repetitive false alarms and loss of your trust in devices", "none" and "other". They could also have selected multiple options from the provided list. We then reached out to the body of highway maintenance and operation community and asked for their participation. We used Google Form for hosting and conducting this survey.

3.3 Statistical Analysis

After receiving responses from our participants, we statistically analyzed the data to investigate our research objectives. For this purpose, we used chi-square test, both goodness-of-fit and independence versions to investigate whether affiliated members and maintenance crew show

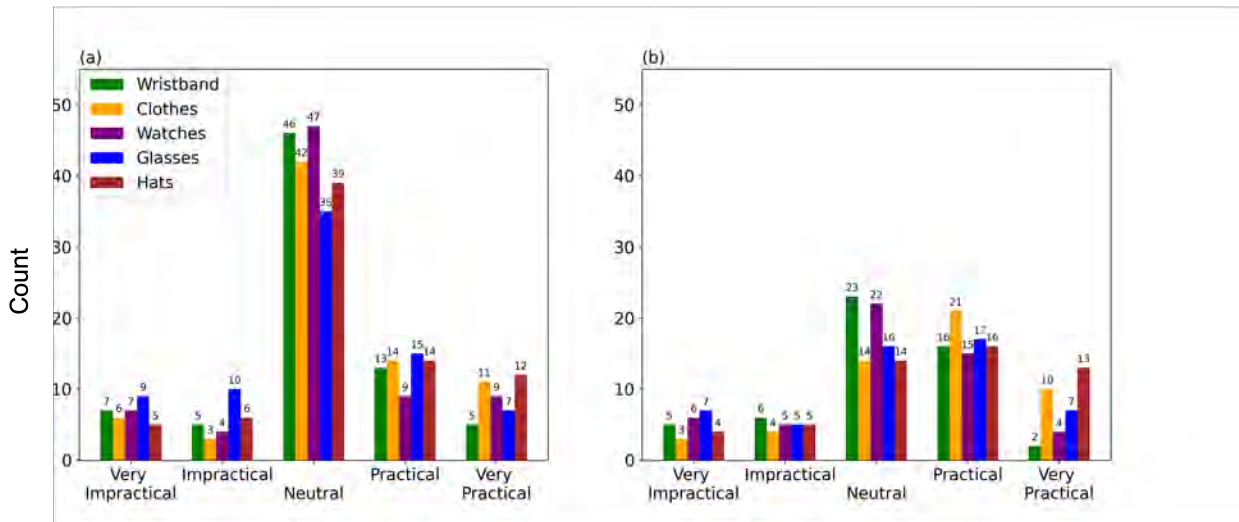


Figure 1. The practicality of the selected wearable options in (a) maintenance crew and (b) affiliated members

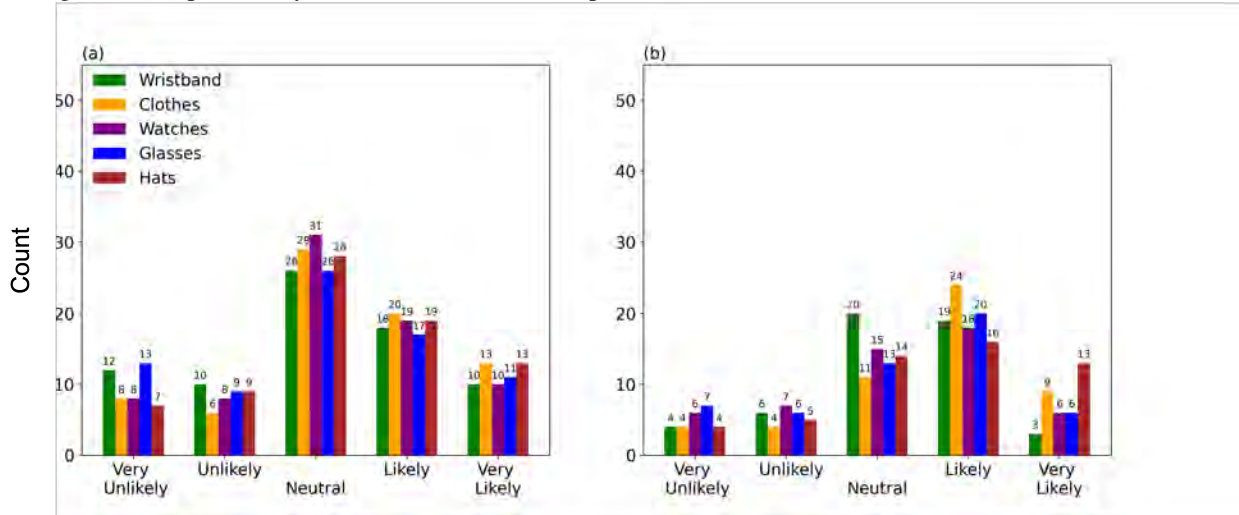


Figure 2. The likelihood of (a) maintenance crew and (b) affiliated members to use the selected wearable options

Table 1. Chi-square goodness-of-fit test on the collected data and the corresponding p-values

Question	Group	Wristband	Clothes	Watch	Glasses	Hats
1	Maintenance Crew	$1.2 * 10^{-16}$	$4.4 * 10^{-13}$	$2.2 * 10^{-17}$	$5.8 * 10^{-13}$	$2.9 * 10^{-10}$
1	Affiliated Members	$5.6 * 10^{-6}$	$2.8 * 10^{-4}$	$9.7 * 10^{-5}$	$1.5 * 10^{-3}$	$2 * 10^{-2}$
2	Maintenance Crew	$1.4 * 10^{-3}$	$1 * 10^{-4}$	$3.2 * 10^{-5}$	$1.8 * 10^{-3}$	$7.8 * 10^{-5}$
2	Affiliated Members	$1.9 * 10^{-5}$	$3.3 * 10^{-5}$	$1.4 * 10^{-3}$	$6.2 * 10^{-4}$	$2 * 10^{-2}$

different behavior toward wearable devices in general and the selected devices in particular.

4 Results

In this section, we will provide the results obtained from our survey and analysis.

4.1 Community Engagement

148 individuals responded to our survey, 76 of which identified themselves as maintenance crew and 52 as affiliated participants with the body of highway work zone community. The rest were general participants. Since the general participants did not belong to our target groups, the corresponding results were excluded from our study. The majority of our participants were from the state of North Carolina (62 from the maintenance crew and the

Table 2. Our participants' major concerns about wearable technology in highway work zones

Concern	Maintenance Crew	Affiliated Members
Impractical for operation in highway work zones	13	12
Unpleasant experience with devices	4	5
Influence on the performance such as vision obstruction	16	21
Slow and painful adaptation to the technology as a routine	4	8
Unreliability of the devices in identifying potential dangers	19	24
Repetitive false alarms and loss of your trust in devices	17	29
None	23	7
Other	11	12

Table 3. Chi-square independence test on the collected data and the corresponding p-values

Question	Wristband	Clothes	Watch	Glasses	Hats
1	0.25	0.013	0.07	0.309	0.09
2	0.26	0.15	0.63	0.39	0.64

34 from affiliated participants). This was followed by the state of Virginia with 6 participants from the maintenance crew and 8 from affiliated ones. The states of Florida, Georgia, Indiana, Pennsylvania, and Texas also had some representatives in this survey. Finally, we received one response from a crew member who mentioned to be based in Canada.

4.2 Questionnaire Results

Among our participants, 28 mentioned that they are at least 55 years old. 78 participants also mentioned that their age is between 36 and 55 years (37 in 36-45 and 41 in 45-55). Moreover, only 10 participants mentioned to be less than 25 years old. Therefore, it is safe to conclude that while we have a fairly acceptable diversity from different age groups, the majority of our participants are either middle-aged or seniors. Additionally, 92 out of 128 participants mentioned that they had been working in the field of highway operation and maintenance for at least 10 years. Only 13 participants mentioned that they were novice with having 0-2 years of experience. 13 people also mentioned that they had been actively involved in this field for at least 3 to 5 years. Therefore, we can also conclude that the majority of our participants were experienced-enough with the hazards of highway work zones. Finally, only 22 participants (roughly 17 percent) mentioned that they had previous experience with wearable technology while they were present in a highway work zones.

Figure 1 illustrates the responses from our participants to Question 1. Figure 1 (a) represents the responses from the maintenance crew and Figure 1 (b) demonstrates the responses from affiliated members. Moreover, Figure 2 visualizes the responses that we received from (a) maintenance crew and (b) affiliated members to Question 2. Furthermore, Table 2 summarizes the major concerns of our participants toward using wearable technology in highway work zones as a safety measure.

4.3 Statistical Analysis

In this section, we statistically analyzed the collected data to further investigate our research questions. We first used goodness-of-fit version of chi-square test on all of the collected samples to investigate whether the difference among the number of votes in each category of collected data for both Question 1 and Question 2 in all provided wearable options are statistically significant. That resulted in 20 tests and the results are summarized in Table 1. The obtained p-values are all less than the traditional 0.05 significance cut-off, and therefore mean that the differences among the collected samples in different categories are statistically significant.

Finally, we used the independence version of the chi-square test to statistically compare the responses from both groups and investigate whether groups show a statistically different behavior toward wearable technology. We used this test 5 times in each question (10 times in total), and the results are summarized in Table 3. This table illustrates that with the exception of practicality of clothes, all obtained p-values are larger than the traditional 0.05 level of significance. This denotes that the collected samples are independent from each other, and the null hypotheses of Chi-squared test, which is the samples are independent, is kept.

5 Discussion

In this section, we are discussing our results. Figure 1 shows that while the majority of the maintenance crew seems to feel neutral about the practicality of the selected wearable devices, affiliated members viewed the practicality of wearable technology more positively. In specific, they think that smart clothes and glasses would be a suitable option in highway work zones. This could be attributed to the potential experience with lighting vests and safety glasses, two common "wearable" safety items

in the highway construction industry. Similarly, we can identify a similar pattern in Figure 2 that illustrates the likelihood of using wearable technology in the future. As observed, maintenance crew participants feel more neutral toward using wearable technology in the future and affiliated members are more likely to use wearables, especially smart clothes and glasses. Lack of prior experience, age, job description, and technology resistance all could possibly contribute toward this neutral feeling of the maintenance crew.

Moreover, Table 2 summarizes the responses of our participants to Question 3. This table indicates that most affiliated members and maintenance crew feel that wearable technology is not reliable and trustworthy enough to be used in highway work zones. Both personas cited repetitive false alarms and the unreliability of devices as their major concerns. Participants also recognized the lack of prior experience, the additional burden of wearable technology on top of the already existing PPE items, and being a potential source of distractions as their major concerns. The cost was another factor that affiliated members raised. Finally, Table 3 indicates that the maintenance crew and affiliated members showed a different behavior toward both the practicality and the likelihood of using wearable technology in the future across almost all of the provided wearable options. This further corroborates our initial assumption that affiliated members and maintenance crew might have different opinions and beliefs toward wearable technology due to the differences in their job descriptions and other possible contributors, and our personification was on point. As we discussed, this difference was also reflected in their perception toward the practicality and the likelihood of using wearable technology illustrated in Figures 1 and 2. Deep investigation of this difference, identifying major contributors to these persona creation, and studying the unique needs of each persona could be interesting future research directions.

6 Conclusion

In this article, we reported our early efforts in investigating the perception and preferences of highway workers toward wearable technology for safety-related purposes. We involved the community of highway maintenance and operation in the research through a survey and actively pursued their perception, thoughts and ideas toward wearable technology. Our results indicate a notable potential for future development and investment in wearable-enabled for highway work zones. Future studies can analyze the perception of highway workers by developing more intensive Technology Acceptation Models (TAMs), early prototypes and other tools to further investigate what highway workers think of wearable technology, and how such devices can be assimilated into the work zones.

7 Acknowledgments

This research is supported by the National Science Foundation under Award Number 1932524. The authors would like to thank the North Carolina Department of Transportation (NCDOT), Virginia Department of Transportation (VDOT), and DBi services for their support.

References

- [1] Facts and statistics—work zone safety. *Federal Highway Administration (FHWA)*, 2018.
- [2] John Hourdos. Portable, non-intrusive advance warning devices for work zones with or without flag operators. 2012.
- [3] John A Gambatese, Hyun Woo Lee, Chukwuma Aham Nnaji, et al. Work zone intrusion alert technologies: Assessment and practical guidance. Technical report, Oregon. Dept. of Transportation. Research Section, 2017.
- [4] Mary Ellen Berglund, Julia Duvall, and Lucy E Dunne. A survey of the historical scope and current trends of wearable technology applications. In *Proceedings of the 2016 ACM international symposium on wearable computers*, pages 40–43, 2016.
- [5] Mareike Kritzler, Martin Bäckman, Anders Tenfält, and Florian Michahelles. Wearable technology as a solution for workplace safety. In *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*, pages 213–217, 2015.
- [6] Ibukun Awolusi, Eric Marks, and Matthew Hollowell. Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices. *Automation in construction*, 85: 96–106, 2018.
- [7] Houtan Jebelli, Byungjoo Choi, and SangHyun Lee. Application of wearable biosensors to construction sites. i: Assessing workers' stress. *Journal of Construction Engineering and Management*, 145(12): 04019079, 2019.
- [8] Chukwuma Nnaji, Ali A Karakhan, John Gambatese, and Hyun Woo Lee. Case study to evaluate work-zone safety technologies in highway construction. *Practice Periodical on Structural Design and Construction*, 25(3):05020004, 2020.
- [9] Sepehr Sabeti, Omidreza Shoghli, and Hamed Tabkhi. Toward wi-fi-enabled real-time communication for proactive safety systems in highway work zones: A case study. In *Construction Research Congress 2022*, pages 1166–1173, 2022.

- [10] Byungjoo Choi, Sungjoo Hwang, and SangHyun Lee. What drives construction workers' acceptance of wearable technologies in the workplace?: Indoor localization and wearable health devices for occupational safety and health. *Automation in Construction*, 84:31–41, 2017.
- [11] Sepehr Sabeti, Omidreza Shoghli, Mohammadreza Baharani, and Hamed Tabkhi. Toward ai-enabled augmented reality to enhance the safety of highway work zones: Feasibility, requirements, and challenges. *Advanced Engineering Informatics*, 50: 101429, 2021.
- [12] Chukwuma Nnaji, Ali Jafarnejad, and John Gambatese. Effects of wearable light systems on safety of highway construction workers. *Practice Periodical on Structural Design and Construction*, 25(2): 04020003, 2020.
- [13] Young Ai Kim, Boong Yeol Ryoo, Yong-Su Kim, and Woon Chan Huh. Major accident factors for effective safety management of highway construction projects. *Journal of construction engineering and management*, 139(6):628–640, 2013.
- [14] Aleck Ian Glendon and Debbie K Litherland. Safety climate factors, group differences and safety behaviour in road construction. *Safety science*, 39(3): 157–188, 2001.
- [15] Ramanath Subramanyam, Fei Lee Weisstein, and Mayuram S Krishnan. User participation in software development projects. *Communications of the ACM*, 53(3):137–141, 2010.
- [16] David Brougham and Jarrod Haar. Employee assessment of their technological redundancy. *Labour & Industry: a journal of the social and economic relations of work*, 27(3):213–231, 2017.
- [17] Anjum Naweed and Janette Rose. Assessing technology acceptance for skills development and real-world decision-making in the context of train driving. *Transportation research part F: traffic psychology and behaviour*, 52:86–100, 2018.
- [18] Diandian Liu, Weisheng Lu, and Yuhan Niu. Extended technology-acceptance model to make smart construction systems successful. *Journal of Construction Engineering and Management*, 144(6): 04018035, 2018.
- [19] Hyojoo Son, Yoora Park, Changwan Kim, and Jui-Sheng Chou. Toward an understanding of construction professionals' acceptance of mobile computing devices in south korea: An extension of the technology acceptance model. *Automation in construction*, 28:82–90, 2012.
- [20] Jesse V Jacobs, Lawrence J Hettinger, Yueng-Hsiang Huang, Susan Jeffries, Mary F Lesch, Lucinda A Simmons, Santosh K Verma, and Joanna L Willetts. Employee acceptance of wearable technology in the workplace. *Applied ergonomics*, 78:148–156, 2019.
- [21] Baozhen Dai, Ebenezer Larnyo, Ebenezer Ababio Tetteh, Abigail Konadu Aboagye, and Abdul-Aziz Ibn Musah. Factors affecting caregivers' acceptance of the use of wearable devices by patients with dementia: An extension of the unified theory of acceptance and use of technology model. *American Journal of Alzheimer's Disease & Other Dementias®*, 35:1533317519883493, 2020.
- [22] Ivan E. Sutherland. A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*, AFIPS '68 (Fall, part I), pages 757–764, New York, NY, USA, 1968. ACM. doi:10.1145/1476589.1476686. URL <http://doi.acm.org/10.1145/1476589.1476686>.
- [23] M. Satyanarayanan. From the editor in chief: Augmenting cognition. *IEEE Pervasive Computing*, 3(2):4–5, April 2004. ISSN 1536-1268. doi:10.1109/MPRV.2004.1316809.
- [24] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies. The case for vm-based cloudlets in mobile computing. *IEEE Pervasive Computing*, 8(4):14–23, Oct 2009. ISSN 1536-1268. doi:10.1109/MPRV.2009.82.
- [25] Suranga Seneviratne, Yining Hu, Tham Nguyen, Guohao Lan, Sara Khalifa, Kanchana Thilakarathna, Mahbub Hassan, and Aruna Seneviratne. A survey of wearable devices and challenges. *IEEE Communications Surveys & Tutorials*, 19(4):2573–2620, 2017.
- [26] Rua M Williams and Juan E Gilbert. Perseverations of the academy: A survey of wearable technologies applied to autism intervention. *International Journal of Human-Computer Studies*, page 102485, 2020.
- [27] Ruxandra Tapu, Bogdan Mocanu, and Ermina Tapu. A survey on wearable devices used to assist the visual impaired user navigation in outdoor environments. In *2014 11th international symposium on electronics and telecommunications (ISETC)*, pages 1–4. IEEE, 2014.

- [28] Mokhinabonu Mardonova and Yosoon Choi. Review of wearable device technology and its applications to the mining industry. *Energies*, 11(3):547, 2018.
- [29] V Hyndavi, N Sai Nikhita, and S Rakesh. Smart wearable device for women safety using iot. In *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, pages 459–463. IEEE, 2020.
- [30] Changbum R Ahn, SangHyun Lee, Cenfei Sun, Houtan Jebelli, Kanghyeok Yang, and Byungjoo Choi. Wearable sensing technology applications in construction safety and health. *Journal of Construction Engineering and Management*, 145(11): 03119007, 2019.
- [31] Yizhi Liu, Mahmoud Habibnezhad, Houtan Jebelli, Somayeh Asadi, and SangHyun Lee. Ocular artifacts reduction in eeg signals acquired at construction sites by applying a dependent component analysis (dca). In *Construction Research Congress 2020: Computer Applications*, pages 1281–1289. American Society of Civil Engineers Reston, VA, 2020.
- [32] Xiao Li, Wen Yi, Hung-Lin Chi, Xiangyu Wang, and Albert PC Chan. A critical review of virtual and augmented reality (vr/ar) applications in construction safety. *Automation in Construction*, 86: 150–162, 2018.
- [33] Carlos M Zuluaga, Alex Albert, and Munir A Winkel. Improving safety, efficiency, and productivity: Evaluation of fall protection systems for bridge work using wearable technology and utility analysis. *Journal of Construction Engineering and Management*, 146(2):04019107, 2020.
- [34] Maxwell Fordjour Antwi-Afari and Heng Li. Fall risk assessment of construction workers based on biomechanical gait stability parameters using wearable insole pressure system. *Advanced Engineering Informatics*, 38:683–694, 2018.
- [35] Riad Kanan, Obaidallah Elhassan, and Rofaida Bensalem. An iot-based autonomous system for workers' safety in construction sites with real-time alarming, monitoring, and positioning strategies. *Automation in Construction*, 88:73–86, 2018.
- [36] Vuzix. <https://www.vuzix.com/products/blade-smart-glasses-upgraded>. 2020.

Integrated Sensor-Based Interface for Human-Robot Collaboration in Construction

X. Wang^a, D. Veeramani^b and Z. Zhu^a

^aDepartment of Civil and Environmental Engineering, University of Wisconsin-Madison, 1415 Engineering Drive, Madison, WI 53706, USA

^bDepartment of Industrial and Systems Engineering, University of Wisconsin-Madison, 1513 University Avenue, Madison, WI 53706, USA

E-mail: xwang2463@wisc.edu, raj.veeramani@wisc.edu, and zzhu286@wisc.edu

Abstract –

Construction robots have the potential to increase construction productivity at job sites and can help overcome industry challenges such as labor shortage and safety risks. User-friendly interfaces are critical for advancing human-robot work collaboration and increasing use of construction robots. However, human-robot interfaces in the context of construction industry applications have been investigated to a limited extent only. This paper proposes a novel sensor-based framework which integrates eye tracking and hand gesture recognition for human-robot interaction in construction. Specifically, it begins with visual detection of construction machines in the first-person views. Then, the machine-of-interest is determined based on the detection results and human gaze points. Finally, a real-time hand gesture recognition system is employed for conveying messages to the machine to guide its operations. So far, the proposed framework was tested in a laboratory setting using a robotic dump truck. The results showed that the proposed framework could serve an effective interface to support the interactions between workers and construction machines.

Keywords –

Wearable Sensors; Eye Tracking; Hand Gesture Recognition; Human-Robot Interface

1 Introduction

The construction industry is facing a unique set of challenges, such as low productivity, poor safety records, labor shortage [1,2], etc. Through years of development, construction robots and autonomous machines have demonstrated their potential to improve the construction industry [3]. They have the functional ability to perform construction tasks that are impossible, undesirable, or unsafe for human workers [4]. Also,

construction robots have the potential to enhance quality and efficiency of job site operations [5].

Recent advances in robotics make it possible for human-robot collaboration on construction sites [6,7]. This collaboration helps workers transfer some of their current duties to robots and instead devote their effort on high-level planning and cognitive work as robot supervisors [8]. Human workers can also benefit from the assistance of robots in performing repetitive physically-demanding tasks [6]. To maximize the benefits from human-robot work collaboration, a user-friendly interface is critical to support their interactions. However, human-robot interfaces in the context of construction is a less explored field [9].

A variety of interfaces, including visual displays, hand gestures, speech language, and eye tracking, have been developed for human-robot interactions in various industries [10–13]. Among them, non-verbal communication, such as hand gestures and eye tracking, is deemed to be an effective channel in noisy construction environments [12]. As natural and intuitive interfaces, they can provide a standard mode for workers from different backgrounds and cultures to convey correct instructions to a robot [14].

There are many research studies proposed for developing human-robot interfaces based on different types of sensors. The employed sensors include electromyography (sEMG) sensors [15], Inertial Measurement Unit (IMU) [16], radar sensors [17], infrared technology [18], etc. These studies aimed to interpret subjects' intentions [16], understand sign language [19], express human emotions [18], etc. They either relied on hand-crafted features [15] or deep neural networks [20]. The results illustrated the potential of deep neural networks for performing recognition with excellent learning ability.

Although the performance of existing interfaces is promising, one significant challenge they face is in dealing with uncertainty and ambiguity that commonly arise in unstructured and dynamic environments such as

construction sites. It has been well noted that one type of sensor data may not be enough to address the uncertainty and resolve unambiguity. This paper proposes a sensor-based framework which integrates eye tracking and hand gesture recognition for human-robot interaction in construction. In this approach, visual detection of construction machines is first conducted using first person view frames. Based on the detection results and gaze points, the machine-of-interest is then defined. Finally, a real-time system is employed for hand gesture recognition. The recognized gesture would be sent to the machine-of-interest. The effectiveness of the framework was tested in a laboratory study to interact with a robotic dump truck. The results showed that the proposed framework can be used to serve as an effective interface for workers to interact with construction machines.

2 Related Work

Various research studies have been conducted to develop human-robot interfaces. They relied on hand gesture recognition, eye tracking, smart glasses, etc. An overview of these studies is provided below.

2.1 Hand Gesture Recognition

Hand gestures, as a common way to express intent, have various applications in human machine interaction due to their simple, yet effective, nature [21–24]. Various research studies have been conducted to achieve hand gesture recognition. They can be classified into two categories, vision-based methods [23,25] and wearable sensors-based methods [26,27], depending on the type of data source they relied on. Vision-based methods generally relied on hand-crafted features, such as Improved Dense Trajectories (iDT) [28] and Mix Features Around Sparse Keypoints (MFSK) [29]. With technical development, the use of deep learning technologies has become mainstream in gesture recognition. For example, Molchanov et al. [24] combined 3D Convolutional Neural Network (CNN) with recurrent layers to perform simultaneous detection and classification of dynamic hand gestures. The recurrent 3D-CNN enabled the gesture classification without requiring explicit pre-segmentation. Cao et al. [30] presented a framework of C3D+LSTM+RSTTM which augmented C3D with a recurrent spatiotemporal transform module. The presented framework could not only capture short-term spatiotemporal features but also model long-term dependencies. Köpüklü et al. [23] proposed a hierarchical CNN structure to realize the real-time hand gesture recognition. The proposed architecture firstly employed a detector which was a lightweight 3D-CNN to detect the existence of hand gestures and then utilized deep 3D-CNNs to classify the

detected gestures.

Motion sensory data provide an alternative data source for hand gesture recognition. For instance, Su et al. [31] presented a robust hand gesture recognition framework based on random forests. The random forests were established using improved decision trees which included the pre-classifiers to avoid the misclassification of gestures with similar features. Côté-Allard et al. [32] applied CNNs on aggregated data from multiple users to identify hand gestures. In their work, CNNs were combined with transfer learning to decrease the data requirement of the training model. Fang et al. [27] designed a new CNN architecture named SLRNet to achieve dynamic gesture recognition. The CNN architecture extracted the features of two hands and fused the features into the fully connected layer. Yuan et al. [20] proposed an improved deep feature fusion network to detect long distance dependency in complex hand gestures. In their work, a LSTM model with fused feature vectors was introduced to classify complex hand motions into corresponding categories.

2.2 Eye Tracking

Conventionally, eye tracking has been regarded as one of the most visible cues for user behavior/intention recognition [33]. There are many efforts dedicated to developing reliable eye tracking-based methods. Zhang et al. [34] presented a novel eye tracking-learning-detection algorithm with tracking feedback. The detection area was adjusted adaptively and narrowed by the tracking feedback to adapt to situations where the human eye was partially blocked or had morphological changes. Santini et al. [35] introduced a novel method named Pupil Reconstructor with Subsequent Tracking (PuReST) for fast and robust pupil tracking. The PuReST consists of three distinct parts: initial pupil detection, shared tracking preamble, outline and greedy tracker. Laddi and Prakash [36] proposed an unobtrusive and calibration-free framework for an eye gaze tracking based interface for a desktop environment. The proposed eye gaze tracking involved a hybrid approach wherein the unsupervised image gradients method computed the iris centers over the eye regions extracted by the supervised regression-based algorithm. Cubero and Rehm [37] relied on eye tracking to obtain eye gazes and then developed an LSTM-based machine learning model to classify human intent. As the technology matures, commercial eye tracking products such as Tobii glasses 3 [38] and Pupil Core [39] are becoming available in the market and have various potential fields of application.

2.3 Smart Glasses

There are other commonly used human-robot

interfaces including augmented reality (AR) / mixed reality (MR) / virtual reality (VR) glasses, etc. For example, Wang et al. [40] presented a method of manufacture assembly fault detection based on AR. The augmented information interactions made the manufacturing and assembly inspection process more visual and intuitive. Du et al. [41] proposed a novel teleoperation method that allowed users to guide robots through a combined form of AR glasses and Leap Motion Controllers. Users could observe the virtual robot from an arbitrary angle, which enhanced the users' interactive immersion and provided more natural human-machine interaction. Wallmyr et al. [42] employed MR interfaces to display information within the excavator operators' field of view, which enhanced information detectability through quick glances. This practice could help lower operator's mental workload together with an improved rate in detection of presented information. However, their main adoption limitation lies in the expensive hardware and training; and also the AR and MR technologies behind those smart glasses are still not mature and/or suitable enough for engineering and construction [43].

3 Proposed Framework

The overview of the proposed framework is illustrated in Figure 1. The framework consists of three components: visual detection, machine-of-interest generation, and hand gesture recognition. Specifically, the visual detection of construction machines from first person view frames is first conducted. Based on the detection results and gaze points, the machine-of-interest is then generated. Finally, a real-time system is employed to achieve hand gesture recognition.

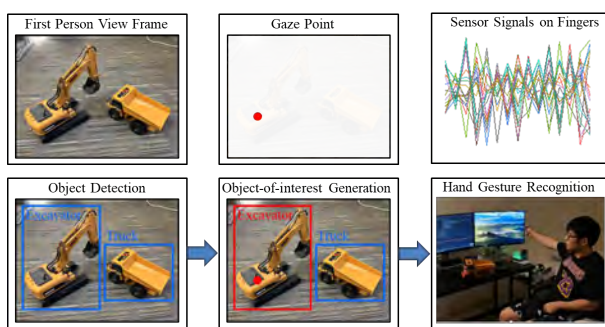


Figure 1. Framework for human-robot collaboration

3.1 Visual Detection

In this component of the framework, an object detection algorithm is employed to extract the regions of construction machines in video sequences. YOLOv3 [44] is selected in this study to detect the construction

machines because many research results have verified the high performance of YOLOv3 in various construction object detection scenarios [45,46]. The YOLOv3 system can be generally divided into two steps: feature extraction and detection. First, Darknet-53 is applied to extract features of the whole image and obtain feature embeddings at different scales. Then, these features are fed into different branches of the detector to get bounding boxes and class information. The coordinates of bounding boxes from the detection results are then used as the input for the object-of-interest generation process.

3.2 Machine-of-Interest Generation

In this component of the framework, the machine-of-interest is generated based on the bounding boxes of construction machines and gaze points. This component can be divided into three steps: synchronization for the bounding boxes and gaze points, determination of the machine-of-interest, and interaction mode triggering. First, the bounding boxes and gaze points are synchronized based on a unified timestamp since they are produced or derived from different sensors. Then, the machine is determined as the machine-of-interest if the gaze point resides in its bounding box. As for the triggering of the interaction mode, if the gaze points stay in the bounding box of the machine-of-interest for a duration longer than a threshold τ , the machine will enter the interaction mode and the hand gesture recognition component can then be applied to convey messages to the machine; otherwise, it means that the framework is not confident regarding which machine the user desires to interact with. It should be noted that the selection of τ depends on how likely the user intends to trigger the interaction mode. Here, based on preliminary trials, τ has been set as 0.3 second.

3.3 Hand Gesture Recognition

The purpose of this component of the framework is to apply a hand gesture recognition system to convey messages to the object-of-interest. Specifically, the accelerometer and gyroscope signals are directly captured from sensors attached on fingers as raw data. Several techniques including sampling rate synchronization and Z score normalization are used to preprocess the raw data. Then, a sliding window approach is designed to achieve real-time classification of hand gestures. With the signals coming in continuously, the window moves through the whole set of signals and the preprocessed data in the latest window are fed into a Fully Convolutional Network (FCN)-based classifier to achieve hand gesture recognition. If the highest probability of the classifier is more than a threshold θ , the identification of the hand

gesture is confirmed. Here, based on preliminary trials, θ has been set as 0.95.

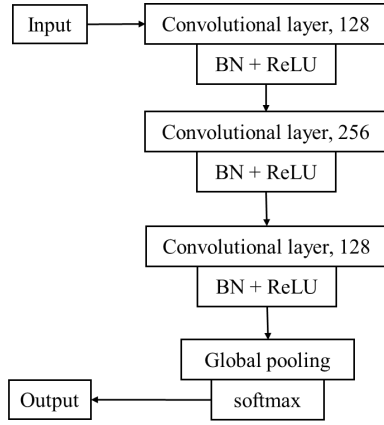


Figure 2. The architecture of FCN classifier

The FCN developed by Wang et al. [47] is selected here since it is superior for multivariate time series classification tasks compared to other deep learning networks [48]. Figure 2 shows an overview of the FCN architecture. It comprises three convolutional blocks where each block contains three operations: a convolution followed by a batch normalization whose result is fed to a ReLU activation function. The result of the third convolutional block is averaged over the entire time dimension which corresponds to the Global Average Pooling (GAP) layer. Finally, a traditional softmax classifier is fully connected to the GAP layer's output.

4 Implementation and Results

4.1 Sensor Selection

Pupil Core [39] is employed as the eye tracking device to get the first person view frames and track the eye gaze points. It is selected since Pupil Core is an open-sourced software, which is conducive for user developments. The structure of Pupil Core is shown in Figure 3. A scene camera is mounted on the front of the eye tracking device to get the first-person view frames. Two eye cameras are facing towards two eyes, separately, to obtain their gaze points.

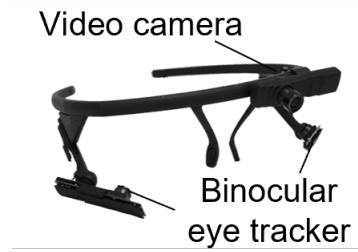


Figure 3. The structure of Pupil Core (adapted from [39])

To capture the hand motions, Tap Strap 2 [49] is selected as the wearable sensor. Compared to other wearable sensors like data gloves which are not easy or comfortable to wear, the Tap sensor is portable, lightweight and easy to wear on the fingers. This is beneficial for the construction workers to complete the tasks using their hands. As shown in Figure 4, the Tap sensor includes five 3-axis accelerometers and one IMU (3-axis accelerometer + 3-axis gyroscope). The five accelerometers are located at five fingers, separately, while IMU is placed on the thumb. In total, there are 21 signal channels captured by the Tap sensor.

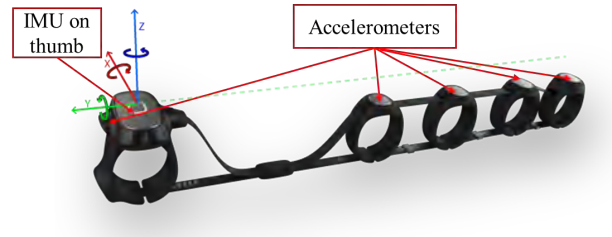


Figure 4. The structure of Tap Strap 2

4.2 Offline Training for Hand Gesture Recognition

The offline training has been conducted on an Ubuntu Linux 64-bit operating system. The hardware configuration is as follows: an Intel® Core™ i7-4820K CPU (Central Processing Unit) @ 3.70 GHz, a 32 GB memory, and an NVIDIA Titan Xp DDR5X @ 12.0 GB GPU (Graphics Processing Unit).

The dataset created in [50] was employed to conduct offline training for hand gesture recognition. The dataset is randomly split into training (66.7%), validation (16.7%) and test (16.6%) sets, resulting in 128 training, 32 validation and 32 test gestures.

For training, the learning rate and the batch size are set as large as possible, i.e., 0.0001 and 16, respectively. When the loss is steady, the learning rate is reduced with a fixed decay factor which is set to 10. Stochastic gradient descent is employed as the optimizer. Table 1

provides a summary of the recognition performance of FCN. The accuracies on validation and test sets are 96.9% and 87.5%, respectively. The inference time achieved on validation and test sets are 0.13 second and 0.14 second, respectively.

Table 1. Recognition performance of FCN

Indexes	Validation set	Test set
Accuracy (%)	96.9	87.5
Inference time (s)	0.13	0.14

4.3 Laboratory Study

A laboratory study was conducted to test whether the proposed framework could serve as an interface to help workers control and/or interact with construction machines. Specifically, the user was asked to stare at the construction machine he/she intended to interact with and then perform hand gestures. The first-person view frames and gaze points were captured by a Pupil core while the hand motions were obtained by a Tap sensor. All these data were transferred to a computer and input into the framework in real time. Based on the recognition results, the corresponding instructions would be sent to a remote controller, where the control signals would be transmitted to operate the truck model remotely.

Figure 5 shows an example of using the proposed framework to remotely control a toy truck to lift its dump box. The user first stared at the truck and made the hand gesture of “hoist” to request the truck model to lift its dump box. The gesture was captured by the framework and the corresponding instruction was sent to the truck model through the remote controller. Following the instruction, the truck model lifted its dump box gradually (Frames 185 and 221). After a short pause, the user stared at an irrelevant place and performed the gesture of “hoist” again. Since the truck did not enter the interaction mode, no recognition results were incurred (Frames 374 and 401).

Although the laboratory study illustrated the feasibility of using the proposed framework as human-robot interface, there are still several technical challenges to be addressed before it can be applied at construction sites. First, the performance of the framework highly depends on visual detection of construction machines. Considering that construction sites are complex and cluttered with tools, materials, workers, etc. The trained detection model needs to be robust to accommodate such complicated characteristics of the environment. Second, the gaze point accuracy is critical for determining which machine the user intends to interact with. However, several complicating factors

at construction sites, such as diverse weather conditions and sunlight intensities, pose challenges for accurate estimation of eye gaze points.

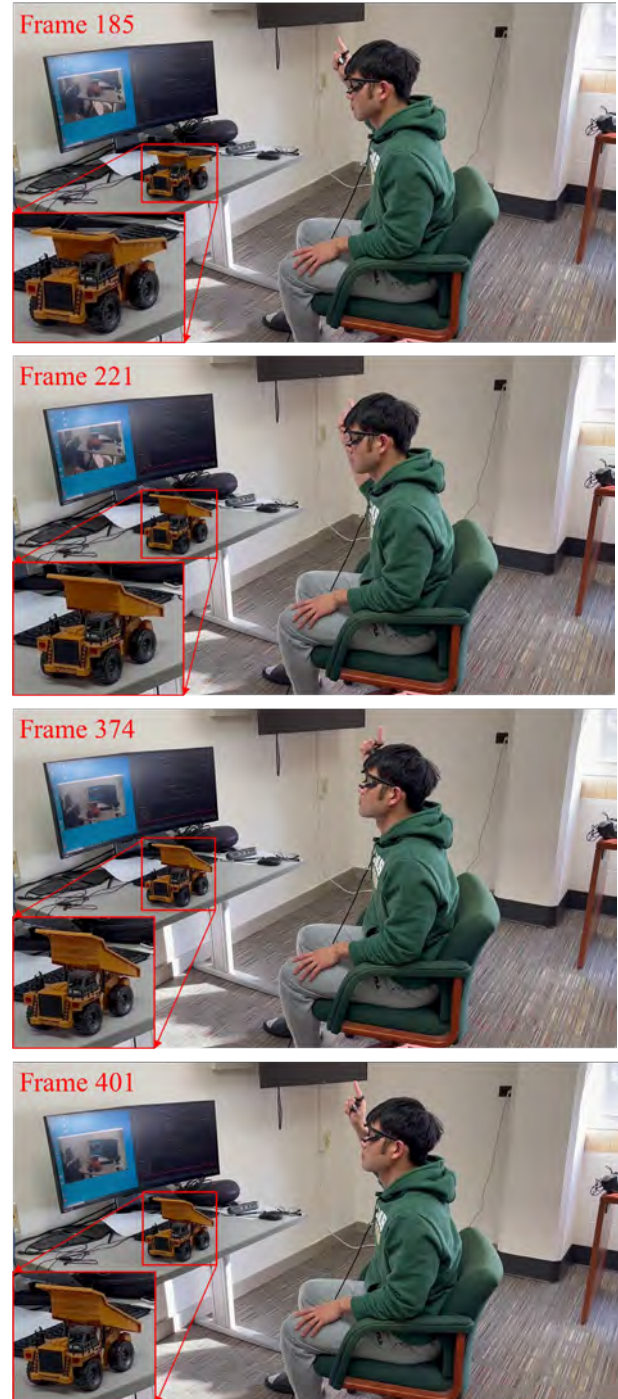


Figure 5. Demonstration of integrated eye-tracking and gesture-based control

5 Conclusions and Future Work

While construction robots have the potential to offer significant benefits to the construction industry, their increased adoption will require user-friendly interfaces for human-robot collaboration. So far, work on human-robot interfaces in the context of the construction domain is limited. This paper has proposed a sensor-based framework which integrates eye tracking and hand gesture recognition for human-robot interaction in construction. The framework comprises three components: visual detection, machine-of-interest generation, and hand gesture recognition. The effectiveness of the framework was tested with a laboratory study to interact with a robotic dump truck. The results show that the proposed framework is suitable for developing an interface to help workers interact with construction machines.

Future work will focus on including more classes of construction gestures into the dataset to make the training and testing of hand gesture classifiers more robust. Additionally, it will investigate the development of a human-robot interaction system using the proposed framework.

Acknowledgment

This paper is based in part upon the work supported by the Wisconsin Alumni Research Foundation (WARF) under Project No. AAJ4872 and the M.A. Mortenson Company. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of WARF or Mortenson.

References

- [1] X. Li, W. Yi, H.L. Chi, X. Wang, A.P.C. Chan, A critical review of virtual and augmented reality (VR/AR) applications in construction safety, *Autom. Constr.* (2018). <https://doi.org/10.1016/j.autcon.2017.11.003>.
- [2] S. Kim, S. Chang, D. Castro-Lacouture, Dynamic Modeling for Analyzing Impacts of Skilled Labor Shortage on Construction Project Management, *J. Manag. Eng.* (2020). [https://doi.org/10.1061/\(asce\)me.1943-5479.0000720](https://doi.org/10.1061/(asce)me.1943-5479.0000720).
- [3] Q. Chen, B. García de Soto, B.T. Adey, Construction automation: Research areas, industry concerns and suggestions for advancement, *Autom. Constr.* (2018). <https://doi.org/10.1016/j.autcon.2018.05.028>.
- [4] H. Ardiny, S. Witwicki, F. Mondada, Construction automation with autonomous mobile robots: A review, in: *Int. Conf. Robot. Mechatronics, ICRoM 2015, 2015*. <https://doi.org/10.1109/ICRoM.2015.7367821>.
- [5] T. Bock, The future of construction automation: Technological disruption and the upcoming ubiquity of robotics, *Autom. Constr.* (2015). <https://doi.org/10.1016/j.autcon.2015.07.022>.
- [6] S. You, J.H. Kim, S.H. Lee, V. Kamat, L.P. Robert, Enhancing perceived safety in human-robot collaborative construction using immersive virtual environments, *Autom. Constr.* (2018). <https://doi.org/10.1016/j.autcon.2018.09.008>.
- [7] A. Bauer, D. Wollherr, M. Buss, Human-robot collaboration: A survey, *Int. J. Humanoid Robot.* (2008). <https://doi.org/10.1142/S0219843608001303>.
- [8] S. You, T. Ye, L.P. Robert, Team Potency and Ethnic Diversity in Embodied Physical Action (EPA) Robot-Supported Dyadic Teams, in: *ICIS 2017 Transform. Soc. with Digit. Innov.*, 2018. <http://aisel.aisnet.org/icis2017/HumanBehavior/Presentations/3/>.
- [9] J. Czarnowski, A. Dąbrowski, M. Maciaś, J. Głowska, J. Wrona, Technology gaps in Human-Machine Interfaces for autonomous construction robots, *Autom. Constr.* (2018). <https://doi.org/10.1016/j.autcon.2018.06.014>.
- [10] P. Majaranta, A. Bulling, Eye Tracking and Eye-Based Human-Computer Interaction, in: 2014. https://doi.org/10.1007/978-1-4471-6392-3_3.
- [11] M.A. Goodrich, A.C. Schultz, Human-robot interaction: A survey, *Found. Trends Human-Computer Interact.* (2007). <https://doi.org/10.1561/11000000005>.
- [12] J. Berg, S. Lu, Review of Interfaces for Industrial Human-Robot Interaction, *Curr. Robot. Reports.* (2020). <https://doi.org/10.1007/s43154-020-00005-6>.
- [13] R.C. Luo, L. Mai, Human Intention Inference and On-Line Human Hand Motion Prediction for Human-Robot Collaboration, in: *IEEE Int. Conf. Intell. Robot. Syst.*, 2019. <https://doi.org/10.1109/IROS40897.2019.8968192>.
- [14] P.E. Hagan, J.F. Montgomery, J.T. O'Reilly, Accident prevention manual for business & industry: engineering & technology, National Safety Council, 2015.
- [15] H. Su, S.E. Ovrur, X. Zhou, W. Qi, G. Ferrigno, E. De Momi, Depth vision guided hand gesture recognition using electromyographic signals, *Adv. Robot.* (2020). <https://doi.org/10.1080/01691864.2020.1713886>.

- [16] T.Y. Pan, C.Y. Chang, W.L. Tsai, M.C. Hu, OrsNet: A hybrid neural network for official sports referee signal recognition, in: *MMSports 2018 - Proc. 1st Int. Work. Multimed. Content Anal. Sport. Co-Located with MM 2018*, 2018. <https://doi.org/10.1145/3265845.3265849>.
- [17] Z. Zhang, Z. Tian, M. Zhou, Latern: Dynamic Continuous Hand Gesture Recognition Using FMCW Radar Sensor, *IEEE Sens. J.* (2018). <https://doi.org/10.1109/JSEN.2018.2808688>.
- [18] J.Z. Lim, J. Mountstephens, J. Teo, Emotion recognition using eye-tracking: Taxonomy, review and current challenges, *Sensors (Switzerland)*. (2020). <https://doi.org/10.3390/s20082384>.
- [19] S.A. Khomami, S. Shamekhi, Persian sign language recognition using IMU and surface EMG sensors, *Meas. J. Int. Meas. Confed.* (2021). <https://doi.org/10.1016/j.measurement.2020.108471>.
- [20] G. Yuan, X. Liu, Q. Yan, S. Qiao, Z. Wang, L. Yuan, Hand Gesture Recognition Using Deep Feature Fusion Network Based on Wearable Sensors, *IEEE Sens. J.* (2021). <https://doi.org/10.1109/JSEN.2020.3014276>.
- [21] G. Yasmeen, S. Arun, J.N. Swaminathan, S.A.K. Jilani, S. Asif, Efficient Hand Gesture Recognition for Traffic Control System Using ti Sensor Tag, in: *2018 Int. Conf. Comput. Commun. Informatics, ICCCI 2018*, 2018. <https://doi.org/10.1109/ICCCI.2018.8441483>.
- [22] Z. Lu, X. Chen, Q. Li, X. Zhang, P. Zhou, A hand gesture recognition framework and wearable gesture-based interaction prototype for mobile devices, *IEEE Trans. Human-Machine Syst.* (2014). <https://doi.org/10.1109/THMS.2014.2302794>.
- [23] O. Köpüklü, A. Gunduz, N. Kose, G. Rigoll, Real-time hand gesture detection and classification using convolutional neural networks, in: *Proc. - 14th IEEE Int. Conf. Autom. Face Gesture Recognition, FG 2019*, 2019. <https://doi.org/10.1109/FG.2019.8756576>.
- [24] P. Molchanov, X. Yang, S. Gupta, K. Kim, S. Tyree, J. Kautz, Online Detection and Classification of Dynamic Hand Gestures with Recurrent 3D Convolutional Neural Networks, in: *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2016. <https://doi.org/10.1109/CVPR.2016.456>.
- [25] O. Koller, C. Camgoz, H. Ney, R. Bowden, Weakly Supervised Learning with Multi-Stream CNN-LSTM-HMMs to Discover Sequential Parallelism in Sign Language Videos, *IEEE Trans. Pattern Anal. Mach. Intell.* (2019). <https://doi.org/10.1109/tpami.2019.2911077>.
- [26] A.A. Neacsu, G. Cioroiu, A. Radoi, C. Burileanu, Automatic EMG-based hand gesture recognition system using time-domain descriptors and fully-connected neural networks, in: *2019 42nd Int. Conf. Telecommun. Signal Process. TSP 2019*, 2019. <https://doi.org/10.1109/TSP.2019.8768831>.
- [27] B. Fang, Q. Lv, J. Shan, F. Sun, H. Liu, D. Guo, Y. Zhao, Dynamic gesture recognition using inertial sensors-based data gloves, in: *2019 4th IEEE Int. Conf. Adv. Robot. Mechatronics, ICARM 2019*, 2019. <https://doi.org/10.1109/ICARM.2019.8834314>.
- [28] H. Wang, D. Oneata, J. Verbeek, C. Schmid, A Robust and Efficient Video Representation for Action Recognition, *Int. J. Comput. Vis.* (2016). <https://doi.org/10.1007/s11263-015-0846-5>.
- [29] J. Wan, G. Guo, S.Z. Li, Explore Efficient Local Features from RGB-D Data for One-Shot Learning Gesture Recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* (2016). <https://doi.org/10.1109/TPAMI.2015.2513479>.
- [30] C. Cao, Y. Zhang, Y. Wu, H. Lu, J. Cheng, Egocentric Gesture Recognition Using Recurrent 3D Convolutional Neural Networks with Spatiotemporal Transformer Modules, in: *Proc. IEEE Int. Conf. Comput. Vis.*, 2017. <https://doi.org/10.1109/ICCV.2017.406>.
- [31] R. Su, X. Chen, S. Cao, X. Zhang, Random forest-based recognition of isolated sign language subwords using data from accelerometers and surface electromyographic sensors, *Sensors (Switzerland)*. (2016). <https://doi.org/10.3390/s16010100>.
- [32] U. Côté-Allard, C.L. Fall, A. Drouin, A. Campeau-Lecours, C. Gosselin, K. Glette, F. Laviolette, B. Gosselin, Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning, *IEEE Trans. Neural Syst. Rehabil. Eng.* (2019). <https://doi.org/10.1109/TNSRE.2019.2896269>.
- [33] D.Y. Cho, M.K. Kang, Human gaze-aware attentive object detection for ambient intelligence, *Eng. Appl. Artif. Intell.* (2021). <https://doi.org/10.1016/j.engappai.2021.104471>.
- [34] J. Zhang, Y. Wu, H. Huang, G. Hou, A New Human Eye Tracking Method Based on Tracking Module Feedback TLD Algorithm, in: *Proc. - 20th Int. Conf. High Perform. Comput. Commun. 16th Int. Conf. Smart City 4th Int. Conf. Data Sci. Syst. HPCC/SmartCity/DSS 2018*, 2019. <https://doi.org/10.1109/HPCC/SmartCity/DSS.2019.8834314>.

- 018.00071.
- [35] T. Santini, W. Fuhl, E. Kasneci, PuReST: Robust pupil tracking for real-time pervasive eye tracking, in: *Eye Track. Res. Appl. Symp.*, 2018. <https://doi.org/10.1145/3204493.3204578>.
- [36] A. Laddi, N.R. Prakash, Eye gaze tracking based directional control interface for interactive applications, *Multimed. Tools Appl.* (2019). <https://doi.org/10.1007/s11042-019-07940-3>.
- [37] C. Gomez Cubero, M. Rehm, Intention Recognition in Human Robot Interaction Based on Eye Tracking, in: *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 2021. https://doi.org/10.1007/978-3-030-85613-7_29.
- [38] Tobii Inc., Tobii Pro Glasses 3, (2021). <https://www.tobiipro.com/product-listing/tobii-pro-glasses-3/> (accessed February 3, 2022).
- [39] Pupul Labs, Pupil Core, (2021). <https://pupil-labs.com/products/core/> (accessed February 3, 2022).
- [40] S. Wang, R. Guo, H. Wang, Y. Ma, Z. Zong, Manufacture assembly fault detection method based on deep learning and mixed reality, in: *2018 IEEE Int. Conf. Inf. Autom. ICIA 2018*, 2018. <https://doi.org/10.1109/ICInfA.2018.8812577>.
- [41] G. Du, B. Zhang, C. Li, H. Yuan, A novel natural mobile human-machine interaction method with augmented reality, *IEEE Access.* (2019). <https://doi.org/10.1109/ACCESS.2019.2948880>.
- [42] M. Wallmyr, T.A. Sitompul, T. Holstein, R. Lindell, Evaluating Mixed Reality Notifications to Support Excavator Operator Awareness, in: *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 2019. https://doi.org/10.1007/978-3-030-29381-9_44.
- [43] J.M. Davila Delgado, L. Oyedele, T. Beach, P. Demian, Augmented and Virtual Reality in Construction: Drivers and Limitations for Industry Adoption, *J. Constr. Eng. Manag.* (2020). [https://doi.org/10.1061/\(asce\)co.1943-7862.0001844](https://doi.org/10.1061/(asce)co.1943-7862.0001844).
- [44] J. Redmon, A. Farhadi, YOLOv3: An incremental improvement, *ArXiv.* (2018). <https://arxiv.org/abs/1804.02767>.
- [45] F. Wu, G. Jin, M. Gao, Z. He, Y. Yang, Helmet detection based on improved YOLO V3 deep model, in: *Proc. 2019 IEEE 16th Int. Conf. Networking, Sens. Control. ICNSC 2019*, 2019. <https://doi.org/10.1109/ICNSC.2019.8743246>.
- [46] X. Luo, H. Li, H. Wang, Z. Wu, F. Dai, D. Cao, Vision-based detection and visualization of dynamic workspaces, *Autom. Constr.* (2019). <https://doi.org/10.1016/j.autcon.2019.04.001>.
- [47] Z. Wang, W. Yan, T. Oates, Time series classification from scratch with deep neural networks: A strong baseline, in: *Proc. Int. Jt. Conf. Neural Networks*, 2017. <https://doi.org/10.1109/IJCNN.2017.7966039>.
- [48] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, P.A. Muller, Deep learning for time series classification: a review, *Data Min. Knowl. Discov.* (2019). <https://doi.org/10.1007/s10618-019-00619-1>.
- [49] Tap Systems Inc., Meet Tap, (2021). <https://www.tapwithus.com/>.
- [50] X. Wang, Z. Zhu, Automatic Recognition of Construction Workers' Hand Gestures Based on Wearable Sensors, *Autom. Constr. In Review* (2022).

Requirements analysis of additive manufacturing for concrete printing – A systematic review

P. Peralta-Abadia and K. Smarsly

Institute of Digital and Autonomous Construction, Hamburg University of Technology, Germany
Email: patricia.peralta.abadia@tuhh.de, kay.smarsly@tuhh.de

Abstract –

The acceptance of concrete printing as a viable construction method is limited because of a lack of expertise and due to the heterogeneous and non-standardized nature of additive manufacturing (AM) data modeling, affecting the reliability and the interoperability of the concrete printing process. To advance standardization of AM data modeling in concrete printing, information exchange requirements must be defined along the digital thread, i.e. the digital workflow that transforms 3D models into printed components. In this paper, a requirements analysis of AM for concrete printing is conducted through a systematic review. The AM process for concrete printing is defined, identifying information exchange requirements. Sources relevant to AM and concrete printing are systematically reviewed, collecting and analyzing attributes of the information exchange requirements for concrete printing. As a result, the requirement analysis serves as basis to standardize the digital thread, in an attempt to advance reliability and interoperability of the concrete printing process.

Keywords –

Additive manufacturing (AM); Concrete printing; Data modeling; Information exchange requirements; Requirement analysis

1 Introduction

In the architecture, engineering, and construction (AEC) industry, research has been conducted to automate construction processes that are based on additive manufacturing (AM). AM allows structures to be built in a layer-by-layer basis, employing computer-controlled processes [1]. Using printable concrete, large-scale building components have been manufactured by deploying concrete-based AM processes, also referred to as concrete printing [2]. In concrete printing, interdependencies of the material and the manufacturing process affect the quality of manufacturing, thus the quality of the printed components. To ensure high-quality

concrete components by successfully conducting manufacturing processes, expertise and a common understanding of concrete printing are required [3]. The acceptance of concrete printing as a viable construction method has been limited due to a lack of expertise and understanding and because of the heterogeneous and non-standardized approaches commonly deployed for AM data modeling, for material testing, and for manufacturing, each of which affecting the reliability and interoperability of the concrete printing process. New data modeling approaches proposed for concrete printing, encompassing the digital workflow to transform 3D models into printed components (i.e. digital thread), have to be developed to improve reliability and interoperability. Synergies between conventional AM methods and concrete printing can be exploited to formally describe AM data modeling for concrete printing.

In concrete printing, actors of heterogeneous domains (e.g. design, engineering, material sciences, and machine operation) collaborate along the digital thread to transform the 3D models into printed components, exchanging information to perform concrete printing tasks and subprocesses. Information exchange requirements describe information of a task or subprocess (i.e. a set of tasks) that are exchanged between the actors to enable downstream tasks or subprocesses [4]. However, information exchange requirements along the digital thread have not been clearly defined, causing a loss of semantic information and a lack of interoperability. Concepts of building information modeling may advantageously be used to aid the standardization of the information exchange requirements, preserving semantic information and improving interoperability along the digital thread, while further advancing the acceptance of concrete printing as a viable construction method.

In this paper, a requirements analysis of AM for concrete printing is conducted through a systematic review. Information exchange requirements are defined, following the methodology developed for information delivery manuals known widely used in building information modeling (BIM), which documents

processes and describes the corresponding information to be exchanged between the actors [4]. The paper is organized as follows. First, the AM process for concrete printing is defined and the information exchange requirements are identified. Second, sources relevant to AM and concrete printing are systematically reviewed to collect attributes of the information exchange requirements along the digital thread in concrete printing. The information exchange requirements are analyzed, using AM design and optimization as well as AM process planning as illustrative examples. The paper concludes with a summary and an outlook on potential future research.

2 Additive manufacturing process for concrete printing

An extract of the digital workflow for concrete printing, from design to print, is represented as a process map shown in Figure 1. A process can be nested, and it may contain subprocesses. A process (as well as a subprocess) is a set of tasks that are interrelated or that interact with one another, transforming inputs into outputs. The process map shown in Figure 1 is based on [5] and follows the business process modeling notation. Four main actors participating in the workflow are identified: designer, engineer, material scientist, and machine operator. The actors develop specific tasks or subprocesses (differentiated by colors), where the data generated in the tasks or in the subprocesses is exchanged among the actors, following a sequence that translates 3D models into printed objects.

As can be seen from Figure 1, the AM process starts with design concepts, from which design specifications are defined. With the design specifications, geometric models are generated considering manufacturability, and manufacturing hardware is selected. Settings for the manufacturing hardware are defined to design the material (concrete) in an iterative process, until satisfying the design specifications. Once the material is designed, material specifications are generated. Then, the geometric models are sliced and toolpaths are planned according to the process data and the material specifications. Within the subprocess of toolpath planning, simulations of the manufacturing process and of the material are carried out. The subprocess of toolpath planning has AM models as outputs. Then, the AM models are evaluated and, if accepted, the AM models are used as basis to generate machine-readable code (CNC code) that provides the instructions for manufacturing.

Data modeling in concrete printing has synergies with the data modeling approaches used for conventional AM methods, where efforts to standardize basic requirements have been described in [6]. However, considerations regarding the interdependencies of the concrete and the

manufacturing process are to be included in the information exchange requirements. A common data model will support interoperability along the digital thread as well as data collection and storage. For the sake of brevity, only the consideration of the material interdependencies in the semantics of AM design and optimization and of AM process planning are reviewed in this paper. In the following section, the requirements of AM for concrete printing are reviewed and analyzed, focusing on AM design and optimization as well as on AM process planning as illustrative examples.

3 Review and analysis of requirements of additive manufacturing for concrete printing

In this section, the systematic review and the analysis of the requirements of AM for concrete printing are presented. Due to the synergies between AM and concrete printing, sources relevant to both areas are systematically reviewed. First, the systematic review of the sources is provided. Then, the requirements analysis according to completeness and interoperability is presented.

3.1 Systematic review

The review methodology comprises three steps, (i) source selection, (ii) data collection, and (iii) data organization. Sources, precisely *standards*, *current research* (i.e. journal papers and conference papers) and *software applications*, are selected for the review to answer the question “what information is necessary for data modeling of concrete printing?” The papers are indexed in the Web of Science Core Collection, in the Scopus database, or in the American Society of Mechanical Engineers digital collection. An initial search using keywords such as “additive manufacturing”, “digital fabrication”, or “3D printing” together with “concrete”, “ontology”, “modeling”, “simulation”, and/or “digital thread” is carried for the period between 2015 and 2021. Publications with sufficient citations and with documentation of parameters relevant to data modeling in AM and in concrete printing are selected. Additionally, software applications commonly used in AM and concrete printing are selected according to availability of user manuals and implementation in documented studies. A total of 30 sources relevant to AM and concrete printing have been selected: 5 standards, 17 journal papers, 6 conference papers, and 2 software applications. From the sources, attributes of the information exchange are collected and organized in information units. In the following paragraphs, an overview of the systematic review is presented.

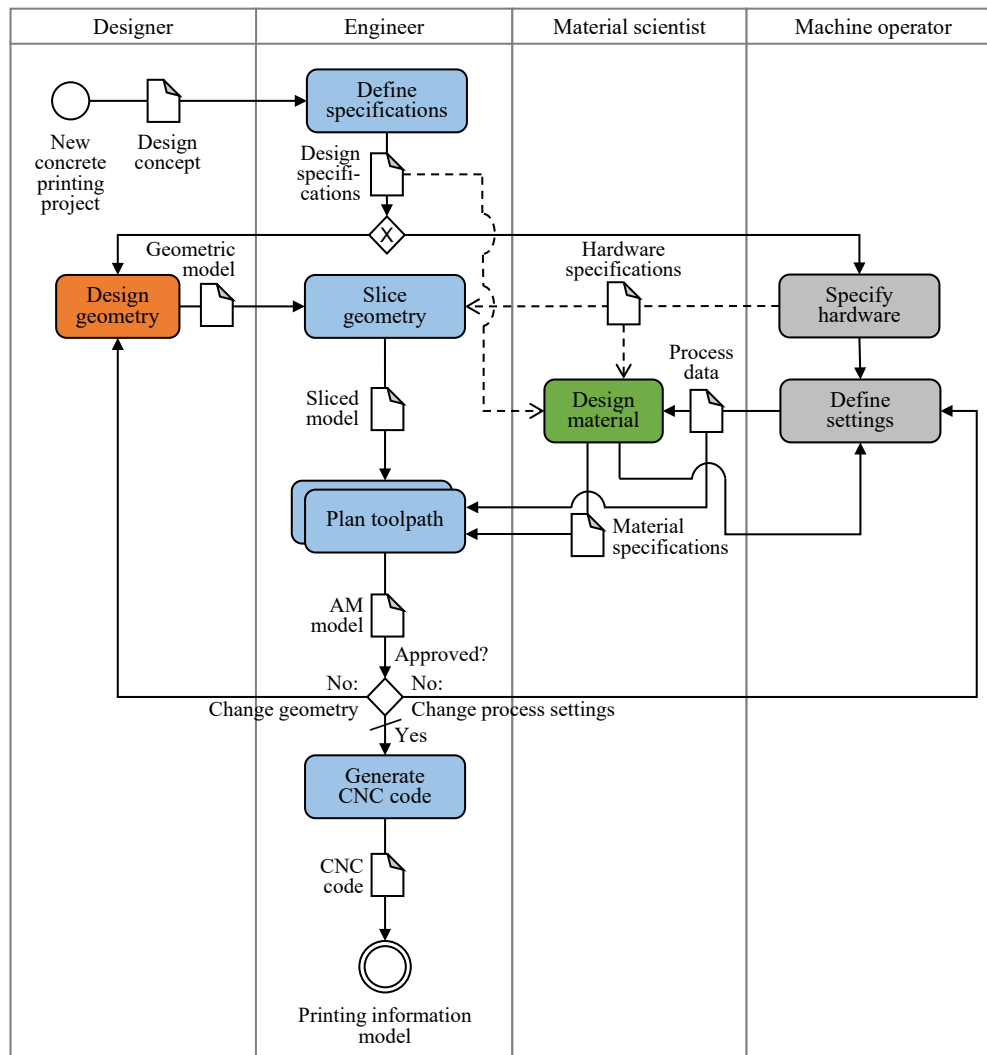


Figure 1. Extract of the process map describing the AM process for concrete printing

Existing *standards* for AM define terminology, data formats, and data models are used to exchange information for geometric representation and for hardware control. The standard terminology for AM technologies is defined in the ASTM F2792 standard [1], categorizing AM technologies. There are three main data format standards for geometry representation: standard tessellation language (STL), additive manufacturing format (AMF), and 3D manufacturing format (3MF). The STL format is the de-facto standard in AM. An STL file is an unordered collection of triangles, vertices, and unitary surface normal vectors in binary or ASCII format [7]. AMF is an ASTM/ISO standard (ISO/ASTM 52915), which extends STL to include dimensions, curved facets, recursive subdivision, color, material, constellation descriptions, and metadata. AMF is based on extensible markup language (XML), providing an XML-based schema definition (XSD) for AM technologies [8]. 3MF is an XML-based data format that provides broad model

information, such as mesh topology, slices, color, material, and texture, allowing multiple objects to be contained within a single archive [9].

For hardware control, there are two main standards: G-code and the Standard for the Exchange of Product model data compliant Numerical Control (STEP-NC). The ISO 6983 standard [10], also known as G-code, is widely used as a numerical control (NC) programming language. G-code supports hardware control in AM processes, defining motion and action commands in sequential lines [7]. The STEP-NC extends the ISO 10303 standard in ISO 14649 and defines a data model for numerical controllers. STEP-NC provides control structures for the sequence of working steps and associated machine hardware functions specified in the EXPRESS language [11].

In the following paragraphs, *current research* in AM and concrete printing, published between 2015 and 2021, is organized according to the research topic. Table 1

presents an overview of the research topics and the related references. Below, a brief overview per research topic is presented.

Relevant to *process and geometry parameters*, the implementation of AM in the AEC industry has been reviewed in [2], while in [3] technical issues in concrete printing have been described. A framework to classify process used in concrete printing has been defined in [12]. Process parameters and the impact on the manufacturing process has been studied in [13]. Strategies to improve the control of the manufacturing process have been proposed by using sensing technologies [14] and by simulating the manufacturing process to optimize process parameters [15]. The interactions between geometry parameters and process parameters, including the manufacturing process, have been studied for concrete printing [3] and for conventional AM methods [16].

Material parameters and the interactions with the manufacturing process have been reviewed in [17]. However, technical issues regarding material parameters are still open points, as discussed in [3]. Research regarding material parameters in concrete printing has been focused mainly on material development [18], on material testing for fresh concrete [19], and on the impact of the manufacturing process on material parameters [20].

Modeling and simulations in AM are used to simulate and optimized the manufacturing process [21], analyze the material [20], and to predict the structural stability of printed components during manufacturing [15]. Numerical modeling and simulations have been used to predict the concrete flow to determine optimal rheological requirements of the concrete [22]. Perrot et al. [23] have discussed the implementation of analytical and numerical tools to assess the concrete printing process as a function of the material properties, the geometry of the components, and the process parameters (e.g. manufacturing hardware settings).

AM-related ontologies are formal descriptions of the field of AM (or subfields), where concepts and relationships are defined. Ontologies contain the current knowledge in AM and can be extended to support future knowledge. The semantics of AM technology have been defined based on STEP-NC [7]. Ontologies have been developed to support manufacturability analysis [24], interoperability for data management [25], and lifecycle data management [26]. Similarly, ontologies have been developed specifically for the digital thread of metal-based AM [5] and for developing BIM-based concrete printing [27].

BIM-related research in AM has been developed to integrate AM into the AEC industry. BIM concepts, providing semantic and geometric information, have been used to digitalize life-cycle information of buildings and infrastructure. Paolini et al. [2] as well as Gradeci et al. [28] have discussed the benefits of coupling AM and

BIM for data modeling. Using an open BIM standard, such as the Industry Foundation Classes (IFC), data can be managed and exchanged between software applications used for AM and software applications used in the AEC industry, maintaining semantic and geometric information. Research has been conducted to couple concrete printing and BIM, showing the potential of BIM-based concrete printing, focusing on data retrieval from BIM models [29] and on IFC-based descriptions of process and material parameters [30].

Table 1. Overview of research topics related to AM and concrete printing

Research topic	Qty. of references	References
Process and geometry parameters	18	[2-3, 5, 7, 12-17, 20-21, 24-27, 29-30]
Material parameters	12	[3, 13, 15, 17-23, 27, 30]
Modeling and simulation	5	[19-23]
AM-related ontologies	7	[5, 7, 24-27, 30]
BIM-related research	6	[2, 16, 27-30]

Software applications used to develop concrete printing projects help determine the information generated in tasks along the digital thread (input and output parameters). Functionalities in the AM software applications support *AM design and optimization* as well as *AM process planning*. Software applications for AM design and optimization are usually based on computer-aided designs and enable geometry optimization. Software applications for AM process planning are commonly used for slicing and toolpath planning, creating manufacturing models. Complex AM software applications for AM process planning also support simulations of the manufacturing process where the effect of toolpaths and hardware settings can be evaluated to ensure buildability. The user manuals of common AM software applications, such as *Cura slicer* [31] and *Slic3r* [32], provide insight into the input and output parameters needed for AM design and optimization and AM process planning.

Due to the nature of concrete printing, vendors have developed proprietary solutions for the manufacturing hardware used for concrete printing (i) by modifying existing AM software or (ii) by developing software tools. In the first case, open-source and proprietary software applications for slicing and toolpath planning have been modified to fit specific manufacturing hardware used for concrete printing. In the latter case, computer-aided

design software applications have been coupled with programming environments to develop specialized software tools for concrete printing. Moreover, simulation software applications commonly used in the AEC industry have been used to conduct simulations of the manufacturing process and of the material behavior to predict the structural performance of printed components, as shown in [19].

To streamline the digital thread in concrete printing, attributes of the information exchange requirements are collected and organized into information units from standards, current research and software applications. In the following subsection, the requirement analysis relevant to AM design and optimization as well as AM process planning is presented.

3.2 Requirements analysis

From the previous review, the attributes of the information exchange requirements have been collected in the following tables and organized into information units. The attributes are analyzed according to *completeness* (i.e. if an attribute is required or optional) and *interoperability*. The attributes have been discussed with experienced users from the domains of robotics, material science, civil engineering, and mechanical engineering. The experienced users, hence, provide insight in the completeness and interoperability of the attributes via a short survey.

Completeness of the information exchange requirements avoids redundancies and ensures the

inclusion of all attributes necessary to manufacture a component. Interoperability is supported when the semantics of the information exchange requirements are preserved when transitioning between actors and tasks. Hence, satisfying completeness and interoperability of the information exchange requirements provides a strong basis for a common understanding between actors, enhancing the reliability of concrete printing. The information exchange requirements are discussed in the following paragraphs, focusing on AM design and optimization as well as on AM process planning for illustration purposes.

In *AM design and optimization*, geometric models are generated from conceptual designs. The main information exchange requirements are design concepts and design specifications (Table 2). The design concepts include the geometry pre-design and information regarding material, print location, structural characteristics, and structural boundary conditions. From the design concepts, design specifications are defined regarding AM process, material requirements, geometric tolerances and geometric requirements. As output, geometrical models (e.g. BIM models) are generated according to design and hardware specifications and can be further be optimized with respect to manufacturability, topology, structural performance, and geometric tolerances. In Table 2, prerequisites of the information exchange requirements for design specifications are highlighted in gray.

Table 2. Information exchange requirements for design specifications

Type of information	Information needed	Rqd.	Opt.
Design concept	The design concept will have been carried out prior to define the basic requirements		
AM process	Requirements for the AM process		
	• AM process type	X	
	• Machine type	X	
	• Reinforcement type		X
Material requirements	Requirements for the materials according to the AM process requirements.		
	• Main material selection (e.g. reinforced high-performance concrete)	X	
	• Support material selection (e.g. plaster)	X	
	• Reinforcement material selection (e.g. carbon fibers)		X
	• Main material minimum strength (e.g. 30 MPa)	X	
Geometric tolerances	Allowed tolerances in the geometric precision of the printed component/structure		
	• Deformation under self-weight	X	
	• Deflection under self-weight	X	
	• Allowed shrinkage	X	
Geometric requirements	Requirements for the geometric details according to the AM process (e.g. machine type) and materials requirements		
	• Maximum overhang angle	X	
	• Minimum feature size	X	

Table 3. Information exchange requirements for process data

Information unit	Attributes	Rqd.	Opt.
Basic requirements	The basic requirements will have been carried out prior to defining the process data		
Hardware specifications	The hardware specifications will have been carried out prior to defining the process data		
Feedback data	Feedback from previous manufacturing processes		
Strategy	Printing strategy for the AM process		
	• Layer-by-layer strategy	X	
	• Layer transition type	X	
	• Infill pattern	X	
	• Infill density	X	
	• Boundary thickness (i.e. number of adjacent filaments)	X	
	• Layer interval time	X	
	• Nozzle height above previous layer		X
Boundary conditions	Boundary conditions of the AM process		
	• Environment temperature	X	
	• Environment humidity	X	
	• Machine boundary conditions (e.g. printing area and flow rate)	X	
Machine parameters	Machine parameters for the AM process		
	• Printing speed	X	
	• Traveling speed (i.e. speed when not extruding)	X	
	• Acceleration	X	
	• Pump pressure	X	

Table 4. Information exchange requirements for material specifications

Information unit	Attributes	Rqd.	Opt.
Hardware specifications	The hardware specifications will have been carried out prior to defining the material specifications		
Process data	The process data will have been carried out prior to defining the material specifications		
Main material design specifications (e.g. Concrete)	Material specifications for the main material to be employed in the AM process. Parameters may vary depending on the material		
	• Concrete type (e.g. C35/45)	X	
	• Design strength	X	
	• Maximum aggregate size requirement	X	
	• Slump requirement	X	
	• Design open time	X	
	• Estimated volume	X	
	• Batching type	X	
	• Pre-process treatment (e.g. contact surface preparation)	X	
	• Post-process treatment (e.g. surface cover and surface uncovered)	X	
Support material design specifications	Material specifications for the support material to be employed in the AM process. Parameters may vary depending on the material	X	
Reinforcement material design specifications	Material specifications for the reinforcement material to be employed in the AM process. Parameters may vary depending on the material	X	

In *AM process planning*, the geometrical models are sliced and the toolpaths are planned. The geometrical models and design specifications, together with hardware specifications, process data and material specifications, provide the basic information necessary for process

planning. The effect of the manufacturing process on the material properties and on the behavior of concrete must be considered when defining the material specifications, hence the process data and the material design are adjusted in an iterative process to satisfy the design

specifications. As an output, AM models are created containing all the information necessary to generate machine-readable code (CNC code). Table 3 and Table 4 present the information exchange requirements for process data and material specifications, where prerequisites are highlighted in gray.

4 Summary and conclusions

In this study, efforts towards standardizing AM data modeling have been presented, and information exchange requirements in AM have been analyzed in the context of concrete printing. The AM process for concrete printing has been defined, and information exchange requirements have been identified. Attributes within information exchange requirements for concrete printing have been collected and analyzed through a systematic review and discussion with experienced users. In conclusion, the information exchange requirements for concrete printing show synergies with the information exchange requirements of conventional AM methods. In particular for concrete printing, the hardening process of concrete has a non-negligible effect on the process parameters (e.g. manufacturing hardware settings, print strategy) and on planning and control of the manufacturing process. Therefore, the interdependencies of the concrete and the manufacturing process have to be considered along the digital thread when advancing reliability and interoperability of the concrete printing process. For illustration purposes, AM design and optimization and AM process planning have been analyzed in detail, identifying the material-related information exchange requirements necessary for concrete printing.

With the information exchange requirements clearly defined, the digital thread can be described as a formal data model. With the data model, collaboration between actors will be enhanced resulting in a smooth workflow to improve the quality of the manufacturing process and the printed components. There is still a need to develop data models that support the digital thread in concrete printing in compliance with current standards used to digitalize the AEC industry, such as open BIM standards. Through standardization, concrete printing may become a more accepted construction method in the AEC industry. Future research may therefore be conducted to further advance standardization in AM data modeling for concrete printing.

5 Acknowledgments

The authors would like to acknowledge the financial support the German Research Foundation (DFG) through grant SM 281/7-1. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of

the authors and do not necessarily reflect the views of DFG.

References

- [1] ASTM International (2012). ASTM F2792-12 Standard terminology for additive manufacturing technologies. West Conshohocken, PA, United States: ASTM International.
- [2] Paolini, A., Kollmannsberger, S. & Rank, E. (2019). Additive manufacturing in construction: A review on processes, applications, and digital planning methods. *Additive Manufacturing*, 30(2019), 100894.
- [3] Buswell, R.A., Leal de Silva, W.R., Jones, S.Z. & Dirrenberger, J. (2018). 3D printing using concrete extrusion: A roadmap for research. *Cement and Concrete Research*, 112(2018), pp. 37-49.
- [4] buildingSmart (2010). Information Delivery Manual: Guide to components and development methods. Accessed on 08/09/2021 from: <https://technical.buildingsmart.org/standards/information-delivery-manual/>.
- [5] Kim, D.B., Witherell, P., Lu, Y. & Feng, S. (2017). Toward a digital thread and data package for metals-additive manufacturing. *Smart and Sustainable Manufacturing Systems*, 1(1), pp. 75-99.
- [6] DIN (2019). DIN SPEC 17071:2019-12 Additive manufacturing - Requirements for quality-assured processes at additive manufacturing centers. Berlin, Germany: Beuth Verlag GmbH.
- [7] Bonnard, R., Hascoët, J.-Y., Mognol, P., & Stroud, I. (2018). STEP-NC digital thread for additive manufacturing: data model, implementation and validation. *International Journal of Computer Integrated Manufacturing*, 31(11), pp. 1141-1160.
- [8] ASTM International (2013). ISO/ASTM 52915:2020 Standard specification for additive manufacturing file format (AMF) version 1.2. West Conshohocken, PA, United States: ASTM International.
- [9] 3MF Consortium (2018). 3D Manufacturing Format Specification. Accessed on 09/09/2021 from: <http://www.3mf.io/>.
- [10] ISO (2009). ISO 6983-1:2009 Industrial automation systems and integration – Physical device control – Data model for computerized numerical controllers – Part 10: General process data. Geneva, Switzerland: ISO.
- [11] ISO (2004). ISO 14649-10:2004 Automation systems and integration – Numerical control of machines - Program format and definitions of address words – Part 1: Data format for positioning, line motion and contouring control systems. Geneva, Switzerland: ISO.
- [12] Buswell, R.A., Leal de Silva, W.R., Bos, F.P., Schipper, H.R., Lowke, D., Hack, N., Kloft, H.,

- Mechtcherine, V., Wangler, T. & Roussel, N. (2020). A process classification framework for defining and describing digital fabrication with concrete. *Cement and Concrete Research*, 134(2020), 106068.
- [13] Bos, F., Wolfs, R., Ahmed, Z. & Salet, T. (2016). Additive manufacturing of concrete in construction: Potentials and challenges of 3D concrete printing. *Virtual and Physical Prototyping*, 11(3), pp. 209-225.
- [14] Wolfs, R.J.M., Bos, F.P., van Strien, E.C.F. & Salet, T.A.M. (2017). A real-time height measurement and feedback system for 3D concrete printing. In: *Proceedings of the 2017 fib Symposium — High Tech Concrete: Where Technology and Engineering Meet*, Maastricht, Netherlands, 06/12/2017.
- [15] Kruger, J., Cho, S., Zeranka, S., Viljoen, C. & van Zijl, G. (2020). 3D concrete printer parameter optimization for high rate digital construction avoiding plastic collapse. *Composites Part B: Engineering*, 183(2020), 107660.
- [16] Correa, F.R. (2016). Robot-oriented design for production in the context of building information modeling. In: *Proceedings of the 33rd International Symposium on Automation and Robotics in Construction*, Auburn, AL, USA, 07/18/2016.
- [17] Mechtcherine, V., Bos, F.P., Perrot, A., Leal da Silva, W.R., Nerella, V.N., Fataei, S., Wolfs, R.J.M., Sonebi, M. & Roussel, N. (2020). Extrusion-based additive manufacturing with cement-based materials – Production steps, processes, and their underlying physics: A review. *Cement and Concrete Research*, 132(2020), 106037.
- [18] Markin, V., Krause, M., Otto, J., Schröfl, C., & Mechtcherine, V. (2021). 3D-printing with foam concrete: From material design and testing to application and sustainability. *Journal of Building Engineering*, 43(2021), 102870.
- [19] Wolfs, R.J.M., Bos, F.P. & Salet, T.A.M. (2018). Early age mechanical behavior of 3D printed concrete: Numerical modelling and experimental testing. *Cement and Concrete Research*, 106(2018), pp. 103-116.
- [20] Wolfs, R.J.M., Bos, F., & Salet, T. (2019). Hardened properties of 3D printed concrete: The influence of process parameters on interlayer adhesion. *Cement and Concrete Research*, 119(2019), pp. 132-140.
- [21] Vaitová, M., Jendele, L., & Červenka, J. (2020). 3D printing of concrete structures modeled by FEM. *Solid State Phenomena*, 309(2020), pp. 261-266.
- [22] Roussel, N., Spangenberg, J., Wallevik, J. & Wolfs, R. (2020). Numerical simulations of concrete processing: From standard formative casting to additive manufacturing. *Cement and Concrete Research*, 135(2020), 106075.
- [23] Perrot, A., Pierre, A., Nerella, V.N., Wolfs, R.J.M., Keita, E., Nair, S.A.O., Neithalath, N., Roussel, N. & Mechtcherine, V. (2021). From analytical methods to numerical simulations: A process engineering toolbox for 3D concrete printing. *Cement and Concrete Composites*, 122(2021), 104164.
- [24] Kim, S., Rosen, D.W., Witherell, P. & Ko, H. (2018). A design for additive manufacturing ontology to support manufacturability analysis. In: *Proceedings of the ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Quebec, Canada, 08/26/2018.
- [25] Sanfilippo, E. M., Belkadi, F. & Bernard, A. (2019). Ontology-based knowledge representation for additive manufacturing. *Computers in Industry*, 109(2019), pp. 182-194.
- [26] Lu, Y., Choi, S. & Witherell, P. (2015). Towards an integrated data schema design for additive manufacturing: Conceptual modeling. In: *Proceedings of the ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Boston, MA, USA, 08/02/2015.
- [27] Smarsly, K., Peralta, P., Luckey, D., Heine, S., & Ludwig, H.-M. (2020). BIM-based concrete printing. In: *Proceedings of the International ICCCBE and CIB W78 Joint Conference on Computing in Civil and Building Engineering 2020*. Sao Paulo, Brazil, 06/02/2020.
- [28] Gradeci, K. & Labonnote, N. (2020). On the potential of integrating building information modelling (BIM) for the additive manufacturing (AM) of concrete structures. *Construction Innovation*, 20(3), pp. 321-343.
- [29] Davtalab, O., Kazemian, A. & Khoshnevis, B. (2018). Perspectives on a BIM-integrated software platform for robotic construction through Contour Crafting. *Automation in Construction*, 89(2018), pp. 13-23.
- [30] Peralta-Abadia, P., Heine, S., Ludwig, H.-M. & Smarsly, K. (2020). A BIM-based approach towards additive manufacturing of concrete structure. In: *Proceedings of the 27th International Workshop on Intelligent Computing in Engineering (EG-ICE)*. Berlin, Germany, 07/01/2020.
- [31] Chakravorty, D. (2020). Cura tutorial: Master Cura slicer settings. All3DP. Accessed on 09/09/2021 from: <https://all3dp.com/1/cura-tutorial-software-slicer-cura-3d/>.
- [32] Hodgson, G., Ranellucci, A., & Moe, J. (2014). Slic3r manual. Slic3r. Accessed on 09/09/2021 from: <https://manual.slic3r.org/>.

Safety Risk Assessment of Drones on Construction Sites using 4D Simulation

Z. Zhu^a, I. Jeelani^a, and M. Gheisari^a

^aM.E. Rinker, Sr, School of Construction Management, University of Florida, U.S.

E-mail: zhuzixian@ufl.edu, idris.jeelani@ufl.edu, masoud@ufl.edu

Abstract –

The use of drones in the construction industry has been dramatically growing in different areas such as building inspection, site mapping, and safety monitoring. The increasing deployment of drones in construction leads to more collaboration and interaction between human workers and drones. This raises novel occupational safety issues, especially for those workers who already work in a hazardous environment. While there is significant research about the benefits of drones for specific construction applications, there is a knowledge gap about the safety risks of integrating such technology into construction sites. This study uses 4D simulation to mimic and visualize virtual construction sites populated with drones to detect safety risks of their presence under different working conditions. The validated 4D simulation can provide a valuable source for safety risk identification and assessment of drone integration on construction sites.

Keywords –

Drones; 4D Simulation; Safety; Risk Assessment; Construction Sites

1 Introduction

The application of drones or Unmanned Aerial Vehicles (UAVs) in the construction industry is continuously increasing in recent years. In 2018, the use of drones in construction rapidly increased by more than 200% compared to the previous year [1]. In 2021, despite the pandemic influence on the global economy, 88% of the present drone adopters in the construction industry were willing to increase or maintain their investment in drone technology [2]. Drones are popular in construction because they can perform tasks more efficiently with less cost, especially in dangerous or inaccessible spaces for human workers. Additionally, drones can carry different sensors, conveniently collecting data and providing comprehensive documentation for site records. Drones in construction have a wide range of application areas, including building inspection, damage assessment, site surveying and mapping, progress monitoring, and safety

inspection [3]. As the application of drones expands in construction, a significant increase in interactions between drones and human workers is expected. Drones are flying robots that share a workspace with humans, equipment, structures, and other objects on construction sites [4]. Therefore, there is always a potential for collision incidents that pose serious safety risks to human workers collaborating with or working around drones. While there are substantial research studies about the application and benefits of drones in construction, limited research has been conducted to analyze the safety challenges of drones on construction sites. For example, Xu et al. [5], Jeelani & Gheisari [4], and Khalid et al. [6] conducted preliminary studies and categorized the safety concerns related to drone applications in construction. Despite these exploratory efforts, there is a dearth of research examining the specific safety risks resulting from varied working conditions based on different drone applications in construction.

2 Background & Motivation

2.1 Construction Safety

Construction is a massive, dynamic, and complicated industry that provides millions of job opportunities worldwide. At the same time, construction work comes with disproportionately higher safety risks and causes more fatal accidents than other sectors [7]. In 2019, there were 1,038 fatal occupational injuries in the construction industry in the United States, accounting for almost 20% of total incidents in all the sectors [8]. According to statistics data from CPWR (The Center for Construction Research and Training), 34.7% of fatal injuries were caused by falls and slips. 22.6% of fatal injuries were caused by contact with objects and equipment, 17.1% were transportation incidents, and 13.4% were caused by exposure to harmful substances or environments [9]. The statistics indicate that workers who work in dangerous locations (such as on heights) and are exposed to automation hazards, including equipment and transportation, are more likely to be exposed to safety

risks. Besides fatal injuries, the rate of non-fatal injuries in the construction industry also remained consistently high. There are over 200,000 injuries reported from construction [4]. Non-fatal injuries can result in severe disabilities, income loss, chronic pain, and ongoing medical expenses, resulting in lower quality of life for the workers. Even less-serious injuries can lead to work time lost, productivity reduction, and increased medical costs [10].

2.2 Drone Application and Safety Challenges in Construction

Drones are unmanned aerial vehicles operated under remote control without a pilot [11]. The increasing uses of drones in the construction industry include aiding with construction structure inspection, mapping and surveying, 3D modeling, progress monitoring, material delivery, and safety inspection [12]. Undoubtedly, drones can provide a more efficient way to perform construction tasks at a lower cost [13]. They can access high-altitude and dangerous working zones, which are difficult to reach by human workers, and provide comprehensive data about construction sites through delicate sensors and processors. However, with more such aerial robots flying on construction sites, the interactions between drones and human workers or other objects (e.g., structures, equipment, materials, and vehicles) will dramatically increase. Furthermore, with the integration of drones in existing construction workplaces, more safety risks are expected for those who already work in high-risk environments. According to fatal injury reports, contact with objects and equipment is one of the top reasons that cause occupational injuries in construction [9]. Integrating drones in construction sites will increase the possibilities of such contact risks. Direct contact with drones includes being struck by flying drones, hit by falling drones, and caught in by drones' moving parts. Indirect accidents include continuous collisions caused by drones contacting other objects and dust and particulate emissions brought by the drones [4].

2.3 Simulation Approaches and Techniques

Simulation can be defined as the art and science of creating a representation of a process or system for experimentation and evaluation [14]. A simulation model is a set of variables and a mechanism for changing those variables dynamically over time [15]. At a systems level, this helps in stimulating the interactions between different modules or objects that constitute a system. As construction activities are dynamic and involve complicated behavior, uncertainties, and dependencies, simulation approaches are beneficial to replicate reality and process information iteratively on construction activities [16], and for quantitative analysis of operations

and processes [17]. In construction, a 4D simulation can link a three-dimensional (3D) model of the building or facility to the dynamic construction activities, allowing the construction process to be visualized over time [18]. One of the goals of simulation approaches is the observation of processes, interactions, and outcomes of those interactions in varying conditions, which help in gaining a better understanding of the situation studied [19]. Drone simulation systems have been used in different fields. For example, an interactive drone flight control system for agriculture sowing is composed of virtual drone models and virtual scenes, and the motor speed was used to change drone altitude and position during simulation [20]. A VR training system for bridge inspectors with an assistant drone used parameters including mass and load, speed, battery capacity, and movement types [21]. Al-Mousa et al. brought up a framework for the drone traffic integration simulation, including aircraft type, dimensions, weight, speed, location, battery charge, and sensing range [22]. In construction, Gilles et al. used a VR-based flight training simulator for drone-mediated building inspections [23].

In this study, a game engine (Unity3D[®]) is employed to help this task by creating a replica of a scenario that mimics construction sites populated with drones while preserving the physical, dynamic, and organizational aspects. Unity3D[®] is a professional game engine with strong rendering capabilities and convenient interactivity, an attractive platform for dynamic visualization and simulations processing [24]. In Unity3D[®], a virtual environment can be developed to simulate a construction site with virtual construction workers, structures, equipment, and other construction entities with their actions and interactions [25].

3 Research Objective

This study aims to develop a 4D simulation of a construction site populated with workers and drones performing different construction-related tasks. This immersive virtual environment will mimic and visualize interactions between drones, workers, and other construction entities and ultimately identify safety risks associated with drone integration in construction under different working conditions. Studying these interactions under varying conditions in the real world is not only dangerous but also impossible on a large scale. Using 4D simulations allowed us to vary multiple conditions and investigate the outcomes of critical situations without any risk.

4 Research Methodology

As illustrated in Figure 1, this study was completed in

two phases: (1) **Scenario Development**: (1-a) identification of simulation scenario characteristics and (1-b) simulation conceptualization. (2) **Simulation Development**: (2-a) simulation parameter determination and (2-b) simulation demonstration and validation.

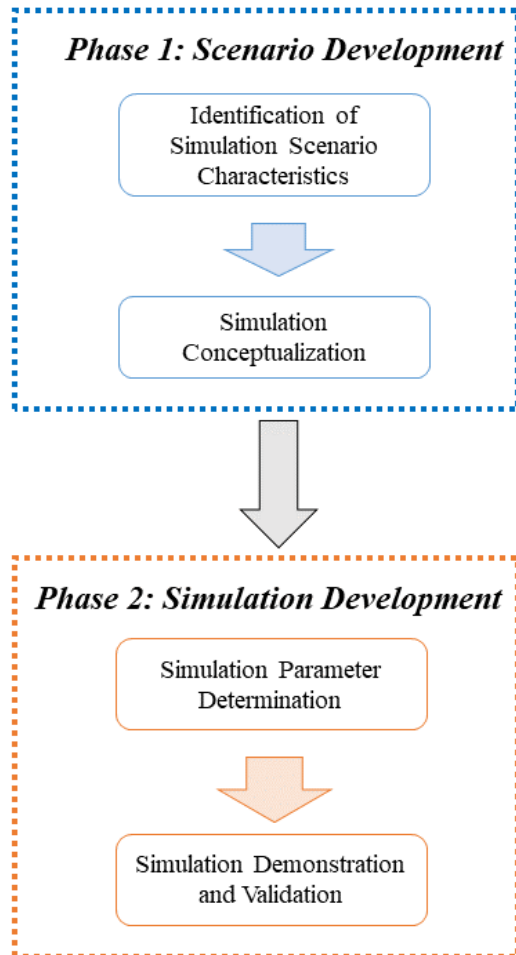


Figure 1. Research Methodology

5 Scenario Development

5.1 Identification of Simulation Scenario Characteristics

The objective of this step was to identify different physical and dynamic characteristics of high-risk scenarios that needed to be mimicked in the simulations. Working with drones is likely to introduce new risks for construction workers, especially those who work on heights. These workers are already at the highest risk of fatalities and are most likely to be affected by drones flying at heights. Therefore, past injury reports were thoroughly analyzed to find common characteristics that provided a base for defining the simulation scenario content.

The CPWR's fatality maps [9] and OSHA's Integrated Management Information System (IMIS) database [26] were explored to analyze the height-related accidents in the last 5 years to identify the frequent construction tasks that resulted in fatalities. The details provided in the investigation reports were used to identify the key characteristics associated with each accident. The analysis indicated that "roof" is one of the primary factors leading to fatal falls in construction. Further investigation of filtered reports of "falls from roof" accidents provided information about workers' tasks, working locations, and material or equipment. After filtering the data, 337 incidents involving falling or physical contact related to the keyword "roof" were analyzed. The most frequent words related to construction accidents and the most frequent falling height were identified by analyzing narrative descriptions in OSHA reports (Table 1.). In the filtered 337 "roof" incidents caused by falling or physical contact, the most frequent tasks for the workers who fall from the roof are installing panels or trusses, and the most frequent falling height is 20 feet.

Table 1. Injury Reports (2015-2018) Analysis Results [9]

No.	Word related to "roof" accidents from injury reports	Count
1	roof	433
2	fell	414
3	concrete	66
4	floor	54
5	metal	44
6	Scaffold	36
7	installing	36
8	residential	33
9	ladder	29
10	skylight	25
11	struck	24
12	lift	22
13	platform	21
14	panels	21
15	trusses	20
The most frequent fall height		
20 ft		

Hence, the common characteristics obtained from this step were

- (1) The working location should be on the roof.
- (2) The building height needs to be 20 feet.
- (3) The workers' task should be installing panels.

5.2 Simulation Conceptualization

The objective of this step was to design dynamic

simulation scenarios. This included (1) the design of static 3D virtual scenarios as the simulation environment and (2) the design of different virtual workers and drones performing their designated tasks within this environment. The common scenario characteristics identified in the previous step formed the basis for designing the 4D simulations in this step. Finally, the drone tasks and flight paths were also incorporated into the identified high-risk scenarios (Figure 2). The area of this virtual construction site is approximately 3,000 square meters (32,000 square feet).

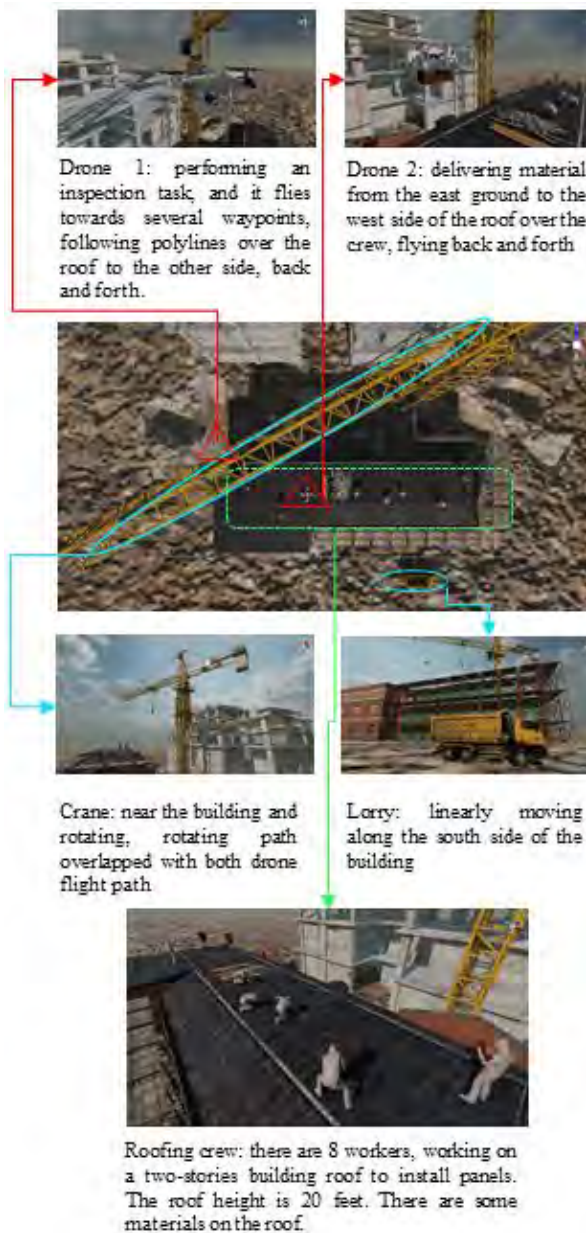


Figure 2. "Roof" Scenario for 4D simulation

6 Simulation Development

6.1 Simulation Parameter Determination

The objective of this step was to identify different simulation parameters necessary to vary the simulation conditions and determine their values to evaluate the physical risks of drones under multiple conditions. Based on previous studies (see section 2.3), speed, altitude, and failure rate are selected as three parameters to mimic different flight conditions for each simulation (Figure 3.). Varying speed helped study the impact of the drone's dynamic movement through the virtual environment on its likelihood of colliding with other entities. Varying altitude can evaluate the impact of relative position between the drone and other virtual objects. Finally, the failure rate was a collective parameter for several characteristics which influence the drone's flight stability but are difficult to quantify. These include operator error, program error, hardware error, and weather conditions.



Figure 3. The Flight Parameters of Virtual drones

The flight parameters of virtual drones in Unity3D scenarios need to be designed based on actual drone flight characteristics in the real construction sites while performing specific tasks to mimic realistic construction sites and drone integration scenarios. According to the regulation for small unmanned systems issued by the Federal Aviation Administration [27] and the memorandum for the use of Unmanned Aircraft Systems in Inspections by the Occupational Safety and Health Administration [28], the drone must not be operated higher than 400 feet above the ground, except when within 400 feet of a structure; the flight speed of drone must not exceed 100 mph. More literature was reviewed to determine the range of flight parameters of virtual drones used in physical contact risks simulation scenarios.

Speed

Usually, the flight speed is fixed for visual data capture when drones are used for monitoring and inspection-related tasks. For example, Ibrahim and Golparvar-Fard [29] set the drone speed to 11 mph (5m/s) to optimize 3D flight templates and generate an algorithm that maximized visual quality and minimized flight execution duration. Similarly, a drone flight speed of

3m/s was used to accomplish surveying applications to a medium-sized building [30]. In a safety monitoring and inspection application case study for a four-story building, the drone performed a task at 3m/s speed outside the boundary of the construction site [31]. In a study analyzing drone photogrammetry's potential application, the drone flight speed was fixed at 2.5m/s to capture visual information in three-dimension and over different epochs for a building zone [32]. In a mapping application experiment study, conducted on a university campus and a residential construction site, the drone speed was suggested to be 4m/s based on the software developer's recommendation [33]. DroneDeploy, a commercial drone flight planning platform, usually sets the flight speed to 13mph (5.8m/s) for construction inspection [34]. Considering the previous literature, a flight speed range of 2-6m/s was selected for this simulation.

Altitude

The Drone flight altitude varies with different construction types and applications. The flight altitude in the context of this study is defined based on the distance from the drone to the surface or boundary of the building. When a drone is performing inspection or monitoring tasks, the distance needs to be such that it ensures that the drone can capture visual information with good quality while avoiding collision with the structure. A study of quality assessment for the drone applications on visual inspection for structure damages indicated that usually under suitable environmental conditions, the detailed image could be captured from 5m to 25m to provide detailed information on structures [35]. In a study for building 3D model reconstruction, the drone was set to keep a distance of 5-meters from the building and 25-meters from the objects to capture clear and sharp images with the required quality [36]. In another study, a distance of 12m away from the building surface was applied to capture images [37]. In a field test for an automatic point cloud registration method for the drone, a two-story building was selected for method validation, and a commercial drone path planner designed the flight height of the drone as 20m (around 10m-14m distance from the roof) [38]. According to the literature review, the distance of the drone from the building surface can range between 5m to 25m. Therefore, the altitude of virtual drones in simulation scenarios should be the height of the building surface, adding 5m to 25m. According to the scenario content developed in the previous step, the virtual building roof's height is 20 feet (6m). Therefore, the altitude range of the virtual drone, which is performing inspection tasks in simulation, should be 11m to 31m.

Failure Rate

Various factors decide the failure rate of the drone. Intrinsic reliability, which stands for system vulnerability is one of the essential factors that cause drone failures. The hierarchy of the intrinsic reliability assessment shows that critical factors are located in different subsystems: ground control system, mainframe, power system, navigation system, electric system, and payload [39]. Human error is another critical factor that causes drone accidents. Human factor issues are related to flight system operation, incorporated automation, and user interfaces [40]. Analysis based on military UAV mishap statistics suggests that most mishaps are caused by unsafe actors of the operators [41]. The top errors related to the operator's behavior are skill-based, decision-making, and cognitive factors. Research also suggests that higher mental demands cause more errors during drone flight operations performed by simulators [42]. The external environment can also lead to drone failures. For example, a bird strike is a potential factor for drone accidents, especially during takeoff and landing [43]. Adverse weather conditions will also impact the success of drone flight and task performance [44].

The failure rate applied in simulation needs to be a combined parameter from several characteristics, including intrinsic reliability, human errors, and external environments, to mimic realistically done failures comprehensively. This study assumes the failure rate of virtual drones ranges from 10% to 50% due to potential internal or external influential factors. Based on the failure rate value, the drones were programmed to suddenly fall while performing tasks in their flight path. The probability of this happening was the same as the failure rate value set for that simulation run.

6.2 Simulation Demonstration and Validation

The objective of this step was to run the simulation with varying simulation parameters and observe drone interaction in these varying conditions. Three levels were set for each parameter in the 4D simulation, representing low, medium, and high levels of value, to mimic different flight conditions in the realistic construction site (Table 2.). Two virtual drones were used in the simulation with drone 1 representing a drone that has close interactions with workers (such as material delivery) and drone 2 simulating an inspection type of drone that does not come close to workers. For each simulation run, the parameter value was randomly made to fall in one of the selected levels. Having 3 parameters with each of the three levels, there were 27 different combinations of parameter levels. For each type of combination, 10 simulations were run resulting in a total of 270 rounds of the simulation conducted in Unity3D.

Table 2. Different levels of parameters

Parameter Type	Level of Parameter		
	Low	Medium	High
Speed (m/s)	1.5-2.5	3.5-4.5	5.5-6.5
Altitude (meter)	15-17	20-22	25-27
Failure Rate	0-0.1	0.2-0.3	0.4-0.5

In Unity3D, the Rigidbody Component is used to provide the object's gravity and could mimic the falling incidents of the drones. The Collider Component is used to detect the collisions between virtual drones and other virtual objects, including workers, equipment, building, and ground, and would collide with other objects. The real drone's rotors may continue working even if it fell but would stop after hitting something. This was mimicked in simulations by adding a delay of 2 seconds after the virtual drone collides with other objects before it ultimately stops working. The number of collisions between each drone with the worker, building, equipment, ground, and another drone was counted.

Table 3 shows the results of 270 simulation runs of this demonstration. The total collisions of drones 1 and 2 with other objects are used to detect drone-related potential safety risks for the system in the virtual scenario.

Table 3. Collisions Counted in Simulation Validation

Collision Type	Count
Total Collisions	773
Collisions Between drones and Building	45
Collisions Between drones and Workers	58
Collisions Between drones and Equipment	17
Collisions Between drones and Ground	653
Collisions Between drones	0

7 Conclusion and Future Work

This study provides a 4D simulation development framework that provides a methodology to mimic and simulate realistic drone-integration construction sites. Such simulations are used to study the safety challenges of workers working with or around drones on

construction sites. First, simulation scenario content was developed based on the analysis of past construction injury reports, which identified the most frequent factors that caused fatal injuries. Furthermore, drone tasks and flight paths are designed to incorporate virtual workers as supplementary scenario content. Second, speed, altitude, and failure rate were selected to mimic different flight conditions of drones in the 4D simulation. These three parameters, which designate how the integrated drones perform their tasks, were defined based on realistic parameter ranges. Finally, this simulation was run multiple times with varying parameters to observe and evaluate the likelihood of drone contact risks under these varying conditions. The 4D simulation detected potential collisions between drones and human workers or other objects. The validated 4D simulation can provide a valuable source for future comprehensive safety risk identification and assessment of drone integration in construction.

Since the interactions with drones and the dynamic construction activities depend highly on the scenario content, the quantitative result from this study can only provide insights for a specific type of drone integrated construction scenario. However, the identified risks and the relationship between different drone parameters and the number of incidents can provide valuable information that applies to other scenarios. Future research should include more scenarios within the 4D simulation to cover more realistic construction site conditions and include more complicated object movements and interactions. For example, the workers might sometimes randomly move in the working area, communicating with each other and exchanging their positions. These need to be captured in future simulations. Besides, other high-risk scenarios such as working on a ladder, or scaffolding, which are also prone to risks posed by drones, should be included in the simulations. Finally, there can be more types of drone flight paths according to their performing tasks and flight scenarios. Comprehensive safety risk assessment of drone integration in construction sites can be conducted through more complicated simulation development based on the preliminary work produced by the current study.

Acknowledgments

This material was produced under the National Science Foundation under Grant No. 2024656.

References

- [1] The Rise of Drones in Construction | DroneDeploy, (n.d.). <https://www.droneDeploy.com/blog/rise-drones-construction/> (accessed December 8, 2021).
- [2] State of the Drone Industry Report 2021 |

- DroneDeploy, (n.d.). <https://www.dronedeploy.com/resources/ebooks/state-of-the-drone-industry-report-2021/> (accessed December 8, 2021).
- [3] S. Zhou, M. Gheisari, Unmanned aerial system applications in construction: a systematic review, *Constr. Innov.* 18 (2018) 453–468. <https://doi.org/10.1108/CI-02-2018-0010>.
- [4] I. Jeelani, M. Gheisari, Safety challenges of UAV integration in construction: Conceptual analysis and future research roadmap, *Saf. Sci.* 144 (2021) 105473. <https://doi.org/10.1016/j.ssci.2021.105473>.
- [5] Y. Xu, Y. Turkan, A.A. Karakhan, D. Liu, Exploratory Study of Potential Negative Safety Outcomes Associated with UAV-Assisted Construction Management, (2020) 1223–1232. <https://doi.org/10.1061/9780784482865.129>.
- [6] M. Khalid, M. Namian, C. Massarra, The Dark Side of the Drones: A Review of Emerging Safety Implications in Construction, in: *EPiC Ser. Built Environ., EasyChair*, 2021: pp. 18–27. <https://doi.org/10.29007/x3vt>.
- [7] X. Li, W. Yi, H.-L. Chi, X. Wang, A.P.C. Chan, A critical review of virtual and augmented reality (VR/AR) applications in construction safety, *Autom. Constr.* 86 (2018) 150–162. <https://doi.org/10.1016/j.autcon.2017.11.003>.
- [8] U.S. Bureau of Labor Statistics, (n.d.). <https://www.bls.gov/home.htm> (accessed January 9, 2022).
- [9] CPWR | Construction Fatality Map Dashboard, CPWR |. (n.d.). <https://www.cpwr.com/research/data-center/data-dashboards/construction-fatality-map-dashboard/> (accessed January 9, 2022).
- [10] D.L. Lucas, J.R. Lee, K.M. Moller, M.B. O'Connor, L.N. Syron, J.R. Watson, Using Workers' Compensation Claims Data to Describe Nonfatal Injuries among Workers in Alaska, *Saf. Health Work.* 11 (2020) 165–172. <https://doi.org/10.1016/j.shaw.2020.01.004>.
- [11] M. Gheisari, J. Irizarry, B.N. Walker, UAS4SAFETY: The Potential of Unmanned Aerial Systems for Construction Safety Applications, (2014) 1801–1810. <https://doi.org/10.1061/9780784413517.184>.
- [12] G. Albeaino, M. Gheisari, B. Franz, A Systematic Review of Unmanned Aerial Vehicle Application Areas and Technologies in the AEC Domain, *Electron. J. Inf. Technol. Constr.* 24 (2019) 381–405.
- [13] M. Gheisari, J. Irizarry, A User-centered Approach to Investigate Unmanned Aerial System (UAS) Requirements for a Department of Transportation Applications, in: 2015.
- [14] P. Klingstam, P. Gullander, Overview of simulation tools for computer-aided production engineering, *Comput. Ind.* 38 (1999) 173–186. [https://doi.org/10.1016/S0166-3615\(98\)00117-1](https://doi.org/10.1016/S0166-3615(98)00117-1).
- [15] C.D. Pegden, Advanced tutorial: Overview of simulation world views, in: *Proc. 2010 Winter Simul. Conf.*, 2010: pp. 210–215. <https://doi.org/10.1109/WSC.2010.5679161>.
- [16] O. Bokor, L. Florez, A. Osborne, B.J. Gledson, Overview of construction simulation approaches to model construction processes, *Organ. Technol. Manag. Constr. Int. J.* 11 (2019) 1853–1861.
- [17] J.C. Martinez, Methodology for Conducting Discrete-Event Simulation Studies in Construction Engineering and Management, *J. Constr. Eng. Manag.* 136 (2010) 3–16. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000087](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000087).
- [18] C. Botton, Supporting constructability analysis meetings with Immersive Virtual Reality-based collaborative BIM 4D simulation, *Autom. Constr.* 96 (2018) 1–15. <https://doi.org/10.1016/j.autcon.2018.08.020>.
- [19] E. Paravizo, D. Braatz, Using a game engine for simulation in ergonomics analysis, design and education: An exploratory study, *Appl. Ergon.* 77 (2019) 22–28. <https://doi.org/10.1016/j.apergo.2019.01.001>.
- [20] Y. Wang, W. Zhang, Four-rotor UAV Virtual Simulation System for Agricultural Sowing, in: *2018 IEEE 4th Inf. Technol. Mechatron. Eng. Conf. ITOEC, IEEE, Chongqing, China*, 2018: pp. 1097–1101. <https://doi.org/10.1109/ITOEC.2018.8740368>.
- [21] Y. Li, M.M. Karim, R. Qin, A Virtual Reality-based Training and Assessment System for Bridge Inspectors with an Assistant Drone, *ArXiv210902705 Cs.* (2021). <http://arxiv.org/abs/2109.02705> (accessed September 20, 2021).
- [22] A. Al-Mousa, B.H. Sababha, N. Al-Madi, A. Barghouthi, R. Younis, UTSim: A framework and simulator for UAV air traffic integration, control, and communication, *Int. J. Adv. Robot. Syst.* 16 (2019) 172988141987093. <https://doi.org/10.1177/1729881419870937>.
- [23] G. Albeaino, R. Eiris, M. Gheisari, R.R. Issa, DroneSim: a VR-based flight training simulator for drone-mediated building inspections, *Constr. Innov. ahead-of-print* (2021). <https://doi.org/10.1108/CI-03-2021-0049>.
- [24] R.-X. Wang, R. Wang, P. Fu, J.-M. Zhang, Portable interactive visualization of large-scale simulations in geotechnical engineering using Unity3D, *Adv. Eng. Softw.* 148 (2020) 102838.

- <https://doi.org/10.1016/j.advengsoft.2020.102838>.
- [25] M. Kurien, M.-K. Kim, M. Kopsida, I. Brilakis, Real-time simulation of construction workers using combined human body and hand tracking for robotic construction worker system, *Autom. Constr.* 86 (2018) 125–137. <https://doi.org/10.1016/j.autcon.2017.11.005>.
- [26] Establishment Search Page | Occupational Safety and Health Administration, (n.d.). <https://www.osha.gov/pls/imis/establishment.html> (accessed January 30, 2022).
- [27] 14 CFR Part 107 -- Small Unmanned Aircraft Systems, (n.d.). <https://www.ecfr.gov/current/title-14/chapter-I/subchapter-F/part-107> (accessed November 16, 2021).
- [28] OSHA's use of Unmanned Aircraft Systems in Inspections | Occupational Safety and Health Administration, (n.d.). <https://www.osha.gov/memos/2018-05-18/oshas-use-unmanned-aircraft-systems-inspections> (accessed November 16, 2021).
- [29] A. Ibrahim, M. Golparvar-Fard, 4D BIM Based Optimal Flight Planning for Construction Monitoring Applications Using Camera-Equipped UAVs, (2019) 217–224. <https://doi.org/10.1061/9780784482438.028>.
- [30] J.G. Martinez, G. Albeaino, M. Gheisari, W. Volkmann, L.F. Alarcón, UAS Point Cloud Accuracy Assessment Using Structure from Motion-Based Photogrammetry and PPK Georeferencing Technique for Building Surveying Applications, *J. Comput. Civ. Eng.* 35 (2021) 05020004. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000936](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000936).
- [31] J.G. Martinez, G. Albeaino, M. Gheisari, R.R.A. Issa, L.F. Alarcón, iSafeUAS: An unmanned aerial system for construction safety inspection, *Autom. Constr.* 125 (2021) 103595. <https://doi.org/10.1016/j.autcon.2021.103595>.
- [32] J. Unger, M. Reich, C. Heipke, UAV-based photogrammetry: Monitoring of a building zone, *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XL-5 (2014) 601–606. <https://doi.org/10.5194/isprsarchives-XL-5-601-2014>.
- [33] J. Irizarry, D. Costa, A. Mendos, Lessons learned from unmanned aerial system-based 3D mapping experiments, in: 2016. chrome-extension://efaidnbmnnpbpcjpcglclefindmkaj/viewer.html?pdfurl=http%3A%2F%2Fascpro0.ascweb.org%2Farchives%2Fcd%2F2016%2Fpaper%2FCPRT157002016.pdf&cLen=503377&chunk=true.
- [34] DroneDeploy, (n.d.). <https://help.dronedeploy.com/hc/en-us> (accessed November 25, 2021).
- [35] G. Morgenthal, N. Hallermann, Quality Assessment of Unmanned Aerial Vehicle (UAV) Based Visual Inspection of Structures, *Adv. Struct. Eng.* 17 (2014) 289–302. <https://doi.org/10.1260/1369-4332.17.3.289>.
- [36] M.N. Zulgaflī, K.N. Tahar, Three dimensional curve hall reconstruction using semi-automatic UAV, *ARN J. Eng. Appl. Sci.* 12 (2017) 3228–3232.
- [37] T. Rakha, A. Gorodetsky, Review of Unmanned Aerial System (UAS) applications in the built environment: Towards automated building inspection procedures using drones, *Autom. Constr.* 93 (2018) 252–264. <https://doi.org/10.1016/j.autcon.2018.05.002>.
- [38] J. Park, P. Kim, Y. Cho, Y. Fang, AUTOMATED COLLABORATION FRAMEWORK OF UAV AND UGV FOR 3D VISUALIZATION OF CONSTRUCTION SITES, in: 2018.
- [39] E. Petritoli, F. Leccese, L. Ciani, Reliability assessment of UAV systems, in: 2017 IEEE Int. Workshop Metrol. Aerosp. MetroAeroSpace, 2017: pp. 266–270. <https://doi.org/10.1109/MetroAeroSpace.2017.7999577>.
- [40] K.W. Williams, A Summary of Unmanned Aircraft Accident/Incident Data: Human Factors Implications, FEDERAL AVIATION ADMINISTRATION OKLAHOMA CITY OK CIVIL AEROMEDICAL INST, 2004. <https://apps.dtic.mil/sti/citations/ADA460102> (accessed November 25, 2021).
- [41] S. Giese, D. Carr, J. Chahl, Implications for unmanned systems research of military UAV mishap statistics, in: 2013 IEEE Intell. Veh. Symp. IV, 2013: pp. 1191–1196. <https://doi.org/10.1109/IVS.2013.6629628>.
- [42] G.G. De la Torre, M.A. Ramallo, E. Cervantes, Workload perception in drone flight training simulators, *Comput. Hum. Behav.* 64 (2016) 449–454. <https://doi.org/10.1016/j.chb.2016.07.040>.
- [43] A. Kumar Jha, S. Sathyamoorthy, V. Prakash, Bird strike damage and analysis of UAV's airframe, *Procedia Struct. Integr.* 14 (2019) 416–428. <https://doi.org/10.1016/j.prostr.2019.05.051>.
- [44] B. Zhang, L. Tang, M. Roemer, Probabilistic Weather Forecasting Analysis for Unmanned Aerial Vehicle Path Planning, *J. Guid. Control Dyn.* 37 (2014) 309–312. <https://doi.org/10.2514/1.61651>.

Progress Estimation of an Excavation Pit

A. Vierling¹, T. Groll¹, D. Meckel¹, K. Heim², D. Walter², L. Scheidhauer²,
K. Körkemeyer² and K. Berns¹

¹Department of Computer Science, Robotics Research Lab, University of Kaiserslautern, Germany

²Department of Civil Engineering, Chair in Construction Management, University of Kaiserslautern, Germany

{vierling,groll,meckel,berns}@cs.uni-kl.de,

{kristina.heim,daniel.walter,lukas.scheidhauer,karsten.koerkemeyer}@bauing.uni-kl.de

Abstract -

This paper presents a method for automated excavation speed and progress estimation. First, a measure for the progress speed of an excavation pit is taken from the literature and evaluated regarding the possibility for automation. For each possible parameter, an automated extraction algorithm is presented. The used system is an autonomous excavator arm of a backhoe loader where the used hardware and software system is described. Experimental evaluation of the presented approach has been done with the autonomous system for a small trench, including multiple digging cycles. The resulting measurements seem to include some systematic errors which could be identified and suitable sanity checks could be implemented, removing the erroneous measurements. The remaining measurements were used to determine the excavation speed of the autonomous excavator arm and compared to the values of experienced and amateur operators.

Keywords -

Excavation; Progress Estimation; Autonomous System; Sensors

1 Introduction & Related Work

In recent years the demand in construction has increased tremendously, and presumably the trend will continue with an estimated growth of 3.5% per year till 2030 [1], where infrastructure construction is deemed as the most rapidly growing subsector with an estimated annual growth of 4% [1] till 2030. However, a major global problem in construction are delays. There exist a multitude of reasons for these delays, but [2] identified a "poor planning and scheduling" as one of the top 3 delay factors. In order to make it possible to react to such mishaps in planning and scheduling, it is necessary to be able to detect delays as early as possible. A continuous comparison between the planned and the real timeline would be needed, but it is not feasible to do such checks by hand. Therefore, these comparisons need to be automated. While the planning data is

usually present in a suitable digital form, the real schedule is much more difficult to extract, and strongly depends on the exact construction task. A specific activity in nearly all construction sites is the excavation of a construction pit.



Figure 1. Excavator with additional sensor system (blue circle)

The volume estimation of such a construction pit is a well-researched problem from the theoretical point of view [3, 4, 5, 6]. However, estimation by real sensor data is not well researched, and if, then by using additional sensors in the surrounding of the pit [7] which might not be applicable for all construction sites. A more promising trend is the advancement in autonomous excavation, which inherently also mandates a pit monitoring. Here current research deals with the lower [8] and higher level [9] control, but also with the environment perception [10] and specifically the modeling of it [11, 12]. Combining the two levels has only recently become possible, as in [13], which uses similar sensors and representations, but does not specifically tackle the estimation of the excavation progress or speed. The paper at hand will also try to solve the problem of an automatic progress estimation utilizing autonomous excavation. An autonomous backhoe loader, more specifically the excavator's arm, is used for a prototypical implementation of an automated excavation state and speed monitoring. The used hardware and already existing software of the backhoe loader is described

in section 2, and the methodology for extracting additional sensor data for a progress estimation is described in section 3. Afterwards and experimental evaluation is done in 4 and discussed in section 5.

2 System & Preliminary Work

The progress estimation of the excavation pit should be based on an autonomous excavation arm of a backhoe loader. Of course, such an autonomous system already yields a lot of helpful intermediate states and systems, which will be beneficial for the estimation, and are therefore described in this section.

2.1 Hardware

The used backhoe loader is a *John Deere 410J TMC* with a front-loader and an excavation arm, an image of the machine can be seen in figure 1. The backhoe loader is already equipped with encoders that give the position of each joint, as well as a control system to control each of the joints via CAN-bus signals. Additionally, a *Microstrain 3DM-GX5-25* Inertial Measurement Unit (IMU) and a *Hemisphere V320* global navigation satellite system (GNSS) is mounted on the top of the cabin, and a corresponding Real Time Kinematic (RTK) base station is placed in the vicinity and connected to the GNSS system. With this, a precise localization of the machine can be achieved. For the perception of the surroundings an *Ouster OS0-128* 3D-laser scanner with 128 lines and a Field-of-View of $90^\circ \times 360^\circ$ and a *IFM O3D301* Time-of-Flight (ToF) camera is mounted next to a simple RGB-webcam on the crowd.

2.2 Software

The higher-level behavior-based control system for the excavation process is mainly based on the architecture already described in [14]. A short summary is given in figure 3. Essentially the complete digging cycle is divided into 4 phases, "move to digging", "digging", "move to dumping", and "dumping" arranged as a state machine. Each phase is further divided into smaller motion primitives. Transitions between phases happen when the right conditions, e.g., bucket angle, are met.

One preliminary perception algorithm which is used in this work is the bucket volume estimation described in [15]. Here, a ToF-camera measures a point cloud of the bucket. Then, a rasterized grid-map is created. Afterward, the difference to a grid-map of the empty bucket is calculated. This difference map allows the estimation of the filled volume inside the bucket. Also other estimation techniques [16] [17] for the bucket volume and fill level exist, but as the one in [15] was specifically implemented

for the excavator at hand it was an obvious choice. Another used result is the classification of rocks inside a rock pile, described in detail in [18]. Here, an RGB-image and a depth-image is used to capture the scene. A watershed-segmentation algorithm extracts pixels belonging to different rocks. With these pixel segmentations, an estimation of the rock sizes becomes feasible. In [18], this approach was used to determine if a rock would need to be crushed for loading, but here it will be used in a slightly different manner, as described later.

3 Approach & Implementation

As a next step, one needs to lay a foundation of how a suitable progress estimation of an excavation pit can be done. There are already methods known to parameterize the progress speed in civil engineering. The most obvious speed estimation would be to measure the excavated volume per time. However, one wants to normalize this excavation speed with respect to the current circumstances. This adapted speed is usually called *performance* $Q_n[\frac{m^3}{h}]$ and the relevant parameters will be explained further. Afterward, new methodologies on how to extract necessary information from the sensor data are given.

3.1 Progress Estimation

The main basis of the progress estimation is based on the works of Girmscheid, especially [19, 20]. An overview of the necessary parameters is given in table 1.

Name	Symbol	Unit	A/H
Nominal Bucket Volume	V_B	m^3	A
Cycle time	t_c	s	A
Dissolving factor	α	1	A
Fill factor	ϕ	1	A
Trench depth factor	f_1	1	A
Rotation angle factor	f_2	1	A
Emptying accuracy	f_3	1	H
Teeth condition	f_4	1	H
Maintenance	f_5	1	H
Operator	η_1	1	H
Operating condition	η_2	1	H
Machine utilization rate	η_G	1	H

Table 1. Progress estimation parameters according to [19, 20]. Names are translated, and symbols are adapted. Column A/H shows whether the parameter will be assumed to be measured automatically (A) or by hand (H).

The performance can then be computed as:

$$Q_n = \frac{V_B}{t_c} \times 3600 \times \alpha \times \phi \times f_1 \times f_2 \times f_3 \times f_4 \times f_5 \times \eta_1 \times \eta_2 \times \eta_G$$

From the needed parameters, the ones that can be measured



Figure 2. Image of the start of the four different digging phases of the excavation process, from left to right: "Digging", "Move to Dumping", "Dumping", "Move to Digging"

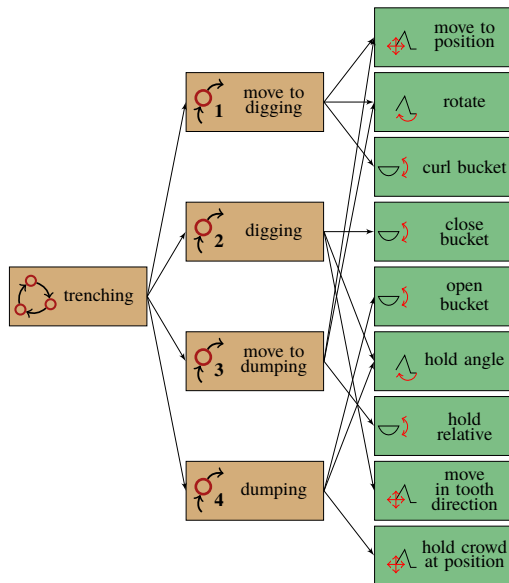


Figure 3. The architecture of the trenching process [14]. Images of each phase can be seen in figure 2

automatically are described first. The datasheet of the machine defines the nominal bucket volume. The cycle time is the time between the start and end of each digging cycle. The dissolving factor is the ratio between the volume of the compacted and the loosened soil. The fill factor measures how much of the fill level of the bucket really corresponds to the volume of the content, e.g., for big rocks, a lot of space is unused compared to fine sand. The trench depth describes the depth of the current digging cycle. The rotation angle is defined as the angle between the digging and dumping positions in each cycle. All other given parameters can currently not be determined automatically. Some are also not relevant for autonomous excavation and can not be calculated in such an excavator. This includes the emptying accuracy, the size of the dumping region (which is always a "point" in our autonomous excavation scenario), as well as the operator factor (skill factor). Similarly, the conditions of the teeth and the maintenance state can be more easily extracted from external sources. In principle it is possible to estimate both with additional sensors, e.g.

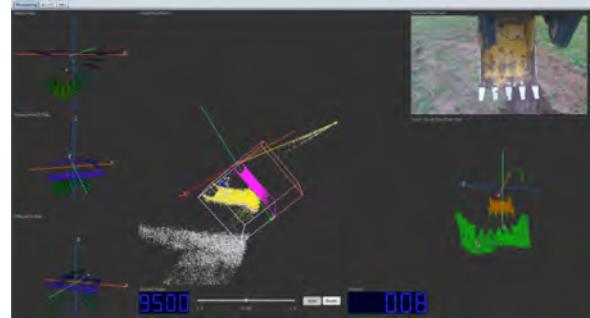


Figure 4. Bucket fill level estimation [15]

the teeth state could be estimated by using the point cloud of the time-of-flight camera and then estimating how sharp the teeth still are. However, in both cases extracting information about the working hours from the software of the manufacturer seems to be the better solution, as this is usually recorded and often accessible to the user. However, an integration is machine and manufacturer specific and therefore omitted here. The operating conditions are not as clearly defined and include the effect of weather, as well as the quality of preparation work and are therefore not measured automatically.

3.2 Data Abstraction

In order to extract the defined parameters, additional evaluation of the sensor and control data has to be done, which is not present in the preliminary work. A valuable help is here the knowledge which action is currently executed by the excavator's arm. The visualization of each phase or action can be seen in figure 2. With this information, the cycle time can be directly calculated by saving the timestamp of the start of the "digging"-phase and calculating the difference till the subsequent "digging"-phase starts. Similarly, the trench depth can be deduced by the z-position of the bucket teeth, which can be calculated by the kinematic model as in [14], at the start of the "move to dumping" position, e.g., the end of the "digging"-phase plus an offset accounting for the bucket width. In the same manner, the rotation angle can be computed by taking the difference of the yaw-values at the start of the "dumping"

and "digging" phases.

The fill and dissolving factors are not as easily calculated. For the fill factor, the RGB-image and depth image at the start of the "digging"-phase is considered, and then a classification of the visible rock sizes, as in [18], is used. The fill factor is then determined according to the rock size in the image and the table 2.

Rock Size (s)	Fill Factor
$s < 63\text{mm}$	1
$63\text{mm} \leq s < 200\text{mm}$	0.85
$200\text{mm} \leq s$	0.75

Table 2. Adapted fill factors. In [19, 20] fill factors for all soil classes (defined by [21]) are given, but the fill factors do not differ much. The used adaptations are a compromise between classification difficulty and effectiveness

In order to calculate the dissolving factor, the estimation of the bucket fill level as in [15] is used, which corresponds to the volume of the loosened soil. To calculate the volume of the compacted soil, the 3D point cloud of the laser scanner is used. The point cloud is rasterized into a grid map, as seen in figure 6. Between the start of two "digging"-phases, the height difference of the grid-cells belonging to the target map, which defines the wanted final form of the pit, are calculated. The compacted volume can be estimated with this depth difference and the grid size. A major disturbance in this method is the soil, which is only moved by the bucket and not loaded into it. Therefore one would like to include all grid-cells in the difference calculation, but this leads to another problem. Figure 6 shows the laser scanner-based pose estimation. The measured points of each scan are fused into the map. Although the terrain around the excavator is flat, the grid map shows increasing terrain further away from the excavator. While the accuracy of the laser scanner reduces with distance, the divergences are mainly due to the fact that the exact pose of the scanner is not known because of the vibrations of the excavator's arm. The pose further deteriorates during the digging process. This pose error is amplified by the distance, which is why it is not directly visible in the vicinity of the excavator but still plays a role in the measurements. Therefore not all grid-cells should be chosen. As a trade-off the cells directly neighboring the target map are additionally used for the difference calculation.

4 Experiments

For the experimental evaluation, a trench with dimensions $2\text{m} \times 1\text{m} \times 5\text{m}$ (Height×Width×Length) has been excavated. An image of the excavated trench can be seen in figure 5. The resulting values can be seen in table 3. In



Figure 5. Final excavated trench

the area of the experiments no bigger rocks were present. Therefore the fill factor was always assumed to be 1 and is excluded from the table.

Some cycles in table 3 include clear measurement errors. For instance, the GPS Signal was lost in cycles 6 and 18. Therefore, the depth offset of the backhoe's position is wrong. For other cycles, e.g., 6 and 19, the estimated compacted volume is bigger than the loosened volume. This can only happen for large rocks. In general, an over-estimation of the compacted volume seems to be a trend. This is probably due to the way the autonomous excavation happens: the excavator's arm often goes too deep into the ground and pulls much more soil out of the ground than can fit in the bucket. This soil will be measured as removed from the pit, but not as loosened volume in the bucket. However, this error should be leveled out, when this soil is scooped up and moved to the correct dumping position. In order to have a reasonable speed estimation at all times, a sanity check taking these two problems into account is proposed. It includes the following rules:

- Skip measurement if $f_1 > 4.9\text{m}$ (maximum digging depth)
- Skip measurement if $\alpha > 1$ and $\phi > 1$ (Reminder: $\phi = 1$ in all provided experiments)

If only the remaining measurements are included, the estimated performance Q_n is resulting in 60% of a proficient excavator operator. In comparison, an amateur would achieve roughly 65% according to [19, 20], which suggests that the measurement methods yield suitable results.

5 Conclusion

The presented approach for the estimation of the excavation speed seems to be promising, as it yields good results in many cases. Even though some measurement errors are present, they do not affect the task's suitability. Still, general improvements in the estimation of the loosened soil could be done. Possible extensions include adapting the size of the grid-cells as well as a more precise localization

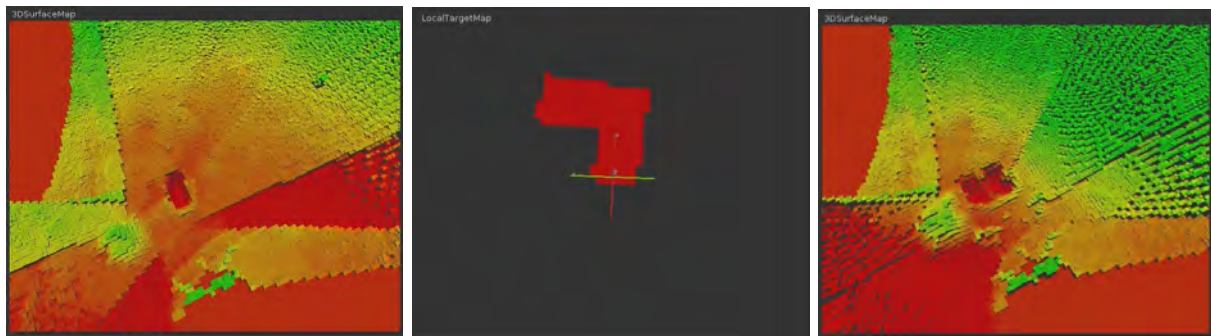


Figure 6. Grid map of two consecutive digging cycles (left and right) and corresponding target map (middle). The cell heights are color-coded. Green indicates a high positive value and red a high negative value. The pit can be seen as dark red cells in the middle of the grid maps

of the laser scanner to achieve a better fusion. Approaches like scene flow or ICP-matching seem suitable for improving this aspect. This would also reduce the effect of a lost GPS signal which is a major disturbance in the current implementation. However, in general, with the advancement of the automation of construction machinery and an increasing number of sensors on a construction site, similar approaches are promising. A detailed picture of the state of the complete construction site could be achieved by combining the knowledge of different vehicles.

6 Acknowledgments

The implementation, methodology and evaluation of this paper were partially funded by the German Federal Ministry of Transport and Digital Infrastructure (Bundesministerium für Verkehr und digitale Infrastruktur **BMVI**) per the *Infra-Bau 4.0* project.

References

- [1] Graeme Harrison and Jeremy Leonard. Future of construction. Technical report, Oxford Economics, September 2021. URL <https://www.oxfordeconomics.com/recent-releases/Future-of-Construction>.
- [2] Youcef J.-T. Zidane and Bjorn Andersen. The top 10 universal delay factors in construction projects. *International Journal of Managing Projects in Business*, 11(3):650–672, Jan 2018. ISSN 1753-8378. doi:10.1108/IJMPB-05-2017-0052. URL <https://doi.org/10.1108/IJMPB-05-2017-0052>.
- [3] Chun-Sung Chen and Hung-Cheng Lin. Estimating pit excavation volume using cubic spline volume formula. *Journal of Surveying Engineering*, 117(2):51–66, 1991. doi:10.1061/(ASCE)0733-9453(1991)117:2(51). URL <https://ascelibrary.org/doi/abs/10.1061/%28ASCE%290733-9453%281991%29117%3A2%2851%29>.
- [4] B. Mukherji. Estimating 3d volume using finite elements for pit excavation. *Journal of Surveying Engineering*, 138(2):85–91, 2012. doi:10.1061/(ASCE)SU.1943-5428.0000059. URL <https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29SU.1943-5428.0000059>.
- [5] Ragab Khalil. Credibility of 3d volume computation using gis for pit excavation and roadway constructions. *American Journal of Engineering and Applied Sciences*, 8:434–442, 04 2015. doi:10.3844/ajeassp.2015.434.442.
- [6] Maxim Tyulenev, O. Litvin, M. Cehlár, Sergey Zhironkin, and M. Gasanov. Estimation of hydraulic backhoes productivity for overburden removing at kuzbass open pits. *Acta Montanistica Slovaca*, 22: 296–302, 01 2017.
- [7] Maximilian Bügler, Gbolabo Ongunmakin, Jochen Teizer, Patricio Vela, and Andre Borrmann. A comprehensive methodology for vision-based progress and activity estimation of excavation processes for productivity assessment. 01 2014. doi:10.13140/RG.2.1.4630.2561.
- [8] Guilherme J. Maeda, David C. Rye, and Surya P. N. Singh. Iterative autonomous excavation. In Kazuya Yoshida and Satoshi Tadokoro, editors, *Field and Service Robotics: Results of the 8th International Conference*, pages 369–382, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg. ISBN: 978-3-642-40685-0, DOI: https://doi.org/10.1007/978-3-642-40686-7_25.

Cycle	t_c [s]	Compacted Volume [m^3]	Loosened Volume [m^3]	α [1]	f_1 [1]	f_2 [1]
01	72,8911	0,2389	0,3385	0,7057	0,2606	62,6429
02	72,9800	0,2929	0,3471	0,8440	0,2269	61,1909
03	66,8381	0,6701	0,3590	1,8666	0,6320	61,4380
04	186,7704	0,3922	0,2905	1,3501	0,7886	62,5649
05	194,0915	0,1275	0,2709	0,4709	0,6929	61,1918
06	690,1479	2,1733	0,2675	8,1247	56,9951	61,2457
07	65,0554	0,0750	0,2419	0,3102	1,1417	61,1202
08	72,5589	0,3927	0,2743	1,4316	1,2213	62,8282
09	370,7839	0,5742	0,3167	1,8129	1,2758	62,3137
10	67,6435	0,0653	0,2808	0,2325	1,3913	61,2390
11	65,2036	0,0733	0,1967	0,3726	1,0822	62,4411
12	139,5601	0,3869	0,1954	1,9805	1,5954	62,2147
13	248,4500	0,5289	0,1807	2,9280	0,9151	60,9772
14	179,0602	0,0992	0,1807	0,5491	0,7188	62,9664
15	85,5514	0,7926	0,2308	3,4334	0,5826	62,8606
16	68,8132	0,4620	0,2484	1,8601	1,4567	62,1681
17	120,2028	0,3718	0,2195	1,6937	0,4379	61,7577
18	1745,1303	0,8514	0,1395	6,1015	123,8112	63,1761
19	54,9628	2,3939	0,2204	10,8624	0,4236	62,5417
20	60,4940	0,5597	0,2977	1,8800	0,3359	62,3002
21	56,5907	0,5973	0,2339	2,5540	0,2400	62,8761

Table 3. Experimental evaluation

- [9] Daniel Schmidt. *Shaping the Future - A Control Architecture for Autonomous Landscaping with an Excavator*. RRLab Dissertations. Verlag Dr. Hut, München, 2016. <http://www.dr.hut-verlag.de/978-3-8439-2816-8.html> ISBN-13: 978-3-8439-2816-8.
- [10] Thomas Emter, Christian Frese, Angelika Zube, and Janko Petereit. Algorithm Toolbox for Autonomous Mobile Robotic Systems. *ATZoffhighway worldwide*, 10(3):48–53, sep 2017.
- [11] Philipp Woock, Nina F. Heide, and Daniel Kuehn. Robotersysteme für die dekontamination in menschenfeindlichen umgebungen. In *Proceeding at Leipziger Deponiefachtagung 2020. Leipziger Deponiefachtagung (LDFT-2020), March 3-4, Leipzig, Germany*. n.n., 2020.
- [12] Nina Felicitas Heide, Philipp Woock, Maximilian Sauer, Timo Leitritz, and Michael Heizmann. Ucsr: Registration and fusion of cross-source 2d and 3d sensor data in unstructured environments. In *IEEE 23rd International Conference on Information Fusion (FUSION)*, 2020.
- [13] Dominic Jud, Simon Kerscher, Martin Wermelinger, Edo Jelavic, Pascal Egli, Philipp Leemann, Gabriel Hottiger, and Marco Hutter. Heap - the autonomous walking excavator. *Automation in Construction*, 129: 103783, 2021.
- [14] Tobias Groll, Steffen Hemer, Thorsten Ropertz, and Karsten Berns. A behavior-based architecture for excavation tasks. In *34th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 1005–1012. International Association for Automation and Robotics in Construction, June 28-July 1 2017. DOI: <https://doi.org/10.22260/ISARC2017/0139>.
- [15] Steffen Hemer, Tobias Groll, and Karsten Berns. Backhoe bucket volume and trench monitoring using a time-of-flight camera. In *7. Fachtagung Baumaschinentechnik 2018: Digitalisierung, Automatisierung, Mensch*, Dresden, September 20–21 2018. Forschungsvereinigung Bau- und Baustoffmaschinen e.V. FVB, Forschungsvereinigung Bau- und Baustoffmaschinen e.V. FVB.
- [16] Liangjun Zhang, Jinxin Zhao, Pinxin Long, Liyang Wang, Lingfeng Qian, Feixiang Lu, Xibin Song, and Dinesh Manocha. An autonomous excavator system for material loading tasks. *Science Robotics*, 6(55): eabc3164, 2021. doi:10.1126/scirobotics.abc3164. URL <https://www.science.org/doi/abs/10.1126/scirobotics.abc3164>.
- [17] Haodong Ding, Xibin Song, Zhenpeng He, and

- Liangjun Zhang. Real-time volume estimation of mass in excavator bucket with lidar data. In Chen Feng, Thomas Linner, Ioannis Brilakis, Daniel Castro, Po-Han Chen, Yong Cho, Jing Du, Semiha Ergan, Borja Garcia de Soto, Jozef Gasparik, Firas Habbal, Amin Hammad, Kepa Iturralde, Thomas Bock, Soonwook Kwon, Zoubeir Lafhaj, Nan Li, Ci-Jyun Liang, Bharadwaj Mantha, Ming Shan Ng, Daniel Hall, Mi Pan, Wei Pan, Farzad Rahimian, Benny Raphael, Anoop Sattineni, Christian Schlette, Isaac Shabtai, Xuesong Shen, Pingbo Tang, Jochen Teizer, Yelda Turkan, Enrique Valero, and Zhenhua Zhu, editors, *Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 696–703, Dubai, UAE, November 2021. International Association for Automation and Robotics in Construction (IAARC). ISBN 978-952-69524-1-3. doi:10.22260/ISARC2021/0094.
- [18] Tobias Groll, Steffen Hemer, and Karsten Berns. Classification of rocks inside a rock pile for task planning. In 8. *Fachtagung Baumaschinentechnik 2020: Maschinen, Prozesse, Vernetzung*, 2020.
- [19] Gerhard Girmscheid. *Leistung im Baubetrieb*, pages 1–18. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. ISBN 978-3-642-13795-2. doi:10.1007/978-3-642-13795-2_1. URL https://doi.org/10.1007/978-3-642-13795-2_1.
- [20] Gerhard Girmscheid. *Faktoren*, pages 249–279. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. ISBN 978-3-642-13795-2. doi:10.1007/978-3-642-13795-2_8. URL https://doi.org/10.1007/978-3-642-13795-2_8.
- [21] DIN 18300. VOB - Vergabe- und Vertragsordnung für Bauleistungen - Teil C: Allgemeine Technische Vertragsbedingungen für Bauleistungen (ATV) - Erdarbeiten, September 2018.

Impact of Passive Back-Support Exoskeletons on Manual Material Handling Postures in Construction

A. Golabchi^a, L. Miller^b, H. Rouhani^c, and M. Tavakoli^d

^a Department of Civil and Environmental Engineering, University of Alberta, Canada

^b EWI Works International Inc., Canada

^c Department of Mechanical Engineering, University of Alberta, Canada

^d Department of Electrical and Computer Engineering, University of Alberta, Canada

E-mail: alirezal@ualberta.ca, lmiller@ewiworks.com, hrouhani@ualberta.ca, mahdi.tavakoli@ualberta.ca

Abstract –

Work-Related Musculoskeletal Disorders (WMSDs) are a leading contributor to workplace injuries in the construction industry, with the lower back being the most affected body part. To mitigate WMSDs, exoskeletons have been developed and recently introduced to industrial job sites to provide workers with assistance and support, reducing exposure to ergonomic risks. Due to the newness of industrial exoskeletons, successful application of this technology in the construction industry requires thorough evaluation of different aspects of its adoption to ensure a successful and effective uptake. As Manual Material Handling (MMH) tasks are the most common cause of lower back injuries, this study aims to evaluate the impact of using exoskeletons when adopting different postures during dynamic and static MMH tasks. An experiment is carried out and data reflecting Rate of Perceived Exertion (RPE), Level of Discomfort (LOD), overall fit and comfort, effectiveness, and limitation and interference levels is collected. Overall, the participants perceived the exoskeleton suit as effective with discomfort being reduced in the lower back and other body parts except the chest. However, the results indicate the importance of considering the specific task at hand (e.g., dynamic vs static MMH) and the posture adopted (e.g., squat vs bend) when evaluating and selecting an exoskeleton for construction tasks.

Keywords –

Exoskeleton; Exosuit; Wearable Robot; Manual Material Handling; Ergonomics; Posture; Construction

1 Introduction

Manual material handling (MMH) involving lifting, carrying, pushing, pulling, lowering, restraining, and holding is the most common cause of occupational

fatigue, lower back pain and lower back injuries [1], leading to high rates of Work-Related Musculoskeletal Disorders (WMSDs) in the construction industry, including 30% of all lost workday cases among construction trades in the US [2]. Recently, exoskeletons, also known as exosuits or wearable robots, are adopted for different industrial applications to mitigate the ergonomic risks associated with physically demanding tasks, especially the ones involving MMH. The use of exoskeletons for such physically demanding tasks have shown to reduce fatigue and the frequency of injuries [3].

Several advancements have been recently made in the development and evaluation of exoskeletons for the construction industry. In a review article, Zhu et al. [4] investigated existing exoskeleton technologies and analyzed their potential for MMH tasks in construction. They generated a map to suggest the appropriate exoskeleton type for each trade while evaluating the benefits and challenges. In another study, Cho et al. [5] designed a wearable exoskeleton to habituate construction workers to safe postures and demonstrated that the developed exoskeleton can effectively assist workers when performing construction tasks. Ogunseiju et al. [6] evaluated a postural assist exoskeleton and its effectiveness for construction tasks involving MMH. They reported improvements in posture when using the exoskeleton over time, although higher perceived discomfort in the lower back was reported due to the pressure applied to the users' back. In another study, Capitani et al. [7] described the development of a passive exoskeleton to assist construction workers in dealing with shotcrete projection tasks. They indicated that the designed exoskeleton preserved adaptability to different lower-limb tasks without reducing its comfort during utilization. Furthermore, Chen et al. [8] presented a bilateral knee exoskeleton to provide kneeling assistance for construction workers. The results showed reductions in knee pressure, potentially leading to decreased WMSD risk for workers when performing kneeling activities on level and sloped surfaces.

While previous studies have provided valuable insight into the potential of adopting exoskeletons in the construction industry, more research is required to evaluate the different aspects of the adoption due to the recentness of using the exoskeleton technology for industrial applications. As one of the important aspects of effective adoption of exoskeletons is the impact of using an exoskeleton on user posture, this study intends to evaluate the effect of a back support exoskeleton on postures adopted when carrying out MMH tasks.

The goal of this study is to compare different postures adopted during dynamic and static MMH tasks, with and without exoskeletons. The experiments are designed to provide feedback on the impact of using exoskeletons on comfort, fatigue, and usability factors, for both male and female users.

2 Methods

2.1 Experimental Design

As passive exoskeletons have shown to be more suitable for industrial applications compared to active exoskeletons due to lighter weight, lower price and simpler maintenance [2], a passive exoskeleton was used for the experiment. Furthermore, since the back is the primary body part affected by WMSDs in construction [9], a back-support exoskeleton was selected. Back support exoskeletons are designed to reduce the load on the low back muscles during bending tasks by redistributing the weight to the legs [10]. The backX exoskeleton was used which weighs 7.2 lbs and can reduce the strain on the user's lower back by an average of 60%.

The experiment was designed to simulate dynamic and static MMH tasks. Participants were asked to carry out the tasks in different scenarios to cover different task types (i.e., dynamic and static), postures (i.e., freestyle, bending, squatting), and the impact of the exoskeleton (i.e., with and without wearing the exoskeleton).

2.2 Participants

For this study, 12 healthy individuals, including 6 male and 6 female, were asked to participate in the experiment. The mean and standard deviation for the age, weight, and height of the participants were 28 ± 6.28 years old, 143 ± 33.87 lb., and $5' 6.8'' \pm 4''$, respectively. None of the participants reported any current or previous musculoskeletal disorder or illness. The detailed process including the objectives, instructions and possible risks were explained to each participant through written and verbal instructions and on-site discussions. Ethics approval was received for the study from the University of Alberta Research Ethics Board (Study ID:

Pro00109264).

2.3 Testing Procedure

The variables of the experiment included freestyle, bending, and squatting lifting postures, existence of the exoskeleton, and the static and dynamic nature of the task. Dynamic MMH involved lifting, carrying, and placing a 20 lb. box multiple times, while each time lifting and placing on a surface with a different height (i.e., on floor and on a table). Static MMH involved moving items from a box and placing them on a table through a static posture.

Prior to the experiment, participants were introduced to the procedure and equipment. Participants were given enough time between each experiment to recover from any fatigue associated with the previous experiment. After completing each experiment, the participants were asked a series of questions including the Rate of Perceived Exertion (RPE), Level of Discomfort (LOD), overall fit and comfort of using the exoskeleton, the extent to which the exoskeleton limits movements and interferes with movements, effectiveness of the exoskeleton, and other general feedback. In total, seven scenarios reflecting different postures and the use of the exoskeleton were tested, as shown in Table 1. Figure 1 shows the experiment setup for some of the scenarios as a sample.

Table 1 Experiment scenarios

Category	Scenario	Trial Type
Dynamic MMH	D1	No Exo + Freestyle
	D2	Exo + Freestyle
	D3	Exo + Bending
	D4	Exo + Squatting
Static MMH	S1	No Exo + Freestyle
	S2	Exo + Bending
	S3	Exo + Squatting

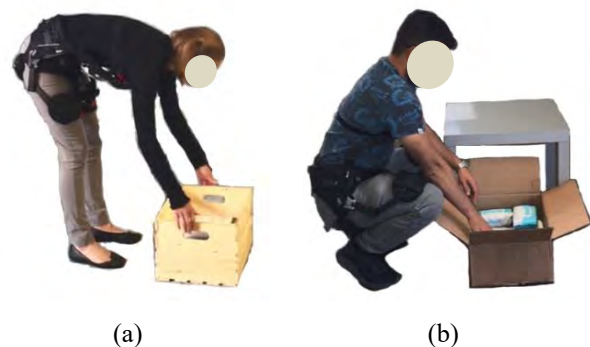


Figure 1. Experiment setup: (a) dynamic MMH,

bending with exoskeleton (scenario D3), (b) static MMH, squatting with exoskeleton (scenario S3)

2.3.1 Participant Response

The participants were asked to rate the level of their perceived discomfort (i.e., LOD) on a Borg CR 10 scale, where 0 indicates no discomfort and 10 shows maximum discomfort [11]. The intensity of perceived discomfort was measured and quantified after conducting each experiment. Furthermore, the participants provided the discomfort ratings separately for each body part including shoulder, chest, lower back, thighs, feet, etc. on a scale of 0 to 10. Also, RPE was rated from 1 (very light activity) to 10 (maximum effort), fit/comfort of the exoskeleton suit was rated from 1 (not satisfactory) to 10 (very satisfactory), limitation/interference was rated from 1 (limits a lot) to 10 (does not limit at all), and effectiveness was rated from 1 (not effective at all) to 10 (very effective). Collected data was explored through descriptive statistical analysis.

3 Results

3.1 Dynamic MMH

The reported RPEs for the dynamic task are shown in Figure 2. As shown in the figure, the average RPE for the different scenarios of the dynamic MMH is fairly close. Overall, it can be concluded that the existence of the exoskeleton and the posture used does not impact the average RPE. It is also worth noting that in cases where the exoskeleton was used, a maximum RPE of 6 was reported by users.

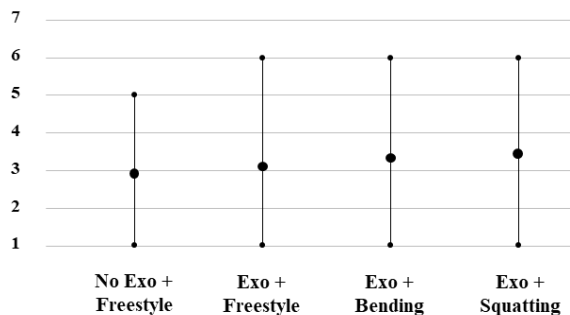


Figure 2. Reported RPE for dynamic MMH

For each scenario, the participants were also asked about the comfort and useability of the exoskeleton. In particular, they were asked to rate the overall fit and comfort level, the extent to which it limits movements and interferes with activities, and the overall effectiveness. The results are shown in Figure 3. The results indicate that the participants feel similar effectiveness in all postures, while they experienced

more comfort during bending. Overall, the participants felt that the exoskeleton moderately limits their movements and can interfere with other tasks.

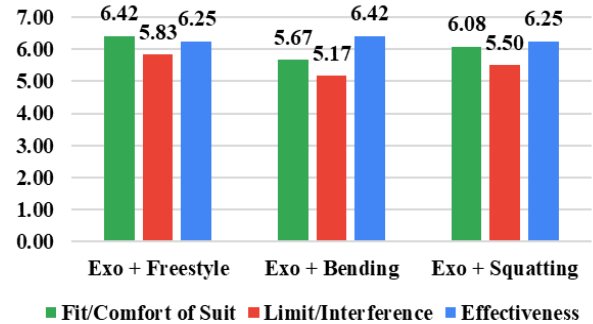


Figure 3. Reported usability for dynamic MMH

The reported body discomforts for the dynamic scenarios are presented in Figure 4. As shown, most of the perceived discomfort is detected in the lower back and legs. While using the exoskeleton substantially reduced the discomfort in the lower back during bending and squatting, using the exoskeleton with a freestyle posture did not have considerable impact in improving the discomfort. It is also observed that the perceived discomfort in legs is much higher when squatting compared to bending.

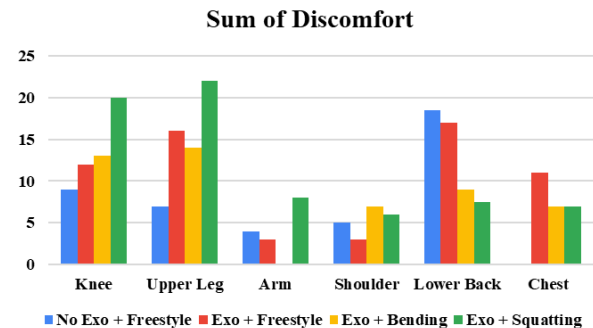


Figure 4. Perceived discomfort of body parts for dynamic MMH

3.2 Static MMH

The reported RPEs for the static task are shown in Figure 5. While minimum and maximum reported RPEs are similar for all scenarios, the average RPE is reported as slightly higher for squatting, which indicates the difficulty of performing the static task in a squatting posture due to the pressure applied to the legs and the need to maintain balance.

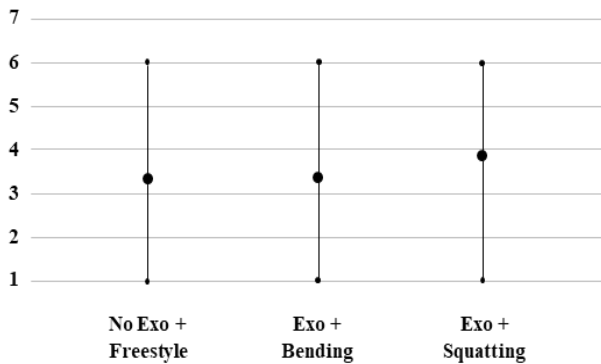


Figure 5. Reported RPE for static MMH

As shown in Figure 6, the overall effectiveness and comfort level is higher in bending compared to squatting. However, higher levels of limitation and interfering with other activities is also reported for bending.

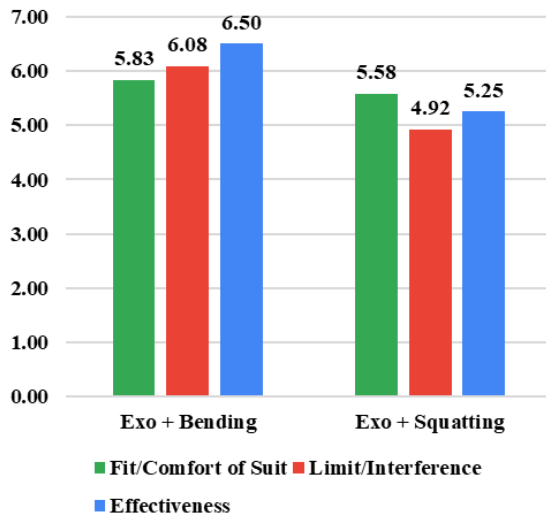


Figure 6. Reported usability for static MMH

Figure 7 demonstrates the reported discomfort for the static scenarios. Since the body is positioned in an awkward position for a prolonged period in the static task, the discomfort levels are generally high without the exoskeleton. Similar to the dynamic tasks, most of the reported discomforts are in the lower back and legs. The use of the exoskeleton has caused higher discomfort levels on the chest during bending, which is due to the existence of the chest pad. While the use of the exoskeleton has reduced the discomfort on the legs in bending compared to squatting, the discomfort in the lower back is much less in squatting.

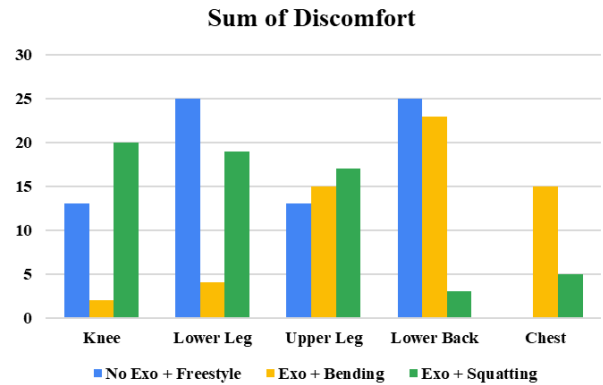


Figure 7. Perceived discomfort of body parts for static MMH

3.3 Dynamic vs. Static MMH

Figure 8 shows a comparison between the results for all scenarios of the dynamic and static bending tasks. While other factors remain the same, higher level of limitation is reported during static tasks. Overall, it can be concluded that the performance of the exoskeleton for bending is similar for both static and dynamic tasks.

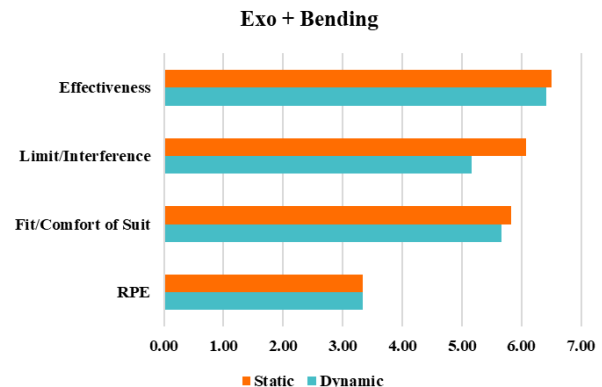


Figure 8. Comparison between static and dynamic scenarios for bending

Similarly, Figure 9 shows the comparison between the results for all scenarios of the dynamic and static squatting tasks. While the overall effectiveness and comfort is higher for dynamic tasks, the limitation level is reported slightly higher.

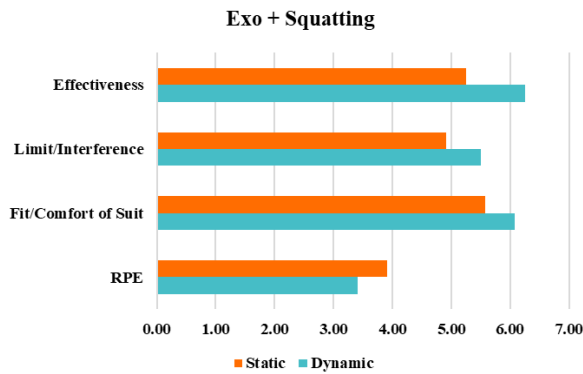


Figure 9. Comparison between static and dynamic scenarios for squatting

3.4 Male vs. Female

Figure 10 shows a comparison between the average RPEs for male and female participants for the dynamic scenarios. As shown, male participants reported a higher RPE in the dynamic scenarios compared to female participants. Also, the perceived exertion is similar for bending and squatting postures among both groups.

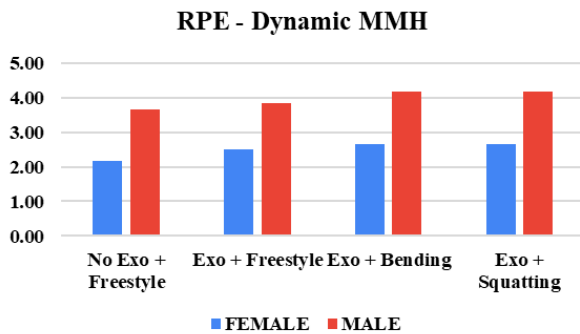


Figure 10. RPE comparison between male and female participants for dynamic scenarios

Similarly, Figure 11 shows a comparison between the average RPEs for male and female participants for the static scenarios. While both groups reported a slightly higher RPE in squatting compared to bending, male participants reported higher RPEs for all scenarios of static MMH. Using the exoskeleton did not improve the exertion levels when using the bending posture.

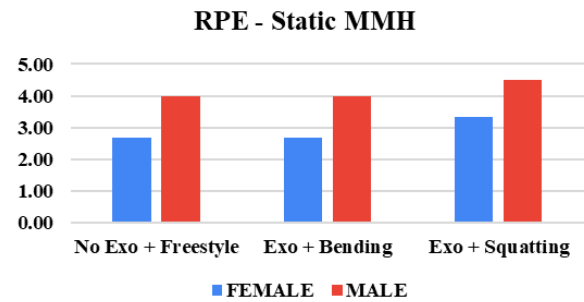


Figure 11. RPE comparison between male and female participants for static scenarios

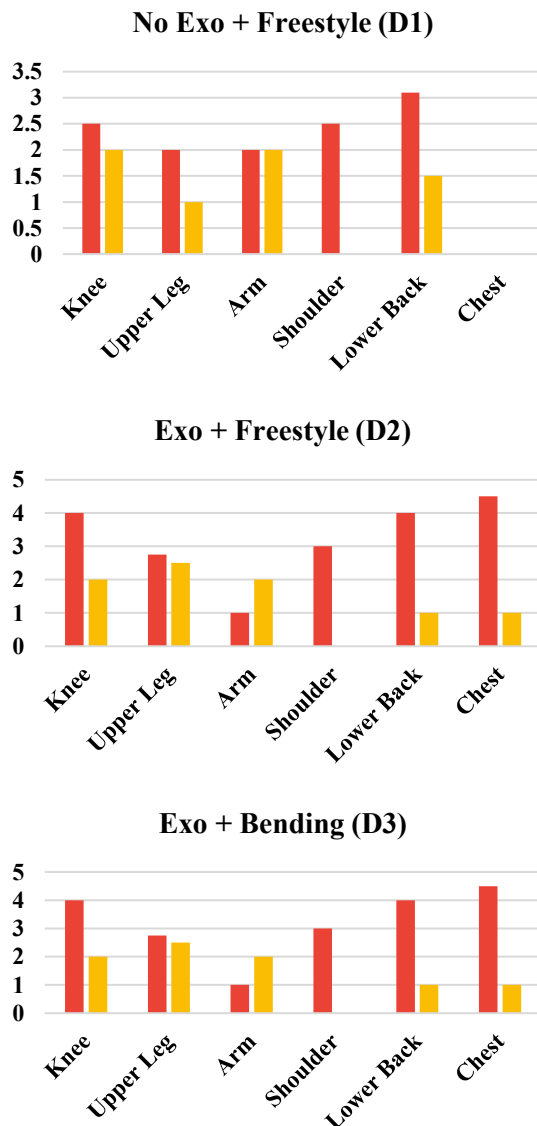
Table 2 shows a comparison of the average responses for the usability factors for male and female participants. As shown in the table, female participants found the exoskeleton more effective in all MMH scenarios, while both groups rated the fit and comfort level fairly similar. On the other hand, female participants rated the limitation factor of the exoskeleton higher than male participants.

Table 2 Comparison of usability responses for male and female participants

Factor	Scenario	Female	Male
Fit/Comfort	D2	6	6.83
	D3	5.5	5.83
	D4	6.17	6
	S2	5.83	5.83
	S3	5.5	5.67
Limit/Interference	D2	6.33	5.33
	D3	5.17	5.17
	D4	6	5
	S2	6.83	5.83
	S3	4.83	5
Effectiveness	D2	6.83	5.67
	D3	6.67	6.17
	D4	7.17	5.33
	S2	7.17	5.83
	S3	5.83	4.67

Figure 12 shows a comparison of the LOD for male and female participants during different dynamic MMH scenarios. Male participants reported higher discomfort when carrying out the dynamic task without the exoskeleton (D1), with the highest discomfort on the lower back. When using the exoskeleton with a freestyle posture (D2), both groups reported discomfort in the chest area, with male participants reporting substantially higher LOD. During dynamic bending (D3), male participants reported most discomfort on the chest, lower

back, and knees, while female participants reported the highest LOD on the upper leg and knees. The highest reported LOD during dynamic squatting (D4) belongs to the chest, lower back, and knees for male participants and upper leg, knees, and arms for female participants. While male participants reported discomfort on the shoulder in all dynamic scenarios, there were no reported LOD for shoulders by female participants. Overall, male participants reported higher LOD for all body parts except arms.



Exo + Squatting (D4)

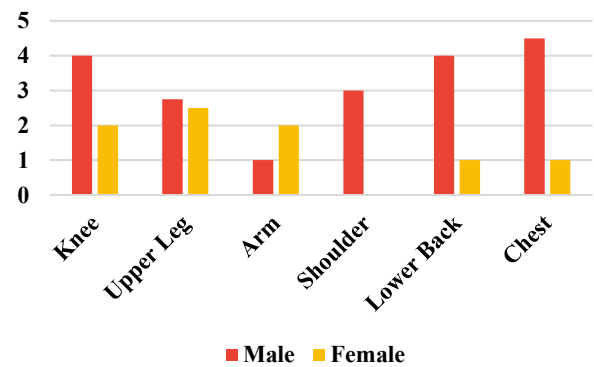
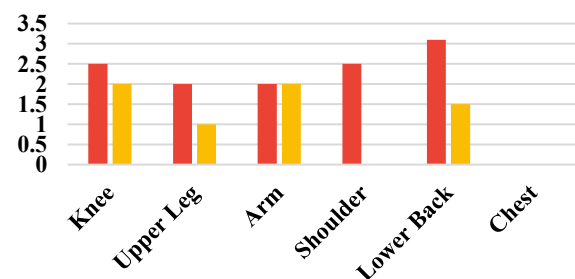


Figure 12. Comparison of LODs between male and female participants in dynamic MMH scenarios

Figure 13 shows a comparison of the LOD for male and female participants during different static MMH scenarios. When carrying out the static task without the exoskeleton (S1), male participants reported most discomfort on lower back, shoulder, and knees, while female participants reported the highest LOD on the arms and knees. The use of exoskeleton for the static bending task (S2) has resulted in higher discomfort in both groups, with male participants reporting highest LOD on the chest, lower back, and knees, and female participants reporting highest LOD on upper leg, knees, and arms. For static squatting (S3), male participants reported higher LOD in all body parts compared to female participants. Chest, lower back, and knees have the highest LOD among male participants, while upper leg, knees, and arms are the highest rated body parts for female participants. Both groups reported similar levels of discomfort in upper legs. Overall, similar to the dynamic MMH scenarios, male participants reported higher LOD for all body parts except arms for all static MMH scenarios.

No Exo + Freestyle (S1)



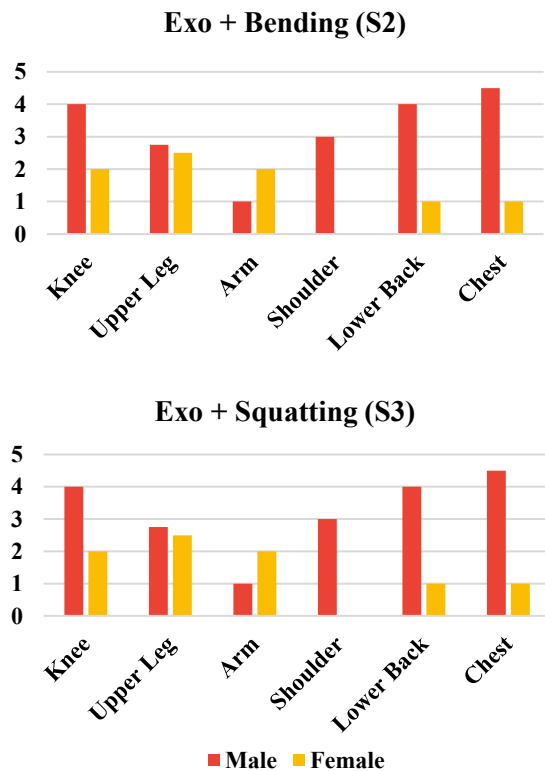


Figure 13. Comparison of LODs between male and female participants in static MMH scenarios

4 Conclusion

Emerging technologies such as exoskeletons have the potential to reduce the high rate of WMSDs in the construction industry. However, their adoption has to be evaluated from different aspects before introducing them to job sites, to ensure a successful and effective uptake. As MMH tasks are among the top contributors to WMSDs in construction, this study aimed to evaluate the impact of a passive back-support exoskeleton on different MMH postures. The results indicate that: (1) the impact of using the exoskeleton is similar for dynamic and static MMH tasks, while it is slightly less effective for squatting during static tasks; (2) using the exoskeleton impacts the level of perceived discomfort on different body parts especially the chest; (3) male participants experience higher discomfort on almost all body parts when wearing the exoskeleton compared to female participants; and (4) majority of the participants rated the exoskeleton suit as providing acceptable usability, while female participants found the suit more effective. According to the reported LOD, the lower back, knees, upper legs, and chest are the most affected body parts by the exoskeleton suit. Meanwhile, it should be noted that the use of exoskeleton reduced discomfort in the

mentioned body parts except the chest.

Based on the results, it can be concluded that passive exoskeleton suits have the potential to be adopted to reduce the rate of WMSDs in construction. However, proper training and supervision is required on the postures adopted by the workers, based on the specific characteristics of the task that is carried out. It is important that exoskeletons are properly selected for the task at hand and is solely used for the identified task.

This study was limited in that the experiments were carried out for a short amount of time. Long-term trials are required to reflect on the impact of using exoskeletons on different factors more accurately. Furthermore, while subjective metrics can be useful for evaluation of exoskeletons from a usability perspective, the lack of objective measures limits the generalization of the analysis. Future studies should also include objective evaluation features [12] for a more comprehensive analysis. Furthermore, the findings of studies such as this one can be used in future studies to assist with improving the design of exoskeletons.

References

- [1] Canadian Centre for Occupational Health and Safety (CCOHS). MMH – Introduction. <https://www.ccohs.ca/oshanswers/ergonomics/mmh/mmhintro.html>, Access: 02/2022.
- [2] Ogunseiju O, Gonsalves N, Akanmu A, Nnaji C. Subjective Evaluation of Passive Back-Support Exoskeleton for Flooring Work. *EPiC Series in Built Environment*. 2021.
- [3] Antwi-Afari MF, Li H, Anwer S, Li D, Yu Y, Mi HY, Wuni IY. Assessment of a passive exoskeleton system on spinal biomechanics and subjective responses during manual repetitive handling tasks among construction workers. *Safety science*. 2021.
- [4] Zhu, Z., Dutta, A. and Dai, F. Exoskeletons for manual material handling—A review and implication for construction applications. *Automation in Construction*, 122, p.103493. 2021.
- [5] Cho, Yong K., et al. A robotic wearable exoskeleton for construction worker's safety and health. *ASCE construction research congress*. 2018.
- [6] Ogunseiju, O., Olayiwola, J., Akanmu, A., and Olatunji, O. A. Evaluation of postural-assist exoskeleton for manual material handling. *Engineering, Construction and Architectural Management*. 2021.

- [7] Capitani, S. L., Bianchi, M., Secciani, N., Pagliai, M., Meli, E., and Ridolfi, A. Model-based mechanical design of a passive lower-limb exoskeleton for assisting workers in shotcrete projection. *Meccanica* 56.1. 2021.
- [8] Chen, S., Stevenson, D. T., Yu, S., Mioskowska, M., Yi, J., Su, H., and Trkov, M. Wearable knee assistive devices for kneeling tasks in construction. *IEEE/ASME Transactions on Mechatronics* 26(4). 2021.
- [9] CPWR. The Construction Chart Book. The U.S. Construction Industry and Its Workers. Sixth Edition. Silver Spring, MD. 2018.
- [10] Kazerooni, H., Wayne Tung, and Minerva Pillai. Evaluation of trunk-supporting exoskeleton. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 63. No. 1. Sage CA: Los Angeles, CA: SAGE Publications. 2019.
- [11] Borg, Gunnar. The Borg CR10 scale folder. *A method for measuring intensity of experience*. Hasselby, Sweden: Borg Perception. 2004.
- [12] Golabchi, A., Chao, A., and Tavakoli, M. A Systematic Review of Industrial Exoskeletons for Injury Prevention: Efficacy Evaluation Metrics, Target Tasks, and Supported Body Postures. *Sensors* 22(7). 2022.

Stag hunt game-based approach for cooperative UAVs

L.V. Nguyen¹, I. Torres Herrera¹, T.H. Le¹, D.M. Phung^{1,2}, R.P. Aguilera¹ and Q.P. Ha¹

¹School of Electrical and Data Engineering, University of Technology Sydney (UTS), Australia

²VNU University of Engineering and Technology (VNU-UET), Vietnam National University, Hanoi (VNU), Vietnam

{vanlanh.nguyen; ignacio.torresherrera; hoang.t.le; manhduong.phung; ricardo.aguilera; quangha}@uts.edu.au

Abstract -

Unmanned aerial vehicles (UAVs) are being employed in many areas such as photography, emergency, entertainment, defence, agriculture, forestry, mining and construction. Over the last decade, UAV technology has found applications in numerous construction project phases, ranging from site mapping, progress monitoring, building inspection, damage assessments, and material delivery. While extensive studies have been conducted on the advantages of UAVs for various construction-related processes, studies on UAV collaboration to improve the task capacity and efficiency are still scarce. This paper proposes a new cooperative path planning algorithm for multiple UAVs based on the stag hunt game and particle swarm optimization (PSO). First, a cost function for each UAV is defined, incorporating multiple objectives and constraints. The UAV game framework is then developed to formulate the multi-UAV path planning into the problem of finding payoff-dominant equilibrium. Next, a PSO-based algorithm is proposed to obtain optimal paths for the UAVs. Simulation results for a large construction site inspected by three UAVs indicate the effectiveness of the proposed algorithm in generating feasible and efficient flight paths for UAV formation during the inspection task.

Keywords -

Unmanned aerial vehicle; Cooperative path planning; Stag hunt game; Payoff-dominant equilibrium; Particle swarm optimization

1 Introduction

The immense development of unmanned aerial vehicles (UAVs) technologies has been drawing significant attention in civilian sectors. In the construction domain, researchers and enterprises tend to seek safer and more efficient solutions for carrying out construction-related tasks. As modern technologies have reduced the cost of UAVs while increasing their dependability, operating time, and maneuverability, smart drone-powered solutions have been introduced as a platform to assist construction activities. They are well-established in numerous construction project stages such as construction site monitoring and

3D mapping [1], building and damage inspection [2], and package delivery logistics [3], demonstrating prospects for wide usage of drones.

Due to the increased quantity and complexity of construction jobs, such as large-site 3D mapping or multiple-package delivery, a single drone with restricted size and capability may sometimes not fulfill the requirements. Consequently, multi-UAVs can be deployed to collaborate as a team for the applications mentioned above in order to optimize processing time and operating possibilities [4].

A hierarchical structure formation system includes three layers: task management, path planning, and task execution, as shown in Figure 1. The task management layer holds and keeps track of the mission's objectives. Based on these objectives, this layer allocates resources and tasks to UAVs and acts as a decision-maker. From mission requirements, the path planning layer generates feasible trajectories for the formation. This layer aims to plan optimal paths for the UAVs moving in a known environment from the start to their target locations. This layer comprises three blocks: real-time trajectory modification, data acquisition, and cooperative path planning, wherein the collaborative path planning block is the primary function of the system and determines the overall optimized path for each quadcopter. Nonetheless, due to uncertainties that may occur along the route in practical applications, the real-time trajectory modification block is combined with the system. The formation can deal with emergencies such as a suddenly appearing obstacle. Generated paths will then be passed down to the task execution layer. This layer directly is connected with the propulsion system of the quadcopter and generates the control laws. To enhance system performance, real-time data, i.e., UAVs' position and velocity, are fed back to the path planning layer to adjust the path, providing a closed control loop.

This paper focuses on cooperative path planning, where the path-planner produces trajectories to fulfill the mission objectives. The mission objectives include formation shape maintenance, minimum path lengths, and threat avoidance. The development is then illustrated in the inspection of a construction site.

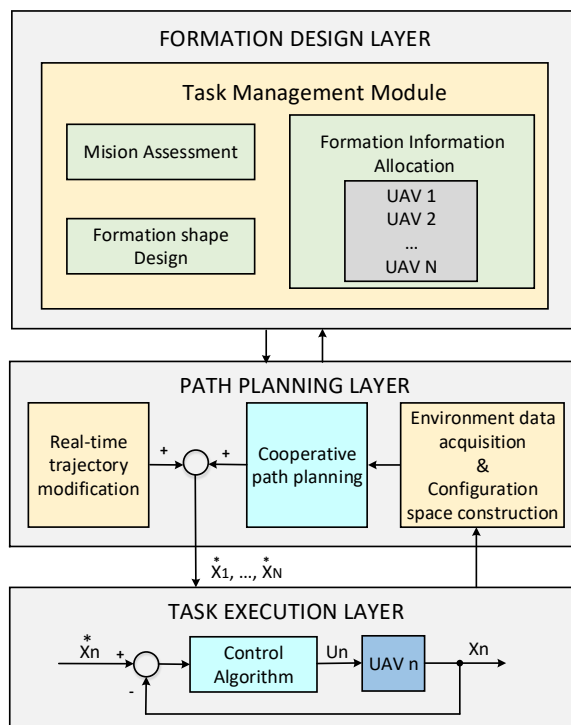


Figure 1. Hierarchical UAV control structure.

2 Related works and Rationale

Path planning for multiple UAVs has been becoming an active research topic recently. The objectives and approaches may be different depending on the application domain.

The artificial potential field is one of the most widely used techniques for UAV cooperative path planning [5]. It considers the operating space as a potential field, with attractive fields surrounding the target and repulsive fields around obstacles. For cooperation, new fields are added, including the internal attractive potential fields to retain the formation configuration and the internal repulsive fields to prevent the UAVs from colliding with each other. The paths are then generated under the total force acting on the UAVs at each position. This technique can produce smooth and continuous paths. It, however, faces local minima problems when the total force is reaching zero.

In another direction, optimal control methods have been used for cooperative path planning by considering a numerical optimization problem subject to multiple constraints [6]. It first finds paths for single vehicles and then attains their cooperation by constraints. To generate paths, optimal control signals for individual quadcopters are computed using the mixed-integer linear programming

(MILP). Although the MILP can solve the optimization problem with different constraints, it involves high computational complexity.

Recently, evolutionary algorithms (EA) have been used to solve multi-UAV cooperative path planning with the capability to find optimal solutions in complex scenarios [7]. They often include two stages, one for individual UAV path planning and the other for path cooperation. Initially, each vehicle is associated with the EA process to generate a feasible path. Those paths are then adjusted via a global cost function to achieve the required cooperation. This approach can generate smooth cooperative paths as it does not discretize the workspace. It however may converge to sub-optimal solutions if cooperative constraints and maneuver properties of UAVs are not properly addressed.

On the other hand, as a theoretical framework for strategic resolving of interactions among competing players, game theory has gained its applications in a variety of fields such as construction bidding [8] and environmental management in the mining industry [9]. In the game, the purpose is to maximize the profit, which depends on not only actions of a player but also other players. Therefore, the best strategy relies on what the player expects others to do. Most games in the literature can be classified as cooperative and non-cooperative [10]. In cooperative games (CGs), several players share a common objective to better achieving it than those working alone. However, a major issue with CGs remains trade-off between the stability of the player groups and the system efficiency [11]. In contrast, each player in noncooperative games (NCGs) has their own properties such as the payoff function, procedural details of the game, intention, and possible strategies which make him more inclusive than in CGs. Among NCGs, Nash Game is widely applied to the situation that all players have to simultaneously make their decisions in symmetric competitions, such as in exploring the public-private partnership investment incentives [12] or seeking strategies for clusters in a distributed system [13]. In these games, each player only has partial information about the choices of others.

In general, Nash games revolve around the prisoner's dilemma (PD), where the stable cooperation is inhibited by its most restrictive conditions: defecting is a dominating option as it always offers a greater payout regardless of how the other player while collaboration is risky [14]. Meanwhile, numerous types of dilemmas are available, the most famous of which is the stag hunt game [15], therein players desire to coordinate, i.e., the preferred choice is to always do as the rival acts. While in such games, mutual defection and mutual cooperation are evolutionary stable strategies, only the mutual cooperation results in a pay-off dominant equilibrium [16].

From the UAVs' cooperation perspective, where profit

from cooperation is much more than profit from individual effort, the stag hunt game can better resolve the conflict between individual UAVs. This is especially the case when examining the influence of group selection, in which social interactions aim to maximize the group's performance.

This paper proposes a stag hunt game based algorithm for UAV cooperative path planning. A cost function is first defined including all requirements on formation, path feasibility, safety and optimality. Unlike existing EA approaches, our method considers cooperative constraints in every individual to maximize the overall profit. The path planning problem is therefore modelled as a game where UAVs are the players. The strategy for each UAV is then formulated and enhanced PSO is introduced to obtain the payoff-dominant equilibrium. As a result, optimal paths can be achieved with the formation being maintained.

This paper is structured as follows. Section 3 formulates the cooperative path planning problem. Section 4 presents the proposed stag hunt based game and enhanced PSO. Numerical simulation results are provided in Section 5. Finally, a conclusion is given in Section 6.

3 Problem formulation

3.1 Multi-vehicle path planning

Consider a team of N drones operating in a given flying area, including numerous obstacles, as shown in Figure 2. The position of the UAV group is determined in an earth frame, xyz , by $P = [P_1^T, P_2^T, \dots, P_N^T]^T$, where N is the total number of drone members and $P_n = (x_n, y_n, z_n)^T$ is the location of the n -th vehicle.

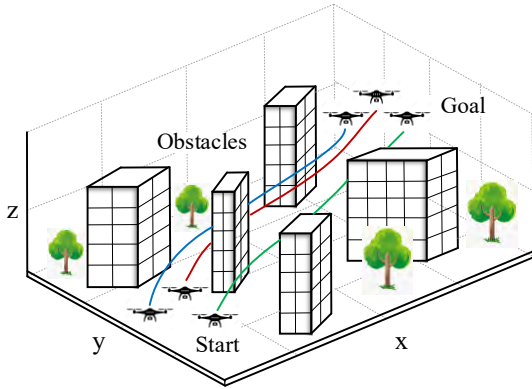


Figure 2. Definition of the path planning problem.

The problem of path planning is to establish a feasible route connecting the start and target positions in a collision-free environment while fulfilling a number of constraints. The problem can be expressed as an optimization process that is subjected to several costs. For a

single UAV planning problem, it can be formulated as

$$P(0) \xrightarrow[s.t. J_s(X(k))]{X(k)} P(end), \quad (1)$$

where $P(0)$ and $P(end)$ are corresponding to the start and the target poses, k stands for the waypoint instant, and $X(k)$ denotes the path of the UAV, including a total of K waypoints subjected to the single-UAV cost $J_s(X(k))$.

In single-vehicle path planning, to attain the most effective path, the cost $J_s(X(k))$ should be optimized, fulfilling constraints on path length, threat avoidance, and turning angle limit. It is defined as

$$J_s(X(k)) = \omega_1 \sum_{k=1}^{K-1} L(k) + \omega_2 \sum_{k=1}^{K-1} \sum_{\tau=1}^{\mathcal{T}} D_{\tau}(k) + \omega_3 \sum_{k=1}^K H(k) + \omega_4 \sum_{k=1}^{K-2} \theta(k) + \omega_5 \sum_{k=1}^{K-1} |\varphi(k) - \varphi(k+1)|, \quad (2)$$

where $L(k)$ is the path length, $D(k)$ is the safety cost concerning \mathcal{T} threats, $H(k)$ stands for the altitude payoff, $\theta_n(k)$ and $\varphi_n(k)$ correspond to the turning angle and climbing angle, ω_i , for $i = \{1, 2, \dots, 5\}$, are the weight coefficients. More details about the single cost function are presented in our previous work [17].

Extending the problem into the multi-vehicle formation path planning, it is written as

$$P_n(0) \xrightarrow[s.t. J(X_n(k), X_n^-(k))]{X_n(k)} P_n(end), \quad n = 1, 2, \dots, N, \quad (3)$$

where $X_n(k)$ and $X_n^-(k)$ are corresponding to the path of UAV _{n} and a set of its neighbours' paths. The multi-vehicle cost function, $J(X_n(k), X_n^-(k))$, consists of a single cost and a formation cost, computed as

$$J(X_n(k), X_n^-) = J_s(X_n) + \beta J_f(X_n, X_n^-), \quad (4)$$

where $J_s(X_n)$ is computed as (2), $J_f(X_n, X_n^-)$ is the formation cost, and β is a weighting factor. The cost function for the formation constraint is determined as follows.

3.2 Formation cost function

We apply here the graph-theoretic approach to represent the structure of the formation and interaction among UAVs. A graph is defined by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, in which $\mathcal{V} = \{v_1, v_2, \dots, v_N\}$ and $\epsilon = (v_n, v_{n'}) \in \mathcal{E}$ represent the vehicles in the group and their interconnections, respectively. To form a formation, there must exist an interconnection in \mathcal{E} between any two vertices $(v_n, v_{n'}) \in \mathcal{V}$. Consider our graph is edge-weighted, i.e., each interconnection in the graph is weighted by $\mu_{nn'}$. The graph incidence matrix \mathcal{D}

has the dimension of $N \times M$, where its uv -th entry is equal to 1 or -1 if the UAV_{*n*} is respectively the head or tail of the v -th edge, and 0 otherwise.

The configured formation shape is defined by a reference position set, $P_r = [P_{1_r}^T, P_{2_r}^T, \dots, P_{S_r}^T]^T$, where an element $P_{n_r} = (x_{n_r}, y_{n_r}, z_{n_r})^T$ is given as the reference position of UAV_{*n*}. The relative reference between two neighbors, n and n' , is then obtained as $P_{nn'} = P_{n_r} - P_{n'_r}$. Accordingly, the formation error between UAV_{*n*} and UAV_{*n'*} is calculated as $P_n - P_{n'} - P_{nn'}$. Denote $\hat{D} = \mathcal{D} \otimes I_3$, where operator \otimes is the Kronecker product. The formation error for UAV_{*n*} can be expressed as

$$E_n = \sum_{n' \in \mathcal{E}} \mu_{nn'} \|P_n - P_{n'} - P_{nn'}\|^2 \quad (5)$$

$$= (P - P_r)^T \hat{D} \hat{W}_n \hat{D}^T (P - P_r) = \|P - P_r\|_{\hat{D} \hat{W}_n \hat{D}^T}^2,$$

where $\hat{W}_n = W_n \otimes I_3$ and $W_n = \text{diag}[\mu_{nn'}]$ is a diagonal matrix of dimension $M \times M$ for the weights $\mu_{nn'}$.

Let $\bar{d}_n(k)$ be the Euclidean distance from UAV_{*n*} to its nearest neighbor at the waypoint k , r_n be the radius of UAV_{*n*}, and d_s be the safe distance. To avoid collision between vehicles, the distance between a UAV and its nearest neighbor needs to be smaller than the sum of a safe distance, d_s , and twice the UAV radius, r_n . Therefore, we incorporate this collision avoidance condition by reformulating the formation error as

$$E_n(k) = \begin{cases} \|P(k) - P_r\|_{\hat{D} \hat{W}_n \hat{D}^T}^2, & \text{if } \bar{d}_n(k) > d_s + 2r_n \\ \infty, & \text{if } \bar{d}_n(k) \leq d_s + 2r_n. \end{cases} \quad (6)$$

The formation cost function is then defined as below:

$$J_f(X_n, X_n^-) = \sum_{k=1}^K E_n(k). \quad (7)$$

4 Stag hunt game-based algorithm for UAV cooperative path planning

Given the cost function $J(X_n, X_n^-)$ defined for each UAV, the cooperative path planning becomes finding paths $X_n, n = 1, 2, \dots, N$ to simultaneously minimize $J(X_n, X_n^-)$. Since this cost depends on not only path X_n generated for UAV_{*n*} but also its rivals' paths X_n^- , finding optimal solutions is a challenging problem.

4.1 The game of stag hunt

To resolve conflicts and interactions among rational decision-makers [18], players can pursue their individual objectives by considering possible goals, behaviors, and countermeasures of other decision-makers to achieve a win-win situation.

The stag hunt game, originated by J.J. Rousseau [19], illustrates a conflict between safety and social cooperation. In the game, two hunters independently decide whether to hunt a stag or a hare without knowing the other's decision. One hunter can catch a hare individually with a high guarantee of success. Meanwhile, the value of a shared stag is far greater than that of a hare, but cooperation between hunters is required to hunt a stag successfully. Therefore, it would be much better for each hunter to choose a more ambitious and far more rewarding goal instead of deciding on a total autonomy and minimal risk strategy. The payoff matrix in Figure 3 illustrates a generic stag hunt with two players.

		Sh	
		Cooperate	Defect
Payoff-dominant equilibrium	Cooperate	4, 4	0, 2
	Defect	2, 0	2, 2
		Risk-dominant equilibrium	

Figure 3. Payoff matrix of stag hunt game

In a game, "Nash equilibrium" is reached when no player can obtain more profits if others do not change their strategies. Formally, a stag hunt is a game with two pure Nash equilibria: risk dominant and payoff dominant. A Nash equilibrium is called "risk dominant" if it has the largest basin of attraction, implying that it is less risky. This means that the more ambiguity players have about the other player's intentions, the more likely they are to pick the plan that best suits them. Meanwhile, the payoff dominant equilibrium is defined as being Pareto superior to all other Nash equilibria in the game.

Pareto optimality is a fundamental concept representing efficiency in a multi-objective optimization problem consisting of several conflicting objectives. A set of alternatives is considered a Pareto optimal solution if no reallocation can further improve any one of the objectives without degrading at least one other. In the stag hunt game, when confronted with a decision among multiple equilibria, all players would vote on the payoff dominant equilibrium since it provides each member with at least as much profit as the other Nash equilibria.

From the mathematical point of view, a formal presentation of the payoff-dominant equilibrium is as follows. Consider a stag hunt game described as $G = (N, S, U)$,

where N is a set of players, $S = (S_1, S_2, \dots, S_N)$ denotes strategy sets, where $S_n = (x_{n1}, x_{n2}, \dots, x_{n\sigma})$, $n = 1, \dots, N$, represents all Σ strategies made by the n -th player, and $U = (U(x_1), U(x_2), \dots, U(x_N))$ stands for a set of players' utility. The allocation $\hat{X} = \{\hat{x}_1^*, \hat{x}_2^*, \dots, \hat{x}_N^*\}$, where $\hat{x}_n^* \in S_n$, is defined as Pareto optimal if it dominates all other re-allocation $X = \{x_1, x_2, \dots, x_N\}$, i.e., both of the following requirements are met:

$$\forall n \in \{1, \dots, N\}, U(\hat{x}_n^*) \geq U(x_n), \quad (8a)$$

$$\nexists n' \in \{1, \dots, N\} : U(\hat{x}_{n'}^*) < U(x_{n'}). \quad (8b)$$

4.2 The game of UAV cooperation

Each vehicle in the cooperative path planning problem has been assigned a cost function $J(X_n, X_n^-)$ defined in (4). These cost functions interact among vehicles. Therefore, it becomes challenging to solve multiple optimal problems simultaneously. As can be seen, the stag hunt game concept aligns well with the cooperative path planning problem and is promising for its solution. As such, this paper proposes a stag hunt game-based approach for cooperative UAVs. The two-step procedure to implement the proposed game-based scheme is as follows.

In the first step, the cooperation of multiple UAVs is formulated as a game to model interactions among the drones, including three key elements: players, strategies, and utility. Each vehicle in the formation is considered as a player, also called a decision-maker. During the game, all UAV players have to simultaneously provide a route, X_n , defined as the player's strategy, without knowledge of the other player's decision. Each player will get his own utility, corresponding to the multi-vehicle cost $J(X_n, X_n^-)$, which is a function of strategies made by himself X_n and his rivals, X_n^- .

Indeed, one UAV can reach its target position alone by solving its single UAV path planning problem to obtain the minimum single cost, $J_s(X_n)$. In a cooperative task, however, this individual optimal solution could result in a significant formation error, $E_n(k)$, leading to a high formation cost $J_f(X_n, X_n^-)$. To successfully perform a cooperative UAV mission with a higher profit, it would be better for each player to choose the more ambitious goal of achieving a far greater reward by providing a formation preserving path in expectation of the other vehicle's cooperation. Therefore, the multi-vehicle cost, $J(X_n, X_n^-)$, combining both single cost $J_s(X_n)$ and formation cost $J_f(X_n, X_n^-)$, should be optimized for all players simultaneously. Accordingly, the game aims to find a payoff-dominant equilibrium.

In the second step, an enhanced PSO-based algorithm is introduced to solve the Pareto optimality, resulting in a payoff-dominant equilibrium as a desired outcome. This step will be presented in the following section.

4.3 Enhanced PSO-based approach for finding payoff-dominant equilibrium

Particle swarm optimization (PSO) is a stochastic optimization algorithm for optimizing a problem by iteratively improving a candidate solution concerning a particular quality measure. It solves a problem by generating a population of possible solutions, known as particles, and relocating them in the search space using a few simple formulae based on the particle's position and velocity. Each particle's movement is guided by its local best-known position and the global best-known pose in the search space, updated when other particles discover better places. This is anticipated to direct the swarm toward the best options.

Formally, consider a d -dimension search space and a swarm consisting of N_{pop} particles, each particle i has a position $X_i \in \mathbb{R}^d$ and a velocity $V_i \in \mathbb{R}^d$. Let Q_i be the best known position of particle i and Q_g be the best known position of the entire swarm. The dynamic algorithm of the swarm is defined as below:

$$V_i(t+1) = c_0 V_i(t) + c_1 r_1 [Q_i(t) - X_i(t)] + c_2 r_2 [Q_g(t) - X_i(t)], \quad (9)$$

$$X_i(t+1) = X_i(t) + V_i(t+1), \quad (10)$$

where c_0 is the inertia weight, c_1 and c_2 are corresponding to self confidence and swarm confidence parameters, and r_1 and r_2 are random values uniformly distributed in the interval $[0, 1]$.

In the UAV path planning problem, the position of a particle is encoded by the flight path X_n . Accordingly, the entire swarm consists of N_{pop} path particles, which are updated for the optimal solution. To speed up the search process, we employ in this study a variant of PSO named spherical vector-based particle swarm optimization (SPSO) developed in our previous work [17]. In the SPSO, waypoints of a flight path are represented in the spherical coordinate system to exploit the corresponding magnitude, elevation, and azimuth components of the variables with speed, turning angle, and climbing slope of the UAV.

To further develop the SPSO for cooperative path planning involving multiple UAVs, we introduce a game-based SPSO to find the payoff-dominant equilibrium. The pseudo-code for the optimization process is described in Algorithm 1. The detail of the algorithm is as follows.

(i) Initialization:

Initially, parameters of the PSO including c_0, c_1, c_2 , number of iterations $maxIt$, and number of particles N_{pop} are first initialized. At this stage, corresponding to $It = 0$, random strategies of the players are also generated and assigned as initial optimal strategies \hat{X}_n, \hat{X}_n^- .

(ii) Evaluation:

At each iteration, from $It = 1$ to $maxIt$, cost values representing the players' profit, $J_i(X_n(It), X_n^-(It))$, are computed as (4), where i denotes a particle in the n -th swarm.

Algorithm 1 Game-based PSO implementation

1. Initialize PSO parameters: $c_0, c_1, c_2, maxIt, N_{pop}$;
2. Set $It = 0$, generate random player's strategies;
3. Obtain the initial optimal strategies \hat{X}_n, \hat{X}_n^- ;
- for** $It = 1 : maxIt$ **do**
 4. Calculate $J(X_n(It), X_n^-(It))$, for $n = 1, 2, \dots, N$;
 - if** $J(X_n(It), X_n^-(It)) \leq J(\hat{X}_n, \hat{X}_n^-), \forall n = 1, 2, \dots, N$ **then**
 5. Update $\hat{X}_n = X_n(It); \hat{X}_n^- = X_n^-(It)$;
 - end if**
 6. Record \hat{X}_n, \hat{X}_n^- ;
 7. Update X_n, X_n^- ;
- end for**
8. Obtain $\hat{X}_n, \forall n = 1, 2, \dots, N$.

(iii) *Optimal strategy:*

The best strategies of all players associated to the particle i at the iteration It are updated if there is more benefit for at least one player without decreasing the other players' profit, i.e., the condition (8) is met. Based on them, the strategy of each player is adjusted for the subsequent iteration according to equations (9) and (10) of the PSO.

(iv) *Termination:*

The algorithm is terminated when exceeding the maximum number of iterations, $maxIt$. Finally, the payoff-dominant equilibrium, or the best allocation, $\hat{X} = \{\hat{X}_1, \hat{X}_2, \dots, \hat{X}_N\}$, is obtained.

5 Simulation results

This section presents simulation results of the proposed path planning algorithm for a group of three drones. The aim is to generate paths for the UAVs flying in an equilateral triangle formation. However, it should be noted that there are no specific constraints on the configuration of formation shapes. The incidence matrix \mathcal{D} is defined as

$$\mathcal{D} = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & -1 \end{bmatrix}. \quad (11)$$

The interconnection weights are set as $W_1 = [1 \ 1 \ 0]$, $W_2 = [1 \ 0 \ 1]$, and $W_3 = [0 \ 1 \ 1]$. Noting that the weights for interconnections are set equal among players since all players play a similar role in the team. In the total cost function, weighting coefficients $[\beta, \omega_1, \omega_2, \omega_3, \omega_4, \omega_5]$ were chosen as $[1, 100, 100, 1, 1, 1]$ for UAV1 to obtain a short collision-free path and $[100, 1, 100, 1, 1, 1]$ for UAV2 and UAV3 to follow their paths while maintaining the formation and avoiding collisions.

In the simulation, we considered two scenarios of construction sites with different sizes of obstacles, illustrating different levels of complexity, to validate the efficiency of the proposed algorithm.

5.1 Construction site with big-size obstacles

In Scenario 1, we considered a construction site with dimensions of $100m \times 100m \times 35m$. The drones were required to travel from the start location to the goal to perform a site monitoring task. The start locations of UAVs were set at: $P_1^{start} = (15; 18.66; 20)$, $P_2^{start} = (10; 10; 20)$, and $P_3^{start} = (20; 10; 20)$. We confirmed the goal poses at $P_1^{goal} = (85; 88.66; 20)$, $P_2^{goal} = (80; 80; 20)$, and $P_3^{goal} = (90; 80; 20)$. The formation reference was obtained at the same as the target position, i.e., $P_{1r} = P_1^{goal}$, $P_{2r} = P_2^{goal}$, and $P_{3r} = P_3^{goal}$. In the construction site, there were two threat areas modeled as yellow cylinders located at (40, 40) and (60, 60) with a radius of 9 m. Parameters of the PSO were set as $c_0 = 0.999$, and $c_1 = c_2 = 1.5$. The PSO run with 2000 particles for 1500 iterations. Aside from the start and goal nodes, each path was established by $K = 10$ waypoints.

The simulation results are shown in Figure 4. The figure shows that collision-free paths are generated. Importantly, an equilateral triangular configuration of UAVs is retained throughout the entire mission, illustrating the efficiency of the proposed approach. The altitude of the UAV team is displayed in Figure 5. The figure demonstrates that the desired height of the vehicles w.r.t. the ground, around 20 m, is achieved and maintained through the routine, further demonstrating the feasibility of the proposed algorithm.

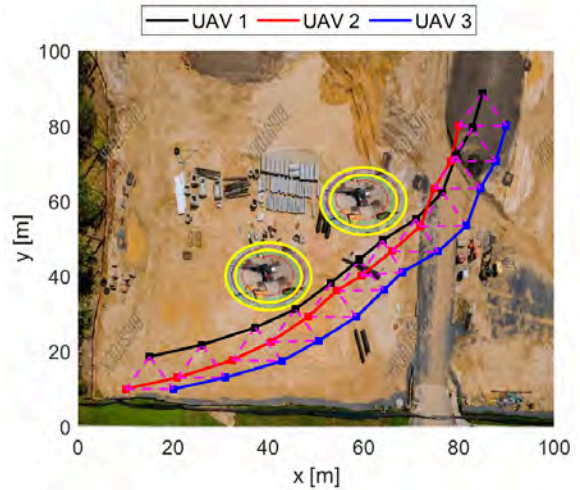


Figure 4. Generated paths for Scenario 1

5.2 Construction site with small-size obstacles

In Scenario 2, we examined the UAV team that has to fly across a construction site to accomplish some tasks. The start and goal locations of the UAVs were chosen to be the same as that of Scenario 1. In Scenario 2, however, two construction cranes as obstacles were located at (40, 48) and (74, 43) with a radius of 4 m. To enhance

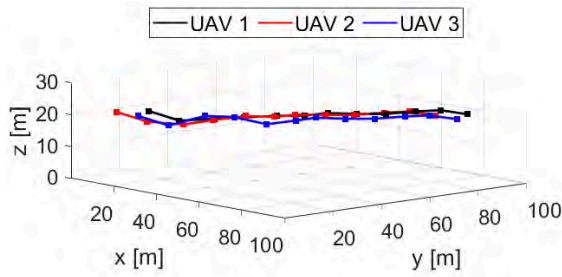


Figure 5. UAVs' altitude

the level of complexity, two more virtual obstacles with the same radius were added at (20, 70) and (70, 60). In this simulation, the performance of the proposed stag hunt game-based algorithm was compared with other available techniques that treat the entire UAV fleet as a rigid body, and path planning was achieved for a virtual drone placed at the centre of the formation [20].

Figure 6 and Figure 7 depict the planned stag hunt game-based path and the planned rigid formation path, respectively. It can be seen that both techniques achieve collision-free and formation-preserving routes. However, the whole group of UAVs, using the rigid formation method, travels around obstacles, resulting in long distances. Meanwhile, the proposed game-based approach presents a capability to split and merge the UAV fleet to avoid small-sized threats and reduce the cost. This illustrates the benefit of the proposed algorithm.

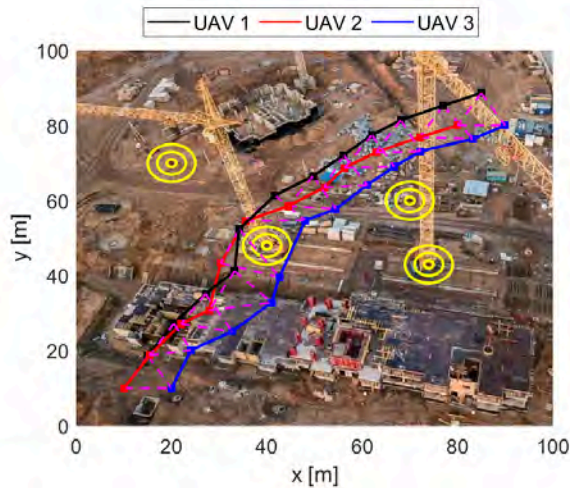


Figure 6. Stag hunt gamed based path

Figure 8 depicts the cost values of all UAVs. As can be seen in the figure, the logarithm of cost values of the UAVs using the stag hunt game-based approach converged to 11.6, 9.7, and 10, after 1000 iterations, implying that the payoff-dominant equilibrium in the UAV game has been achieved. Compared to the logarithm of the cost

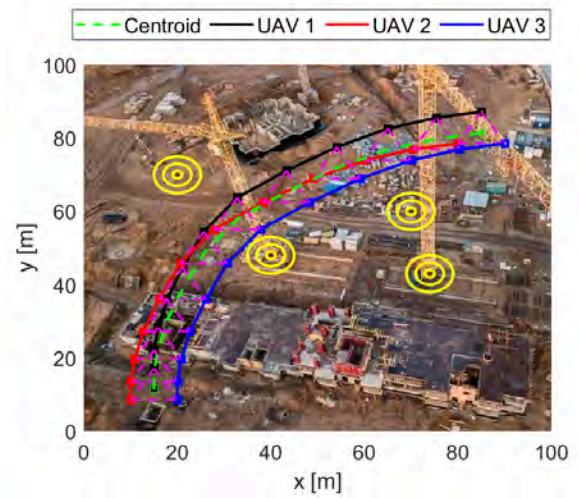


Figure 7. Rigid formation path

value of the virtual UAV using the rigid formation technique, which was about 12.62, all three UAVs achieved better utility. This confirms that the obtained game-based strategy dominates the rigid formation strategy. Since the proposed method requires a larger number of particles and iterations to achieve the payoff-dominant equilibrium, resulting in a higher computational cost. However, the path planning algorithm is implemented offline, this cost is worth for a better Pareto optimal result.

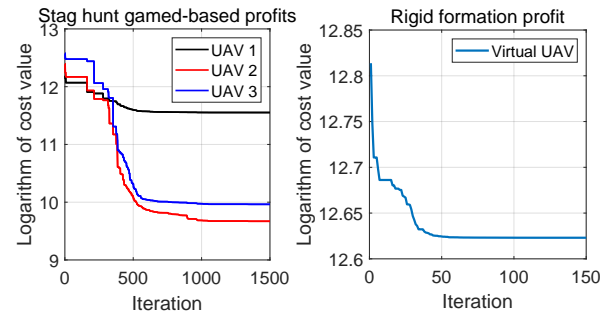


Figure 8. Cost values

6 Conclusion

This study has introduced a novel method based on the stag hunt game theory and game-based particle swarm optimization for cooperative UAVs navigating to assemble into a desired formation configuration. The UAV collaborative path planning problem is solved by finding the payoff dominant equilibrium of the game. An optimization framework using PSO was integrated to find the Pareto optimality by minimizing all cost functions simultaneously. Extensive simulation has been conducted to evaluate the performance of the proposed method for various scenarios

in monitoring construction sites with multiple UAVs. Our future work will develop receding horizon game theory-based platforms for cooperative UAVs in a dynamic construction environment.

Acknowledgements

This work is supported, in part, by the Vingroup Science and Technology Scholarship (VSTS) Program for Overseas Study for Master's and Doctoral Degrees.

References

- [1] Alexey Bulgakov, Daher Sayfeddine, Thomas Bock, and Awny Fares. Generation of orthomosaic model for construction site using unmanned aerial vehicle. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, volume 37, pages 900–904. IAARC Publications, 2020.
- [2] N Bolourian, MM Soltani, AH Albahria, and A Hamad. High level framework for bridge inspection using lidar-equipped uav. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, volume 34. IAARC Publications, 2017.
- [3] Jacek Grzybowski, Karol Latos, and Roman Czyba. Low-cost autonomous uav-based solutions to package delivery logistics. In *Advanced, Contemporary Control*, pages 500–507. Springer, 2020.
- [4] Fran Real, Ángel R Castaño, Arturo Torres-González, Jesús Capitán, Pedro J Sánchez-Cuevas, Manuel J Fernández, Manuel Villar, and Aníbal Ollero. Experimental evaluation of a team of multiple unmanned aerial vehicles for cooperative construction. *IEEE Access*, 9:6817–6835, 2021.
- [5] Thi Thoa Mac, Cosmin Copot, Robin De Keyser, and Clara M Ionescu. The development of an autonomous navigation system with optimal control of an uav in partly unknown indoor environment. *Mechatronics*, 49:187–196, 2018.
- [6] Farid Kendoul. Survey of advances in guidance, navigation, and control of unmanned rotorcraft systems. *Journal of Field Robotics*, 29(2):315–378, 2012.
- [7] Rahul Kala. Multi-robot path planning using co-evolutionary genetic programming. *Expert Systems with Applications*, 39(3):3817–3831, 2012.
- [8] Muaz O Ahmed, Islam H El-Adaway, Kalyn T Coatsney, and Mohamed S Eid. Construction bidding and the winner's curse: Game theory approach. *Journal of Construction engineering and Management*, 142(2):04015076, 2016.
- [9] Benjamin C Collins and Mustafa Kumral. Game theory for analyzing and improving environmental management in the mining industry. *Resources Policy*, 69:101860, 2020.
- [10] José Moura and David Hutchison. Game theory for multi-access edge computing: Survey, use cases, and future trends. *IEEE Communications Surveys & Tutorials*, 21(1):260–288, 2018.
- [11] Zhu Han, Dusit Niyato, Walid Saad, Tamer Başar, and Are Hjørungnes. *Game theory in wireless and communication networks: theory, models, and applications*. Cambridge university press, 2012.
- [12] Shiau-Jing Ho, Sheng-Lung Lin, and Hui-Ping Tserng. An analysis using game theory on the investment incentive of ppp projects. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, volume 34. IAARC Publications, 2017.
- [13] Xianlin Zeng, Jie Chen, Shu Liang, and Yiguang Hong. Generalized nash equilibrium seeking strategy for distributed nonsmooth multi-cluster game. *Automatica*, 103:20–26, 2019.
- [14] Michael W Macy and Andreas Flache. Learning dynamics in social dilemmas. *Proceedings of the National Academy of Sciences*, 99(suppl 3):7229–7236, 2002.
- [15] Brian Skyrms. *The stag hunt and the evolution of social structure*. Cambridge University Press, 2004.
- [16] Herbert Gintis. *Game theory evolving*. Princeton university press, 2009.
- [17] Manh Duong Phung and Quang Phuc Ha. Safety-enhanced uav path planning with spherical vector-based particle swarm optimization. *Applied Soft Computing*, 107:107376, 2021.
- [18] Ankang Ji, Xiaolong Xue, QP Ha, Xiaowei Luo, and Minggong Zhang. Game theory-based bilevel model for multiplayer pavement maintenance management. *Automation in Construction*, 129:103763, 2021.
- [19] Jean-Jacques Rousseau. Discourse on the origin and basis of inequality among men. *Rousseau J.-J. Traktaty*, pages 31–109, 1754.
- [20] YongBo Chen, JianQiao Yu, XiaoLong Su, and GuanChen Luo. Path planning for multi-uav formation. *Journal of Intelligent & Robotic Systems*, 77(1):229–246, 2015.

Reinforcement learning with construction robots: A preliminary review of research areas, challenges and opportunities

X. Xu¹ and B. García de Soto¹

¹ S.M.A.R.T. Construction Research Group, Division of Engineering, New York University Abu Dhabi (NYUAD), Experimental Research Building, Saadiyat Island, P.O. Box 129188, Abu Dhabi, United Arab Emirates

E-mail: xx927@nyu.edu; garcia.de.soto@nyu.edu

Abstract

The interest in the use of robots in the construction industry has been around for several decades; however, the advancement of technology for related applications has been slow. Considering that most construction robots are not fully automated and require extra guidance and instructions from the operators, an autonomous way for robots to understand how to execute specific construction tasks is needed. Reinforcement learning (RL) is a possible solution to this problem. Instead of explicitly detailing the solution to a problem, RL algorithms enable a robot to autonomously discover an optimal behavior through trial-and-error interactions with its environment. It constructs a learning model to solve various sequential decision-making problems. RL algorithms could help construction robots establish a learning process based on the feedback from the construction site and lead to an optimal strategy to finish the sequential construction work. Nowadays, on-site construction robots still require tedious work in preprogramming. Many other areas have used RL algorithms on different tasks, such as dexterous manipulation, legged locomotion, or pathing planning; thus, there is great potential for combining construction robot applications with RL algorithms.

Despite these achievements, most works investigated single-agent settings only. However, many real-world applications naturally comprise multiple decision-makers that interact simultaneously, such as traffic modeling for autonomous vehicles and networking communication packages for multi-robot control. These applications have faced significant challenges when dealing with such high-dimensional environments, not to mention the challenges for the on-site construction robots. The recent development of deep learning has enabled RL methods to drive optimal policies for a sophisticated high-dimensional environment. To the best of our knowledge, there is currently no extensive survey of the applications of RL techniques within the construction industry. This study can inspire future

research into how best to integrate powerful RL algorithms to achieve a higher-level autonomous control of construction robots and overcome resource planning, risk management, and logistic challenges in the industry.

Keywords –

Construction robot; Deep reinforcement learning; Multi-agent; Task allocation; Path planning

1 Introduction

With the expansion of technologies, the construction industry has seen significant attempts at robotization. Some single-purpose robots were designed to conduct specific tasks in highly controlled environments. These applications have shown promising results for single repetitive tasks. However, all these robotic applications do not have the capabilities to adapt their operations to unique work environments. Their motions implement mobility, and manipulations are preprogrammed by the expertise considering the characteristics of tasks and work environments, which limited the wide usage of such robotic systems. Thus, instead of preprogramming all the details, a way for the robot to learn by itself and adopt different scenarios on the construction site is necessary to develop construction robot applications. Reinforcement learning (RL) seems to be a promising solution to these problems.

RL and optimization control theory are used to solve a wide range of tasks using a simple architecture where the agent operates in an environment that models the task it needs to fulfill. RL studies how to use past data to enhance the future manipulation of a dynamical system that adequately adapts to environmental changes [1]. Over the past decades, advances in RL have led robotics to be highly automated and intelligent, with safe operation instead of manual work for many challenging tasks. As an essential branch of machine learning, RL can realize sequential decision-making and has made a series of significant breakthroughs in robot applications. It has

led to a wide range of impressive progress in various domains, such as industrial manufacturing [2], board games [3], robot control [4], and autonomous driving [5]. Although so much success has occurred in various robotic domains, few studies have focused on robots in construction applications.

The lack of attention for RL in the construction industry is due to its complexity [6]: with too many uncertainties on the job site and massive data from the dynamic environment, it is tough for the robot to model the tasks, set up the simulator, and get predictions from past results. However, considering the development of the digitalization of the construction industry, with the more standardized construction process and more modular products coming into the market, the construction work is far more predictable than it used to be. Besides, with BIM models becoming the norm in most job sites, the numerous information stored in the digitalized model can provide reliable data for simulation of the robot's behavior. Thus, this study aims at establishing the connections between RL and construction robotic applications, trying to advance the functionality and reliability of the current construction robot systems.

Therefore, a review of reinforcement learning applications in construction is presented herein to guide subsequent research. This paper aims to help researchers working in this field quickly place their work within the current spectrum, bearing in mind the current challenges and potential. The rest of this paper organizes as follows: Section 2 summarizes the state of the art of the algorithms and gives a brief overview of RL applications. Section 3 provides a framework to figure out how to apply these algorithms to construction robot applications. Section 4 presents a simple case study to adopt the RL algorithms and explains how to set up the algorithms and the environment. Section 5 discusses the results, and Section 6 summarizes the findings and provides further advancement for future research.

2 State-of-the-art RL

2.1 RL Overview

RL is about training an agent to interact with its environment. The agent arrives at different scenarios known as states by performing actions. Actions lead to rewards which could be positive or negative [8].

Some key terms that describe the essential elements of an RL problem are as shown in Figure 1:

- Environment* — Physical world in which the agent operates
- State* — Current situation of the agent S_t , Next situation S_{t+1} after taking an action.
- Action* — Step A_t taken by the agent when in a particular state.

- Reward* — Feedback from the environment R_t .
- Policy* — Method to map an agent's state to actions.
- Q-Value* — Expected return $q_\pi(s, a)$ starting from state S_t , following policy π , taking action A_t , used to determine how good an action is.

The RL process is a cycle that begins with the observation of a current state, choosing an action, observing the received state, and updating the evaluation of its value function based on the action taken. After that, the next cycle begins.

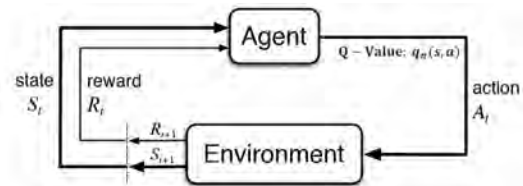


Figure 1. Reinforcement learning scheme

If the action's highest quality value (Q-value) is selected in every state, it results in the optimal policy. Q-learning finds an optimal policy in the sense of maximizing the expected value of the total reward over any successive steps, starting from the current state.

Policy Gradient is another popular method. It does not calculate the Q-value but instead uses a policy. The policy learns a mapping from every state to act, and its objective is to find which actions lead to higher rewards and increase their probability. The policy gradient observes the environment, acts in it and keeps updating its policy based on its rewards. After multiple iterations, the policy converges to a maximum value.

In section 4, our example that illustrates how to adopt Deep Reinforcement Learning (DRL), and Multi-agent Reinforcement Learning (MARL) on construction robots, will mainly focus on these two most used methods.

2.2 DRL to handle the dynamic environment

In situations with extensive state and action spaces, it is not feasible to learn the policy for each state and action pair. DRL allows agents to make decisions from high-dimensional and unstructured input data [9] using neural networks to represent policies. DRL combines both the technique of giving rewards based on actions from RL and the idea of using a neural network for learning feature representations from deep learning [10]. DRL is among the most promising algorithms when a predefined training dataset is required, which suits robotic manipulation and control tasks [11].

DRL is firstly used in video games and simulated control, which does not connect with the constraints of learning in natural environments. In 2015, Mnih et al. [7] used a Deep Q-Network (DQN) structure to create an agent that outperformed a professional player in 49 classic Atari games. With the development of algorithms, DRL has also demonstrated promises in enabling

physical robots to learn complex skills in the real world [12]. For example, DRL algorithms allow robots to learn tasks such as door opening [13] and to grasp and assemble objects using manipulators [14] with low-dimensional state observations.

With the exact logistics, by adopting the DRL algorithms to construction robots, robots can be trained in a dynamic environment. With simulation or execution in the real world, the robot will generate its dataset autonomously by accumulating its experience. And then, with iterations of training, the robot could find the optimal actions to take from the dataset generated when given an initial state.

2.2.1 MARL to solve multi-agent situations

As real-world problems have become increasingly complicated, there are many situations where a single DRL agent cannot cope. Traditional RL approaches such as Q-Learning or policy gradient are poorly suited for multi-agent environments. One issue is that each agent's policy changes as the training progresses, and the environment becomes non-stationary from any individual agent's perspective [15]. Learning among the agents sometimes causes changes in an agent's policy and can affect the optimal policy of other agents. The estimated potential rewards of action would be inaccurate, and therefore, good policies at a given point in the multi-agent setting could not remain so in the future. To solve this problem, based on the current literature, generally, two ways are used chiefly to solve this multi-agent framework.

The centralized learning approach assumes a joint model for all the agents' actions and observations [16]. A centralized policy maps the joint observation of all the agents to joint action. A significant drawback of this approach is that it is centralized in both training and execution and leads to an exponential growth in the observation and actions spaces with the number of agents.

The centralized paradigm is usually a beneficial choice for exchanging information between agents, and problems like non-stationarity, partial observability, and coordination can be diminished [17]. In 2016, Mnih et al. [18] introduced a lightweight approach based on actor-critic architecture by simultaneously training multiple agent-environment instances. As a result, many other works have started to extend the state-of-the-art single-agent asynchronous advantage actor-critic algorithm to enable multi-agent training to accomplish more sophisticated tasks. Elfakharany et al. [19] proposed another framework using the policy gradient method known as Proximal Policy Optimization (PPO) to allocate path planning globally in a shared environment among multiple robots.

The second method uses decentralized learning. In the decentralized execution paradigm, the agents make decisions independently according to their individual policies [16]. Multiple agents work together to learn a

homogeneous, distributed policy toward a common goal without explicitly interacting. One of the advantages of this approach is that it makes the learning of heterogeneous policies easier. It can be beneficial in domains where agents may need to take on specific roles to coordinate and receive a reward. For example, Sartoretti et al. [20] proposed an approach that relied on decentralized policy in an entirely visual system and succeeded in multi-agent-based brick construction. With a similar actor-critic architecture. However, training individual policies does not scale to many agents. Because the agents do not share the experience, this approach adds additional sample complexity to the RL task. Second, as the agents are learning and adjusting their policies, the change in the policies make the environment dynamics non-stationary. Stored experiences can be quickly meaningless due to the changing dynamics of other agents.

In addition to these two generally used methods, to speed up the learning procedure, another method can be used as a supplement by adding prior knowledge to guide the multi-agent learning process. Prior knowledge can dramatically help guide the learning process. These approaches significantly reduce the search space and, thus, speed up the learning process [21]. For example, in [22, 23], both the prior knowledge and prior rules are used to improve the DQN algorithm to solve the multi-robot path-planning problem.

3 RL with construction robots

RL has been widely adopted in different industries such as video games, marketing, public services, and manufacturing; however, there is no review of combining RL techniques with the construction industry. Aiming to inspire future research into solving the single-agent and multi-agent construction task execution problem, this section presents a survey discussing the possible challenges, opportunities, and research areas for the adoption of RL with construction robots based on previous studies and applications in relevant areas.

3.1 Challenges of construction robot training

3.1.1 Hard to build up the training environment

As an essential branch of machine learning, In comparison to supervised and unsupervised learning, few works have been published using the RL algorithm in the construction research field. The main reason for this may be that RL is a trial and error-based algorithm, which is hard for researchers to build up the training scheme in construction scenarios [24]. It is hard to initialize the agent(s) and the environment. To address that, RL research has generally been divided into two types of works [10]. The first type uses simulated control by testing virtual data in a simulated environment to verify the feasibility of robotic manipulation in the real world.

The second type is to ask the robot to sense the natural environment and learn directly in a real-world environment through trials and errors.

Due to the existence of many complex and random factors on the construction site, it is often impractical to develop an explicit construction environment based on the real-world situation. Construction sites are dynamic. Tracking changes and simulating real-world physics is hard in real-time. Besides, the simulation of robot models is also time-consuming. Generally, the simulation can only mimic the situation in small-scale scenarios, which is insufficient for large-scale construction operations. For real-world RL learning, errors or collision happens, and construction robotic systems are fragile and expensive. Therefore, it is not wise to test the robot directly on a dangerous construction site.

3.1.2 Construction tasks are hard to simulate

As the construction site evolves, it is hard to model the dynamics between states and actions and set up constraints and reward functions. In addition, construction tasks are unique for the construction application. No generalized building principles can be applied to all tasks or projects. Understanding the mapping between actions and states requires much expertise and prior knowledge to solve the physical equations and write down the scripts to control the robot, not to mention writing out the temporally and spatially coupled operational constraints or setting up specific goals for a characterized structure.

Moreover, it is challenging to solve large-scale construction work in real-time. When optimization is needed, these methods must compute all possible solutions entirely or partially and choose the best one, so the computation process is time-consuming when the solution space is vast [25].

3.1.3 Multi-workers working in the same environment with different roles

The construction industry is overwhelmed with resource planning, risk management, and logistic challenges, resulting in design defects, project delivery delays, cost overruns, and contractual disputes. Systematically putting all these factors into robotic learning is also a challenging job.

A construction job is a typical scenario that involves interaction among multiple agents, where emergent behavior and complexity arise from agents evolving together. For example, in collaborative construction activity, different agents have different tasks in the same environment to complete the activity. We need to define the working area for each individual and show them the path of their work routine to avoid physical collisions. Besides, because of the resource or material limitations, we need to wisely allocate the resources, arrange the sequence of operations for each agent, and build up a scheduling network. All these happen in a multi-agent

domain. Successfully scaling RL to environments with multiple agents is crucial to improving construction robots' eligibility.

3.2 Opportunities

3.2.1 Mega data generated and stored in the construction industry

The Building Information Model (BIM) is a crucial contribution to the construction industry. The BIM consists of a three-dimensional graphical reproduction of the building geometry and a related database in which all data, properties, and relations are stored [26]. BIM provides digital models for RL simulation and massive data stored to represent components on the job site. With this information fed to the robot, we could build up a real-time simulator for RL training, and a more accurate construction process could be simulated to represent the physical feedback. The virtual infrastructure represented in augmented reality (AR) will improve the user interaction with intelligent infrastructure for specific applications such as maintenance, training, and wayfinding. This could solve the first challenge that the construction industry lacks task planning simulators.

3.2.2 Development of RL algorithms

As stated in section 2, combined with developments in deep learning, DRL has emerged as an accelerator in related fields. DRL can help solve the problem of finding the optimal policy between state and actions through neural networks. Also, it can speed up the training process to make the simulation of construction tasks feasible. Besides, from the well-known success in single-agent DRL, such as Mnih et al. [7], we now witness a growing interest in its multi-agent extension to simulate and find the optimal strategy in a higher dimension. This can better represent the real situations on the job site and show the potential to combine different roles of robotic workers to collaborate just like human beings.

3.2.3 Human-robot interaction

As stated in Section 2, one can ease the training process by combining it with the proper operation in advance, just as the supervised learning, but somehow it will not influence the general structure of the DRL algorithm. The only contribution is to speed up the convergence and add the correct episode into the simulation process with a user-defined map. As [27] stated, there has been a rising demand to provide human feedback in the agent training process. Another more feasible method is proposed by creating a goal map, which integrates human strategies before the training process [28]. RL algorithms provide other choices in human-robot collaboration with only limited prior knowledge from human beings. New formats such as augmented reality (AR) or game-engine-based simulation can be further developed to teach the robot

how to initialize the learning process, prevent unsafe behavior, and define the goal or reward accordingly.

3.3 Research areas

After identifying some of the challenges and opportunities for RL-based construction robot applications, we put forward some possible research areas to guide future research on such innovative systems.

3.3.1 Simulator set up

First, as we identified in the Challenges and Opportunities sections, capturing the real-world environment into simulated augmented ones is a priority for RL-based construction robots to run the simulation. BIM-based digital twins could provide detailed information and the possibility to track real-time changes during construction. The communication between such models with the robotic system needs a platform to connect robotic training scripts with the BIM environment. A possible solution is integrating the ROS system with game engines such as Gazebo, Unreal or Unity to render the environment and imitate the physics from the real world. The combination between ROS and BIM (digital twin) allows the user-defined scripts for robotic control and visualization in real-time. This could be a critical topic for future research.

3.3.2 DRL algorithm advancing

Second, developing RL algorithms for robotic control is always a trend for advancing such systems. For example, algorithms that allow a mobile robot to reach target positions and navigate safely are open research fields [29]. Besides, many researchers are working on robotic manipulation, such as grasping and door-opening robots. In this way, there is great potential for us to develop the RL algorithms based on construction activities. Many researchers have succeeded in implementing the single repetitive work of construction robots, such as bricklaying and drilling. With the formula representing the construction dynamics, we can quickly adapt the RL algorithm structure to allow construction robots to fulfill such tasks in a more dynamic environment. The robot will then learn how to modify its actions to adapt to the changes in the environment, making the construction robot brighter and more easily implemented in the industry.

3.3.3 MARL algorithm advancing

Finally, instead of low dimensional observation on the single-agent repetitive tasks, a higher dimensional network for MARL will automate construction tasks in collaborative ways just like human beings. As [30] stated, a higher level of on-site automated robots should fulfill perception, mobility, and manipulation tasks. Based on this and the essential characteristics of construction work, we can categorize the MARL algorithms into two parts [19]. Multi-Robot Task Allocation (MRTA) and Multi-

Robot Path Planning (MRPP) as two separate steps, each with its own set of algorithms. The MRTA algorithm assigns each robot to a task to determine the optimal way to allocate tasks and resources, follow the constraints and manipulate specific tasks under a predefined schedule. The MRPP algorithm guides each robot through the environment towards the assigned goal position while avoiding both static and dynamic obstacles such as temporary structures and moving laborers or other robot agents.

4 Case study

This section presents a path planning mobile robot application developed based on the [31] to set up a foundation for RL-based construction robots in single-agent and multi-agent scenarios.

4.1 Simulator set up

We developed an experimental test for goal-oriented navigation and obstacle avoidance tasks using a TurtleBot3 Waffle Pi in the Gazebo simulation environment (Figure 2a) for behavior learning in autonomous agents.

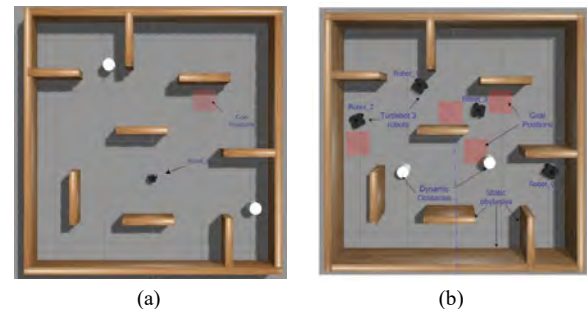


Figure 2. Gazebo environment for single-agent (a) and multi-agent robot environment path planning (b)

In the Gazebo simulation, to define the state of the robot, we need to observe the environment and describe the spatial relationship between them. Besides, to measure the right movement of the robot, we also need to sense the position of the virtual markers. Thus, with 24 sensors embedded in the robot model in the simulated environment, we define the state size for this single robot setting to be 26, 24 LDS (Laser Distance Sensor) values, along with the other two values: distances to goal and angle to goal.

Where LDS denotes the (24) values that the LIDAR sensor emits, the Distance represents the distance to the goal, and the Angle is the angle between the robot heading and vector to the goal.

4.2 DRL algorithm for single-agent robot navigation

The reinforcement training algorithms were run on a GPU with NVIDIA GeForce GTX 2060 (the CPU was an

8-Core Intel (R) Xeon (R) CPU E7) and implemented using the OpenAI Baselines package with DQN and DDPG. The DDPG and DQN algorithms were modified to process a weighted reward.

The goal is to train a Deep Q-Networks (DQN) agent to learn an optimal policy to navigate the robot from the initialized position to a goal position (user set up) with minimum effort.

Besides, we also adopted the Deep Deterministic Policy Gradient (DDPG) agent-based training algorithm, which allows continuous control of a robot. In our case we have linear velocity (0 to 0.22 m/s) and angular velocity (-1 to 1 rad/s) as outputs.

The robot has five actions that can act on depending on the type of state (Figure 3). The robot has a fixed linear velocity of 0.15 m/s, which determines the angular velocity. The linear velocity has discrete actions, as shown in Figure 3a. The angular velocity is determined by the state and the linear velocity, as shown in Table 1. The corresponding values are shown in Figure 3b. Table 1 also shows the reward functions.

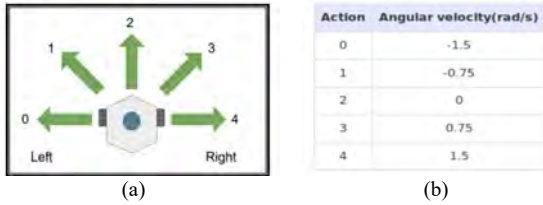


Figure 3. Direction of the different actions (a) and corresponding angular velocity (b)

Table 1. Reward function equation and parameters

Reward function	Parameter
$\theta = \frac{\pi}{2} + act * \frac{\pi}{8} + \varphi$	θ : Angle from goal φ : Yaw of robot
$R_\theta = 5 * 1^{-\theta}$	R_θ : Reward from angle
$R_d = 2^{\frac{D_c}{D_g}}$	R_d : Reward from distance D_c : Current distance from goal D_g : Absolute distance from goal
$R = R_\theta * R_d$	R : Total reward $R_\theta \geq 0$ else < 0

4.3 MARL algorithm for robot navigation

For the MARL algorithm, we used the DQN network to train four TurtleBot3 robots in the same working environment to achieve different goals (Figure 2b). We used the centralized learning or decentralized execution paradigm as discussed in the literature, in which each robot has a copy of the policy network, and each robot collects its data (O_i^t, a_i^t, r_i^t) from the environment. Each robot (i) at time step (t) receives an observation (O_i^t) and calculates the output action (a_i^t) that drives the robot from the start position towards the goal position. The observation for each robot is composed of several parts:

The first part is the relative positions of all the goals in the robot's local polar coordinates. The second part is the relative positions of all the goals in the other moving robots' local polar coordinates. Last but not least is the laser scanner measurements. The reward function r_i^t is designed to ensure each robot move towards the unique zone set up separately. It penalizes getting near obstacles, colliding with other robots, or reaching the wrong goal position.

After each episode, it sends data rollouts to a centralized copy of the policy. The gradients are then calculated on the centralized policy, and the centralized policy is updated. After that, each robot receives a copy of the updated policy weights to start collecting a new batch of data. The episode ends when either the robots have a collision, when all the robots reach all the goal positions, or when the episode duration is exhausted.

5 Results

5.1 Single-agent DRL

To compare DDPG and DQN, it was necessary to define metrics. The reward was chosen as the primary indicator. Simulation results in a world of fixed obstacles and random start and end positions are shown in Figure 4. DQN achieved the average target score somewhere after 200 episodes, with each episode consisting of a maximum of 120 time steps (Figure 4a). Training a DDPG generally tends to take more time compared to DQN. In this case, it took about 600 episodes to achieve the average score (Figure 4b).

One of the reasons is that the number of parameters to deal with in DDPG is much higher than in DQN and requires more computation resources. DQN is a value-based learning method, whereas DDPG is an actor-critic method. With DDPG, we have to fine-tune not one but two neural network models. Moreover, since the performance of the actor model strongly depends on the critic's performance, they both must have proper stable growth, which is quite challenging to assure.

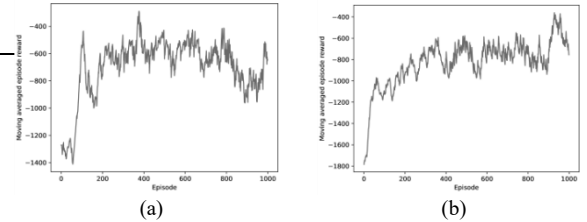


Figure 4. DQN reward (Temporal difference) (a) and DDPG reward (Actor-critic) (b)

The second reason is that DQN is training under discrete action space, with less uncertainty than the DDPG ones and continuous action selections. Although DQN was successful in more significant dimensional problems, such as the Atari game, the action space is still discrete. However, for many tasks of interest, especially

the physical control tasks of robots, the action space is mostly continuous.

5.2 Multi-agent DRL (MARL)

5.2.1 Decentralize learning

Figure 5b shows that the cumulative reward increases over time and reaches a stabilized value, which means the decentralized training succeeds in collaborative path planning. The episode length, Figure 5a, shows the length of each episode in the environment after all agents reach the goals. The agents struggle to complete the tasks within the episode limit for the initial few episodes, but we observe a drop in the episode length as the agents learn. Finally, it reaches a platform, which indicates that it takes 20 s for all the robots to reach the assigned goals in each episode. For decentralized learning, it is easy to get to the desired reward with limited time steps. The steps to take are gigantic, requiring almost 4 million episodes of training periods; this is because multi-agent collaboration leads to an exponential growth in the observation and actions spaces with the number of agents.

Nevertheless, the results are faster converge than the centralized training. Because all the agents do not need to share the experience, which eliminates additional sample complexity to set up the neural networks, it will take shorter to find the optimal strategies. Besides, the final result seems robust with a static platform. Less interface leads to a less dynamic environment, which simplifies the whole process.

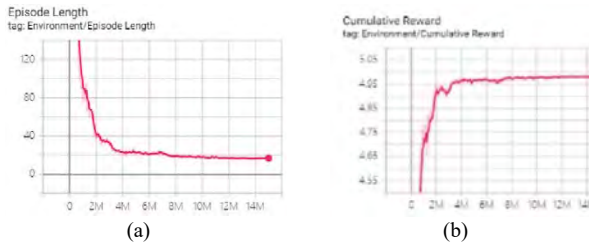


Figure 5. Decentralized training result (DQN) Episode Length (a) and Cumulative Reward (b)

5.2.2 Centralized learning

Unlike decentralized learning, centralized learning could not reach the expected platform after a long training period in the first 2 million episodes (Figure 6a). The robot could not find the desired goals because the robot needed to find the desired joint observations in the same joint actions. This is because agents do share the experience, add additional sample complexity, and have significant computational burdens to set up the neural networks, which means it will take longer to initialize the simulation and find the optimal strategies. Thus, it takes longer for the robot to initialize the mapping between states and action, which takes longer for the robot to find the optimal trajectory in the beginning. As shown in Figure 6a, it takes almost 1,000 s for the first 3 million episodes to start. We can also see that the result from

centralized learning is not as good as the one in decentralized learning. After a long training period, 5 million episodes (Figure 6b), the final result still does not converge. This is because some mapping between state and actions is not correctly build-up and the interface between different agents makes it harder to map the states to the actions. The change in the policies makes the environment's dynamics change, which will cause deviations in finding the optimal strategy and may take a longer time for the robot to adjust to reach its goals.

The summary of the results (metrics used for comparison purposes) for each scenario is shown in Table 3. In general, for a single-agent reinforcement learning task, it is accessible for the robot to find the optimal solution in a short period. However, as more agents come into the same environment, the variables such as states and actions lead to explosive growth, which results in significant computational burdens and sample complexities. One of the robust ways to solve this problem is to ask the robots to share their experiences and get feedback in the same reference domain. If not all information is handled and transformed, the robot could also learn, but it is pretty hard to find the optimal strategy.

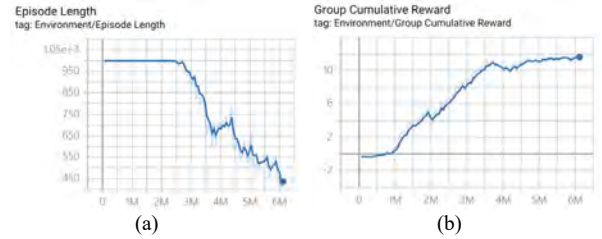


Figure 6. Centralized training result (DQN) Episode Length (a) and Cumulative Reward (b)

Table 3: Results for the metrics from each scenario.

Metric	DRL (DQN)	MARL (Decentralized)	MARL (Centralized)
Episode to converge	400	4e6	5e6 (Not converge)
Episode length	20 s	20 s	450 (Not converge)

6 Conclusions and outlook

This study gives an overview of RL and DRL algorithms, discusses the challenges, opportunities, and research areas of the integration of RL with construction robot applications, and implements a part of them in a simple case study to illustrate how they can be used in construction robots applications for path planning. This study shows that, as expected, for simple tasks such as navigation, single-agent RL algorithms (e.g., DQN and DDPG) effectively solve the problem, with few iterations and durations to train. However, the computational resource (training period, episode length) grows

exponentially when more agents come into the same environment (more realistic for construction tasks). Using a decentralized learning algorithm is feasible to train multi-agents in those cases. The only problem is efficiency, as it takes a long training time. Still, the results are robust because this method provides a suitable evolving environment without too many changes to deal with, unlike centralized training. Therefore, it can be said that for the case study investigated, the decentralized learning approach could be better suited for complex construction task simulations.

Ongoing work by the authors includes the investigation of other RL algorithms for single task manipulation and execution, such as bricklaying, painting, door installation, etc., to prove the eligibility of such a combination of construction robot applications with RL. Our future goals include: 1) Adopting and developing current DRL and MARL algorithms for construction robot applications. 2) Adopting the multi-robot framework to solve complex tasks (e.g., laying bricks to build up a user-defined structure) through the cooperation of individual agents. 3) Assign different characters that allow different robots to work in different roles, such as manipulator, inspector, material delivery, etc. 4) Conducting the simulation in game engines such as Gazebo, Unreal, or Unity, and setting up communication among different platforms by using BIM models, this set up a great foundation for future verification in a real environment. 5) Benchmarking for efficiency of other algorithms to see which one is the best fit for the construction robot applications.

References

- [1] Recht B., A tour of reinforcement learning: the view from continuous control. *Ann Rev Control Robot Autonom Syst* 2019; 6: 253–279.
- [2] Mahadevan S. and Theodorou G. Optimizing production manufacturing using reinforcement learning. In: *FLAIRS Conference*, 18 May 1998, pp. 372–377. AAAI Press. 3.
- [3] Silver D., Hubert T., Schrittwieser J., A general reinforcement learning algorithm that master's chess, shogi, and go through self-play. *Science* 2018; 362(6419): 1140–1144. 4.
- [4] Kober J., Bagnell J.A., and Peters J. Reinforcement learning in robotics: a survey. *Int J Robot Res* 2013; 32(11): 1238–1274. 5.
- [5] Isele D., Rahimi R., Cosgun A., Navigating occluded intersections with autonomous vehicles using deep reinforcement learning. *ICRA*, Brisbane, QLD, Australia, 21–25 May 2018, pp. 2034–2039.
- [6] Yayin X., Ying Z., Przemyslaw S., Lieyun D., Machine learning in construction: From shallow to deep learning, *Developments in the Built Environment*, Volume 6, 2021, 100045.
- [7] Mnih, V., Kavukcuoglu, K., Silver, D. et al. Human-level control through deep reinforcement learning. *Nature* 518, 529–533 (2015).
- [8] Shopov V., and Markova V., “A Study of the Impact of Evolutionary Strategies on Performance of Reinforcement Learning Autonomous Agents”, *ICAS 2018*, p.56-60, 2018.
- [9] Dresch-Langley B., Ekseth O.K., Fesl J., Gohshi S., Kurz M., Sehring H.W., Occam's Razor for Big Data? On detecting quality in large unstructured datasets. *Appl. Sci.* 2019, 9, 3065.
- [10] Ibarz J., Tan J., Finn C., Kalakrishnan M., Pastor P., Levine S., How to train your robot with deep reinforcement learning: lessons we have learned. *The International Journal of Robotics Research*. 2021; 40(4-5):698-721.
- [11] Rongrong L., Florent N., Philippe Z., Michel D.M., Birgitta D., (2021). Deep Reinforcement Learning for the Control of Robotic Manipulation: A Focussed Mini-Review. *Robotics*. 10. 1-13.
- [12] Zhang T., Mo H.. Reinforcement learning for robot research: A comprehensive review and open issues. *Journal of Advanced Robotic Systems*. May 2021.
- [13] Gu S., Holly E., Lillicrap, T., Levine S. (2017) Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates. *ICRA*, 3389–3396.
- [14] Haarnoja T., Pong V., Zhou A., Dalal M., Abbeel P., Levine S., Composable deep reinforcement learning for robotic manipulation. *ICRA* (2018a).
- [15] Lowe R., Wu Y., Tamar A., Multi-agent actor-critic for mixed cooperative-competitive environments. *arXiv preprint arXiv:1706.02275*, 2017.
- [16] Lyu, X et al., Contrasting Centralized and Decentralized Critics in Multi-Agent Reinforcement Learning, 20th *AAMAS*. 2021
- [17] Gronauer S., Multi-agent deep reinforcement learning: a survey. *Artif Intell Rev* (2021).
- [18] Mnih V., Badia A. P., Mirza M., Graves A., et al., Asynchronous methods for deep reinforcement learning. In *International Conference on Machine Learning*, pages 1928–1937, 2016.
- [19] Elfakharany A., Towards Multi-Robot Task Allocation and Navigation using Deep Reinforcement Learning 2020 *J. Phys.: Conf. Ser.* 1447 012045
- [20] Sartoretti G., Wu Y., Paivine W., Distributed reinforcement learning for multi-robot decentralized collective construction. *Springer*, Cham, 2019: 35-49.
- [21] Kober J., Bagnell J. A., and J. Peters J., 2013. Reinforcement learning in robotics: A survey. *Int. J. Rob. Res.* 32, 11, 2013, 1238–1274.
- [22] Yang Y., Juntao L., Lingling P., (2020), multi-robot path planning based on a deep reinforcement learning DQN algorithm. *CAA Trans. Intell. Technol.*, 5: 177
- [23] Li B., Liang H., Multi-Robot Path Planning Method Based on Prior Knowledge and Q-learning Algorithms. 2020 *J. Phys.: Conf. Ser.* 1624 042008
- [24] Chung, H. et al. Brick-by-Brick: Combinatorial Construction with Deep Reinforcement Learning, *35th NeurIPS 2021*
- [25] Yu L. et al., "Multi-agent deep reinforcement learning for HVAC control in commercial buildings", *IEEE Trans. Smart Grid*, vol. 12, no. 1, 407–419, Jan. 2021.
- [26] Will S., (2019). Deep Reinforcement Learning Algorithms in Intelligent Infrastructure. *Infrastructures*. 4. 52.
- [27] Long P., Fan T., Liao X., Liu W., Zhang H., and Pan J., “Towards Optimally Decentralized Multirobot Collision Avoidance via Deep Reinforcement Learning,” in *ICRA*, 2018.
- [28] Zhou B., Khosla A., Lapedriza A., Oliva A., and Torralba A., Learning deep features for discriminative localization. In *CVPR*, pages 2921–2929, 2016.
- [29] Xu X. and García de Soto B., "On-site Autonomous Construction Robots: A review of Research Areas, Technologies, and Suggestions for Advancement," 37th *ISARC*, Kitakyushu, Japan, 2020.
- [30] Christiano P. F., Leike J., et al.. Deep reinforcement learning from human preferences. *Advances in Neural Information Processing Systems*, pages 4302, 2017.
- [31] ROS Packages for TurtleBot3 Machine Learning, https://github.com/ROBOTIS-GIT/turtlebot3_machine_learning.git

Constrained Control Scheme for the Manipulation of Heavy Pre-fabricated Elements with Lightweight Robotic Arm

M. Ambrosino¹, F. Boucher², P. Mengeot², and E. Garone¹

¹Service d'Automatique et d'Analyse des Systèmes, Université Libre de Bruxelles, Brussels, Belgium.

²NV BESIX SA, Brussels, Belgium.

Michele.Ambrosino@ulb.ac.be, fabian.boucher@besix.com, pierre.mengeot@besix.com, egarone@ulb.ac.be

Abstract -

Building activities involving heavy suspended elements are one of the construction activities with the highest level of danger. Typically, during these activities, one or two masons work in conjunction with a machine such as a crane or a lifting machine. Several robotics solutions have been proposed to replace the masons during these hazard operations. In this work, we propose to use a lightweight robotic arm to handle and place a heavy suspended object ensuring a high level of precision during the planned operations. To control the resulting robotic system, we propose a constrained control scheme based on Explicit Reference Governor (ERG) theory, an add-on unit that modifies the applied reference in such a way that the trajectory of the system always fulfills the constraints of the system. The simulation results show the efficiency of our approach by testing it against other solutions proposed so far.

Keywords -

Robotics; Cooperative Robotic Systems; Building Activity; Robots for Construction; Constrained Control Scheme.

1 Introduction

Various construction activities are based on the handling and positioning of placement of prefabricated heavy elements, such as renovation of facades [1], or construction of walls [2]. These types of activities are carried out through the use of machines, such as cranes or lifting mechanisms, in which the object to be positioned is suspended through a cable. Then, the machine moves the object near its final position and a human being (e.g. a bricklayer), manually finishes the operation by guiding the suspended object for the last few centimeters. The presence of the bricklayer is necessary to ensure a high level of precision in the operations. However, these operations involve heavy suspended objects, which represent possible causes of accidents for the bricklayer (sometimes fatal) [3].

Several approaches have been proposed so far in the literature with the aim of using a robotic solution for this type of construction activity [4, 5]. For a complete overview

of the drawbacks and the benefits of the proposed robotic solutions, please refer to [6].

Among the various solutions proposed so far, in this work we focus on the one discussed in [6, 7] in order to improve the preliminary proposed control strategies. The authors show how the cooperation between a crane and a industrial robotic arm is able to perform the positioning of heavy blocks in order to guarantee a high level of precision during masonry activities. The control laws proposed in these two researches are based on an 'ad-hoc' trajectory for the positioning operations, and as highlighted by the authors themselves these control schemes should be reinforced with a governing unit that is capable of managing the constraints that are present in this type of operations. In particular, the main constraints that must to be considered concern the torque required to the actuators of the robotic arm used during operations. In fact, the robot having to handle a payload much heavier than the maximum permissible payload could find itself in an overload situation which would affect the robot itself.

In this paper we propose the design of a lightweight constrained control scheme for a robotic arm that unlike those proposed so far in the literature: *i*) it does not require any offline pre-evaluation of a feasible trajectory; *ii*) it does not solve any online optimization problems. Moreover, the proposed control scheme always fulfills the constraints of the system.

A general purpose control solution that is able to handle constraints in real-time is Model Predictive Control (MPC) [8]. MPC provides an optimal control strategy through the solution of an optimization problem at each sampling time. However, this kind of control schemes have a high computational cost with respect to simpler control schemes, especially when applied to nonlinear systems, therefore, in practice, its application is still limited [9]. A promising alternative to MPC is to consider a first inner controller to stabilize the system, and then, 'augment' the system with constraint-handling capabilities. This idea is the core of the Reference Governor (RG) schemes [10]. The RG is an add-on unit that filters the desired reference in such a way that the trajectory of the system always fulfills the

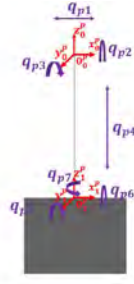


Figure 1. Suspended element configuration.

constraints. However, as the MPC scheme, RGs rely on on-line optimization as well. To overcome this problem, in the past few years, a novel constrained control scheme that does not require on-line optimization was proposed, called the Explicit Reference Governor (ERG) [11, 12].

The aim of this paper is to implement a trajectory-based ERG for the handling activity of a heavy suspended elements with a lightweight robotic arm. In the first part of this paper, a mathematical model of system under consideration will be derived. Next, the constrained control scheme used in this work is analyzed. At the end of this work the performances of the proposed control scheme are shown through simulations and compared to those proposed in the previous works in the literature.

2 Modeling

This section provides the general dynamic model of the system under consideration by systematically combining two types of model: a *suspended object*, and a *robotic arm*. The dynamic model of the two parts will be discussed separately. Then, the dynamic model of the entire system will be derived. More details about this kind of modeling can be found in [7].

2.1 Dynamic Model

The dynamic of the suspended element can be treated as that of a 7-DoF pendulum, see Fig.1. In particular, the configuration of the suspended object can be described by seven variables, $q_p \in \mathbb{R}^7$, where $q_p = [q_{p1}, q_{p2}, q_{p3}, q_{p4}, q_{p5}, q_{p6}, q_{p7}]^T$. Where q_{p1} is the displacement along the x-direction, q_{p4} is the length of the rope, q_{p2} and q_{p3} are the radial sway and the tangential pendulation respectively, and q_{p5}, q_{p6}, q_{p7} are the orientations of the block w.r.t the cable. It is worth noting that this kind of mechanical system is an underactuated system, having fewer independent actuators than system degrees of freedom (DoFs). In our work we consider the possibility of moving the object along the x axis and the z



Figure 2. Robotic arm configuration.



Figure 3. Frame configuration

axis based on the fact that in real scenarios machines such as cranes can perform these two movements.

The robotic arm used in this paper is a 7-DoF robotic, and more precisely a KUKA IIWA14 R820 [13]. The robot configuration is described by the joint variables vector $q_r \in \mathbb{R}^7$, with $q_r = [q_{r1}, q_{r2}, q_{r3}, q_{r4}, q_{r5}, q_{r6}, q_{r7}]^T$, see Fig.2.

As shown in [6], the dynamic model of the whole system can be obtained introducing a set of closed-chain constraints that come when the robot has already grabbed the suspended object (see Fig.3). Therefore, the dynamic model of the constrained mechanical system can be compactly rewritten considering as state vector for the entire system the vector $q \in \mathbb{R}^{14}$, $q = [q_r, q_p]^T$, and using the equations:

$$B(q)\ddot{q} + C(q, \dot{q})\dot{q} + F\dot{q} + G(q) = u + A(q)^T \lambda - J(q)^T h_e \quad (1)$$

$$\text{s.t.} \quad A(q)\dot{q} = 0, \quad (2)$$

where $A(q) \in \mathbb{R}^{6 \times 14}$, is the so-called *Jacobian* of the constraints, the matrices $B(q) \in \mathbb{R}^{14 \times 14}$, $C(q, \dot{q}) \in \mathbb{R}^{14 \times 14}$, $F(q) \in \mathbb{R}^{14 \times 14}$, and $G(q) \in \mathbb{R}^{14}$ represent the inertia, centripetal-Coriolis, friction matrix, and gravity term, respectively. Moreover, $\lambda \in \mathbb{R}^6$ is the vector of the *Lagrange*

multipliers, $h_e \in \mathbb{R}^6$ represents the vector of the forces generated by contact with the environment, $J(q) \in \mathbb{R}^{6 \times 14}$ is the manipulator geometric Jacobian [14], and $u \in \mathbb{R}^{14}$ is the vector of the control input of the system. As we highlighted before, the suspended object is modelled as an underactuated system, therefore the control input u is

$$u = [\tau_r, \tau_x, 0, 0, \tau_l, 0, 0, 0]^T, \quad (3)$$

where $\tau_r \in \mathbb{R}^7$ is the vector of the robot control input, τ_x is the object control input for the displacement along the x axis, and τ_l is the object control input for the displacement along the z axis. According to [6], the model (1) can be rewritten eliminating the Lagrangian multipliers as follows:

$$B(q)\ddot{q} = (I - A^T(q)A^*(q))(u - J(q)^T h_e - m(q, \dot{q})) - B(q)A^*(q)\dot{A}(q)\dot{q}, \quad (4)$$

where, $m(q, \dot{q}) = C(q, \dot{q})\dot{q} + F\dot{q} + G(q)$, and $A^*(q)$ is the inertia-weighted pseudo-inverse of the constraint Jacobian A defined as

$$A^*(q) = B^{-1}(q)A^T(q)(A(q)B^{-1}(q)A^T(q))^{-1}. \quad (5)$$

2.2 Control objective

The main goal of this paper is to propose a constrained control scheme for the system (4). This scheme must allow the system to follow a piece-wise constant reference $r(t) \in \mathbb{R}^9$ while ensuring that

- i. for any piece-wise continuous reference $r(t)$, the control law guarantees constraint satisfaction of the state constraints;
- ii. safe cooperation between the two sub-units is ensured, *i.e.* the robot will never be overloaded and the robot actuators torque limits are never violated;
- iii. if the reference r is constant and steady-state admissible, the closed-loop system satisfies $\lim_{t \rightarrow \infty} q(t) = r(t)$.

In particular, in the development of our control law, we consider the following constraints:

- joint range constraints:

$$\begin{cases} q_{min,r,i} \leq q_{r,i} \leq q_{max,r,i} \\ q_{min,x} \leq q_x \leq q_{max,x} \\ q_{min,l} \leq q_l \leq q_{max,l} \end{cases} \quad (6)$$

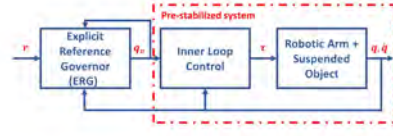


Figure 4. Control Scheme

for some lower and upper joint range limits $q_{min,r,i}, q_{max,r,i} \in \mathbb{R}, i = 1, \dots, 7$ for the robotic arm, and $q_{min,x}, q_{max,x}$ and $q_{min,l}, q_{max,l}$ for the actuated joint of the suspended object.

- Maximum joint velocity constraints:

$$\begin{cases} \dot{q}_{min,r,i} \leq \dot{q}_{r,i} \leq \dot{q}_{max,r,i} \\ \dot{q}_{min,x} \leq \dot{q}_x \leq \dot{q}_{max,x} \\ \dot{q}_{min,l} \leq \dot{q}_l \leq \dot{q}_{max,l} \end{cases} \quad (7)$$

for some lower and upper joint velocity limits $\dot{q}_{min,r,i}, \dot{q}_{max,r,i} \in \mathbb{R}, i = 1, \dots, 7$ for the robotic arm, and $\dot{q}_{min,x}, \dot{q}_{max,x}$ and $\dot{q}_{min,l}, \dot{q}_{max,l}$ for the actuated joint of the suspended object.

- Actuator saturation on the input:

$$\begin{cases} \tau_{min,r,i} \leq \tau_{r,i} \leq \tau_{max,r,i} \\ \tau_{min,x} \leq \tau_x \leq \tau_{max,x} \\ \tau_{min,l} \leq \tau_l \leq \tau_{max,l} \end{cases} \quad (8)$$

for some lower and upper actuator limits $\tau_{min,r,i}, \tau_{max,r,i} \in \mathbb{R}, i = 1, \dots, 7$ for the robotic arm, and $\tau_{min,x}, \tau_{max,x}$ and $\tau_{min,l}, \tau_{max,l}$ for the actuated joint of the suspended object.

3 Control Scheme

In this section, we describe the constraint control scheme used in our work. This scheme decouples the stabilization of the system and the satisfaction of the constraints: an *internal controller* fulfills the former task, whereas a *governing unit* modifies the reference fed to the system in such a way that constraints are fulfilled at all times. An illustration of this control architecture can be found in Fig. 4.

3.1 Internal Control Layer

As previously mentioned, the goal of the inner control loop is to stabilize the system without taking into account system constraints. This task can be performed

using a common control scheme based on a Proportional-Derivative (PD) action with gravity compensation. In particular,

- for the robotic arm:

$$\tau_r = K_{Pr}\tilde{q}_r - K_{Dr}\dot{q}_r + g_r(q_r), \quad (9)$$

- for the displacement along the x-axes:

$$\tau_x = K_{Px}\tilde{q}_x - K_{Dx}\dot{q}_x, \quad (10)$$

- to stabilize the length of the cable:

$$\tau_l = K_{Pl}\tilde{q}_l - K_{Dl}\dot{q}_l + g_l(q_p), \quad (11)$$

where $\tilde{q}_r = q_{vr} - q_r$, $\tilde{q}_x = q_{vx} - q_{p1}$, and $\tilde{q}_l = q_{vl} - q_{p4}$, with $q_v = [q_{vr}, q_{vx}, q_{vl}]^T \in \mathbb{R}^9$ being the vector of the applied reference associated to the commanded reference r . $K_{Pr} \in \mathbb{R}^{7 \times 7}$ and $K_{Dr} \in \mathbb{R}^{7 \times 7}$ are positive definite diagonal matrices, instead K_{Px} , K_{Dx} , K_{Pl} , and K_{Dl} are positive scalar gains. Instead, $g_r(q_r) \in \mathbb{R}^7$, and $g_l(q_p)$ represent the gravity compensation relating to the robotic arm and the cable length. It is possible to prove that the control laws (9)-(11) is able to pre-stabilize the system (4).

Lemma 1. Consider the system (4), the controller (9)-(11) makes every equilibrium configuration $q = [q_{vr}, q_{vx}, 0, 0, q_{vl}, 0, 0, 0, 0]^T$, $\dot{q} = 0 \in \mathbb{R}^{14}$ Globally Asymptotically Stable (GAS).

Proof. Please refer to [7].

3.2 Governing Unit: Trajectory-Based ERG

While the controller (9)-(10) ensure the asymptotically stability of the system, it is unable to manage the constraints (6)-(8). Therefore, in this paper we propose to augment the first inner control loop with a Governing Unit to be able to deal with the system constraints. In particular, the proposed governing unit belongs to the theory of the Explicit Reference Governor (ERG) [11]. In particular, rather than pre-computing a suitable trajectory $q_v(t)$, the ERG achieves these objectives by continuously manipulating the derivative of the applied reference as the product of two terms: the Navigation Field (NF) $\rho(q_v, r)$, and the Dynamic Safety Margin (DSM) $\Delta(q, \dot{q}, q_v)$. In particular, $\rho(q_v, r)$ is a vector field that generates the desired steady-state admissible path, and $\Delta(q, \dot{q}, q_v)$ is a scalar that quantifies the distance between the predicted transient dynamics of the pre-stabilized system and the constraint boundaries if the current $q_v(t)$ were to remain constant. More formally,

$$\dot{q}_v = \rho(q_v, r)\Delta(q, \dot{q}, q_v). \quad (12)$$

Since the set of admissible references is convex, the NF can be designed using an attraction and repulsion field [15],

$$\rho(q_v, r) = \rho^{\text{att}} + \rho^{\text{rep}}, \quad (13)$$

where the attraction field is

$$\rho^{\text{att}}(q_v, r) = \frac{r - q_v}{\max(\|r - q_v\|, \eta)}, \quad (14)$$

and where $\eta > 0$ is a smoothing parameter ensuring ρ^{att} is a class C^1 function.

The repulsion field for the problem at hand can be expressed as the sum of two repulsion field, one for the position constraints and the second one for torque constraints, *i.e.*

$$\rho^{\text{rep}} = \rho_{pos}^{\text{rep}} + \rho_{\tau}^{\text{rep}}, \quad (15)$$

where

$$\rho_{pos}^{\text{rep}} = [\rho_{1,pos}^{\text{rep}}, \dots, \rho_{9,pos}^{\text{rep}}]^T, \quad (16)$$

with

$$\rho_{i,pos}^{\text{rep}} = \max\left(\frac{\zeta - |q_{v,i} - q_{min,i}|}{\zeta - \delta}, 0\right) - \max\left(\frac{\zeta - |q_{v,i} - q_{max,i}|}{\zeta - \delta}, 0\right), \quad (17)$$

and

$$\rho_{\tau}^{\text{rep}} = [\rho_{1,\tau}^{\text{rep}}, \dots, \rho_{9,\tau}^{\text{rep}}]^T, \quad (18)$$

with

$$\rho_{i,\tau}^{\text{rep}} = \max\left(\frac{\zeta - |\tau_i - \tau_{min,i}|}{\zeta - \delta}, 0\right) - \max\left(\frac{\zeta - |\tau_i - \tau_{max,i}|}{\zeta - \delta}, 0\right), \quad (19)$$

where $\delta > 0$ is the static safety margin of all the joint angles and $\zeta > \delta$ is the influence margin.

Among the different design tools for generating a suitable Dynamic Safety Margin (DSM), in this work we propose to use the trajectory based approach. The idea behind the trajectory-based ERG is to compute the trajectories of the pre-stabilized system $\hat{q}(t|q, \dot{q}, q_v) = \hat{q}_{t'=t}$, under the assumption that the current applied reference $q_v(t)$ is kept constant. The DSM is then characterized as the minimum distance of the trajectory to each of the constraints. The

trajectory can be obtained by simulating the forward dynamics of the system (4). We first initialize at the current time t the states that will be predicted, \hat{q} and \dot{q} ,

$$\begin{cases} \hat{q}_{t'=t} = \dot{q}(t), \\ \hat{q}_{t'=t} = q(t), \end{cases} \quad (20)$$

Then we simulate the system dynamics using numerical integration as follows

$$\begin{cases} \hat{\tau}_{r,t'+dt} = K_{Pr}(q_{vr} - \hat{q}_{r,t'}) - K_{Dr}\hat{q}_{r,t'} + g_r(\hat{q}_{r,t'}) \\ \hat{\tau}_{x,t'+dt} = K_{Px}(q_{vx} - \hat{q}_{x,t'}) - K_{Dx}\hat{q}_{x,t'} \\ \hat{\tau}_{l,t'+dt} = K_{Pl}(q_{vl} - \hat{q}_{l,t'}) - K_{Dl}\hat{q}_{l,t'} + g_l(\hat{q}_{l,t'}) \\ \ddot{q}_{t'+dt} = B(\hat{q}_t')^{-1} \left(I - A^T(\hat{q}_t')A^{*T}(\hat{q}_t') \right) \\ \quad \left(\hat{u} - J(\hat{q}_t')^T h_e - m(\hat{q}_t', \dot{\hat{q}}_t') \right) - B(\hat{q}_t')A^*(\hat{q}_t')\dot{A}(\hat{q}_t')\dot{\hat{q}}_t' \\ \hat{q}_{t'+dt} = \hat{q}_{t'} + \dot{\hat{q}}_{t'+dt}dt, \\ \dot{\hat{q}}_{t'+dt} = \dot{\hat{q}}_{t'} + \ddot{\hat{q}}_{t'+dt}dt. \end{cases}$$

Once the trajectory $\hat{q}(t', \dot{q}, q_v)$ has been simulated over a sufficiently long horizon T , it is possible to obtain the distances to the position and velocity constraints as

$$\begin{aligned} \Delta^{pos} &= \min_{t' \in [t, t+T]} \left\{ \min_{i \in \{1, \dots, n\}} \left\{ \hat{q}_{t',i} - q_{min,i}, q_{max,i} - \hat{q}_{t',i} \right\} \right\}, \\ \Delta^{vel} &= \min_{t' \in [t, t+T]} \left\{ \min_{i \in \{1, \dots, n\}} \left\{ \dot{\hat{q}}_{t',i} - \dot{q}_{min,i}, \dot{q}_{max,i} - \dot{\hat{q}}_{t',i} \right\} \right\}, \\ \Delta^\tau &= \min_{t' \in [t, t+T]} \left\{ \min_{i \in \{1, \dots, n\}} \left\{ \hat{\tau}_{t',i} - \tau_{min,i}, \tau_{max,i} - \hat{\tau}_{t',i} \right\} \right\}. \end{aligned} \quad (21)$$

respectively.

The overall trajectory-based DSM, Δ_T , can be computed as

$$\Delta_T(q, \dot{q}, q_v) = \min \{ k^{pos} \Delta^{pos}, k^{vel} \Delta^{vel}, k^\tau \Delta^\tau \}, \quad (22)$$

with positive real gains k^{pos} , k^{vel} , k^τ . These design parameters can be used to scale arbitrarily the impact of the distances in the final computation of the DSM.

To ensure that constraints are never violated, in line of principle, the predict horizon should be extended to infinite. To do this, it is sufficient to ensure that, from $t+T$ onward, the closed-loop system dynamics will not exceed a terminal energy constraint. Therefore, we can considered for the system (4), the following energy function [7]:

$$\begin{aligned} V(t) &= \frac{1}{2} \dot{q}^T B(q) \dot{q} \\ &+ mg(q_{pp4} - q_{pp4} C_{qp2} C_{qp3} + l_p - l_p C_{qp5} C_{qp6}) \\ &+ \frac{1}{2} \tilde{q}_r^T K_{Pr} \tilde{q}_r + \frac{1}{2} K_{Pl} \tilde{q}_l^2 + \frac{1}{2} K_{Px} \tilde{q}_x^2. \end{aligned} \quad (23)$$

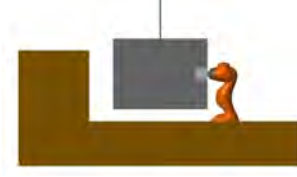


Figure 5. Simulation scenario.

Then, we can compute the terminal energy of the system evaluating (23) at the end of the predicted trajectory, and we can add this cost in the evaluation of the overall Dynamic Safety margin considering

$$\Delta_V(q, \dot{q}, q_v) = k^{E_{term}} (E_{term} - V(\hat{q}_{t'}, \dot{\hat{q}}_{t'}, q_v)), \quad (24)$$

where $k^{E_{term}}$ is a positive real arbitrary scaling factor and $E_{term} > 0$ is a suitable threshold value.

Gathering all of the above, the overall DSM for an infinite time horizon is,

$$\Delta(q, \dot{q}, q_v) = \max \{ \min(\Delta_T, \Delta_V), 0 \}. \quad (25)$$

As can be seen in (25), in case the next applied reference violates the system constraints, the DSM will be equal to 0. This means that the internal control loop will be fed with the previous applied reference which guarantees satisfaction constraints.

4 Simulation Results

The aim of this section is to show that: *i*) the proposed control scheme is able to perform correct and safe manipulation of a heavy suspended element; *ii*) the proposed controller works better than the preliminary control scheme proposed in [7].

The simulation scenario we considered is the following: the robotic arm has to move the suspended object with a weight of **100kg**, along the x-axis of 0.2cm and along the z-axis of 0.1cm, in order to place the object in its final position. The robotic arm used in the simulations, a KUKA IIWA14 R820, can handle a payload of up to **14kg**, [16]. However, due to the fact that the mass of the block is sustained almost entirely by the cable, we can still use a lightweight robotic arm to perform the foreseen operations. In this case, the robot assumes the role of precision unit during the positioning of the suspended element.

In the following treatment, the angles will be expressed in degrees while the lengths in centimeters.

We set the initial condition of the system to

$$\begin{aligned} q_r(0) &= [0, 0, 0, -90, -90, 90, 0]^T \\ q_p(0) &= [0, 0, 0, 0.1, 0, 0, 0]^T. \end{aligned}$$

For the desired configurations, we set a first desired reference for the x displacement as

$$r_1 = [-14.4, 6.6, -13.7, -83.2, -88.2, 62, -0.45, 0.2, 0.1]^T$$

and after 25 second, the reference is set so that the object is lowered

$$r_2 = [-13.9, 8.3, -14.2, -95.8, -95.3, 62.3, 15.6, 0.2, 0.2]^T$$

imitating the approach and the placement of a block.

Moreover, the values of the actuator saturation (8), operating region (6), and speed limitation (7) are set as (see [16] for more details regarding the robotic arm used in this work),

$$\begin{aligned} -\tau_{min,r} &= \tau_{max,r} = [320, 320, 170, 170, 110, 40, 40]^T \text{ Nm}, \\ -\tau_{min,x,l} &= \tau_{max,x,l} = 2 \times 10^3 \text{ Nm}, \\ -q_{min,r} &= q_{max,r} = [170, 120, 170, 120, 170, 120, 175]^T \text{ deg}, \\ -q_{min,l} &= 0 \text{ m}, q_{max,l} = 4 \text{ m}, \quad -q_{min,x} = q_{max,x} = -2 \text{ m}, \\ -\dot{q}_{min,r} &= \dot{q}_{max,r} = [85, 85, 100, 75, 130, 135, 135]^T \text{ deg/s} \\ -\dot{q}_{min,l} &= \dot{q}_{max,l} = 5 \text{ m/s}, \quad -\dot{q}_{min,x} = \dot{q}_{max,x} = 3 \text{ m/s}. \end{aligned}$$

The PD control gains in (9)-(10) were chosen as

$$\begin{aligned} K_{Pr} &= \text{diag}([300, 300, 300, 300, 300, 300, 300]), \\ K_{Dr} &= \text{diag}([10, 10, 10, 10, 10, 10, 10]), \\ K_{Px} &= K_{Pl} = 1 \times 10^3, \\ K_{Dx} &= K_{Dl} = 5 \times 10^2, \end{aligned}$$

The smoothing parameter of the attraction field (14) is $\eta = 1e^{-4}$. This value has been chosen in order to eliminate numerical noise in the attraction field that can occur when q_v is very close to r . The parameters of the repulsion field (17)-(19) are $\zeta = 0.3$, and $\delta = 0.01$. The parameters of the Dynamic Safety Margin defined in (22) are chosen as $k^{pos} = 0.8$, $k^{vel} = 0.5$, $k^\tau = 0.01$, $k^{E_{term}} = 15$. These gains are chosen such that the various DSMs of the active constraints have the same order of magnitude. The prediction sampling time to simulate the dynamic of the system is fixed to 1 ms, and with 100 prediction samples we predict over a time horizon of 100 ms. The terminal energy constraint is set to $E_{term} = 150 \text{ J}$. It is important to remember that since the weight of the block is supported by the cable, the internal control of the robot has been developed to ensure high precision in positioning and to avoid elastic deformations of the links of the lightweight robotic arms used in the following simulations.

As one can see in Fig.6-7 the ERG modified the commanded reference r in order to avoid that the constraints are violated and the inner controller is able to move the

system robot+object through the applied reference. In particular, as one can see in Fig.10, the commanded reference is modified in such a way that the speed constraints are never violated. Moreover, the non-actuated variables for the suspended object are damped by the robot during the movement (see Fig.8). It is important to notice that, despite the limitations of payload that can be managed by the robot, the ERG modifying the applied reference taking into account the constraints, it ensures that the robot is never overloaded. In fact, as one can see in Fig.9, the torque required to the robotic actuators are well within the joint torques limits. Finally, we show why a governing unit is necessary by analyzing the behaviour of the control scheme proposed in [7] considering the same simulation scenario discussed so far. In Fig.11 is shown the time evolution for the joint robot torque related to the control scheme proposed in [7]. This control scheme is based on a trajectory pre-evaluated off line and for this reason, it does not take into account any constraints (and in particular the torques robot saturation). Therefore, if no off-line trajectory is calculated, when there are sudden changes in the reference (*i.e.* at the beginning and at about 25 seconds), the torques required by the robot actuators violate the maximum (or the minimum) limits. This leads to an overload of the robot and makes the operation unfeasible.

5 Conclusions

This paper proposed a constrained control scheme based on the ERG framework for the control of a mechanical system to place heavy prefabricated elements. Based on the previous results already present in the scientific community, the aim of this paper is to propose a control scheme that is able to overcome the limitations of the works proposed so far by enforcing an inner loop controller with an external control unit that is able to manage the constraints of the system. In the first part of this paper a mathematical model of the overall system is derived, and then the constrained control scheme is described. Simulation results show that the proposed control scheme is able to fulfill the constraints of the system while moving the suspended object to its final position. It is worth noting that no off-line trajectory has been calculated that it is the control law itself that decides how to move the reference to avoid constraints violations. Moreover, future research will focus on the possibility of mounting the robot on a mobile base in order to guarantee easy movement in the working environment.

6 Acknowledgments

This research has been funded by The Brussels Institute for Research and Innovation (INNOVIRIS) of the Brussels Region through the Applied PHD grant: Brickiebots -

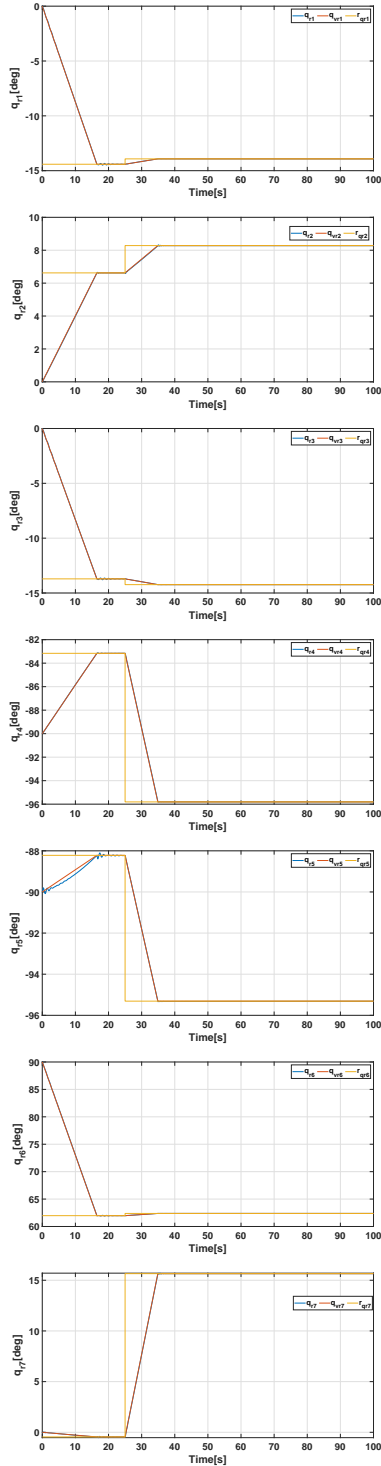


Figure 6. Time evolution of joint robot positions.

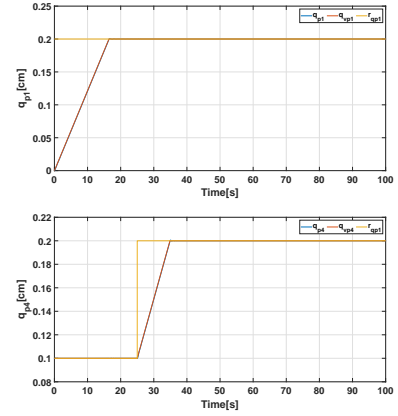


Figure 7. Time evolution of actuated joints of the suspended element positions.

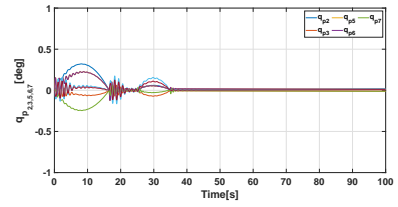


Figure 8. Time evolution of non-actuated joints of the suspended element positions.

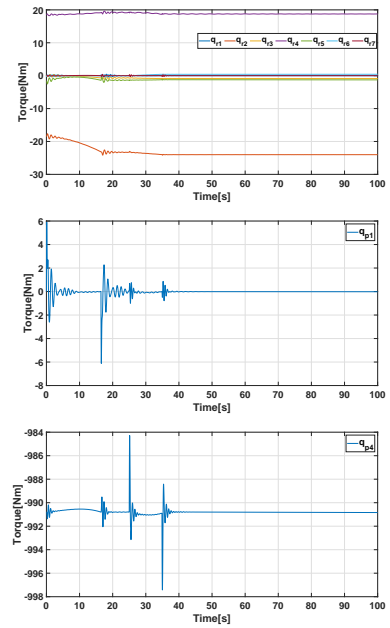


Figure 9. Time evolution of the required torques.

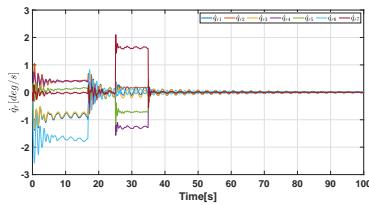


Figure 10. Time evolution of joint robot speed.

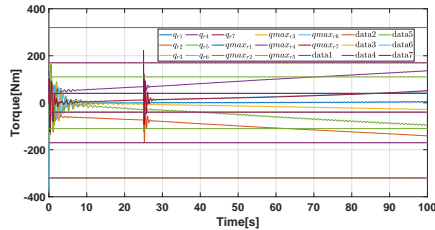


Figure 11. Time evolution of joint robot torques for the control scheme [7].

Robotic Bricklayer: a multi-robot system for sand-lime blocks masonry (réf : 19-PHD-12).

References

- [1] Sung-Min Moon, Jaemyung Huh, Daehie Hong, Seunghoon Lee, and Chang-Soo Han. Vertical motion control of building facade maintenance robot with built-in guide rail. *Robotics and Computer-Integrated Manufacturing*, 31:11–20, 2015.
- [2] Marwan Gharbia, Alice Chang-Richards, Yuqian Lu, Ray Y Zhong, and Heng Li. Robotic technologies for on-site building construction: A systematic review. *Journal of Building Engineering*, 32:101584, 2020.
- [3] Constructon site accidents. <https://www.attorneystevelee.com/our-library/types-of-construction-site-accidents/>.
- [4] J. Andres, Thomas Bock, F. Gebhart, and W. Steck. First results of the development of the masonry robot system rocco. *Proceedings of the 11th ISARC*, pages 87–93, 01 1994.
- [5] Marwan Gharbia, Alice Chang-Richards, Yuqian Lu, Ray Y Zhong, and Heng Li. Robotic technologies for on-site building construction: A systematic review. *Journal of Building Engineering*, page 101584, 2020.
- [6] Michele Ambrosino, Philippe Delens, and Emanuele Garone. Control of a multirobot bricklaying system. *Advanced Control for Applications: Engineering and Industrial Systems*, 3(4):e90, 2021.
- [7] Michele Ambrosino, Philippe Delens, and Emanuele Garone. Modeling and control of multi-units robotic system: Boom crane and robotic arm. In *Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 789–796. International Association for Automation and Robotics in Construction (IAARC), November 2021. doi:10.22260/ISARC2021/0107.
- [8] S Joe Qin and Thomas A Badgwell. An overview of nonlinear model predictive control applications. *Nonlinear model predictive control*, pages 369–392, 2000.
- [9] Michael Nikolaou. Model predictive controllers: A critical synthesis of theory and industrial needs. 2001.
- [10] Emanuele Garone, Stefano Di Cairano, and Ilya Kolmanovsky. Reference and command governors for systems with constraints: A survey on theory and applications. *Automatica*, 75:306–328, 2017.
- [11] Marco M. Nicotra and Emanuele Garone. The explicit reference governor: A general framework for the closed-form control of constrained nonlinear systems. *IEEE Control Systems Magazine*, 38(4):89–107, 2018. doi:10.1109/MCS.2018.2830081.
- [12] Michele Ambrosino, Arnaud Dawans, and Emanuele Garone. Constraint control of a boom crane system. In *Proceedings of the 37th International Symposium on Automation and Robotics in Construction (ISARC)*. International Association for Automation and Robotics in Construction (IAARC), October 2020. ISBN 978-952-94-3634-7. doi:10.22260/ISARC2020/0069.
- [13] Günter Schreiber, Andreas Stemmer, and Rainer Bischoff. The fast research interface for the kuka lightweight robot. In *IEEE workshop on innovative robot control architectures for demanding (Research) applications how to modify and enhance commercial controllers (ICRA 2010)*, pages 15–21. Citeseer, 2010.
- [14] Gustavo M Freitas, Antonio C Leite, and Fernando Lizarralde. Kinematic control of constrained robotic systems. *Sba: Controle & Automação Sociedade Brasileira de Automatica*, 22:559–572, 2011.
- [15] Elie Hermand, Tam W Nguyen, Mehdi Hosseinzadeh, and Emanuele Garone. Constrained control of uavs in geofencing applications. In *2018 26th Mediterranean Conference on Control and Automation (MED)*, pages 217–222. IEEE, 2018.
- [16] Kuka. <https://www.kuka.com/>.

Importance of a 5G Network for Construction Sites: Limitation of WLAN in 3D Sensing Applications

H.J. Lee^a, A. Krishnan^a, S. Brell-Cokcan^a, J. Knußmann^b, M. Brochhaus^b, R.H. Schmitt^c, J.J. Emontsbotz^d, J. Sieger^d

^a Chair of Individualized Production (IP), RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany

^b Fraunhofer Institute for Production Technology IPT, Steinbachstraße 17, 52074 Aachen, Germany

^c Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany

^d Institute of Mineral Resources Engineering, RWTH Aachen University, Wüllnerstraße 2, 52062 Aachen, Germany

office@ip.rwth-aachen.de

Abstract -

Teleoperated construction machinery dominates construction sites, as it can, with relatively little effort, prevent operators from working in dangerous conditions by keeping them in the control loop. However, the operators usually have to execute tasks with limited situational information due to poor depth perception from 2D camera images, reducing local accuracy and work efficiency. Thus, 3D sensing technology such as depth cameras is used more and more in combination with the teleoperated construction machinery. As these depth cameras are preferably mounted on a mobile robot to prevent the occlusion and observe the remote work place from several viewpoints, the corresponding on-site Information and Communication Technologies (ICTs) that can cover the required data transmission are of great importance for further developments. This paper presents a robotic platform capable of navigating and providing the 3D point cloud data of the remote work place from different viewpoints according to the operator's input. The captured information is transferred to the operator using the standard network Wireless Local Area Network (WLAN). To that end, first, the limitation of the WLAN in 3D sensing applications and the needs of the Fifth-Generation (5G) of mobile networks are jointly analyzed within the presented use case. Finally, the characteristics of 5G that address the test results are identified.

Keywords -

Robot-Assisted Construction; Construction Robots; Auto-mated Construction; 5G Network

1 Introduction

In the last decade, teleoperated heavy machinery has become an essential element for construction or disaster sites since it directly extends the operator's sensory-motor facilities. Remote teleoperation, however, has a major drawback, compromised human perception due to the decoupling of the human operator from the physical envi-

ronment. Poor perception have often detrimental effects on safety and efficiency especially in dynamic environments like construction sites [1].

A safe operation of teleoperated machines requires the operator to have good spatial awareness of the environment at all times. Various image representations of the environments have been developed to facilitate this [2, 3, 5]. These often include using exteroceptive and proprioceptive sensors, mounted directly onto the construction machine, coupled with panoramic visual feedback with birds eye view. Typical control stations include multiple 2D camera views from different perspectives to increase the telepresence by projecting the obtained sensory information to a remote place. However, such approaches have shown to be problematic concerning the operator's divided attention, and overloaded network [6].

RGB-D cameras, which can capture geometrical information in 3D, are often employed to improve perception in remote environments. In [7], it has been shown that providing the operator an additional 3D representation of the environment resulted in increased task performance when compared to monocular RGB views. Moreover, changing the sight-of-view is another essential aspect for enhancing the manipulation capabilities of a teleoperated system, as confirmed in the study of Huang et al. [8]. These sensors are often mounted on an external mobile robot to provide a dynamic perspective of the ongoing operations. The gathered information needs to be sent to the remote control station over a reliable network connection. In the past, wired networks have been used to remotely operate the machines. There are multiple reasons for using a tethered mobile robot for exploration and remote sensing. It could serve as high bandwidth and low latency communication channel while maneuvering underwater or underground mines or to power the onboard sensors in search and rescue operations. However, without computationally expensive path planning algorithms [4], such tethers strongly restrict the maneuverability of the mobile robot in unstructured environments like construction sites,

thereby increasing the chance of damage and decreasing the robustness of the system. Therefore, a further investigation into the wireless network is of great importance, especially for construction sites, as more mobile robotic applications are being integrated into construction sites, which has pushed recent developments in the direction of wireless networks [5, 9].

But the usage of WLAN comes at the cost of limited bandwidth and latency, which in turn affects the efficiency of the operation. Some methods have been developed to circumvent this such as sensor scheduling or lowering the frame rate of the sensor feedback [10]. These approaches while effective in resource-constrained systems, are not ideal for continuous operation on a construction site. Data compression methods are also often employed to shrink the amount of the data, allowing it to be stored and sent via a low-bandwidth channel. Such compression techniques aim to reduce data size by finding and removing statistical redundancy while keeping the original data [11] [12]. However, the compressed data size for wireless data transport remains rather large in multiple megabytes, and information loss still often occurs, necessitating appropriate network technology.

In line with this problem, several research initiatives are now focusing on the on-site usage of a more sophisticated network, such as the 5G of mobile networks (5G.NAMICO,2022). The basic characteristics for 5G have been detailed as: higher transmission rate, shorter latency, higher reliability, and more User Equipment (UE) connection. A recent study [13], has evaluated the benefits of using 5G technology in construction sites by comparing the performance of video streams from static cameras on the construction site. Yet, further investigations in line with the automated systems are needed to appropriately plan the deployment and usage of the 5G network on construction sites.

In this paper, we first present the robotic setup which can capture the remote working scene in a 3D point cloud and transmit it to the operator using the standards WLAN network. The mobile robot is based on a commercially available platform. It is further equipped with a 3D depth camera to provide 3D visual feedback from an arbitrary viewpoint of the remote workplace. Effective collaboration with the operator requires reliable data transfer of the captured visual information. Therefore, a speed test was conducted using the commonly used network, WLAN. Based on the test results, the limitations of the existing telecommunications technologies are highlighted. Then, the required characteristics and functionalities of the new 5G networks are identified. Furthermore, some of the possible challenges for installing 5G networks on the construction sites are discussed.

2 Mobile Robot Platform

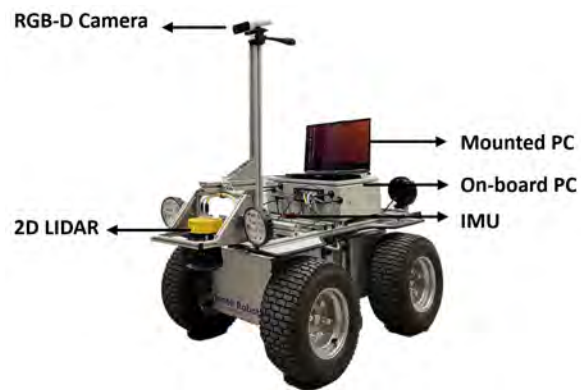


Figure 1. The mobile robot used in this work. It carries several sensors such as a IMU, 2D LIDAR and a RGB-D camera.

The mobile robot used in this work is based on a commercial platform from *Innok Robotics*. For observability of the environment and the systems states, various sensors are integrated in the robot. An inertial measurement unit (IMU) *Xsens MTi-30-2A8G4* is placed in the middle of the base and provides the rotational speed and acceleration information. The robot is also equipped with a 2D laser scanner *Sick microScan 3* to detect obstacles or create a 2D environment map for localization tasks. In order to visualize the working place in 3D, we use a RGB-D camera *Microsoft Azure Kinect DK* at the front of the mobile robot. It is worth mentioning that the used *Azure Kinect DK* is not designed to work outdoors as the camera relies on its infrared sensor to collect the depth information. The accuracy of the measurements are heavily influenced by the infrared interference from the sun [14]. The performance in outdoor environments might be guaranteed by using a stereo depth camera. However, such a device was not available at the time of this work.

The mobile robot is equipped with two computers, one on-board-PC with Intel Core i5-5350 CPU and 8GB RAM and another PC with Intel Core i7-9850 CPU and 16GB RAM are both mounted on the base (see Fig.1). The on-board-PC is responsible for the basic functionality such as controlling the wheels through the SPS controller and sending the most sensor data i.e., 2D laser scanner and IMU. The other PC is only responsible for processes related to the RGB-D camera which require more computational power. A radio remote controller is integrated into the mobile platform. A remote controller with joysticks is used to control the mobile platform. As an alternative, the on-board PC can communicate with another PC by using the integrated WLAN-router. Here, the communication is established in the Robot Operating System (ROS) envi-

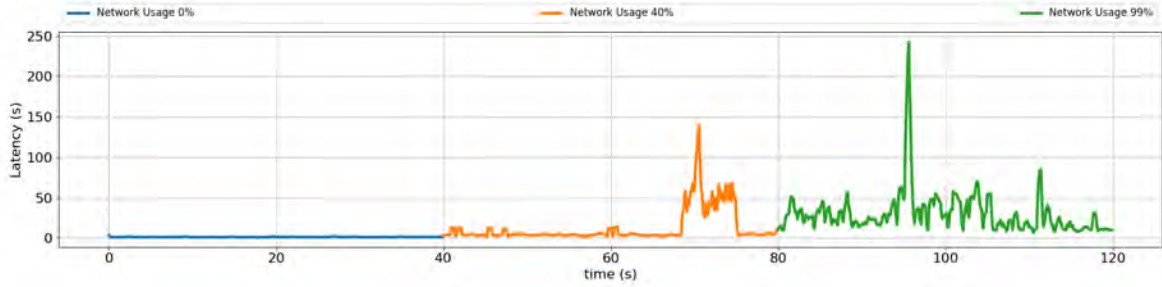


Figure 2. Latency measures for control commands of the size 17 KB from Experiment 3.1: blue (0% network usage), orange (40% usage) and green (100% usage).

ronment [15]. To capture the remote working space from an arbitrary perspective, the operator can send the control commands (x_c, y_c, ψ_c) to move the robot in the XY-plane and rotate around z-axis. As a time discrepancy between these two PCs results in an inconsistent environment measurements, we use ROS and a time management software [16] to synchronize the clocks.

3 Experiments

To evaluate the capabilities of the frequently used WLAN in 3D sensing applications, we separately investigated the network latency and the data rate that the sensor suite on the mobile robot needs to transfer the data.

3.1 Latency in relation to network overload

The network latency is first obtained by estimating the time for command signals to arrive at the mobile robot from the control PC. The mobile robot platform used in this work is programmed in the ROS environment. Therefore, we transmit ROS control messages with a size of 17 KB at 30 Hz. We measure the time discrepancy between the transmission and arrival for each of these messages. We use the WLAN with the 802.11ac standard deployed in our lab facility for this data transmission. The router is positioned directly next to the control PC, and the mobile robot is positioned around 10 m away from the control PC and the router. To highlight the limitation of the WLAN, we measure the latency in three different settings: i) 0% usage of the network's full bandwidth, ii) 40% usage (approx. 60Mbps), and iii) 99% usage (approx. 200Mbps).

For each setting, the message is transmitted for 40 seconds. After 40 seconds, the network usage is changed. Fig. 2 shows the latency measurements during the experiment. In the first case, when the network is free, the average of the measured latency is around 1 ms. This value increases to 5 ms with the network usage of 40%. The figure clearly shows that the network stability deteriorates as the latency increases up to around 150 ms. In the last

case, when other tasks use the full bandwidth of the network, the network is, in general, very unstable, and massive peaks in the latency constantly occur.

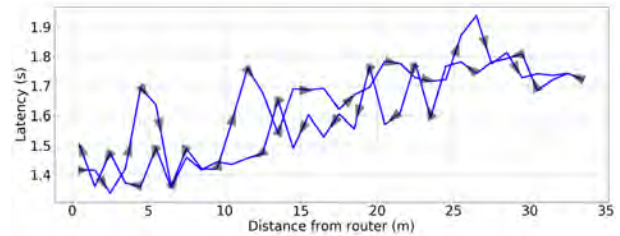


Figure 3. Latency measures in relation to distance from router, Experiment 3.2.

3.2 Latency in relation to distance from router

Due to the nature of the dynamic construction sites, the workspace for robots constantly changes. In this experiment, we demonstrate the impact of the distance between the WLAN router and the mobile robot, highlighting the limitation of the WLAN in the mobile robotic application. For the test, the mobile robot drives from its start position, 0.5 m away from the router, 34 meters straight forwards, and returns to its start position. While driving, we measure the latency that a point cloud data needs to travel from the mobile robot to the control PC. We use a pre-captured point cloud of size 29 MB to keep the data size constant throughout the test. The timestamp stored in the point cloud when leaving the mobile robot is compared with the arriving time to measure the latency.

Fig. 3 depicts the overall instability of the WLAN network. As the point cloud size requires the full bandwidth of the network, the latency fluctuates strongly and increases depending on the distance up to 35% from 1.4 s to 1.9 s within the test trajectory of 34 m. This is, in fact, critical, as due to the dynamic nature of construction sites, the on-site operations require large coverage

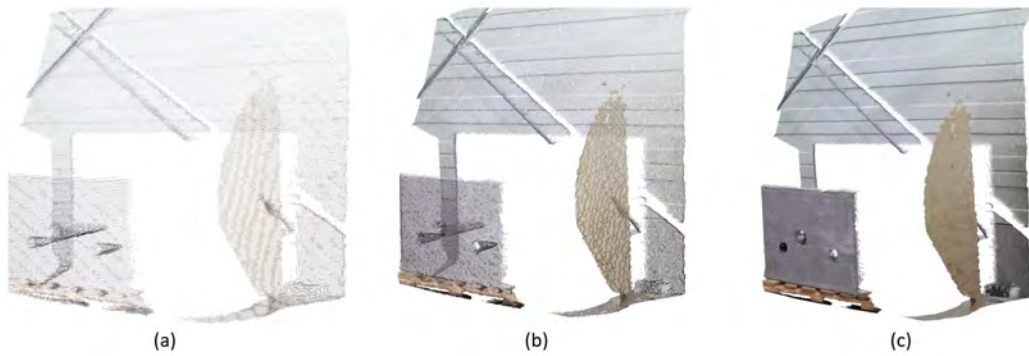


Figure 4. Captured 3D point clouds in different resolutions: a) 1 cm, b) 0.5 cm and c) 0.2 cm.

Type	Resolution [cm] or [px x px]	Frequency [Hz]	Msg Size [MB]	Req. data rate [Mbps]
Point Cloud (Low)	1	22	0.2	35.2
Point Cloud (Med.)	0.5	22	0.6	105.6
Point Cloud (High)	0.2	22	1.2	211.2
3D map	1	1	1.1	8.8
RGB Views (Comp.)	640 x 576	22	0.12	21.12

Table 1. Measured message sizes and corresponding data rate for different compression levels and data type

of the network. However, if the latency time varies, as the experiment shows, it can lead to the malfunction of on-site construction robots and cause a safety issue. The WLAN coverage area can be expanded by adding more access points, but this raises another issue: how to mitigate the impact of momentary communication instability between the access points.

3.3 Capturing 3D information

We capture the scene in 3D using the RGB-D camera mounted on the mobile platform. The captured information is then converted into point clouds. As raw point clouds are in the size of several tens of megabytes, we first compress the point clouds using the point cloud library (PCL) [17] and send them to the host PC (i.e., the user) via the WLAN network. In this experiment, the robot drives three times to the object with varying compression levels for point clouds: a) 1 cm, b) 0.05 cm c) 0.02 cm (see Fig. 4). We demonstrate the quality and the required network throughput trade-off. To control the motion of the mobile robot, the ROS navigation stack is used [18]. The estimated control inputs are forwarded to the robot controller by using the driver library provided by the Innok Robotics [19].

The final column of Fig. 5 clearly identifies the distinct features of the visual feedback. However, as Table 1 illustrates, the full bandwidth of the WLAN network (~200Mbps) is needed to transmit the point cloud in the high-level compression without any delay. While the 3D point clouds provide depth information and support

the telepresence, the 2D camera images generally have greater resolution than the 3D point clouds (see Fig. 6). Consequently, the 2D camera pictures acquired by the resolution 640 x 576 are also transmitted parallel to the 3D point clouds as baseline information, requiring an extra data rate of 21.12 Mbps. Although the 2D and 3D camera views give comprehensive scene information, the operator cannot analyze the complete remote scene, as the captured data disappears as soon as the robot continues to drive. Real-time appearance-based mapping (RTAB-Map) [20] is used to perform the online processing of the 3D map by detecting the loop closure with the extracted key features and building a global 3D map with it. Here, we transmit the optimally generated 3D map every 1 second, generated in the fixed 1cm resolution. Although the needed data rate is recorded at 8.8 Mbps throughout the drive, it needs to be recognized that this data rate will steadily rise as the map grows.

3.4 Summary and Discussion

In this paper, a WLAN ac 802.11ac network was used to send and receive data. The experiment findings clearly illustrate that the quality of the visual feedback largely relies on the compression level (i.e., resolution). However, the findings also reveal that the network's capacity is completely utilized to provide the information to the user wirelessly. The problem is aggravated if 2D camera views and the 3D map data are transmitted at the same time. Although the tests were done in an ideal indoor setting, the limitations of the present WLAN network in terms of the

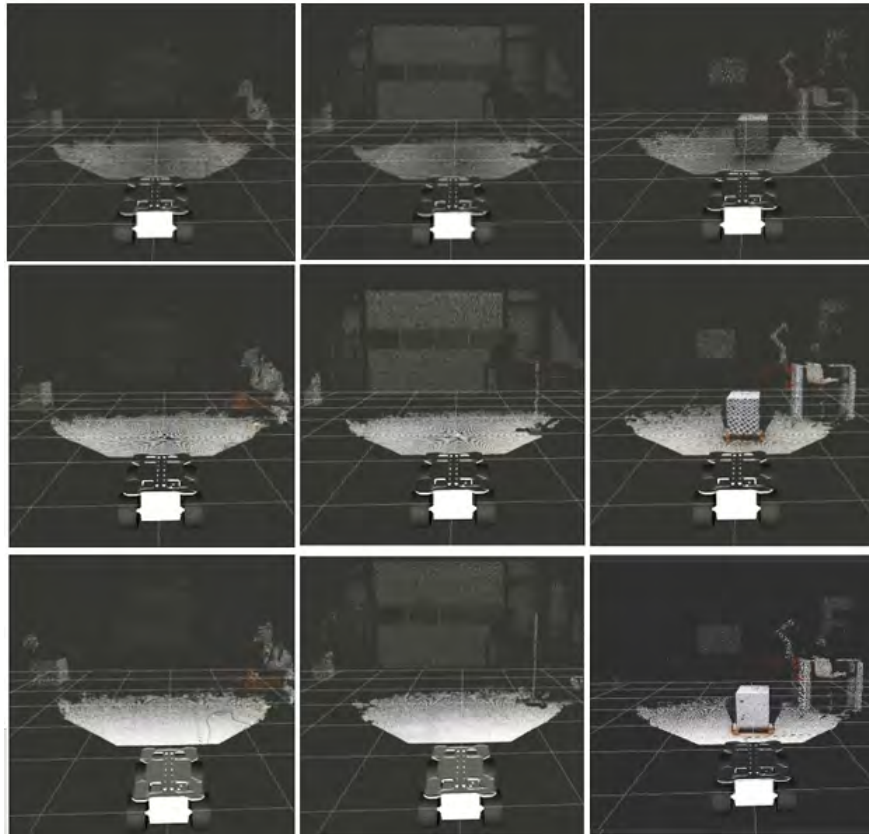


Figure 5. During the drive towards the concrete object, 3D point clouds were compressed in different resolutions and transmitted to the operator via the WLAN network (the first row: 1 cm, the second row: 0.5 cm and the third row: 0.2 cm).

bandwidth and the accompanying latency could be shown. Without these ideal conditions, we expect that the issue would be compounded on large construction sites with multiple equipment and machines using the same network which may result in reduced work performance and critical safety issues.

4 Chances and Challenges for 5G

4.1 Chances for 5G

One possibility to overcome the described shortcomings and fulfill the criteria for teleoperated construction machinery is the usage of the 5G mobile networking standard. 5G was developed to meet industrial requirements and thereby serves as an enabler for the digitalization of various industrial applications [21, 22]. While previous generations of wireless network standards did not provide many opportunities for adaption, development for 5G and its individual releases follows a different approach. One of the main principles of 5G is its service-based architecture (SBA) that offers a lot of customization potential [23].

Instead of a uniform solution for all use cases, modifications in network architecture allow users to take a variety of heterogeneous requirements of different use cases into account [24]. The three most important performance characteristics 5G offers are:

- Ultra-Reliable Low Latency Communication (URLLC)
- Enhanced Mobile Broadband (eMBB)
- Massive Machine-Type Communication (mMTC)

URLLC enables 5G networks to have a less than 1 ms end-to-end latency with the reliability of more than 99.999%. eMBB enables the transfer of a high data rate through wireless communication with up to 20 Gbit/s for downlink (DL) and up to 10 Gbit/s for uplink (UL). mMTC enables the setup of communication networks with up to 1 million devices per square kilometer [25]. Current measurements of 5G implementations (Release 15/16) in an industrial setup show the potentials of 5G: With background traffic but no specific load tests, measurements show that for a 1 kB message with URLLC the

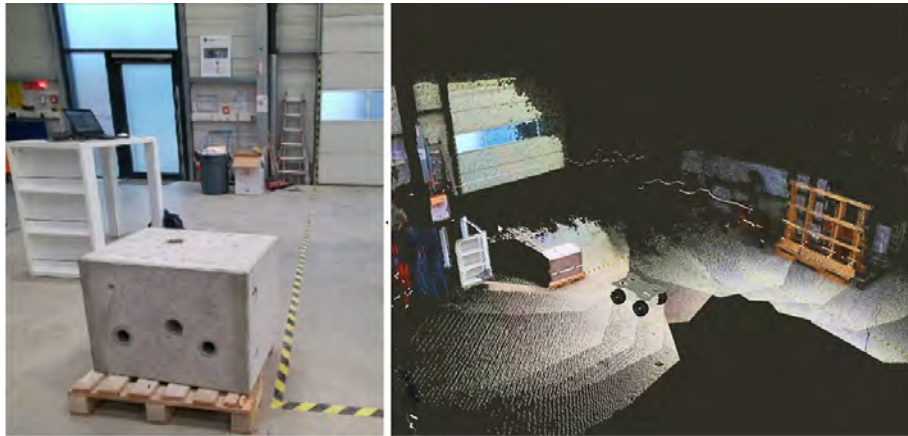


Figure 6. The visual feedback visualized to the user in the control PC (Left: 2D camera views in 640 x 576, Right: 3D map generated with the captured point clouds)

latency of 1 ms is kept with 99.9%. In the setup, however, the median latency is increased from 12 ms up to 25 ms in an eMBB configuration [26].

At this point, it is important to mention that fulfilling all three performance characteristics at once is not possible in the same network. Optimizing a network towards one characteristic leads to a degradation in the other characteristics. 5G uses network slicing to adapt to the needs of the respective use cases. With network slicing, 5G builds different logical networks on the same physical infrastructure. These logical networks meet the different needs of the use cases regarding the quality of services (QoS) characteristics like latency or data rates and also take into account different security needs [27].

Examples of use cases requiring different QoS characteristics are the transmission of critical control messages needing a high reliability but only low data rates, in comparison to the transmission of camera images needing high data rates but a not so high reliability [28].

For application scenarios with high data rates, 5G offers an experienced data rate of 1 Gbps for DL and 500 Mbps for UL for an indoor hotspot [29]. Furthermore, for a message with the size of 15 kB to 250 kB, 5G provides the target values for a transfer interval (latency) of 10 ms to 100 ms [30]. Since the message size of the ROS control messages for our experiment is 17 kB, the transfer interval can be assumed to be somewhere between 10 ms and 20 ms. Both, the data rate and the latency represent a clear improvement over the measured values using the WLAN network. Another advantage of 5G is the possibility to guarantee the specified latency through prioritizing traffic in the network [31]. This allows sending the critical ROS control messages with low latency even if the full bandwidth of the network is used. Therefore, we assume that 5G is suitable for high-bandwidth and time-critical appli-

cations on construction sites.

Further, 5G supports higher mobility than WLAN. If the connected device, the mobile robot, drives from one access point to another, it needs to change its connection between different access points. The device using WLAN needs to authenticate itself leading to a momentary loss of connection or significantly higher latency [32]. This limitation is critical for safe on-site operations, given the present trend in the construction sector, which involves many digital sensor systems and robotic machinery. As a result, it is vital to investigate the implementation and use of 5G on construction sites, which can provide high-bandwidth time-critical communication capabilities.

4.2 Challenges for 5G

So far, the use of 5G for wireless communication on construction sites is not yet widely explored. One main problem of the existing research gaps lies in the on-site installation of a 5G network. Unlike most industrial production environments, the workspace of construction sites is subject to continuous changes in infrastructure. In building projects, the workspace keeps changing as the building grows. Also, the environment where the network is deployed is changing as more building components such as walls can cause signal interference. Thus, very high adaptability is required for the deployed 5G network. Another challenge is the possible adverse environmental conditions that a 5G network might be facing on a construction site. Dust, dirt, and water are potential sources of disturbances for equipment and transmitted signals, so their influence on the installation and usage of the network might be challenging.

5 Future Works

This section describes possible research directions related with the use of the 5G network on construction sites:

- To address the aforementioned challenges, future research aims to investigate the requirements of an on-site 5G network and implement the required 5G network architecture first on the reference construction site on Aachen West. Standard network technologies such as WiFi, LoRaWAN, and MQTT are already deployed on the test site. Therefore, the development and benchmark of the designed 5G network can be performed under conclusive and realistic conditions.
- Another line that future research can focus on is the transferability of the developed solution. Most of the time, the developed solutions in research works are isolated without considering the transferability of the solution to other use cases. To avoid this, we aim at investigating the 5G uses in underground and above-ground construction works at the same time. Particularly in the underground works, the investigation of how 5G signals behave in the individual sections of the underground, especially in the expanding last mile, and how it can be propagated over spatially limited distances under adverse conditions, such as mineral dust, humidity, are of great interest. Using the synergy effects of these similar but different environments, we aim to maximize the transferability of the developed 5G solution.

6 Conclusion

Technology such as robotics presents significant opportunities for the construction industry and involved on-site human workers. One fundamental requirement for such emerging technology is that the communication to the robot is stable with a minimum latency and data communication is realized with a minimum delay. In this paper, the clear limitation of the existing WLAN is demonstrated in the context of using a mobile robot for the remote sensing application. In this way, the required specification for the new 5G technology was first defined, and corresponding core characteristics (URLLC, eMBB, and mMTC) were described. Next, the 5G technology's chances for the construction industry and the challenges that might arise from the installation and usage of the 5G technology on the construction sites were shortly described. In the last part, an outlook into future works was established to address the potential challenges.

7 Acknowledgments

This work has been supported by the North Rhine-Westphalia Ministry of Economic Affairs, Innovation,

Digitalisation and Energy of the Federal Republic of Germany under the research intent 5G.NAMICO: Networked, Adaptive Mining and Construction. The reference construction site on RWTH Campus West is run by the Center Construction Robotics and the first living lab to test new construction processes, products, automated machines and robots in Europe.

References

- [1] Chen, J.Y., Haas, E.C. and Barnes, M.J., "Human performance issues and user interface design for teleoperated robots." 2007 IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 37(6), pp.1231-1245.
- [2] Sato et al., 2013, December. Spatio-temporal bird's-eye view images using multiple fish-eye cameras. In Proceedings of the 2013 IEEE/SICE international symposium on system integration (pp. 753-758). IEEE.
- [3] James et al., 2011, August. Tele-operation of a mobile mining robot using a panoramic display: an exploration of operators sense of presence. In 2011 IEEE International Conference on Automation Science and Engineering (pp. 279-284). IEEE.
- [4] Kim S., Bhattacharya S. and Kumar V. , "Path planning for a tethered mobile robot," 2014 IEEE International Conference on Robotics and Automation (ICRA), 2014, pp. 1132-1139.
- [5] Kamezaki, M., Yang, J., Iwata, H. and Sugano, S., 2014. A basic framework of virtual reality simulator for advancing disaster response work using teleoperated work machines. Journal of Robotics and Mechatronics, 26(4), pp.486-495.
- [6] Nielsen CW, Goodrich MA, Ricks RW (2007) Ecological interfaces for improving mobile robot teleoperation. IEEE Transactions on Robotics 23(5):927–941
- [7] Mast et al., (2013). Teleoperation of Domestic Service Robots: Effects of Global 3D Environment Maps in the User Interface on Operators' Cognitive and Performance Metrics. 8239. 392-401.
- [8] Huang, K., Chitrakar, D., Rydén, F. et al. (2019) Evaluation of haptic guidance virtual fixtures and 3D visualization methods in telemanipulation—a user study. Intel Serv Robotics 12, 289–301
- [9] Larsson, J., Broxvall, M. and Saffiotti, A., 2010, May. An evaluation of local autonomy applied to teleoperated vehicles in underground mines. In 2010 IEEE In-

- ternational Conference on Robotics and Automation (pp. 1745-1752). IEEE.
- [10] Mourikis, A.I. and Roumeliotis, S.I., 2006. Optimal sensor scheduling for resource-constrained localization of mobile robot formations. *IEEE Transactions on Robotics*, 22(5), pp.917-931.
 - [11] Sun, X., Ma, H., Sun, Y., Liu, M. (2019). A Novel Point Cloud Compression Algorithm Based on Clustering. *IEEE Robotics and Automation Letters*, 4, 2132-2139.
 - [12] Nguyen, D.T., Quach, M., Valenzise, G., Duhamel, P. (2021). Learning-Based Lossless Compression of 3D Point Cloud Geometry. *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 4220-4224.
 - [13] Xiang et al., 2020. 5g meets construction machines: Towards a smart working site. *arXiv preprint arXiv:2009.05033*.
 - [14] Tölgyessy, M., Dekan, M., Chovanec, L., Hubinský, P. Evaluation of the Azure Kinect and Its Comparison to Kinect V1 and Kinect V2. *Sensors* 2021, 21, 413.
 - [15] Quigley et al., "ROS: an open-source Robot Operating System," in *Proc. of the IEEE Intl. Conf. on Robotics and Automation (ICRA) Workshop on Open Source Robotics*, 2009.
 - [16] Chrony. Introduction. <https://chrony.tuxfamily.org/>, Jan 2022
 - [17] R. B. Rusu and S. Cousins, "3D is here: Point Cloud Library (PCL)," 2011 IEEE International Conference on Robotics and Automation, 2011, pp. 1-4.
 - [18] E. Marder-Eppstein, E. Berger, T. Foote, B. Gerkey and K. Konolige, "The Office Marathon: Robust navigation in an indoor office environment," 2010 IEEE International Conference on Robotics and Automation, 2010, pp. 300-307.
 - [19] Heerklotz A (2014) Innok robotics. https://github.com/innokrobotics/innok_heros_driver
 - [20] Labbé, M. and Michaud, F. (2017). Long-term online multi-session graph-based slam with memory management. In: *Autonomous Robots*.
 - [21] Rao, S.K., Prasad, R. Impact of 5G Technologies on Industry 4.0. *Wireless Pers Commun* 100, 145–159 (2018).
 - [22] J. Sachs and K. Landernäs, "Review of 5G capabilities for smart manufacturing," 2021 17th International Symposium on Wireless Communication Systems (ISWCS), 2021, pp. 1-6.
 - [23] Hou, Xinli; Xia, Liang (2020): 5G E2E TECHNOLOGY TO SUPPORT VERTICALS URRLC REQUIREMENTS. *Verticals URLLC Use Cases and Requirements*. Hg. v. Next Generation Mobile Networks e.V.
 - [24] Foukas, Xenofon; Patounas, Georgios; Elmokashfi, Ahmed; Marina, Mahesh K. (2017): Network Slicing in 5G: Survey and Challenges. In: *IEEE Commun. Mag.* 55 (5), S. 94–100.
 - [25] ITU-R Radiocommunication Sector of ITU (Hg.) (2015): IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond. Recommendation ITU-R M.2083-0. International Telecommunication Union.
 - [26] Ansari, Junaid; Andersson, Christian; Bruin, Peter de; Farkas, János; Grosjean, Leefke; Sachs, Joachim et al. (2022): Performance of 5G Trials for Industrial Automation. In: *Electronics* 11 (3).
 - [27] Ericsson (Hg.) (2018): 5G Deployment Considerations.
 - [28] Mendoza et al., (2021): 5G for Construction: Use Cases and Solutions. In: *Electronics* 10 (14), S. 1713.
 - [29] 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects: 3GPP TS 22.104 V16.5.0 (2020-07) Technical Specification. Service requirements for cyber-physical control applications in vertical domains; Stage 1 (Release 16).
 - [30] 3rd Generation Partnership Project Technical Specification Group Services and System Aspects: 3GPP TS 22.261 V16.16.0 (2021-12) Technical Specification. Service requirements for the 5G system; Stage 1 (Release 16).
 - [31] 5G Alliance for Connected Industries and Automation (5G-ACIA) (2021): 5G for Industrial IIoT: Capabilities, Features, and Potential. 5G-ACIA White Paper. Hg. v. ZVEI e. V.
 - [32] VDMA (Hg.) (2020): 5G in mechanical and plant engineering. Guide for the integration of 5G in product and production.

Path Generation for Foam Additive Manufacturing of Large Parts with a Cable-Driven Parallel Robot

E. Paquet¹, M. Metillon¹, K. Subrin¹,
B. Furet¹ and S. Caro¹

¹ Nantes Université, Ecole Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes, France

elodie.paquet@ls2n.fr, marceau.metillon@ls2n.fr, kevin.subrin@ls2n.fr, benoit.furet@ls2n.fr, stephane.caro@ls2n.fr

Abstract -

In this paper, a framework for foam printing with a Cable-Driven Parallel Robot (CDPR) is described for building large parts. Compared with the traditional robotic systems, CDPRs have a large workspace that can include the printing area. In addition, the potential reconfigurability of CDPRs is an asset to get rid of the collisions between the cables and the environment during the execution of the printing task. The printing feasibility is verified through the process identification where key parameters are used in the proposed control law. The features to be taken into account through a framework to make a successful printing with a CDPR are described. Finally, advantages and drawbacks of CDPRs for additive manufacturing are discussed and future work is presented.

Keywords -

Additive Manufacturing; Cable-Driven Parallel Robot; Innovative Construction

1 Introduction

3D printing solutions are generally based on Cartesian robots with a 2D motion in the horizontal plane and a 1D motion for the support table, corresponding to a three degrees of freedom machines [1][2][3]. For printing large parts, several materials can be used such as: plastic materials [5], polymer foams [1][4][5] concrete [2][6][7], or materials based on metallic particles [8][9]. The combination of automatic system and process has allowed multiple advances for large-scale 3D-printing, particularly house walls in construction fields. For instance, BatiPrint3D™ construction technology, composed of a Staubli poly-articulated arm and a B2A Systems Automatic Guided Vehicle (AGV), was proposed by Nantes University, France (Figure 1). In 2017, Yhnova demonstrator became a 95m² social dwelling built for the social landlord Nantes Métropole Habitat (NMH) [4].

The 3D-printing demonstrator BatiPrint3D™ technology focuses on the construction of house walls through the deposition of two layers of polymer foam used as a formwork for a third concrete layer inside. To build large parts



Figure 1. Additive foam manufacturing: BatiPrint3D, Yhnova's house 3D Printing with an AGV and a serial robot

without interruption in the printing process that would be related to the concatenation of the different workspaces of a mobile robot, the use of CDPRs for additive manufacturing in a short term is conceivable [7][10][11][12]. CDPRs are a potentially suitable replacement for very large-scale applications [8][13]. They easily achieve a large workspace without requiring massive equipment and machinery [14, 15].

This paper introduces a framework to adapt the 3D Printing process to a CDPR. A printing head with polymer foam for the production of large parts is managed by this architecture. The 3D foam printer performance is demonstrated through the construction of two different large parts, with an accuracy equal to 1 cm. The advantages and drawbacks of the novel 3D printer are then discussed. The paper is organized as follows: Section 2 introduces the process and the robotic architecture and a way to identify key parameters to adapt the process on a CDPR. Section 3 highlights the experimental validation and proposes a framework to fulfill the need. Section 4 draws some conclusions and future work.

2 Spraying End-Effector mounted on a CDPR

2.1 Spraying end-Effector

The process of deposition is managed by an air-actuated motor which controls the spraying nozzle enabling or disabling the foam deposition. Two tanks containing the Polyol and Isocyanate materials, respectively, are placed in the robotic cell and mixing them together. The polymer foam is then obtained.

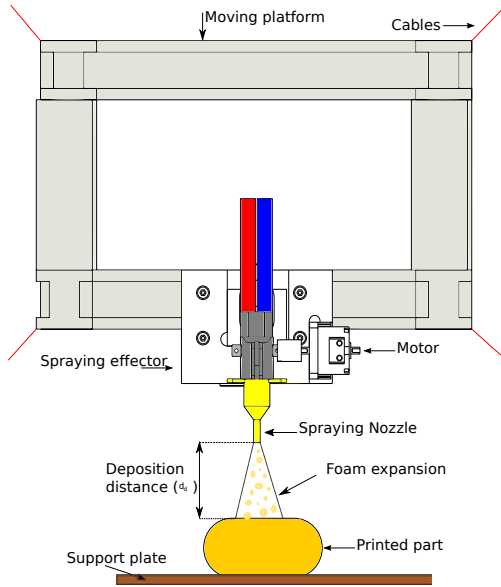


Figure 2. The moving-platform with the spraying gun and its actuation system

The mixture is performed in a static mixer and the compound is then sprayed on the surface of the slab or on the sub-layers as described in Figures 2 and 5. The material expands and acquires a sufficient stiffness in about six seconds. This time depends on the mix temperature and on the compound reactivity. The foam density is bounded between 35 kg/m^3 and 45 kg/m^3 . Its thermal conductivity is equal to 0.027 W.M.K and the Young modulus is equal to 7 MPa . The mean width of the wall created by foam spreading is around 72 mm obtained by adapting different parameters such as flow rates of Isocyanate and Polyol, the distance between the spraying nozzle, the speed of displacement of the nozzle. We now present the robotic architecture i.e. the CDPR which moves the platform.

2.2 CDPR presentation

A CDPR is a robotic system composed of at least 6 cables, reeled in and out by winches, which connect together a frame and a MP. It belongs to a particular class of parallel robots where a MP is linked to a base frame

using cables [2][16]. Motors are mounted on a rigid base frame and drive winches. Cable coiled on these winches are routed through exit points located on the rigid frame to anchor points on the MP as shown in Figure 3. By controlling the cable lengths in a synchronous manner, the load can be steadily translated and rotated in a large space with a good dexterity and stability. CDPRs have a large workspace and reach high dynamic performance. Figure 3 depicts the overall architecture of the cable-suspended 3D foam printer with its main components, namely, winches, exit-points and the MP. The winches control the eight cable lengths, which move and actuate the MP. Cables are routed through exit-points located on the rigid frame and connected to anchor-points located on the MP. The printed part is located on the ground so that the MP can access to the top layer of the printed part.

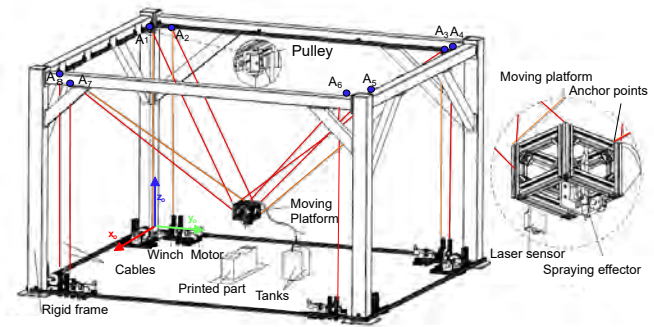


Figure 3. Schematic of a CDPR dedicated to additive foam manufacturing

The extruder is controlled by an output of the controller of the CDPR. The 3D foam printer, named CRAFT and displayed in Figure 4, has the following dimensions: $2.4 \text{ m} \times 3.67 \text{ m} \times 2.76 \text{ m}$ ($l \times b \times h$). Furthermore, it is re-configurable and their different reconfiguration strategies were studied in [17], [18], [19].

2.3 Process identification

The maximum possible speed of the flow rate and thus the maximum axial speed of the foam spraying effector is 210 mm/s . This system provides relatively robust position control with a PID position controller and a dynamic positioning error equal to 0.05 mm . An important parameter for printing with the proposed device is the relationship between the spraying nozzle position and the foam flowrate. A model of the foam 3d printing process is expressed in terms of the nozzle height d_d , the velocity of the spray end-effector v_d and the layer height h_c .

The flowrate and the geometrical dimension of the layer were found at different positions and at different deposition speed by measuring on different section of layers. Figure 6 represents the height of foam deposited during a certain

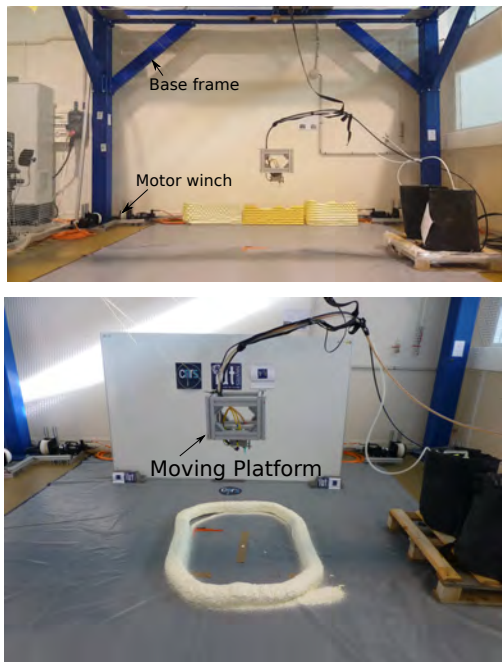


Figure 4. A CDPR, named CRAFT, used for foam additive manufacturing at LS2N, Nantes, France

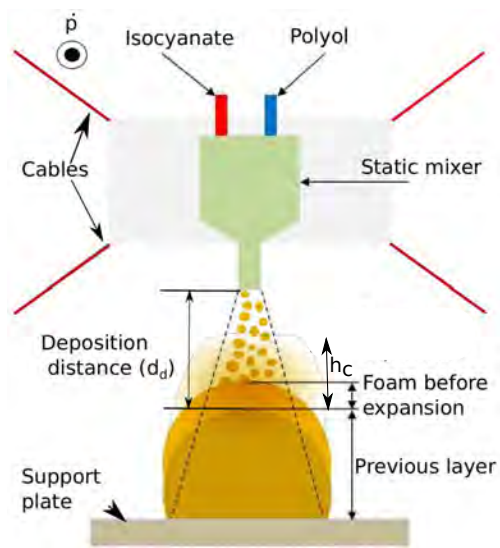


Figure 5. Description of the foam additive manufacturing process and associated variables

period of time.

Table 1 gives the measured layer height h_c as a function of v_d and d_d . Figure 7 shows the relationship between the three variables h_c , d_d and v_d governing the foam printing process. It is noteworthy that the printing process under study is non-linear.

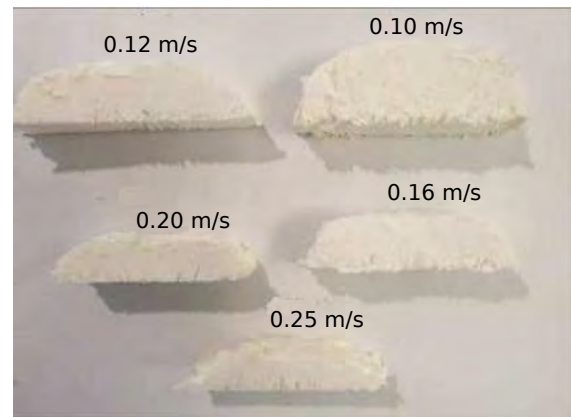


Figure 6. Sectional view of printed foam samples as a function of the linear velocity v_d of the spraying end-effector for deposition distance d_d equal to 150 mm

Table 1. Measured layer height h_c (m)

		d_d (m)		
		0.10	0.15	0.20
$v_d(\text{m.s}^{-1})$	0.050	0.071	0.068	0.064
	0.075	0.057	0.056	0.045
	0.100	0.043	0.043	0.034
	0.150	0.036	0.027	0.021
	0.250	0.024	0.018	0.013

The foam process control model was developed using a non-linear multivariate regression and is illustrated in Figure 7. The blue dots on the response surface are the values obtained from the tests conducted during the identification of the printing process on the CDPR.

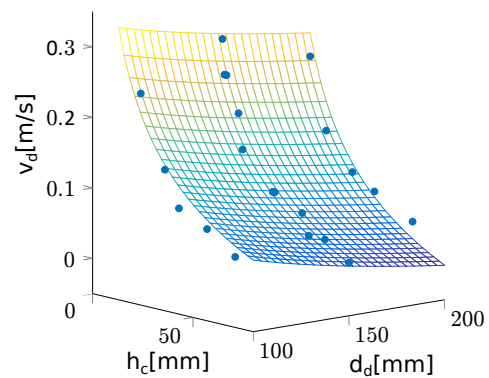


Figure 7. Process modeling characterized by a response surface expressed in Eq.(1) obtained based on a non-linear multivariate regression of data given in Table.1

The foam printing process is modeled as follows:

$$v_d = b_1 + b_2 h_c^{b_3} + b_4 d_d^{b_5} \quad (1)$$

with $b_1 = -1.743$, $b_2 = 1.464$, $b_3 = -0.314$, $b_4 = 1.830$ and $b_5 = -0.054$ being computed with ©Matlab *fitnlm* function from the values given in Tab. 1.

2.4 CDPR Control

The actuators of the CDPR are controlled using a computed torque feed-forward control scheme with a joint space feedback corrector. The integration of the model is performed inside the control law of the CDPR. The control scheme is shown in Fig. 8. The desired velocity of the spray end-effector v_d , namely the one of the moving-platform, along the deposition axis is obtained using Eq. (1). The desired moving-platform twist \mathbf{t}_d is computed by projecting the deposition velocity profile onto the deposition axis. The desired moving-platform pose \mathbf{x}_d and acceleration $\dot{\mathbf{t}}_d$ are obtained using the desired twist first time integrative and derivative respectively. The Inverse Geometric Model (IGM) and the Inverse Kinematic Model (IKM) are used to define the desired joint position \mathbf{q}_d and velocity $\dot{\mathbf{q}}_d$ respectively. Using the measured joint position from the encoder readings, the joint position error \mathbf{e}_q and the joint velocity error $\dot{\mathbf{e}}_q$ are used to define a correction torque $\mathbf{\Gamma}_{PID}$ of the motor in a Proportional Integrative Derivative corrector. A friction compensation torque $\mathbf{\Gamma}_f$ is computed using a viscous and Coulomb friction model with the desired joint velocity to anticipate the actuation torque lost in friction. A feed-forward term accounts for the dynamic and static wrenches exerted on the moving-platform to determine a compensation torque $\mathbf{\Gamma}_{FF}$. The friction compensation torque, the PID corrector torque and the feed-forward torque are summed up to define the actuation torque $\mathbf{\Gamma}$.

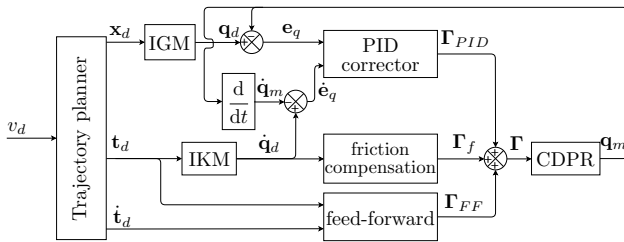


Figure 8. Control scheme of the CDPR used for foam additive manufacturing, the trajectory planner being defined based on the foam additive model (1)

3 Experimental validation

First, some layers were printed using a polyurethane foam with a density of 35kg/m³. The first tests with the

trajectory planner used in the CDPR allowed to test different parameters, in particular the printing speed versus its accuracy. The proposed framework for foam 3D printing with a CDPR is illustrated in Fig 9.

The developed measuring framework is used to improve the flatness surface and achieve the foam process through the implementation of a closed feedback loop. The principle is to keep the geometrical error along the z-axis smaller than 10 mm for one layer thickness. A first approach is to continuously vary the layer thickness proportionally to the measured 'z-error' by using the foam modeling described in Fig 7. This can be achieved either by defining the new operating parameters. It is also possible to adapt the nozzle spraying velocity v_d or to change the deposition height d_d of the nozzle online. The propose framework for foam printing with a CDPR is described in Fig. 9 for a dedicated high dimension parts to be printed. The quality of the foam to be deposited should satisfy some buildability (mechanical characteristics), geometry (shape) and flatness performance. Accordingly, the trajectory to be followed by the spraying end-effector should be defined and well followed.

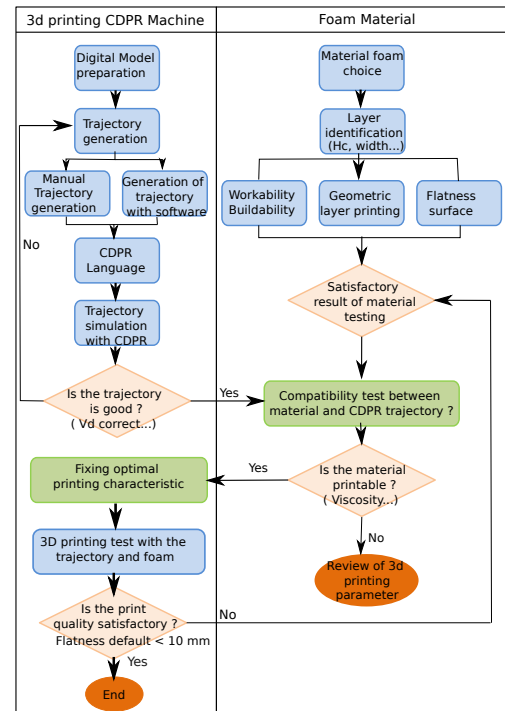


Figure 9. The proposed framework for foam printing with a CDPR

The first tests consisted in printing polylines in order to see the constitution of the robot to follow a straight trajectory, and to test the stacking of the layers. The CDPR well behaved at the tested speeds (up to 0.3m/s) in the

sense that the trajectory remains linear along a line of 1m long and the height d_d remains constant so that the layer thickness is homogeneous at 32 mm, and the layer width is also constant at 70mm (Figure 10 a). After, this experimentation shows a good repeatability of the CDPR trajectories regardless of the printing location in the robot workspace (Figure 10c). We note some widths of non homogeneous layer section, due to the accumulation of material formed at the starting and stopping points of the lines (Figure 10 b).

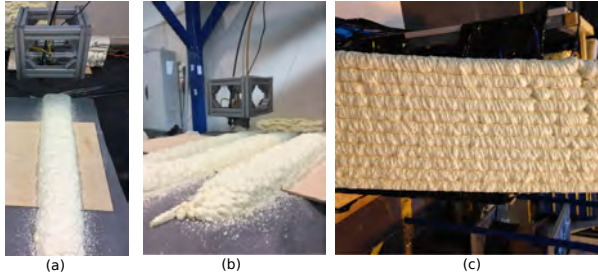


Figure 10. First experimentation polyline printing (a) , polyline with non homogenous layer section at the begin and end of layer (b), Layers overlay printing (c)

The second test involved a curved trajectory serving as the pattern for a curved wall. The pattern is 0.8m long and 0.25m wide with the oval shape.

The profile is set to the variables d , e and r , which correspond to the width, length of the profile and the radius of the rounding (see Figure 9) in different frames such as :

$R_{imp} = (O_{imp}, x_{imp}, y_{imp}, z_{imp})$: Printing frame whose origin is the first printing point.

$R_{str} = (O_{str}, x_{str}, y_{str}, z_{str})$: Structure frame whose origin is the center of the printed structure.

$R_p = (O_p, x_p, y_p, z_p)$: Moving-platform (MP) frame whose origin is the gravity center of the MP.

$R_b = (O_b, x_b, y_b, z_b)$: Base frame or Global frame related to the fixed frame.

To define the trajectory, the user defines:

- The desired time to move the effector from its initial position (at rest) to the 1st printing point
- The nominal speed of the effector for printing v_{dmax} (which should be kept during printing)
- Number of layers to print
- The height difference between two successive layers

The profile to be printed is defined in R_{imp} . The user places this local coordinate system in the structure frame R_{str} and then in the base reference regarding the robot R_b (Fig.10). To print this experimentation, two types of geometries are used: lines and circular curves. The methodology used for the structure definition is the following :

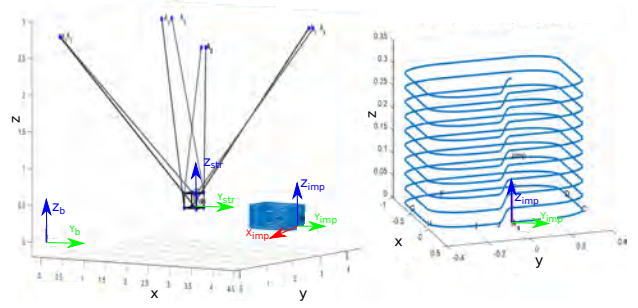


Figure 11. The shape trajectory inside the CDPR Workspace (a) , the printing shape trajectory (b)

- Define the position and orientation of the different frames
- Define the transformation matrixes between the different frames
- Define the shape to be printed in the printing frame R_{imp}
- This shape is first projected in the structure frame R_{str}
- It is then projected in the global frame R_b
- The movement of the moving-platform (MP) is defined by the trajectory.

When the trajectory is defined, the moving platform is suspended in its initial position, which was identified by a laser-tracker. The moving platform goes from the center of the structure frame to the initial printing point. It starts with zero velocity and acceleration and arrives to point A with the desired tracking velocity v_{dmax} (0.3 m/s). The spraying effector based on the moving platform follows the predefined structure shape. There is a vertical transition from one layer to the following. The velocity norm is kept constant, along the printing phase. When printing is finished, the MP goes from the last printed point to the structure center.

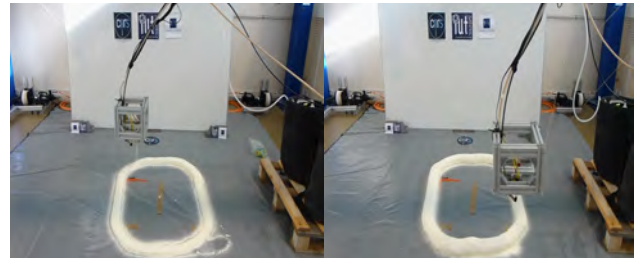


Figure 12. Construction of a foam shape part with the Craft CDPR

The path height and path width used for constructing the shape foam is 70 mm, respectively, the polyurethane foam characteristics. These dimensions were chosen based mainly on the objective to build as large an object as possible, in a reasonable amount of time and the quantity of foam in material tanks. With these dimensions, the shape wall is built in 10 layers and the flatness default is approximately 10 mm. However, other factors include a desire to maximize construction resolution and optimize foam gun performance.

4 Discussion on the results

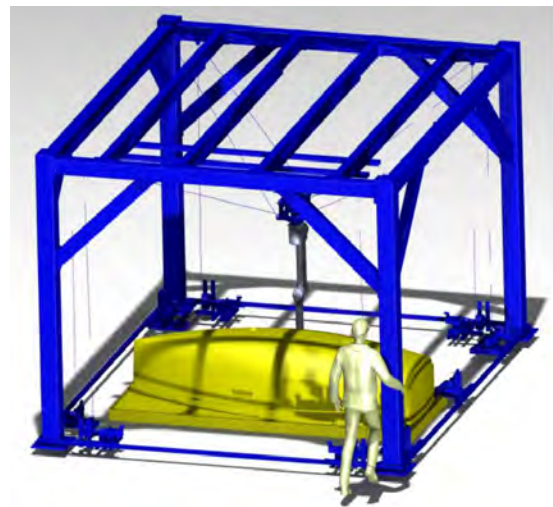
A prototype system, built at Nantes University, is presented on the CRAFT platform, and data from this system shows its suitability for large-scale printing. In order to scale this out to full-size deployment there are, however, different challenges associated with workspace shape and make correction of flatness surface print are identified as targets for future research. The success of this system demonstrates the feasibility of CDPR for large additive manufacturing systems for construction field.

The proposed novel robotic solution allows the association of the CDPR with a polyarticulated robot, embedded on its mobile platform, on which various end-effectors dedicated to additive manufacturing end-effector based on expanding material, edge finishing tool, measurement/control tool would be mounted (Fig.12).

Many research works have been conducted on the design, modeling and control of CDPR over the last fifteen years because of their numerous potential applications, especially in industry, and their potential improved performances, especially in terms of accuracy. It should be noted that the national platform "XXL Robotics" of the Equipex + TIRREX project (Technological Infrastructure for Robotics Research of Excellence 2021-2028) will be located at LS2N, Nantes University, site Mlab XXL in Saint-Aignan de Grandlieu. This platform will include a reconfigurable cable-driven parallel robot of size 24m x 14m x 6m, equipped with an embedded robotic arm on its moving-platform. This experimental platform will allow us to tackle many industrial applications related to manufacturing robotics and construction robotics such as 3D printing of large parts, namely, adding material, but also removing material such as machining.

5 Acknowledgements

This work was supported by the ANR CRAFT project, grant ANR-18-CE10-0004, <https://anr.fr/Project-ANR-18-CE10-0004> and the EquipEx + TIRREX project, grant ANR-21-ESRE-0015.



(a)



(b)

Figure 13. (a) Concept of a new CDPR used as a foam large scale printer; (b) reconfigurable CDPR (Nantes University and IRT Jules Verne) used in TIRREX-XXL platform

References

- [1] Eric Barnett and Clément Gosselin. Large-scale 3d printing with a cable-suspended robot. *Additive Manufacturing*, 7:27–44, 2015. jul.
- [2] Viktor Mechtcherine, Venkatesh Naidu Nerella, Frank Will, Mathias Näther, Jens Otto, and Martin Krause. Large-scale digital concrete construction—conprint3d concept for on-site, monolithic 3d-printing. *Automation in Construction*, 107:102933, 2020.

- 2019.
- [3] Steven Keating. Towards Site-Specific and Self-Sufficient Robotic Fabrication on Architectural Scales. URL https://www.media.mit.edu/publications/dcp_scirobotics/.
- [4] Benoît Furet, Philippe Poullain, and Sébastien Garnier. 3D printing for construction based on a complex wall of polymer-foam and concrete. Additive Manufacturing, 28:58–64, 2019. Num Pages: 7.
- [5] Elodie Paquet, Alain Bernard, Benoit Furet, Sébastien Garnier, and Sébastien Le Loch. Foam additive manufacturing technology: main characteristics and experiments for hull mold manufacturing. Rapid Prototyping Journal, 2021.
- [6] Jean-Baptiste Izard, Alexandre Dubor, Pierre-Elie Hervé, Edouard Cabay, David Culla, Mariola Rodriguez, and Mikel Barrado. On the improvements of a cable-driven parallel robot for achieving additive manufacturing for construction. In Cable-driven parallel robots, pages 353–363. Springer, 2018.
- [7] Julian Kaduk. Development of a computer vision based real-time feedback system for closed-loop control in 3d concrete printing, 2021.
- [8] Phillip Chesser. Design of a cable-driven manipulator for large-scale additive manufacturing. 2021.
- [9] Chad E. Duty, Vlastimil Kunc, Brett Compton, Brian Post, Donald Erdman, Rachel Smith, Randall Lind, Peter Lloyd, and Lonnie Love. Structure and mechanical behavior of Big Area Additive Manufacturing (BAAM) materials. Rapid Prototyping Journal, 23(1):181–189, January 2017. ISSN 1355-2546. doi:10.1108/RPJ-12-2015-0183. URL <https://doi.org/10.1108/RPJ-12-2015-0183>.
- [10] D Gueners, Helene Chanal, and BC Bouzgarrou. Stiffness optimization of a cable driven parallel robot for additive manufacturing. In 2020 IEEE International Conference on Robotics and Automation (ICRA), pages 843–849. IEEE, 2020.
- [11] Edouard Cabay, David Culla, Mariola Rodriguez, and Mikel Barrado. On the improvements of a cable-driven parallel robot for achieving additive manufacturing for construction. In Cable-Driven Parallel Robots: Proceedings of the Third International Conference on Cable-Driven Parallel Robots, volume 53, page 353. Springer, 2017.
- [12] Jean-Baptiste Izard, Alexandre Dubor, Pierre-Elie Hervé, Edouard Cabay, David Culla, Mariola Rodriguez, and Mikel Barrado. Large-scale 3d printing with cable-driven parallel robots. Construction Robotics, 1(1):69–76, 2017.
- [13] Ishan Chawla, PM Pathak, Leila Notash, AK Samantary, Qingguo Li, and UK Sharma. Workspace analysis and design of large-scale cable-driven printing robot considering cable mass and mobile platform orientation. Mechanism and Machine Theory, 165: 104426, 2021.
- [14] Phillip Chesser, Peter Wang, Joshua Vaughan, Randall Lind, and Brian Post. Kinematics of a cable-driven robotic platform for large-scale additive manufacturing. Journal of Mechanisms and Robotics, pages 1–17, 2021.
- [15] Lorenzo Gagliardini, Stéphane Caro, Marc Gouttefarde, and Alexis Girin. Discrete reconfiguration planning for Cable-Driven Parallel Robots. Mechanism and Machine Theory, 100:313–337, 2016. doi:10.1016/j.mechmachtheory.2016.02.014. URL <https://hal.archives-ouvertes.fr/hal-01400440>.
- [16] Damien Gueners, Belhassen-Chedli Bouzgarrou, and Hélène Chanal. Cable behavior influence on cable-driven parallel robots vibrations: experimental characterization and simulation. Journal of Mechanisms and Robotics, 13(4):041003, 2021.
- [17] Etienne Picard. Modeling and Robust Control of Cable-Driven Parallel Robots for Industrial Applications. PhD thesis, École centrale de Nantes, 2019.
- [18] Zane Zake, Stéphane Caro, Adolfo Suarez Roos, François Chaumette, and Nicolò Pedemonte. Stability analysis of pose-based visual servoing control of cable-driven parallel robots. In International Conference on Cable-Driven Parallel Robots, pages 73–84. Springer, 2019.
- [19] Marceau Métillon, Philippe Cardou, Kévin Subrin, Camilo Charron, and Stéphane Caro. A cable-driven parallel robot with full-circle end-effector rotations. Journal of Mechanisms and Robotics, 13(3):031017, 2021.

Automated material selection based on detected construction progress

T. Stephans¹, A. McClymonds¹, R. Leicht¹ and A.R. Wagner¹

¹The Pennsylvania State University, USA

tbs5111@psu.edu

Abstract -

One of the many ways in which automation may help the construction industry is on-site material management. This paper presents an automated process where materials are selected for staging by detecting construction progress from site images. The materials are then delivered to their respective workface locations by a robot. The effectiveness of the material selection process is assessed using a simulated and physical construction site. We demonstrate that our process is successful under a number of different conditions and environments. Our system contributes to the feasibility of autonomously managing materials on a construction site and reveals potential avenues for future research.

Keywords -

Automated material delivery; Progress detection; BIM; Robot

1 Introduction

This paper seeks to create a process that allows a robot to: a) recognize the current state of a construction project, b) predict which materials will be needed, and c) deliver those materials to their appropriate workface. The work presented here focuses on the first two of these components. Material management is an important part of any construction project. Consisting of the transportation, storage, identification, handling, and usage of materials, material management is directly related to efficiency, safety, and waste during construction. One case study found that the cost of inefficiencies resulting from improper material management was 5.7 times greater than the cost of implementing material management [1]. With respect to safety, Lipscomb et al. observed that 11.5% of slipping and tripping accidents on construction sites occurred while the worker was carrying an object [2]. Additional studies have found that human error, accidents, and damage during transportation to the site or on the site were all sources of construction material waste [3].

The use of robots for construction site material delivery could maximize the tracking of materials while also minimizing waste. Robots used at night or moving along dedicated pathways could also increase safety, provided

no workers are present at night. Hence, the goal of this research is to develop a system that can automatically detect construction progress and deliver the materials that will be needed for the next step in a construction plan.

The task of automating the tracking of construction progress is referred to as automated progress monitoring [4]. Research in this area typically compares a 3D model of the structure to be built (as-planned model) with another model generated using sensor data of the construction site (as-built model). The as-planned model is derived from a building information model (BIM), and contains physical and functional information about the structure. By comparing these models, individual elements of the structure such as walls, beams, or columns are recognized and their progress assessed. The expected progress is derived from a critical path schedule. In this paper however, no expected amount of progress is prescribed, and progress monitoring is used as a means to automate material selection for delivery via mobile robots. As such, we refer to this problem as automated progress *detection* as opposed to automated progress *monitoring*.

This paper presents an automated process where materials are selected for staging based on images collected by a mobile robot on a construction site. The system developed for selecting materials makes use of a pre-existing technique for automated progress monitoring. The output of this process may then be used to instruct a robot to deliver materials to specific workface locations where the material will be needed. The primary contribution of this paper is an automated process for guiding the timely distribution of construction materials.

In the next section we review related work before describing the process itself. Next, experiments demonstrate our process followed by an analysis of the experimental results. The paper concludes with directions for future work.

2 Related work

Recent advances in robotics have begun to make the use of construction robots a possibility [5]. Mobile robots however must cope with the cluttered and dynamic environments found in construction sites, which research in obstacle detection [6] and localization using BIM [7]

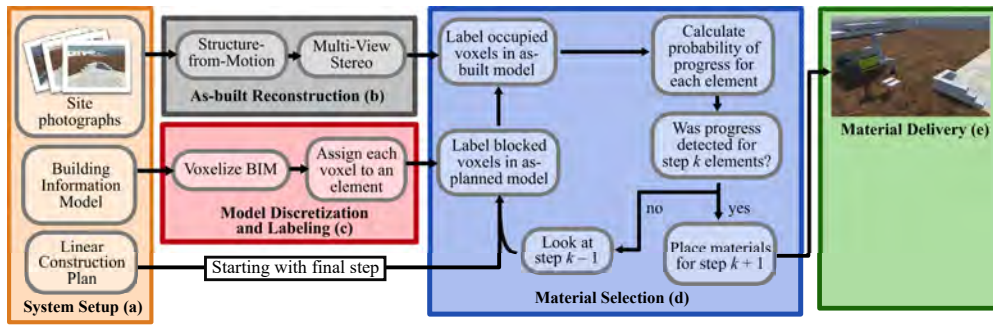


Figure 1. System Diagram: Input photographs, BIM, and construction plan are used to generate a list of materials to be delivered by a robot.

have begun to address. The capability of robots to place construction elements has also been investigated [8], and mobile manipulators capable of moving materials on construction sites have been demonstrated [9].

A great deal of research has focused on automated progress monitoring for construction sites [4, 10]. Attempts have been made to automate construction progress monitoring using either laser scanners [11, 12] or images [13, 14, 15] of the site to generate an as-built model. The model must then be aligned/registered with the BIM in order to compare them for determining progress. This typically involves a degree of manual manipulation, although studies have demonstrated automated techniques [16]. Some methods utilize machine learning to help in monitoring progress [13, 11, 15]. Others incorporate temporal information to take advantage of relationships between construction elements [11, 14]. Bayesian models [13] and computer vision have also been used to recognize construction materials [15].

Our work uses a method developed by Golparvar-Fard et al. [13]. In their process, unordered site images and a 4D BIM are used to automatically monitor the construction progress of a site. Images are used to reconstruct the as-built model by first using structure-from-motion (SfM) to create a sparse point cloud of feature points. Multi-view stereo (MVS) techniques are then used to create a dense point cloud. Once the model is aligned with the BIM, the model space is discretized into voxels to support the inference of progress over small areas of the construction site. The voxels are labeled for both the as-planned and as-built models based on their occupancy and visibility from the site images. These labels are then used to determine the probability of observing progress for each construction element by applying a Bayesian probabilistic approach. Finally, by rephrasing the problem as a linear classification problem, a support vector machine (SVM) is trained to calculate a probability threshold for each element and classify it as either having progressed or not progressed.

For the purposes of this paper, Golparvar-Fard et al.'s process is simplified by removing the SVM.

3 System description

Our process is depicted in Figure 1 and consists of five parts. These are system setup, as-built reconstruction, model discretization and labeling, material selection, and material delivery.

3.1 System setup

The inputs for our system can be found in box (a) in Figure 1, and are 1) site photographs, 2) a BIM, and 3) a linear construction plan. These inputs serve as the basis for generating the as-planned and as-built point cloud models, and when combined with a construction plan allow for materials to be selected via progress detection.

The critical characteristic of the BIM input is that it is a 3D model composed of uniquely named construction elements. Each element also has a material label, which can be used for generating a list of materials for a robot to deliver. As a medium for design, BIM does not always contain the level of detail assumed by this paper. For example, a brick wall would be a single 3D model with the individual bricks visually represented using a mapped texture. Before discretizing the BIM, we specify a volume of interest on the model space using Cartesian bounds $\mathbf{b}_{\text{lower}}$ and $\mathbf{b}_{\text{upper}}$. All 3D objects outside these bounds are removed from the as-planned model, such that only construction elements important for progress detection remain. The result is an as-planned model consisting of the set of elements $\theta = \{\theta_1, \dots, \theta_{N_{\text{elements}}}\}$.

To generate a list of materials for delivery, a plan is also required. When a construction schedule is generated, it is targeted for craft labor tasks, and thus it normally contains temporal information in the form of a task schedule, where multiple tasks may be performed concurrently. Our process assumes a more granular linear plan consisting of

a series of steps $k = 1, \dots, N_{\text{steps}}$, where one or more construction elements are placed at each step. In this paper, the word *step* specifically refers to a step in a linear plan and not some component of the material selection process.

The input images must provide enough information to discern the current step based on the plan without having to infer based on relationships between construction elements. In general, this means that the collection of images used will feature at least one element placed during the previous step and one space where an element will go in the following step. Additionally, the images must be devoid of any unplanned occlusions blocking construction elements. The only permissible occlusions are other construction elements. When collecting images in simulation, the precise location and orientation for each image was also provided. This was done to remove the need for manually aligning the as-built point cloud model with the as-planned model by selecting corresponding points, as was done for the physical experiments.

3.2 As-built reconstruction

In this paper, as-built reconstruction is the process of using site images to create a 3D point cloud of the construction site. As indicated in Figure 1 by box (b), the as-built point cloud model, denoted from here onward as as-built PC, is needed to detect progress and determine which construction elements have been placed.

The process of reconstructing the as-built PC from the provided camera images and poses is handled by COLMAP [17], an open source structure-from-motion (SfM) and multi-view stereo (MVS) pipeline. SfM compares identical features for a scene across multiple images to generate the scene's 3D geometry. The outputs from SfM are then used with MVS to map normal and depth information onto the images. This results in a dense point cloud P_b representing the as-built construction site, and is used as the as-built PC. Notice that the subscript b indicates a variable's association with the as-built PC.

3.3 Model discretization and labeling

The as-planned model is discretized into finite, fixed volume voxels of width δ for labeling (see Figure 1 box c). By dividing the as-planned model into smaller volumes, visibility can be reasoned per voxel for each element, thereby informing the progress detection which voxels are expected to contain points from P_b .

To voxelize the as-planned model, each element is converted into a point cloud P_e , where $e = 1, \dots, N_{\text{elements}}$ is an index for each construction element. Doing so allows each voxel to be associated with a particular element based on the P_e with the most points inside the voxel. This adds a requirement that each P_e must have the same density ρ of equally spaced points.

The as-planned model space is discretized based on combinations of P_e for all e with

$$V = f_{\text{voxelGrid}}\left(\bigcup_{e=1}^{N_{\text{elements}}} P_e, \delta, b_{\text{lower}}, b_{\text{upper}}\right) \quad (1)$$

where $f_{\text{voxelGrid}}()$ generates voxels of width δ , bounded between limits b_{lower} and b_{upper} , from a union of P_e point clouds. The result is a voxel grid V where voxels V_j for $j = 1, \dots, N_{\text{voxels}}$ are only placed where at least one point exists. Each V_j is provided two labels, for the as-planned model and as-built PC, and will have values of either empty (E), occupied (O), or blocked (B).

3.4 Material selection

The as-built PC is compared with the as-planned model at various steps per the linear plan. By starting with the final step and analyzing in reverse order, as shown in box (d) in Figure 1, the step which matches the as-built PC is detected. This process informs the current construction progress and determines what materials need to be delivered for the subsequent step. Due to this relationship, the terms *progress detection* and *material selection* are used mostly interchangeably in this paper.

To begin detecting progress, a new voxel grid $V_k \subseteq V$ for step k must be created. For details on how a voxel V_j is assigned to a voxel grid V_k , see Figure 2.

3.4.1 Label blocked voxels

Next, progress detection requires the blocked voxels in the as-planned model to be labeled. Only voxels expected to contain points from P_b should be checked for occupancy, so the voxels labeled as blocked are discounted.

Voxels are traversed and labeled in a manner similar to [13]. Readers are directed to the method outlined by Golparvar-Fard et al. for further details on voxel traversal and labeling blocked voxels based on the as-planned model. However, unlike Golparvar-Fard et al., we do not use the radiance of a voxel's projected pixels to evaluate if a voxel is occupied. This removes the need for a visibility constraint, and thereby enables a unique voxel label solution to be found even when considering one image at a time. Therefore, the order of voxel traversal is defined by

$$d_{i,j} = \|\mathbf{x}_j - \mathbf{x}_i\| \quad (2)$$

where \mathbf{x}_j and \mathbf{x}_i are the center of $V_{k,j}$ and image i respectively in the world coordinate frame. The voxels in V_k are then traversed in order of increasing distance $d_{i,j}$.

The as-planned label for $V_{k,j}$ is specified as (O_p) or (B_p) based on the number of reprojected pixels of $V_{k,j}$ with an evaluation of 1 as described by [13]. Here, the subscript p indicates a variable's association with the as-planned model. Further, $V_{k,j}$ can be labeled as (B_p) if $d_{i,j}$

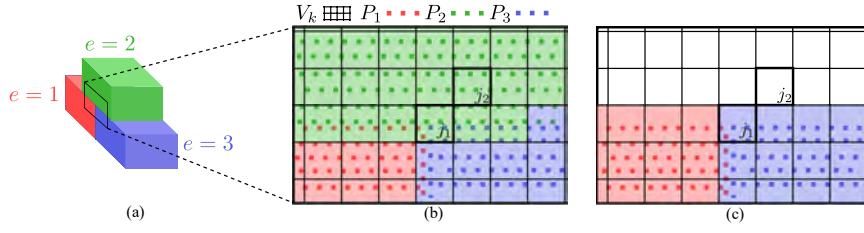


Figure 2. Model discretization and labeling visualized using a structure consisting of three elements: (a) Red, green and blue construction elements with $e = 1, 2, 3$ respectively. Current step is k . (b) 2D zoom of voxel grid V_k with point clouds P_1 , P_2 and P_3 . Voxels j_1 and j_2 have more points from P_2 than P_1 or P_3 , leading to $e_{j_1,k} = (2, 3, 1)$ and $e_{j_2,k} = (2)$. (c) 2D zoom of voxel grid V_{k-1} where element 2 was removed from the voxel grid and from $e_{j_1,k-1}$ and $e_{j_2,k-1}$, giving $e_{j_1,k-1} = (3, 1)$ and $e_{j_2,k-1} = ()$. Voxel j_1 now belongs to element $e = 3$.

is greater than some distance threshold as well. This was done primarily to reduce computation time.

3.4.2 Label occupied voxels

The voxels not labeled as blocked in the as-planned model, are labeled for occupancy based on the as-built PC. By counting the number of occupied and empty voxels for each element, the probability of observing each element is determined.

Because the assumption was made that no unplanned occlusions were present in the images used to reconstruct the as-built PC, all voxels labeled with (B_p) are also labeled (B_b) . Each voxel $V_{k,j}$ not labeled (B_b) is checked to see if it contains a point from P_b . If so, then it is labeled occupied (O_b) . However, when V is generated, it is possible for some voxel faces to be very close to an element face. Due to the limited precision of the reconstruction process, scenarios where points in P_b fall just outside of a voxel can be common. To address this problem for some empty voxel $V_{k,j}$, all voxels adjacent to $V_{k,j}$, which are not part of V_k , are also checked for occupancy. Here, two voxels are considered adjacent if they share a common face. If a point is found in any of the adjacent voxels, then $V_{k,j}$ is labeled (O_b) . Otherwise it is labeled empty (E_b) .

3.4.3 Plan traversal for material selection

Given a step in the linear plan, the probability of observing each element in that step is calculated. The first step found where every element placed in that step is expected to be observed, is said to be the current state of construction. By knowing the construction progress, the materials required are known and may be delivered.

Progress detection is performed to recognize element $\theta_{k,e}$ in θ_k , starting with $k = N_{\text{steps}}$. As with [13], progress is evaluated by comparing a probability to a threshold. Provided evidence of occupancy η_e , Golparvar-Fard et al.

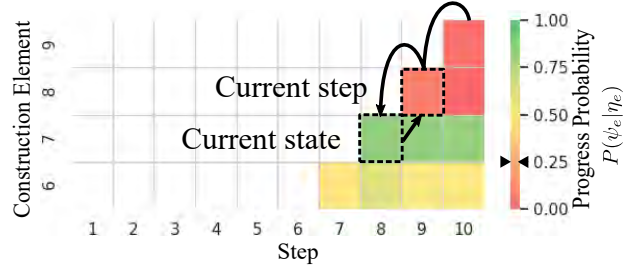


Figure 3. Example progress board from physical experiment. Columns represent steps of the plan, and rows represent construction elements organized by step. Each band of color within a cell, is the $P(\psi_e|\eta_e)$ of element e for whatever step k the system is on. No color (white) indicates that element is not visible. Here, each step involves placing one element, so there is a single band of color per cell. The process of determining the current step is visualized by the black arrow and outlined cells.

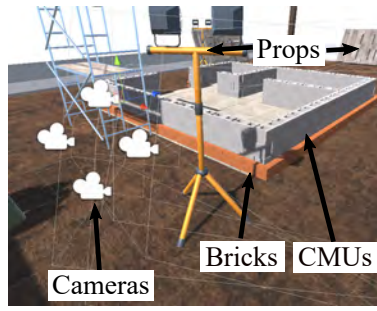
define the probability of ψ_e , the event of observing an expected element, as

$$P(\psi_e|\eta_e) = \left[\frac{\sum O_b}{\sum O_b + \sum E_b} \right]_{\theta_{k,e}} \quad (3)$$

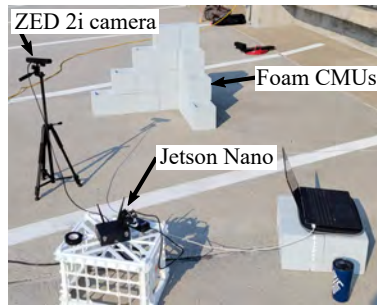
Given our assumptions, and a linear construction plan rather than a general schedule, $P(\psi_e|\eta_e)$ is satisfactory to detect $\theta_{k,e}$. If $P(\psi_e|\eta_e)$ is greater than the threshold, then the element is marked as having progressed.

After evaluating the progress for all elements in θ_k , the current step is said to be $k + 1$ if progress was detected for all elements placed in step k . Otherwise, progress detection is performed on step $k - 1$ as shown in Figure 1.

We now introduce the **progress board** in Figure 3 to better explain this process. A progress board functions as



(a) Simulation experimental setup



(b) Physical experimental setup

Figure 4. BIM, experiment environment, and example image capture for physical and simulation experiments

a recording of all steps the material selection analyzes. To find the elements placed on some step, we search the grid diagonal, moving from the top right cell to the bottom left, as shown in Figure 3.

Once the current step is determined, the elements for that step are chosen to be delivered by a robot. Because each element is associated with a material, the result is the robot receiving a list of materials to deliver.

4 Experiments

4.1 Simulation experiment

The BIM used in simulation was designed to resemble the foundation of a small structure which can be seen in Figure 4(a). Images were captured of the simulation at an interim step in the plan, and progress detection was used to determine the current step. Using the progress detection results, a list of materials for delivery was output. This process was evaluated at every step in the plan.

The BIM was generated as an FBX file, with unique names for each construction element. Materials were distinguished using specific tags in the element names. For this research, the two materials used were standard bricks and concrete masonry units (CMUs). The linear plan was designed such that elements were distributed across steps

Table 1. Camera properties

camera	width (px)	height (px)	focal length f (px)	principal pt. o_x (px)	principal pt. o_y (px)
simulation	2048	1536	1525.1	1024	768
ZED 2i	2208	1242	1063.6	1107.2	632.1

based on wall face and course. In total, the plan contained 57 steps to place 447 elements.

Unity was used for the material selection simulation. Unity offers a capable physics engine, cameras for simulating image capture, and importation of many assets into the simulation. Additional textures and construction props were added, as shown in Figure 4(a), to make the simulation look more realistic and improve the performance of the as-built reconstruction. It should be noted that if a prop was ever occluding a construction element in the camera view, it was removed for that as-built reconstruction. The lighting in the simulation was configured to represent noon at Penn State in early summer to balance realistic lighting and ease of reconstruction.

Uncompressed site images were captured using camera objects arranged in Unity. The cameras were arranged in groups of 4 for each reconstruction, as seen in Figure 4(a). The camera intrinsics were chosen to resemble a smartphone camera and are provided in Table 1. It should be noted that the simulated camera parameters differ from the physical camera because the simulation experiments were conducted before the ZED 2i camera was acquired. These camera parameters and poses were imported into COLMAP [17].

During reconstruction, COLMAP's default "low" quality settings were used to reduce computation time. Our process performed the material selection and returned a list of materials for delivery. When discretizing the BIM, the voxel width δ and point cloud density ρ were chosen to be 0.0254 m (1 in) and 15,500 pts/m² respectively. The threshold for progress to be detected for an element was set to 0.25, and the distance threshold for a voxel to be considered visible was set to 4 m. All parameters were chosen empirically.

4.2 Physical experiment

The BIM used for the physical experiments consisted of a small stack of CMUs arranged in a corner shape. The completed construction can be seen in Figure 4(b). Construction was conducted on top of a parking deck to best approximate the large concrete flooring found on many construction sites. Images were then captured using the Stereolabs ZED 2i camera while varying the current step,

Table 2. Material selection results across all elements in θ_k for k = current state, where $P_k(\psi_e|\eta_e)$ is the probability of an element placed during step k being observed

Setting	# reconstructions	Min $P_k(\psi_e \eta_e)$	Avg. $P_k(\psi_e \eta_e)$	Max $P_k(\psi_e \eta_e)$	# mistaken steps	# bricks not selected	# CMUs not selected
morning	8	0.429	0.744	0.995	0	0	0
afternoon	18	0.502	0.830	1	0	0	0
evening	8	0.497	0.702	0.881	0	0	0
simulation	57	0.289	0.831	1	0	0	0
sim. occluded	2	0	0.585	1	2	1	0

time of day, and position of camera. Using the collected images, progress detection was performed to assess if the correct materials could be selected based on progress.

Once again, the BIM for the physical experiments was generated as an FBX file. The actual CMUs used were foam blocks (Figure 4(b)). Foam CMUs were chosen for ease of handling in future robot experiments. Unlike the simulation experiments, the linear plan for the physical experiments only had 1 C MU placed per step, with a total of 10 steps.

The physical experiments were varied in three ways: lighting, camera position, and construction step. Lighting was varied by taking images in the morning, afternoon, and late evening. Additional variation due to weather occurred naturally. Sets of images were captured around the CMUs at approximately 90° apart. For each lighting condition and position, the C MU corner was constructed to steps 6 and 10. Images were also captured and processed for every step, but these were only gathered in the afternoon at a single position around the C MU corner. In total, 34 variations of physical experiment were performed.

A ZED 2i camera was mounted on a tripod to capture uncompressed images, and its rectified intrinsics can be found in Table 1. The ZED 2i is factory geometrically calibrated, and so no manual calibration was performed prior to the experiments. For each as-built reconstruction, the tripod's center column was extended in 0.06 m increments a total of 0.3 m to capture left and right images. This yielded 12 images per reconstruction, which will now be referred to as a set of images. A Python script was written to record the images, and was run using an NVIDIA Jetson Nano Developer Kit-B01. An example of an experiment setup is in Figure 4(b).

When processing the real images, COLMAP was set to its default "high" settings to increase the final point density of the reconstruction. Although the poses of the camera were not measured for use in COLMAP, the camera intrinsics were still provided using the ZED camera Python API. Again, SfM and MVS was performed using COLMAP's command line interface.

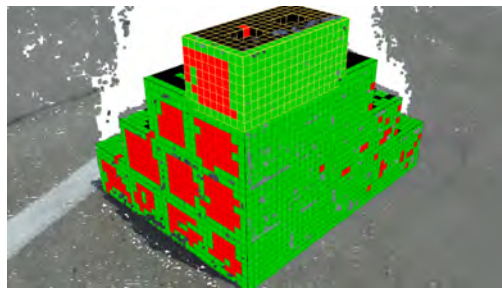
After reconstruction, because the camera poses were

not recorded, the as-built PC had to be aligned to the same coordinate frame as the BIM. CloudCompare's Align tool was used to generate a transformation matrix using the C MU vertices as a reference point. The matrix was used to rotate and scale both the as-built PC and reconstructed camera poses during progress detection. No changes were made to the material selection process. All parameters were left at their previous values.

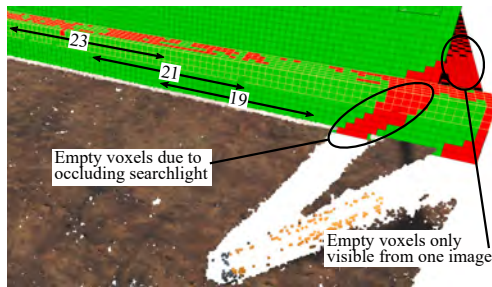
5 Results and discussion

Evaluation of our system in both simulation and physical environments resulted in the successful selection of the correct materials for delivery across all of the construction steps, camera locations, and lighting differences tested. As seen in Table 2, all minimum probabilities were above the threshold of 0.25 when there were no occlusions, so no materials were improperly selected. In practice, this means the material selection system will determine the correct materials to deliver for the construction site, provided ample lighting and images, and the workspace is free of obstructions. While the system was successful, some experiments do demonstrate how failures could occur.

Correct materials are not selected for delivery when elements placed during the current state of construction have too many voxels labeled as empty. In the experiments conducted through this research, voxels were incorrectly labeled as empty primarily because of poor lighting and occlusions. A representative example of the effect of lighting on progress detection can be seen in Figure 5(a). Even though the sun was nearly directly overhead, the brick faces in the shade were reconstructed noticeably worse than the faces in the sun. Table 2 shows the probability $P_k(\psi_e|\eta_e)$ that an element placed during the current state is observed was generally higher during the afternoon than morning or evening, although no materials were mistakenly selected. However, Table 2 shows both simulated reconstructions where occlusions were introduced resulted in the wrong steps being determined. Figure 5(b) depicts voxels labeled as empty (red) due to a spotlight leg occluding some of the bricks from the cameras. This resulted in the current state



(a) Foam CMU corner at noon



(b) Simulated experiment with an occlusion

Figure 5. As-built reconstruction point cloud and voxel labels: red = empty, black = blocked, green = occupied, yellow grid lines = current construction elements

being detected as 19 instead of 23. The reason for this can be observed in Figure 6. Looking at step 23 (actual current state) on the diagonal, red bands can be seen in the cell. There are multiple bands in each cell because multiple elements are placed per step. The red bands indicate that some elements were not detected, and explains why progress detection continued searching for the current step. Further down the diagonal, the elements in steps 21 and 19 were detected fairly well. In Figure 5(b) it is shown that these steps represent the brick courses directly below step 23. Therefore, step 19 was deemed the current state because it was reconstructed enough to be detected as opposed to step 21. The consequence of this error is that the number of bricks selected for delivery would be based on step 19 instead of step 23. For these two steps, this is only the difference between delivering 8 and 7 bricks, but the error could be greater or smaller if the mistake was between a different pair of steps.

Two other limitations of this material selection system are its reliance on detailed construction schedules and as-planned to as-built model alignment. As stated in subsection 3.1, this paper uses a BIM and linear plan with an uncommon level of detail. For a system such as the one presented here to be used on a real construction site, detailed schedules for element placement will need to be

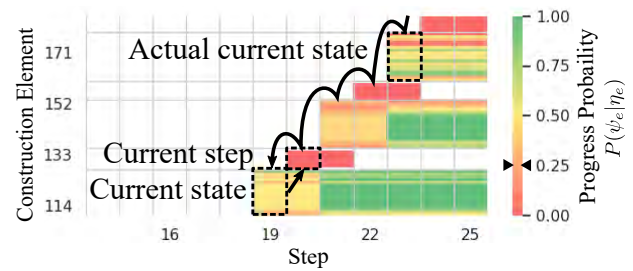


Figure 6. Progress board for simulation at step 23

incorporated into BIM, or methods will need to be developed to reverse engineer the information from current models. Assuming such scheduling data is available, care must still be placed into aligning the as-built PC with the BIM coordinate frame, as errors will be magnified by small misalignments for large models. This concern can be reduced by collecting more images and by automating as-built model alignment using GPS or registering images with on-site control markers [16].

6 Conclusion

This paper has demonstrated a process for automated material selection based on detecting construction progress from robot captured images. Our system was evaluated in simulation and at a simulated, physical construction site. The material selection process was shown to be robust to lighting and is capable of detecting any step in a linear plan despite limitations related to model alignment and occlusions. This paper contributes a progress detection system capable of selecting the necessary materials for further construction progress. The system presented serves as a template for future advancement, which will allow workers to spend less time documenting progress and moving materials, and more time helping the construction industry reach its productivity goals.

Future research aims to address the noted limitations and investigate the challenges of detecting progress on a real construction site. These challenges include scalability, large varieties of construction elements, non-linear construction schedules, and the generally more complex and varied environment that a construction site poses.

7 Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant Number 1928626. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We would like to thank Adith Calix for her help in conducting the physical experiments.

References

- [1] H. Randolph Thomas, Victor E. Sanvido, and Steve R. Sanders. Impact of material management on productivity—a case study. *Journal of Construction Engineering and Management*, 115(3):370–384, 1989. doi:10.1061/(ASCE)0733-9364(1989)115:3(370).
- [2] Hester J. Lipscomb, Judith E. Glazner, Jessica Bondy, Kenneth Guarini, and Dennis Lezotte. Injuries from slips and trips in construction. *Applied ergonomics*, 37(3):267–274, 2006.
- [3] B. A. G. Bossink and H. J. H. Brouwers. Construction waste: Quantification and source evaluation. *Journal of Construction Engineering and Management*, 122(1):55–60, 1996. doi:10.1061/(ASCE)0733-9364(1996)122:1(55).
- [4] Marianna Kopsida, Ioannis Brilakis, and Patricio Antonio Vela. A review of automated construction progress monitoring and inspection methods. In *Proc. of the 32nd CIB W78 Conference 2015*, pages 421–431, 2015.
- [5] Mi Pan, Thomas Linner, Wei Pan, Hui-min Cheng, and Thomas Bock. Influencing factors of the future utilisation of construction robots for buildings: A hong kong perspective. *Journal of Building Engineering*, 30:101220, 2020.
- [6] Pileun Kim, Jingdao Chen, Jitae Kim, and Yong K. Cho. SLAM-driven intelligent autonomous mobile robot navigation for construction applications. In *Workshop of the European Group for Intelligent Computing in Engineering*, pages 254–269. Springer, 2018.
- [7] JeeWoong Park, Yong K. Cho, and Diego Martinez. A BIM and UWB integrated mobile robot navigation system for indoor position tracking applications. *Journal of Construction Engineering and Project Management*, 6(2):30–39, 2016.
- [8] Chen Feng, Yong Xiao, Aaron Willette, Wes McGee, and Vineet R. Kamat. Vision guided autonomous robotic assembly and as-built scanning on unstructured construction sites. *Automation in Construction*, 59:128–138, 2015.
- [9] Petr Štibinger, George Broughton, Filip Majer, Zdeněk Rozsypálek, Anthony Wang, Kshitij Jindal, Alex Zhou, Dinesh Thakur, Giuseppe Loianno, Tomáš Krajník, and Martin Saska. Mobile manipulator for autonomous localization, grasping and precise placement of construction material in a semi-structured environment. *IEEE Robotics and Automation Letters*, 6(2):2595–2602, 2021. doi:10.1109/LRA.2021.3061377.
- [10] Jun Yang, Man-Woo Park, Patricio A. Vela, and Mani Golparvar-Fard. Construction performance monitoring via still images, time-lapse photos, and video streams: Now, tomorrow, and the future. *Advanced Engineering Informatics*, 29(2):211–224, 2015.
- [11] Changmin Kim, Hyojoo Son, and Changwan Kim. Automated construction progress measurement using a 4D building information model and 3D data. *Automation in Construction*, 31:75–82, 2013.
- [12] Frédéric Bosché. Automated recognition of 3D cad model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction. *Advanced engineering informatics*, 24(1):107–118, 2010.
- [13] Mani Golparvar-Fard, Feniosky Peña-Mora, and Silvio Savarese. Automated progress monitoring using unordered daily construction photographs and IFC-based building information models. *Journal of Computing in Civil Engineering*, 29(1):04014025, 2015. doi:10.1061/(ASCE)CP.1943-5487.0000205.
- [14] Alexander Braun, Sebastian Tuttas, André Bormann, and Uwe Stilla. Automated progress monitoring based on photogrammetric point clouds and precedence relationship graphs. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, volume 32, page 1. IAARC Publications, 2015.
- [15] Jacob J. Lin, Jae Yong Lee, and Mani Golparvar-Fard. Exploring the potential of image-based 3D geometry and appearance reasoning for automated construction progress monitoring. In *Computing in Civil Engineering 2019: Data, Sensing, and Analytics*, pages 162–170. American Society of Civil Engineers Reston, VA, 2019.
- [16] Sebastian Tuttas, Alexander Braun, André Bormann, and Uwe Stilla. Acquisition and consecutive registration of photogrammetric point clouds for construction progress monitoring using a 4D BIM. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 85(1):3–15, Feb 2017. ISSN 2363-7145. doi:10.1007/s41064-016-0002-z.
- [17] Johannes Lutz Schönberger and Jan-Michael Frahm. Structure-from-motion revisited. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4104–4113, 2016.

Solution Kits for automated and robotic façade upgrading

K. Iturralde^a, P. Zimmermann^a, R.G.B. Santos^a, S. Das^a, W. Shen^a, M. Malik^a, S. Maio^a, A. Hidri^a, J. Shen^a, T. Bock^a

^a Chair of Building Realization and Robotics, School of Engineering and Design, Technical University of Munich, Germany

E-mail: kepa.iturralde@br2.ar.tum.de

Abstract –

In order to reach zero energy consumption of the current building stock, adding a new insulating envelope with Renewable Energy Sources onto the existing building is necessary. This can be achieved by using prefabricated modules, automation and robotics. Since the topic is complex, three main subcategories were defined: Data Flow, Off-site Manufacturing, and On-site Installation. Latest studies suggest that there are still gaps in the way of achieving economically feasible solutions. In other words, there must be a reduction in the working time achieved in each sub-category. In this paper four different solutions are explained: 1) online data processing of the building, 2) automated layout definition, 3) accurate measurement with targets and 4) Automated CAM generation and Manufacturing. The solutions are still being developed but the current results are promising.

Keywords –

Renovation; Prefabrication; Accuracy; Time

1 Introduction

One of the measures to cope with the Climate Change is by reducing globally the necessary amount of energy. Over recent years, strategies for achieving zero-energy consumption of buildings focused on insulating existing buildings and installing Renewable Energy Sources (RES) on top [1]. Applying this process manually implies two hazards: building user's intrusiveness and disturbances and risky activities carried out at heights. To avoid such inconveniences, prefabricated modules are manufactured off-site. These modules include insulation, RES, and other necessary components such as windows or waterproofing elements [2]. In previous research of automated, robotic façade renovation with prefabricated module aspects [3–5], three sub-categories were defined, as explained in Figure 1.

- SC1. Information or Data Flow. This section is related to the data acquisition of the existing building, the processing of the data and the

definition of the prefabricated modules that are attached to the existing building. This information is used for the manufacturing of the modules, basically for creating a CAM (Computer Aided Manufacturing) for an off-site, manufacturing process. Besides, on-site, it is necessary to mark the location of the connectors on the existing wall.

- SC2. Off-site Manufacturing of the modules, which refers to the off-site automated manufacturing of the modules. Two phases can be differentiated; SC2.1, cut and machine the elements of the prefabricated modules, and SC2.2, assembly the module with the elements cut on previous point.
- SC3. On-site Installation of the modules. It deals with the process that spans from the arrival of the module to the site and the fixation of the modules to the existing building-wall. In robotic installation processes, the robot must be set up [6]. Apart from that, the connectors must be fixed and finally the modules need to be placed on top.

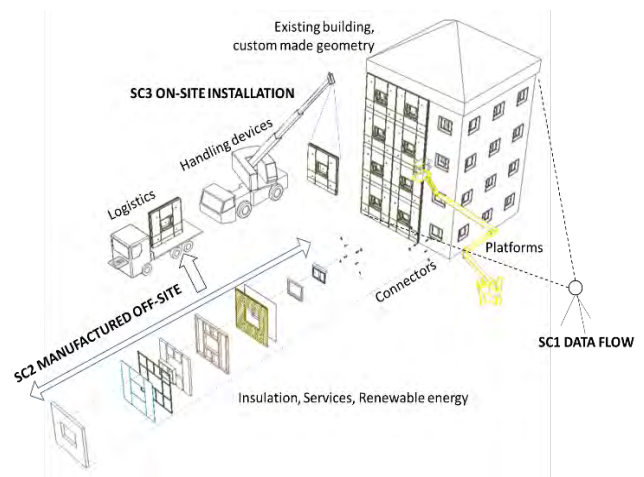


Figure 1. Scheme of the three sub-categories [6].

However, processes with prefabricated elements have not become competitive comparing to manual methods, which is the largest obstacle towards marketization. For this reason, a 90% reduction in data acquisition and

processing was envisioned as one of the goals of the ENSNARE research project [7] where the research explained in this paper is based on. Shortening the current manufacturing and installation processes is an additional objective too.

However, working time reduction cannot come at the expense of losing quality, and regulation and standards must be fulfilled [8]. Data flow errors can lead to manufacturing modules with deviations. Minor deviations can lead to water and heat leaks, while larger deviations may cause collisions, or even jeopardize the installation of a module in its assigned location. Therefore, deviations of less than 2 mm are necessary to obtain an optimal result.

2 Research Gaps

Robotics and automated solutions in construction are only marketed when these are economically feasible [9–12]. The question of how to manage technology and make it ready for the market must also be addressed [13–15]. This topic has been approached in some different research projects such as BERTIM [16], and HEPHAESTUS [17]. In previous instances [18], up to 15 Research Gaps (RG) were detected. The research in this paper focuses on four RGs and their solutions. With respect to Data Flow, it was determined necessary to:

- RG1.1: Avoid recurrent data acquisition of the building. Online data for the initial measurements of the building should be explored, in order to avoid excessive visits to the existing building, especially on preliminary stages.
- RG1.2: Define automatically the layout definition of the modules. Normally, in building renovation, the preliminary definition of the module layout is defined manually, and it can take up to 15-20 hours in a low-rise building.
- RG1.3: Discriminate “unnecessary” information of the building, and determine the locations of the connectors using photogrammetry.

On the manufacturing process, the remaining RG considers how to:

- RG2: automate agile robot path adjustment depending on the layout-CAD of the modules. It is of a vital importance to generate a parametric adjustment of the robot path based on the CAD file of the module.

The remainder of this paper focuses on the developed solutions to these RGs.

3 Solution Kits

In order to improve the current process, a set of Solution Kits (SK) are being developed in the ENSNARE project. SK1 deals with online data acquisition, building definition, and online module layout definition. On the

other hand, SK2 is related to the manufacturing process.

3.1 SK 1.1. Online data processing of the building

Online tools, such as *OpenStreetMaps*, were used for a preliminary semi-automated data acquisition and the subsequent initial building modelling. For that purpose, an algorithm was developed to generate semi-automatically the shape of the building (see Figure 2). Computational design tools and software, such as *FreeCAD* [19] were used to merge information taken from online databases.

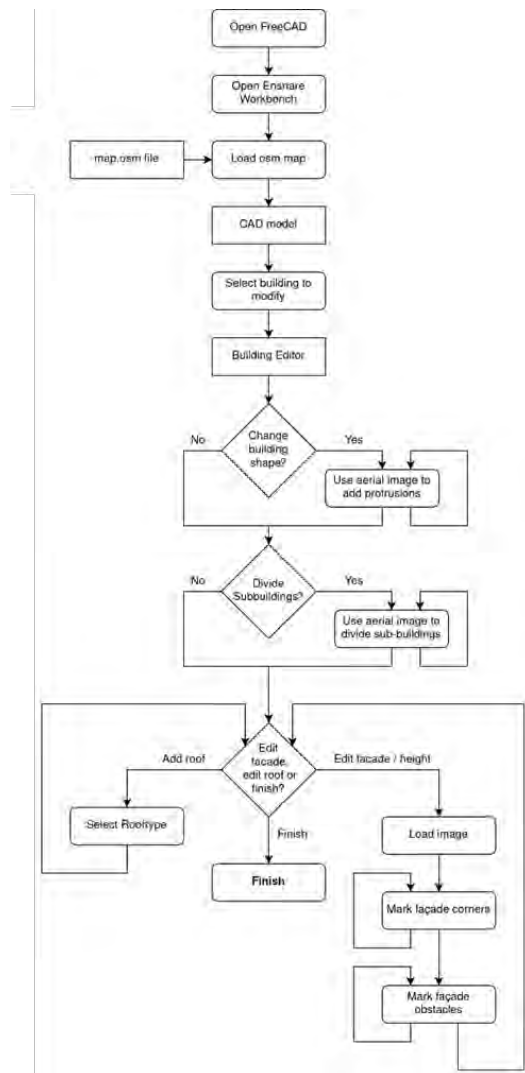


Figure 2. Scheme of the online data processing.

First, the user exports a map section from *OpenStreetMaps* as an XML file. The user then opens *FreeCAD* and uses the Load .osm file command to select the exported map file. The command creates a CAD object for each building in the OSM file with the correct

layout (Figure 3). However, at this point the height of the building and the roof shape are unknown and height is estimated from the number of floors.



Figure 3. A map section in *OpenStreetMaps* (left) and the same map section after import in *FreeCAD* as 3D model (right).

In step two, the user selects a façade in the 3D view and uses the *Adjust façade* command. The user then selects a previously taken photo of the façade which is opened in a new window. In this window the corners of the façade can be marked by the user. With the corners given and by assuming a rectangular façade, we are able to locate the vanishing points. Through a series of geometrical operations, we are then able to revert the projection of the façade and determine the original proportion between width and height. Since the length of the façade is known from the *OpenStreetMaps* data we can multiply it with the proportion to calculate the building height. The height is then adjusted in the 3D model (see Figure 4).

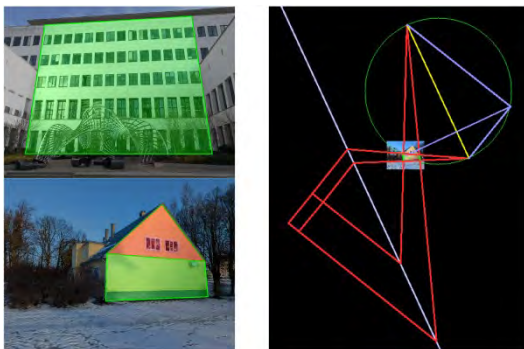


Figure 4. Left, façade photos with marked façade and gable area. Right, the algorithm for computing the height-width proportion from a façade image.

In step three, a perspective transformation is performed on the façade area to make it rectangular. In this transformed façade image, the user draws bounding boxes around façade objects and selects the appropriate object type, for example door or window. A grid

selection mode is available for marking many objects in one step since objects like windows are frequently positioned in a grid. When the user has marked all objects on the façade, they can finish the step and the objects are added to the 3D model (see Figure 5).

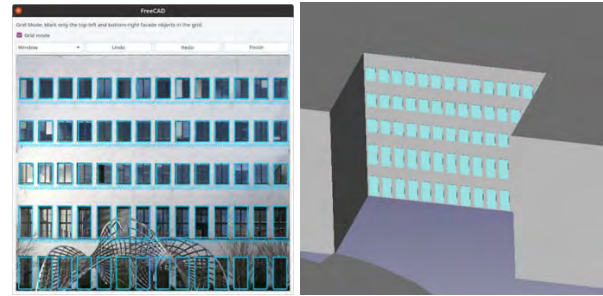


Figure 5. Marked windows with the grid function on transformed façade image.

When the user marks more than four corners for the façade, the additional points are interpreted as roof points. The roof shape is then extruded along the length of the building to create the roof in the 3D model (see Figure 6).



Figure 6. Different types of marked façade elements and pitched roof.

With the process described above, the building model can be defined. Several tests have been carried out, with satisfactory results, as shown in Figures 5 and 6. The difference between the obtained data and the actually measured building sizes ranges between 20 and 40 cm, depending on the size of the building. The differences are tolerable because SK1.1 is supposed to be used at the first stages of the façade renovation project. The measurements at this stage are used only for an estimation. With this building model, the layout of the modules can be determined automatically, as described in the next step.

3.2 SK1.2. Automated layout definition

The main focus of the prefabricated facade module model is that it is parametric. This means that a library of solutions can be used for various facade typologies and adjust accordingly. In order to have more flexibility,

there are also criteria that the parameters must fulfill. These define the constraints of the model, such as:

- floor height of the given building
- window dimensions (if a window is present)
- vertical distance of the window from the floor
- solar panel dimensions

The foundations of the developed *FreeCAD* model are a sketch and a spreadsheet. In the sketch an abstracted two-dimensional drawing defines and visualizes the placement and size of the elements of a module. The desired placement and size is controlled through adjustable parameters, which in *FreeCAD* are addressed as constraints. Most of these constraints are defined in the spreadsheet. The rest are referenced in the sketch. Constraints (such as the height of the window) are first accessed through the spreadsheet, where they are given a value, then referenced with an alias, and finally linked in the constraints of the sketch in the form of a formula. This way of linking the data offers a parametric workflow. For modifying the constraints, it is only necessary to change the value of that constraint in the spreadsheet, and the model will adjust accordingly.

Regarding the solar panel's dimensions, several panel sizes are selected for this project. Their dimensions are listed in the spreadsheet. From case to case the right size of the solar panel for the module needs to be chosen; afterwards its dimensions need to be selected as values for the constraints.

The three-dimensional module model is comprised of several building elements belonging to the back frame and the panels on top. This is optimized for the later module assembly.

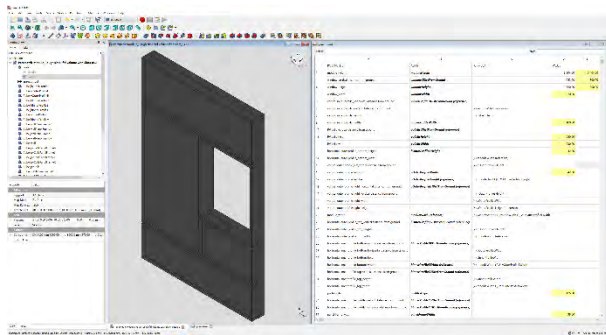


Figure 7. Different types of marked façade elements and pitched roof.

Currently the scenario in which the facade symmetrically admits modules which are all geometrically equivalent to each other is being developed. In this scenario the problem simplifies to placing a solar panel and accompanying registration area within just one module since this will be used to reconstruct the rest of the facade. Tackling more complex facades is a future

task.

It is necessary to abstract the facade, module, solar panel, registration, and window as 2D regions in the plane. Under this abstraction the placement of the solar panel and registration area is a packing problem with the additional goal of maximizing the area of placed solar panel.

A module is represented as a rectangular region in the plane whose lower left vertex is at the origin as shown in Figure 8. The other features (the window, the solar panel, and the registration area) are rectangular regions subjected to the following constraints:

- A feature must be contained within the module.
- No two features may overlap.
- The registration area and solar panel must maintain some proximity.
- The registration area must be placed in a way that it can be reached by a neighboring module.

The approach is a greedy method that finds the first feasible configuration for a given solar panel. To ensure maximal surface area the approach starts with the solar panel of maximal area, and then moves on to the solar panel of second maximal area, and so on until all possible solar panels have been exhausted.

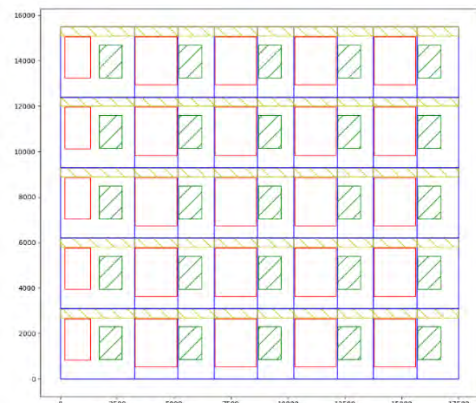


Figure 8. Automated maximization of the solar panels (red) and sub-division of the modules.

The implementation of the approach is done using python (version 3.6.9) without any external libraries. The solution is found by the function *find_optimal_placement* which takes the following inputs:

- the length, and height of the module
- the local coordinates of the bottom left corner of the window, its length, and its height
- a list of dimensions of potential solar panel
- a list of dimensions of possible registration areas
- a margin that specifies distance between features,

and distance features and the boundary of the module

- a step size which controls how many configurations are tested.

With the definition of the building module in SK1.1 and module layout definition described in this SK 1.2, the information is determined with high tolerances (see Figure 9).

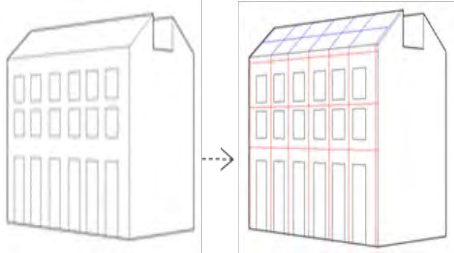


Figure 9. Automated generation of the layout.

The next section SK1.3 is based on the on-site accurate data acquisition of the building.

3.3 SK1.3. Accurate measurement with targets

In solution SK1.3, the so-called Matching Kit [20] is being redeveloped and improved. The Matching Kit consists of using the effectively measurement, markers, and using customized interfaces for adapting to existing buildings geometry. It has three main parts.

Part1 consists of targets that are placed on the existing building and gives information not only on the location, but also on the inclination of the wall. The Part1s will be placed on the existing building according to the preliminary layout developed in SK1.2 with an accuracy of 20 mm.

Part2 is located onto the prefabricated module depending on the location of Part1. In between both parts, an interface is placed which corrects the possible deviation of Part1 in regard to the planned position.

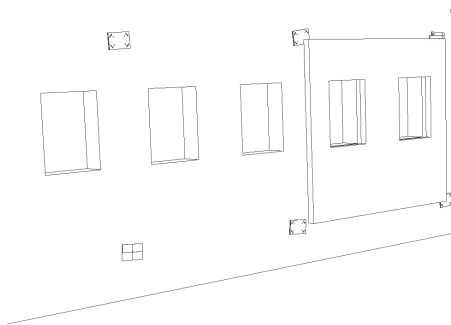


Figure 10. Part1s (markers) and the prefabricated modules.

The objective is to place and measure the markers in less than 8 hours for a low-rise building. In previous instances, Part1 was placed manually. In the project ENSNARE, a solution is being developed to place the Part1s by using drones, as shown in Figure 11.

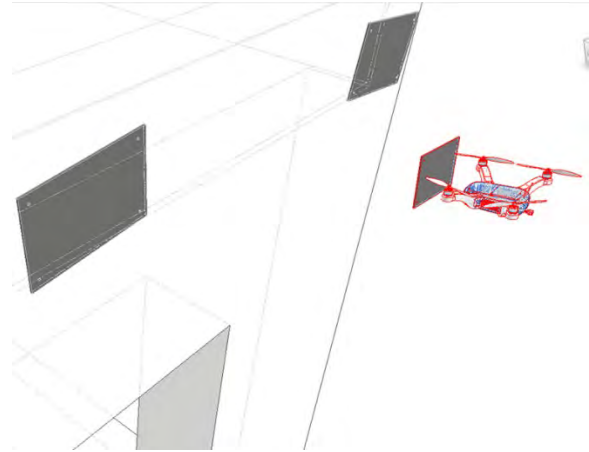


Figure 11. Placement of targets with drones.

In order to obtain the exact locations and angles of the Part1s, *AprilTag* markers are used and these are measured with photogrammetry methods. The 3D coordinates of the targets are estimated by an algorithm, with which the location and direction of each Part1 are obtained. Our goal is to measure the exact position and angle of each marker, and the error of the marker should not exceed 1mm at a distance of about 10m away from the building.

The following shows the entire process:

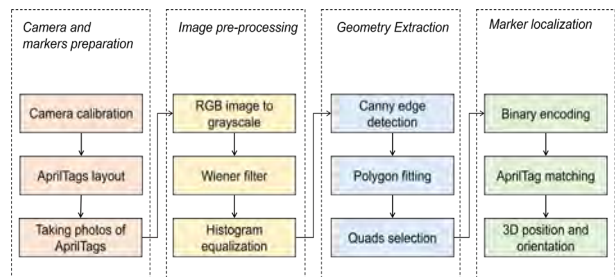


Figure 12. Scheme of the Part1s detection.

Before the measurement, the calibration of the camera is required. Zhang's calibration method [21] and Matlab's Camera Calibrator application to calibrate the camera [22] were used. Firstly, several photos of a 10 by 7 chessboard are taken by a digital camera from different angles. Then, those photos are imported into the calibrator, and after computing, the focal length, principal points, and distortion of the camera as a calibration matrix are determined. Next is the target localization. *AprilTag* is a visual targeting system that is

useful for various tasks, including augmented reality, robotics, and camera calibration. Targets can be created with a common printer, and the *AprilTag* detection algorithm calculates the precise 3D position, orientation, and identity of the tag relative to the camera [23]. The detection algorithm has following steps: the color photos with targets are converted to grayscale, then the noises is reduced by an adaptive low-pass Wiener filter [24]. To make sure that the images are not too bright or too dark, the contrast using histogram equalization is enhanced. After that, we use the Canny edge detection algorithm to get the edges [25], fit the resulting edges into several closed polygons, and filter out the unqualified polygons, leaving only the quads. The candidate quads are binary encoded and if they match successfully with the *AprilTag* library, their IDs and the pixel positions of the 4 corners can be output.

Finally, we use PNP (perspective-n-point) method to solve the 3D position and orientation of each marker according to the 4 corners, marker size, and the camera calibration matrix.

With the information obtained in SK1.3, the data from SK1.2 is adjusted to a more accurate building description and therefore, low tolerance manufacturing can be achieved as explained on the next section SK2.

1.1 SK2: Automated CAM generation and Manufacturing

Once the data is processed and the prefabricated modules are defined as described in SK1.2 and SK1.3, this information is transformed to the robotic workstation that is used for producing and assembling modules. The idea behind it is to adjust the robot's path to the different geometry of each module.

As explained before, the geometry of the prefabricated modules varies, all are different. For this reason, for each of the prefabricated modules, a specific file layout information is used to define the coordinates of picking and placing of the robot.

First, for the CAM generation, a workbench named *viz. ENSNARE_CAD_CAM* was created. In this workbench, properties (placement and center of mass) of individual profiles of an entire module can be transported as a single *json* file.

The *FreeCAD* simulation window is shown in Figure 13 when some of the profiles of the module have been selected.

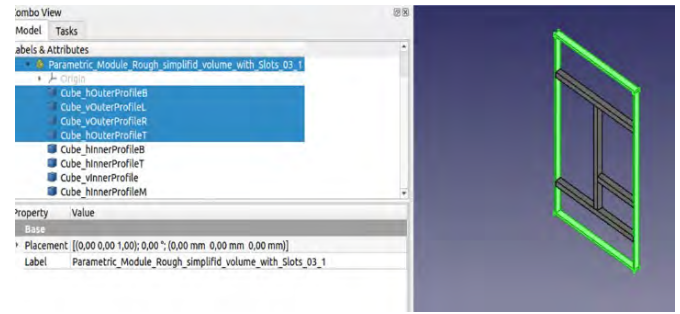


Figure 13. *FreeCAD* simulation.

Then, a Robot Operating System (ROS [26]) workspace was created which serves the purpose of providing the sequence of assembly of the profiles and the placement of individual profiles.

The process of retrieving the assembly is an automated process, which involves applying a simple nearest neighbor algorithm to select the next profile to be assembled. Once the sequence in which the profiles need to be assembled is known from the previous service, the placement, i.e. the position, orientation and the center of mass of individual profiles is obtained through this service.

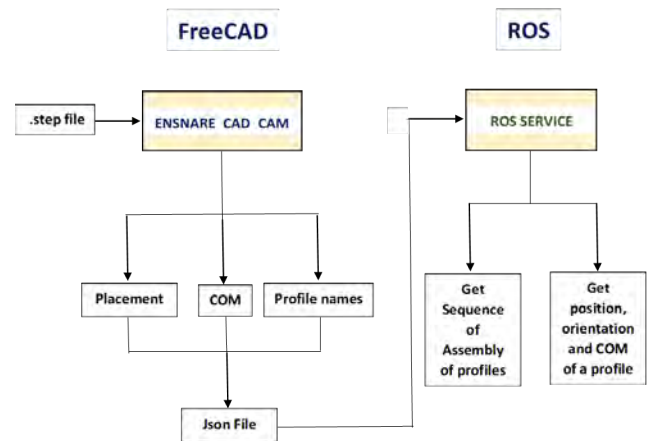


Figure 14. CAD-CAM synchronization scheme.

The robotic system mainly consists of a robotic arm (Universal Robot UR10e [27]) and a linear axis system. The base of the robotic arm is attached to the linear axis system which extends the reachability of the robotic arm. The linear axis system uses a stepper motor in combination with a gearbox to have enough torque to move the manipulator.

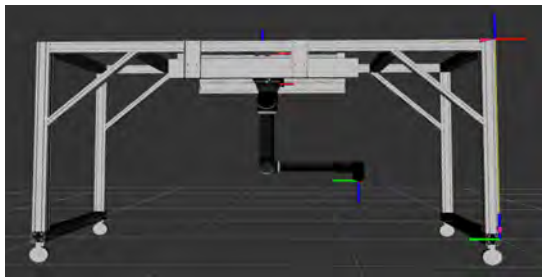


Figure 15. ROS simulation.

A Beckhoff PLC (Programmable Logic Controller) [28] is connected to the motor and offers an integrated position controller. Additionally, some sensors are used to define the limits of the linear axis. The below shows the configuration of the system.

The trajectory of the robotic arm and the pick and place operations are planned and executed by the motion planner *MoveIt* in ROS, which also includes collision checking. While the control of the manipulator is achieved by *MoveIt*, the motor in the rail system is controlled with a Beckhoff PLC. To synchronize the motion of both parts, a communication interface is created, where motion commands can be sent from a ROS node to the PLC machine, and the feedback position of the motor can be fetched from the PLC for a more accurate planning in *MoveIt*. The following Figure 16 shows the control scheme of the system.

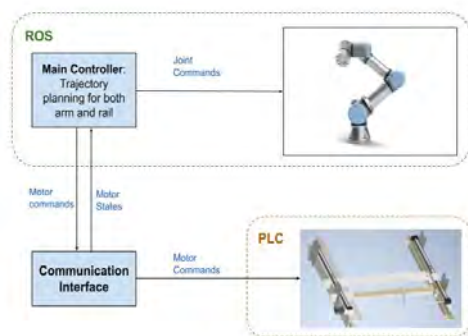


Figure 16. Control system of the robotic workstation

4 Conclusion

In this paper, a set of solutions for the automation of the façade renovation process have been explained. The solutions described in this paper are being developed currently. However, SK1.1 is almost finished and the building model can be generated by using online tools. This is a remarkable advance regarding building renovation with prefabricated models. The rest of the SKs, such as the robotic workstation in SK2 in Figure 17 still need more development and experimentations.

Moreover, the research gaps defined in [18] still need further revision and update.



Figure 17. Robotic workstation.

However, it must be remarked that if the solutions are developed as planned, close to market technology will be ready.

Acknowledgements



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 958445.

References

- [1] Garay Martinez, R., Astudillo Larraz, J., Performance assessment of façade integrated glazed air solar thermal collectors, *Energy Procedia*. (2017) pp 353–360.
- [2] Barco-Santa, A.F., Vareilles, É., Gaborit, P., Aldanondo, M., Building renovation adopts mass customization: {Configuring} insulating envelopes, *Journal of Intelligent Information Systems*. (2017) pp 119–146.
- [3] Tsai, I.C., Kim, Y., Seike, T., Decision-making consideration in energy-conservation retrofitting strategy for the opening of existing building in Taiwan, *AIJ J. Technol. Des.* (2017). <https://doi.org/http://doi.org/10.3130/aijt.23.963>.
- [4] Du, H., Huang, P., Jones, P.J., Modular facade retrofit with renewable energy technologies: the definition and current status in Europe, *Energy and Buildings* (205). (2019). <https://doi.org/https://doi.org/10.1016/j.enbuild.2019.109543>.
- [5] D'Oca, S., Ferrante, A., Ferrer, C., Perneti, R., Gralka, A., Sebastian, R., Op't Veld, P., Technical, financial, and social barriers and challenges in deep building renovation: Integration of lessons learned from the H2020 cluster projects, *Buildings*, 8. (2018) pp 174.

- <https://doi.org/https://doi.org/10.3390/buildings8120174>.
- [6] Iturralde, K., Study on Automated and Robotic Renovation of Building Façades with Prefabricated Modules, Technical University of Munich, 2021.
- [7] ENSNARE Consortium, No Title, (2021). <https://www.ensnare.eu/> (accessed July 27, 2021).
- [8] DIN 18202 Toleranzen im Hochbau-Bauwerke, 2013.
- [9] Skibniewski, M., Hendrickson, C., Analysis of Robotic Surface Finishing Work on Construction Site, *Journal of Construction Engineering and Management*. 114 (1988).
- [10] Balaguer, C., Abderrahim, M., Trends in robotics and automation in construction., in: Intechopen, 2008. <https://doi.org/10.5772/5865>.
- [11] Warszawski, A., Economic implications of robotics in building, *Building and Environment*. 20(2) (1985) pp 73–81. [https://doi.org/https://doi.org/10.1016/0360-1323\(85\)90001-0](https://doi.org/https://doi.org/10.1016/0360-1323(85)90001-0).
- [12] Hu, R., Iturralde, K., Linner, T., Zhao, C., Pan, W., Pracucci, A., Bock, T., A Simple Framework for the Cost-Benefit Analysis of Single-Task Construction Robots Based on a Case Study of a Cable-Driven Facade Installation Robot, *Buildings*. 11 (2021) pp 1–8. <https://doi.org/10.3390/buildings11010008>.
- [13] Pan, M., Linner, T., Pan, W., Cheng, H., Bock, T., Structuring the context for construction robot development through integrated scenario approach., *Automation in Construction*. (2020). <https://doi.org/https://doi.org/10.1016/j.autcon.2020.103174>.
- [14] Pan, M., Linner, T., Pan, W., Cheng, H., & Bock, T., A framework of indicators for assessing construction automation and robotics in the sustainability context, *Journal of Cleaner Production*, 182. (2018) pp 82–95. <https://doi.org/https://doi.org/10.1016/j.jclepro.2018.02.053>.
- [15] Pan, M., Linner, T., Pan, W., Cheng, H.M., Bock, T., Influencing factors of the future utilisation of construction robots for buildings: A Hong Kong perspective, *Journal of Building Engineering*. (2020). <https://doi.org/https://doi.org/10.1016/j.jobe.2020.101220>.
- [16] BERTIM, D2.5. Efficient mass manufacturing and installation of prefabricated modules, 2016.
- [17] Hephaestus Consortium, Hephaestus - EU H2020 Project, (2017). <https://www.hephaestus-project.eu/> (accessed October 5, 2021).
- [18] Iturralde, K., Gambao, E., Bock, T., Compilation and assessment of automated façade renovation, in: *Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC)*, Dubai, UAE, 2021: pp. 797–804. <https://doi.org/10.22260/ISARC2021/0108>.
- [19] The FreeCAD Team, FreeCAD Your own 3D parametric modeler, (n.d.). <https://www.freecadweb.org/>.
- [20] Iturralde, K., Linner, T., Bock, T., Matching kit interface for building refurbishment processes with 2D modules, *Automation in Construction*, 110. (2020). <https://doi.org/https://doi.org/10.1016/j.autcon.2019.103003>.
- [21] Zhang, Z., A flexible new technique for camera calibration, *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 22 (2000) pp 1330–1334. <https://doi.org/10.1109/34.888718>.
- [22] Bouguet, J.Y., Camera Calibration Toolbox for Matlab, (n.d.). http://www.vision.caltech.edu/bouguetj/calib_doc/.
- [23] Olson, E., AprilTag: A robust and flexible visual fiducial system, in: *2011 IEEE International Conference on Robotics and Automation*, 2011: pp. 3400–3407. <https://doi.org/10.1109/ICRA.2011.5979561>.
- [24] Treitel, S., THE COMPLEX WIENER FILTER, *GEOPHYSICS*. 39 (1974) pp 169–173. <https://doi.org/10.1190/1.1440419>.
- [25] Rong, W., Li, Z., Zhang, W., Sun, L., An improved Canny edge detection algorithm, in: *2014 IEEE International Conference on Mechatronics and Automation*, 2014: pp. 577–582. <https://doi.org/10.1109/ICMA.2014.6885761>.
- [26] Robot Operating System, Robot Operating System Homepage, (2021). <https://www.ros.org/> (accessed October 5, 2021).
- [27] Universal Robots, UR10, (2021). <https://www.universal-robots.com/products/ur10-robot/> (accessed September 30, 2021).
- [28] Beckhoff, TwinCAT 3, (2021). <https://www.beckhoff.com/en-en/products/automation/twincat/> (accessed October 5, 2021).

Concept of a Robot Assisted On-Site Deconstruction Approach for Reusing Concrete Walls

H.J. Lee^{a*}, C. Heuer^{a*} and S. Brell-Cokcan^a

^a Chair of Individualized Production (IP), RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany

* These authors contributed equally to this work.

lee@ip.rwth-aachen.de, heuer@ip.rwth-aachen.de, brell-cokcan@ip.rwth-aachen.de

Abstract -

Construction and Demolition Waste (CDW) is one of the major waste streams in the EU by mass, accounting for 374 million tons in 2016 (excluding excavated soil), and is made up of a variety of components. Many of them can include dangerous materials and pose specific concerns to the environment and human occupants if not separated at source, but they also have a high resource value and great potential for recycling and reuse if extracted through a more controlled deconstruction process. Current deconstruction methods are ineffective in terms of being minimally invasive (air and noise pollution, destruction from tremors of explosions or abrupt demolition using explosives), safe or efficient. Furthermore, conventional methodologies fail to integrate modern technology (robots, remote sensing, and so on) in a systematic manner. This research work explores the shortcomings and strengths of previous approaches and provides conceptual approaches for robot-assisted deconstruction using the example of concrete walls.

Keywords -

robot-assisted deconstruction, construction robots, reusing construction material, automated deconstruction

1 Introduction

Construction and demolition waste (CDW) is one of the most voluminous waste streams generated in the EU. About 30% of all waste comes from construction and demolition (CD) [1]. By 2030, over two-thirds of all in-use material stock in the building sector will be 50 years old [2]. In the following decade, they will reach the end of their lives. The end-of-life product chain must be improved to guarantee that these materials may be used as secondary resources in another cycle. However, the performance of the proposed approach is limited in terms of working cubic meters per hour compared to the conventional methods.

Currently, despite the enormous share of CDW in the global waste stream, in some countries, the disposal rates of the mineral wastes amount up to over 90%, indicating a significant upside potential for recycling and reuse of

these high-valued materials. Moreover, as depicted in Fig. 1, most of the CDW consists of valuable metal and mineral waste (concrete and ceramics). So far, the demand for high manual labor has restricted the recycling and reusing option.

Building deconstruction and demolition are two separate processes: While demolition is the science and engineering of breaking down structures safely and quickly, deconstruction is the process of dismantling a structure while keeping elements and materials for reuse. The building elements are shredded into little bits by large demolition equipment in the conventional demolition process. The deconstruction process for material reuse is more careful and time-consuming: The removal of hazardous materials has to be ensured comprehensively for their later utilization. Furthermore, to optimize the possibility for reuse, the concrete blocks cannot be shredded in a rough and quick manner but must be cut precisely and without damage.

Teleoperated hydraulic devices with various attachments are often utilized on building sites. They have a large number of applications, are mobile, and have a quick turnaround time. Material modification, however, is not feasible owing to the lack of local precision. Furthermore, substances hazardous, such as asbestos, are released during demolition, implying a significant health danger to employees. Consequently, only a few current deconstruction techniques address the reuse option, even though several research studies have previously found the potential for reuse of building components.

In line with these problems, this work investigates the possible advance in the current material chain by increasing the reusing rate of construction components via semi-automated processes with deconstruction robots. Advanced and proven robotic technology will be used to enable the accurate and automated cutting process of components. For this purpose, existing demolition machines will be adapted contrary to common methods where completely new robot systems are developed from scratch. In this way, proven robotic technology will be integrated more easily into reliable construction machines so that deconstruction tasks can be semi-automated with fewer ef-

forts and risks.

2 State of the art

2.1 Demands for robotically controlled deconstruction

The changes in social and environmental requirements necessitate efficient appraisal of existing buildings' potential for refurbishment, environmental risk, and recycling. While current research into sustainable construction considers deconstruction a critical design factor, the most significant part of existing buildings was not optimized for this life-cycle stage. Demolition waste accounts for almost 30% of all waste produced in the EU and consists of numerous materials, including concrete, bricks, gypsum, wood, glass, metals, plastic, solvents, asbestos, and excavated soil, many of which can be recycled. Despite the intrinsic material value, the potential for material harvest is not fully exploited (recycling and recovery vary between 10% and 90% across the EU [3]). Moreover, deconstruction of old building materials often comes with significant health risks for construction workers. Current strategies are struggling to handle dust and hazardous materials in an affordable manner that sufficiently protects the workers involved. This adds to the rising issue of the construction industry as an unattractive employer with low worker retention.

Especially in marginal, low participation areas, the housing sector's degrowth of 3.8% is expected until 2030. Up to 2030, there will be a vacancy rate of 1.5 million buildings. In recent years 370.000 buildings in the east of Germany were demolished, as they were not used. Since 2001 the German government is actively supporting the program called "Program Städtebau" to demolish these buildings. However, as the areas suffer from the low participation rate, the question of how to execute the planned construction efforts remains unanswered.

On the other hand, the big European cities suffer from the degradation of construction efforts. As European Union statistics show six million people, or 5% of EU population, suffer from severe housing deprivation [4]. The issue of repurposing is also becoming increasingly important, as new regulations for environmental pollution (noise, dust, waste, and so on) in these big cities are being introduced.

2.2 Concrete Walls

Every year, approximately 55.5 million tonnes of valuable mineral wastes (i.e., bricks, sand-lime bricks, porous and concrete) are generated from construction industry. Here, concrete plays a vital role, as around 42% of used materials in the construction industry is made of concrete. At the same time, approximately 12 million tonnes of con-

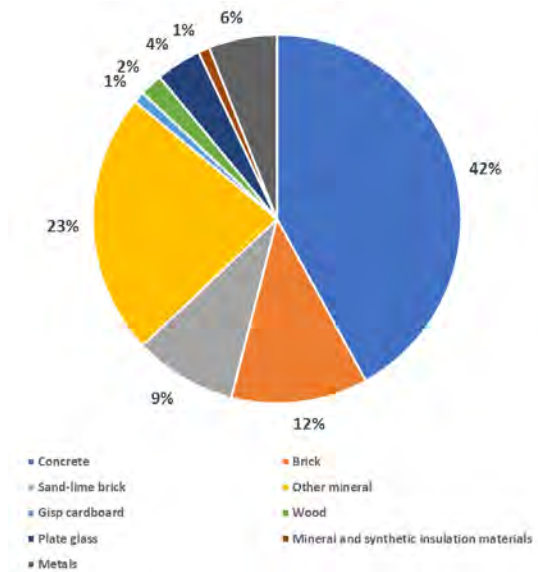


Figure 1. Used material resources in the construction industry [5]

crete waste are generated every year. Although most of the concrete wastes are recycled in the highway construction and paving, a significant amount of resources are consumed for the down- and upcycling.

2.3 Deconstruction of Concrete Walls

The following section describes the different methods currently used in deconstruction and renovation. Only those technologies are considered that are suitable for non-destructive dismantling or reuse. Here, experimental methods, which are still under development but have importance for future processes, are distinguished from conventional methods commonly employed on the market.

2.3.1 Conventional Deconstruction Methods of Concrete Walls

In Fig. 2, the conventional deconstruction machines for cutting construction materials are listed, which are already finding broad applications.

The summary in Fig. 2 clearly shows the limitation of the existing deconstruction methods: despite the continuous developments in hardware components, the level of autonomy has been stagnant, as most of the operations still have to be manually done.

2.3.2 Automated Deconstruction of Concrete Walls

First attempts in the 1970's from Japan, ranging from asbestos removal processes from surfaces to a cutting ma-

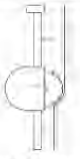
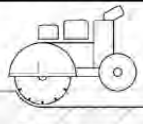

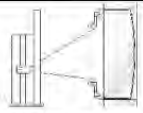
	Functionality	Materials	Degree of automation
 Wall saw	The system consists of three main elements: i) the concrete saw blade (saw blade diameter between 60 and 160 cm), ii) a drive for the saw blade (hydraulic or electric), and iii) a rail along which the driver travels. The rail can be mounted on the wall horizontally and vertically. The saw slides along the rail with the help of an integrated motor, ensuring a clean and precise cut.	The saw can be used for concrete, brick, foam concrete.	Some saws, for instance, from the company Hilti have integrated an assistance programs that can automate the cutting process. The saw assembly is still done manually, which is very time-consuming and cost-intensive.
 Floor saw	The floor saw is cutting equipment that rolls over the floor on rollers. Unlike the wall saw, the sawing device is mounted on a trolley. A saw is guided along a surface by rollers, similar to a wall saw, and can be lowered into the material to process the cutting step by step.	The saw can be used for concrete and asphalt.	Some machines can adjust their speed automatically, as they have electric guidance.
 Mounted wall saw	This particular type of wall saw is developed for use with a hydraulic manipulator. Unlike the wall saw, the saw is attached to the retaining plate of the manipulator. The operator guides the saw by remotely teleoperating the manipulator.	Materials such as concrete, stone, and metals can be cut.	The machine is manually teleoperated.
 Wire saw	A diamond saw wire is driven by a motor and tensioned with high pressure on the surface to be sliced. Pulleys guide the diamond wire in a closed circuit where the wire tension is kept constant.	This machine can be used to cut large components from reinforced concrete.	Guiding the wire saw requires a high degree of manual control. Currently, only the tension of the saw is adjusted automatically.

Figure 2. Summary of conventionally available deconstruction machinery

chine using a water jet, demonstrate that there is a great economic and ecological potential for robotic use in deconstruction to raise the recycling rates and employ new concepts of building component reuse and material harvest instead of total demolition and disposal as waste [6]. However, introducing the first robotics system on the deconstruction site resulted in several issues, some coming from the primitive human-machine-interface (HMI) and highly specified custom designs of the developed systems. To address this issue, researchers investigated the potential of the industrial robot-aided deconstruction approach [7, 8] and also the impact of a new HMI method, for instance based on a laser designation.

However, industrial robots typically suffer from a low payload/weight ratio. Also, they are mostly designed for indoor environments limiting the usage of processes with industrial robots on construction sites. Here, the sky factory approach was introduced to ease the integration of industrial robots by changing the working environment into a factory-like setting [10], which, however, can be applied only to a limited extent on a construction site due to its high cost and complexity.

The idea of automating the deconstruction process with new emerging technology has drawn the attention of many researchers. Recently, the researchers developed a selective deconstruction method based on electrodes generating electro-pulses for removing and cutting concretes, which mainly aims at the mining industry [11]. How-

ever, the additional safety issues and scalability still remain open.

3 Robot-Assisted Deconstruction Approach

This study introduces an alternate deconstruction strategy that attempts to offer an example of a practical solution for the controlled deconstruction of concrete walls which is developed at the conceptual level in this work.

This work aims to develop a robot-assisted deconstruction process, which minimizes human risks and maximizes efficiency and accuracy by utilizing robot technology. To achieve this goal, three main subsystems are to be considered:

Subsystem 1: Robotic system for cutting tasks. In order to increase the level of task automation, the corresponding level of machine autonomy is first considered. Here, the development goal is that the robotic system perceives the environment, considers its hardware capability, and considers the input from the human operator (i.e., the desired geometry of the cut element) to generate the appropriate motions automatically.

Subsystem 2: Vision System. As the overall goal is to minimize human risks and automate the deconstruction process, the information from the environment has to be collected and transferred to the

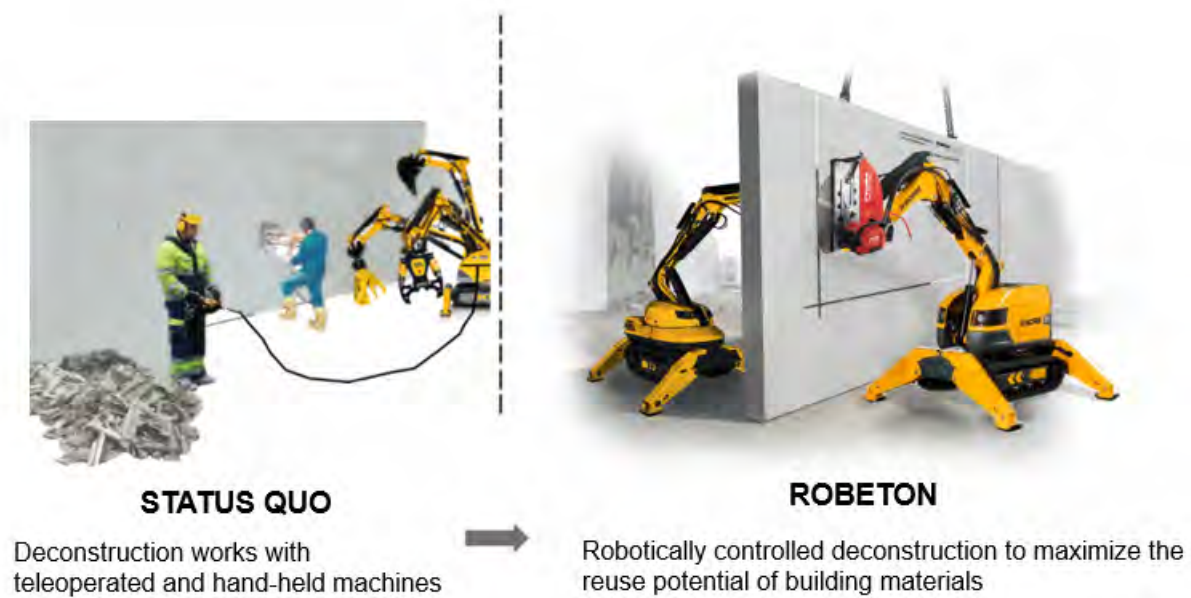


Figure 3. The overall vision of this work is to maximize the reuse potential by increasing the level of autonomy in deconstruction tasks.

robotic system. By collecting the visual feedback from the remote working place, one can improve the telepresence of the operators and avoid the scenarios where the operators have to be present on the construction sites.

Subsystem 3: Robotic system for stabilizing the cutting process and holding the cut elements. While the first robotic system is responsible for automatic cutting, this other robot supports the first robot so that the element that is being cut does not swing back and forth and collides with the environment during and after the cutting process.

3.1 Deconstruction Robot

The first subsystem focuses on the automated, exact cutting of concrete walls, which is one of main difficulties connected to controlled deconstruction. Currently, teleoperated hydraulic machines are commonly used for this task. However, the operation is difficult even for experienced operators since many joints of the machine must be manipulated simultaneously with joysticks.

In this work, we aim to increase the autonomy level of an existing teleoperated hydraulic deconstruction machine rather than developing a new deconstruction robot from the scratch, as it offers various benefits:

1. Deconstruction of concrete walls is a heavy-payload process. The machinery used for the process must be

very dependable, durable, and resistant to force. Hydraulic machines are a better fit in this situation than electric machines, such as industrial robots, which have a limited payload and are generally error-prone in outside conditions (dirt, humidity, etc.).

2. The robustness of the selected hydraulic machine has been proven in different disaster places [12]. The system can drive on non-flat or soft ground and can perform different tasks by changing the end-effector.
3. The commands for the control system and the actuators are already interacting using electrical signals since the system is teleoperated. The whole system does not need to be electrified in order to increase the level of autonomy.

The deconstruction machine's precise motion control is critical as it determines whether the deconstructed material can be directly reused without losing its intrinsic value or has to be recycled through down- and upcycling. In our previous work [13], we investigated the possibility of converting the teleoperated hydraulic machine BROKK 170 into a programmable robot that can be controlled by a motion controller implemented on an onboard computer. As most teleoperated hydraulic machines just like BROKK 170 are not equipped with any motion sensors that can estimate the disparity between actual and commanded motion. Thus, motion sensors such as encoders were added to the machine. Then, a communication link between an onboard computer and the machine was built by analyzing



Figure 4. Snapshots of the BROKK 170 tracking a trajectory in the TCP-level. The grey tool can be interactively positioned and rotated by the operator defining the desired tool pose. According to the desired tool pose, a TCP trajectory and the corresponding joint trajectories (upper row) are generated. The joint trajectories' joint angle changes are then converted into the PWM signals to move each actuator along the desired joint trajectory (lower row).

the communication between the original input device and the machine. Based on this result, a general CLIK technique was used to convert the desired motion descriptors represented in task space to corresponding joint motions. The joint motions were then converted to CAN bus messages, which generated the corresponding PWM signals for the valve system. Instead of controlling each axis independently with the joystick, the suggested technique allowed the teleoperated machine to be programmed in task space. In the previous work, the resulting accuracy was 2 cm for the position and 0.015 rad for a point-to-point movement. Due to the nonlinearities in the hydraulic system, the error in Root Mean Square Error (RMSE) while tracking TCP path was relatively larger with 6 cm in the TCP space.

The current system setup interacts with the operator using an interactive tool, as the snapshots of Fig. 4 visualizes. The operator defines the higher-level command, i.e., desired tool pose at the task level. The corresponding joint trajectories are generated using the Robot Operating System (ROS) environment [14] and the related motion planning framework, MoveIt! [15] to generate a collision-free motion by fulfilling the pre-defined requirements such as the step size and the goal time. To allow the command from the operator, as depicted in Fig. 4, the operator has to be aware of the local circumstances of the remote environment, i.e., the geometry of the target object. Here, we deploy another robotic platform capable of providing the visual feedback of the remote workplace from different viewpoints according to the operator's input.

3.2 Vision System

Typical control stations incorporate numerous 2D camera views from various viewpoints to boost telepresence by presenting the gathered sensory data to a remote loca-

tion. Such techniques, on the other hand, have been found to be troublesome in terms of the operator's fragmented attention and overloaded network.

This work uses a mobile robot with a depth camera to capture the remote working scene in a 3D point cloud and transmit it to the operator and the robot. Especially, the 3D visualizing technique has shown its effectiveness in improving the telepresence and operator's manipulation capabilities in different works [16] [17]. The geometrical information collected in 3D will be used together with normal 2D camera images to increase the operator's telepresence and maximize the understanding within remote environments.

The mobile robot used in this work is based on a commercially available platform from Innok Robotics. It is further equipped with a 3D depth camera to provide 3D visual feedback from an arbitrary viewpoint of the remote workplace. Additionally, it is equipped with different sensors such as an inertial measurement unit (IMU) *Xsens MTi-30-2A8G4* and a 2D laser scanner *Sick micro Scan 3* which will be utilized for localization tasks on construction sites.

3.3 Supporting System

The cutting process has to be stabilized by an additional system. The concrete walls often weigh up to multiple tons and are often up to several meters long. When the cutting process is halfway done and the cut element is partially dissolved from the wall, there is the risk that the element can fall down on the robots or the ground. To prevent this risk, we utilize a supporting system that consists of two different machines with distinct purposes:

- **Additional hydraulic robot:** Additional robot will be mainly used on the other side of the cut element



Figure 5. Conceptual visualization of the planned usecase: Top-down deconstruction approach with tower crane

to prevent the swing that might be generated when the element is dissolved from the wall. Thus, it is essential to estimate the forces applied by the first deconstruction robot to compensate them. Here, the communication between these multiple systems will be enabled in the ROS environment.

- **Tower crane:** In this case, the crane is used to compensate for the heavyweight of the cut element during the process and carries it to the temporary storage point after cutting is finished.

4 Demonstrator

The major goal of this project is to enable robotically controlled deconstruction, which will maximize the reusability of building materials. The envisaged demonstrator involves the following steps: (i) 3D point cloud methods for visualizing the remote scene to the operator and extracting useful visual features to support for decision making, (ii) Adaptive robotic path planning considering the local information such as the geometry of the concrete wall and the input from the human operator, (iii) Development of suitable end-effectors for cutting the concrete walls into directly reusable pieces.

Here, the experimental setup will be installed on the reference construction site on Aachen Campus West. The robotic system described in the previous sections will be evaluated within the scope of the scenario (see Fig. 5).

4.1 Top-down deconstruction approach with tower crane

The test scenario will demonstrate how the envisioned robotic technologies can be applied to the adaptive reuse of concrete walls. The top of the structure is presumed to have been removed at this point, as visualized in Fig. 5.

First, the demolished wall structure will be captured in 3D, and the acquired data will be communicated to the operator over the on-site wireless network. The deconstruction and support robots will then be teleoperated and placed in front of the wall. It's worth mentioning that the robot's autonomous driving isn't covered in this research. The ideal position for the robot to reach and handle the wall will be determined and shown to the operator based on the acquired geometric information of the wall. The operator defines the required geometry for the cut element after the robots are in place. An adaptive tool-path trajectory is developed based on this input and the gathered 3D knowledge about the wall structure. The deconstruction robot's low-level controller then accurately executes the generated trajectory. The tower crane is mostly utilized in this scenario to stabilize and transport the cut part to a temporary storage location.

4.2 System requirement for the selected application area

This section briefly describes the necessary system requirement that the envisaged robotic system described in the previous section has to fulfill. Although our previous research [13] about the deconstruction robot is rather primitive and leaves further questions, for instance, in terms of accuracy, the technical findings are used in this section to build the bottom line of the development direction.

4.2.1 Reuse of concrete elements

The potential of reusing construction components has been intensively analyzed in the project Superlokal in Kerkrade [18]. Four high-rise concrete buildings located in Kerkrade were deconstructed, renovated, recycled and reused as experimental attempts to analyze the potentials of the materials from the old housing. One of the subproject was aimed at the reuse of the harvested material into a new building. The concrete blocks as main body, the concrete staircase, the kitchen core, the paving material, the wood doors and windows, the aluminium handrails and so on were all reused without post-processing and have made up to 95% of the overall building materials. Especially, the idea of reusing mineral materials such as concrete which makes up to 80% of the total CDW waste stream has the potential to dramatically contribute to the sustainability in the construction industry.

Table 1. Indicators for the performance assessment

	Indicator
Automated cutting process	Improvement of the accuracy of the cutting tasks
	Reduces the time required for the cutting tasks
	Reduces the number of human workers working on-site for the cutting tasks
	Workers feel more safer during the cutting tasks
Cost	Additional costs for extra machines/ hardware
	Reduced time/costs
	Increased work productivity
Sustainability	Savings of the grey energy
	Material savings/ increased resource reuse rate

Table 2. Evaluation of the tracking in the TCP-level

	PTP Movement	TCP-Trajectory Tracking
Current RMSE	2 cm	6 cm
Goal RMSE (DIN 18202)	1.2 cm	1.2 cm

However, the manual process of measurement, positioning, machines operation and removal of toxic substances has increased the required efforts and the complexity to the project. As a result, the need for automated technology to compensate for inefficiency and safety issues throughout the deconstruction process has been established.

Certain technical tolerances, such as cut precision, must be ensured during the automated deconstruction process to maximize the reuse potential of cut elements. Here, we refer to the regulations of DIN 18202 [19], which describe permissible tolerances for the manufacture of components and the execution of structures in building construction: For a structural component, i.e., a wall with a length of 3 m, the deviation must not exceed 12mm.

As stated in Sec. 3.1 and summarized in Table 2, the current error in PTP-movement and tracking TCP-trajectory lies in 2 cm and 6 cm, respectively, whereas the allowed deviation in material processing lies in 12 mm. Thus, the first step to deploying the planned deconstruction robotic system in the actual process is the precise motion control of the hydraulic manipulator. However, as the nonlinearities in the hydraulic system greatly affect the high-precision control, further research is required in this direction.

4.3 Usecase Assessment

The increase of reusable concrete material due to automated deconstruction controlled cutting process will be assessed and compared to regular procedures. Primarily, the demonstrator will be evaluated by the following indicator categories: (i) Automated cutting process, (ii) Cost, and (iii) Sustainability.

As the primary goal of this work lies in improving the autonomy level of the cutting task, the new robotically controlled deconstruction approach will be compared with the existing deconstruction/demolition approaches to demonstrate the improved accuracy, significant reduction energy, and human resources consumed for cutting tasks. Besides improving cutting tasks, productivity and worker safety on-site will also be evaluated as one of the major constraints for automated deconstruction is usually the cost. The financial aspect, i.e., the economic viability of this new robotic approach, will be investigated. Finally, the ecological viability, increased reuse potential, corresponding material savings, and related grey energy will be investigated.

5 Acknowledgments

This work has been supported by the Federal Institute for Research on Building, Urban Affairs and Spatial Development and Zukunftsbau of Germany under the research intent ROBETON and project number 10.08.18.7-21.11. The reference construction site on RWTH Campus West is run by the Center Construction Robotics and the first living lab to test new construction processes, products, automated machines and robots in Europe.

6 Conclusion

Automated deconstruction will remain the main research topic in the following decades due to the constantly rising demand for sustainability and the permanent shortage of skilled workers. In line with this issue,

this work investigated the current limitation of existing methods and identified the lack of automated solutions for deconstructing concretes. Given this result, this work presented a conceptual approach with different subsystems for the robot-assisted on-site deconstruction process. Next, the challenges that might arise from deploying and using the robotic systems on the construction sites and the corresponding demonstrator for further investigation were jointly analyzed and introduced. Finally, possible assessments to effectively evaluate the planned demonstrators were defined.

References

- [1] Directive 2008/98/EC on waste (Waste Framework Directive), 2008/98, Official Journal of the European Union (2013)
- [2] Krausmann et al.(2017) Global socioeconomic material stocks rise 23-fold over the 20 th century and require half of annual resource use. *Proc. Natl. Acad. Sci.*114, 1880-1885
- [3] The European Demolition Industry Report, European Demolition Association (2015)
- [4] Eurostat (2019) Severe housing deprivation rate by degree of urbanisation - EU-SILC survey Retrieved from https://ec.europa.eu/eurostat/databrowser/view/ilc_mdho06d/default/table?lang=en
- [5] Kaiser O., VDI Technologiezentrum GmbH (2019) VDI ZRE Kurzanalyse Nr. 26: Rückbau im Hochbau – Aktuelle Praxis und Potenziale der Ressourcenschonung
- [6] Bock T, Linner T (2017) Construc- tion Robots - Elementary Technologies and Single-Task Construction Robots. 10.1017/CBO9781139872041.
- [7] Lublasser E, Hildebrand L, Vollpracht A, Brell-Cokcan S (2017) Robot assisted deconstruction of multi-layered façade constructions on the example of external thermal insulation composite systems. *Construction Robotics*. 1. 10.1007/s41693-017-0001-7.
- [8] Lublasser E, Iturralde K, Linner T, Brell-Cokcan S Bock T (2016) Automated refurbishment & end-of-life processes research approaches in German and Japanese construction. (DE-HGF)8
- [9] Corucci, F., Ruffaldi, E. (2015). Toward Autonomous Robots for Demolitions in Unstructured Environments. In *Intelligent Autonomous Systems 13* (pp. 1515-1532). Springer International Publishing.
- [10] Bock T, Linner T (2016) Site Automation: Automated/Robotic On-Site Factories. 10.1017/CBO9781139872027.
- [11] Otto et al. (2021) Selektiver Rückbau mittels Elektro-Impuls-Verfahren: Grundlagenermittlung zur Adaption eines innovativen Abbruchverfahrens aus dem Bergbau (EIV) als neue Bautechnologie für selektiven Rückbau in sensiblen Bereichen. BBSR-Online-Publikation 30/2021, Bonn, Dezember 2021.
- [12] Hiramatsu Y, Aono T, Nishio M (2002) Disaster restoration work for the eruption of Mt Usuzan using an unmanned construction system. *Adv. Robot* 16(6):505–508
- [13] Lee HJ, Brell-Cokcan S (2021) Cartesian coordinate control for teleoperated construction machines. *Constr Robot* 5, 1–11. <https://doi.org/10.1007/s41693-021-00055-y>
- [14] Quigley et al. (2009) ROS: an open-source Robot Operating System, in *Proc. of the IEEE Intl. Conf. on Robotics and Automation (ICRA) Workshop on Open Source Robotics*
- [15] D. Coleman, I. A. Sucan, S. Chitta and N. Correll, "Reducing the Barrier to Entry of Complex Robotic Software: a MoveIt! Case Study", *Journal of Software Engineering for Robotics*, 5(1):3 16, May 2014.
- [16] Mast M. et al. (2013) Teleoperation of Domestic Service Robots: Effects of Global 3D Environment Maps in the User Interface on Operators' Cognitive and Performance Metrics in Social Robotics. *ICSR* 2013
- [17] Huang et al. (2019) Evaluation of haptic guidance virtual fixtures and 3D visualization methods in telemanipulation—a user study. *Intel Serv Robotics* 12, 289–301
- [18] Ritzen et al. (2019) Circular (de)construction in the Superlocal project. *IOP Conference Series: Earth and Environmental Science*. 225. 012048. 10.1088/1755-1315/225/1/012048.
- [19] Construction Index (2020) URL: <https://www.bauindex-online.de/regelwerke/din-18202/> (last access on 26.04.2022)

Subjective Evaluation of Passive Back-Support Wearable Robot for Simulated Rebar Work

N. J. Gonsalves, M. Khalid, A. Akinniyi, O. Ogunseiju and A. Akanmu

Virginia Polytechnic Institute and State University, USA

E-mail: gonsnihar@vt.edu, khalidm21@vt.edu, abiola@vt.edu, omobolanle@vt.edu, abiola@vt.edu

Abstract –

Work-related musculoskeletal disorders continue to be a severe problem in the construction industry. Rebar workers are exposed to ergonomic risks such as repetitive stooping and forward bending resulting in low back injuries. Wearable systems such as back support exoskeletons are emerging as potential solution to reducing the risk of low back injuries. User acceptance of exoskeletons is necessary to facilitate adoption of the technology in the construction industry. Exoskeletons could have unintended consequences such as discomfort and interference with work. This paper presents an assessment of a commercially available back-support exoskeleton for rebar work in terms of usability and perceived discomfort. Ten student participants performed rebar tasks with and without the back-support exoskeleton. Participants completed usability (ease of use, learning, and comfort) and level of perceived discomfort questionnaires after the task. Findings indicate that the back-support exoskeleton is easy to learn and use but reduced the participant's comfort. Although, the exoskeleton triggered increased discomfort, there was a reduction in the level of perceived discomfort at the lower back and lower leg. This study contributes to existing discourse on the influence of perceived usability and level of discomfort when using exoskeleton on user acceptance.

Keywords –

Wearable robot; Rebar work; Usability; Discomfort

1 Introduction

Work-related musculoskeletal disorders have emerged as one of the top causes of non-fatal occupational injuries, as well as a critical health and safety factor in the U.S. construction industry [1]. Work-related musculoskeletal disorders (WMSDs) is a severe and widespread problem in the construction industry, with an incidence rate of 1.7 times the overall industry

average [2]. A variety of construction operations involve a high level of physical exertion of the body. Most WMSD injuries are caused by overexertion efforts, including repetitive physical motion, excessive force, and unusual postures [3]. Moreover, construction tasks often require extended standing, bending, and stooping postures, which directly impose varying stresses on the worker's musculoskeletal system [4] and have been identified as some of the primary triggers of WMSDs [5].

The number of cases of WMSD-related injuries among reinforcing iron and rebar workers have alarmingly risen by a staggering 400% in the last two years [6]. This might be considerably higher, given that research estimates that approximately one-fourth (27%) of construction injury cases go unreported due to lack of proper reporting [7]. Compared to other occupations, construction employees who undertake rebar-related tasks are more likely to be subjected to stressful and physically demanding work conditions [8]. Rebar work typically entails placing and tying rebars prior to concrete pouring. These tasks are typically performed by assuming non-neutral trunk postures for extended periods [9]. Rebar workers spend approximately 40- 48% in the forward trunk bending position [10], exposing them to a heightened risk of low-back disorders compared with other construction trades[11]. One of the implications of back disorder is increased absenteeism amongst workers and in severe cases, premature disability [12]. Furthermore, construction project profitability can be significantly hampered as a result of WMSDs. WMSDs can also lead to a surge in early retirement [13] resulting in labour shortage in the industry. In addition to the direct expenses of WMSDs, organizations may also incur various indirect expenditures (e.g., absent wage, cost of lost time, and reduced productivity) [12].

In recent years, wearable robots or exoskeletons are increasingly being explored as a potential solution to WMSDs in the construction industry [14]. An exoskeleton is a wearable device that can augment a wearer's physical abilities, thereby reducing the load and risk of WMSDs on the supported body part [15]. Among the types of exoskeletons (i.e., active and passive

exoskeletons), passive back-support exoskeletons are increasingly attracting industry interest because they are cheaper and lighter than active exoskeletons and do not require external power source [5].

Commercially available exoskeletons are designed for rehabilitation purposes, where the technologies are intended to support disabled people. Some exoskeletons have also been developed for healthcare and military applications. Given the benefits of reduced muscular activity and improved productivity experienced from the use of exoskeletons in these industry sectors [16], it is important to explore the applicability of exoskeletons in construction.

To support the uptake of exoskeletons in the construction industry, it is important to investigate its usability for construction work. Usability is a key intervention concept that can be used to decide whether a technological solution can be successful in the workplace. However, there are scarce studies exploring the usability of exoskeletons for trade-specific applications in construction. The integration of passive back-support exoskeleton could introduce unintended consequences, such as deviated working posture and discomfort [17]. Since construction tasks may have different physiological exposures, it is also crucial to understand the level of discomfort from the use exoskeletons. End-users (such as rebar workers) will be reluctant to utilize the technology if it lacks acceptable usability or if there is an elevated level of discomfort, leading to a counterproductive intervention.

Therefore, the objective of this study is to assess the usability and level of discomfort of a passive back-support exoskeleton for rebar work.

2 Background

Studies have showcased the potential of passive back-support exoskeletons to reduce physical demands on the back [1]. Laevo, BackX, FLx ErgoSkeleton and SPEXOR are some of the passive back-support exoskeletons being explored in different industry sectors such as healthcare and automobile industries. The adoption of the exoskeletons depends on user acceptance of the device.

Researchers have conducted usability studies to assess user acceptance by employing structured questionnaires addressing ease of use, donning and doffing, comfort and perceived discomfort [3]. FLx ErgoSkeleton was assessed for patient transfer tasks wherein participants showcased a good level (76.2/100) of usability [5]. Another study tested SPEXOR for twelve different functional tasks where participants reported a reduced low back discomfort [5].

Of all these exoskeletons, Laevo was identified as one of the most promising back support devices due to its

ability to allow axial rotation of the upper body [7]. Usability studies on Laevo showcased moderate to good levels of user acceptance and reduced discomfort. For example, participants identified lower discomfort at the waist and moderate usability during material handling task, while using Laevo compared to the BackX exoskeleton [2]. Lee, Yang [3] explored Laevo for industrial and functional tasks and reported a good level of usability (75.4/100). Ogunseiju, Gonsalves [5] evaluated Laevo for flooring tasks and identified reduction in perceived discomfort of the back (28%). Alemi, Madinei [18] reported moderate level of usability for Laevo during repetitive lifting tasks. The tasks performed in above studies involved forward bending and twisting actions which are similar risks that rebar workers are exposed to during placing and tying tasks. Thus, one can envision similar exoskeleton being useful for rebar tasks. Despite the high occurrences of WMSDs amongst rebar workers, scarce studies have explored the suitability of back support exoskeleton for rebar work. Thus, this study evaluates a commercially available passive back support exoskeleton for the rebar work in terms of usability and perceived discomfort.

3 Methodology

This section presents the approach employed in this study (Figure 1), including an overview of the back-support exoskeleton, participants involved, simulated rebar task, study design, and data collection and analysis.



Figure 1. Overview of the methodology

3.1 Participants

The study employed a convenience sample size of 10 male student participants. The subjects signed the informed consent form approved by the Institutional Review Board at Virginia Tech prior to commencing the experiment. None of the participants reported any musculoskeletal injuries affecting their ability to perform physical tasks. The mean and standard deviation of the demographics of the participants was: age = 23yrs \pm 1.99, weight = 155.70 lbs. \pm 22.51 and height = 173.40 cm \pm 4.97.

3.2 Exoskeleton

A commercially available passive back-support exoskeleton called Laevo V2.56 was employed in this study which weighs 2.8 kgs. Laevo is a rigid wearable device which is designed to provide support to users' back muscles while they perform forward bending tasks. The exoskeleton (shown in Figure 2) consists of a chest pad, thigh pads, adjustable hip pad, metal torso, smart joint and comes with three torso sizes (i.e., small, medium and large) to fit different body types. Once switched on, the energy-storing spring system in the smart joint is activated and generates the torque, which provides support to workers when they assume stooping postures through the chest pad, thereby relieving stress from the back muscle and transferring the load to the legs via the thigh pad.

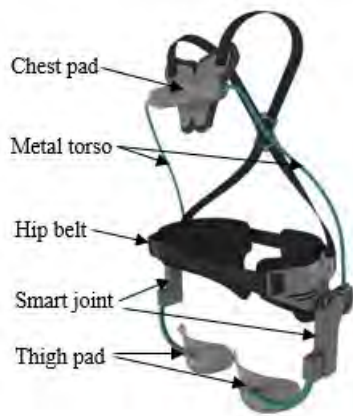


Figure 2. Laevo exoskeleton

3.3 Experimental Task

The participants of the study performed rebar task (Figure 3) consisting of placing and tying subtasks which were simulated using metal bars in the form of prefabricated gates. Each participant performed a total of four cycles for each experimental condition (i.e., with and without the exoskeleton). The subjects placed the prefabricated gates, comprising of #11 bars at 2" centre to centre in both directions, on the floor and used a plier to tie six of the joints with pre-cut ties. Combination of placing and tying subtasks was considered as one cycle.

3.4 Study Design

The participants signed the informed consent form after which they were introduced to the rebar task. The participants were allocated 20 minutes to practice the rebar task until they were comfortable. Thereafter, the participants performed the rebar tasks without the exoskeleton. Following that, the students were allowed to

rest for 15 mins. to avoid fatigue and were asked to complete a level of perceived discomfort questionnaire (section 3.5). Subsequently, the functioning of the exoskeleton was explained to the subjects. Once they were confident with the exoskeleton, the participants were asked to don the exoskeleton during which they were timed. Afterwards, the students performed the rebar task with the exoskeleton. After completing the task, the participants were timed while doffing the exoskeleton. After both experimental conditions (with and without exoskeleton) were completed, the participants were asked to fill the usability and level of perceived discomfort questionnaire (section 3.5).

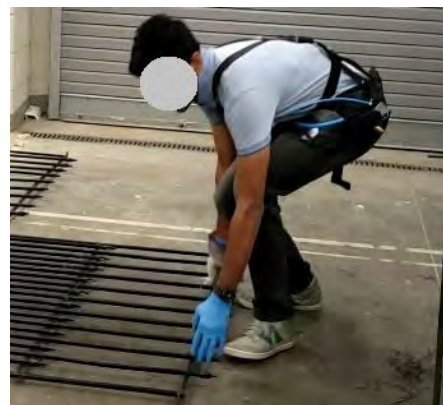


Figure 3. Participant performing rebar task

3.5 Data Collection and Analysis

The usability and level of perceived (LOD) were measured using structured questionnaires. The questionnaire, designed for assessing the usability of back-support exoskeleton, consisted of 16 questions addressing three criteria (i.e., ease of use, ease of learning and comfort). The ease of use the responses from the participants were recorded using a 5-point Likert's scale (varying from 1- Strongly Disagree to 5 - Strongly Agree). The LOD questionnaire presented the participants with eight different body parts (i.e., hand/wrist, upper arm, shoulder, lower back, thigh, lower leg, neck and chest). Using the Borg CR 10 scale, varying from 0 - 'Nothing at all' to 10 - 'Maximal', the participants provided the level of discomfort experienced at each body part.

The usability data, and donning and doffing timings were analysed using descriptive statistics such as mean and standard deviation. Two-way Analysis of Variance (ANOVA) was used to analyse the collected LOD data. The dependent variable was the participants' ratings whereas the body parts and experimental conditions were the independent variables. Separate bar charts were plotted for the results of the usability and LOD to understand the trends of the data.

4 Results

4.1 Usability

The usability questionnaire covered three aspects which included, ease of use (Figure 4), ease of learning (Figure 6) and comfort (Figure 7) of the exoskeleton. Furthermore, the durations of the donning and doffing were recorded for each participant (Figure 5).

4.2 Ease of Use

The participants provided an average rating of moderate to high (3.5 ± 0.29) for the overall ease of use of the exoskeleton. The average time registered by the participants to don and doff the Laevo was less than a minute (Figure 4) whereas moderate to high rating (3.7 ± 0.30) for donning and doffing the exoskeleton was recorded. When asked about 'ease of adjustment', the participants provided an above average rating of 3.8 ± 1.17 . The ability of the exoskeleton to help the participants in task accomplishment, as well as meeting user requirements received moderate rating. The participants' preference to use the Laevo for rebar tasks received the lowest rating of 3.1 ± 1.04 whereas the participants' ability to use the exoskeleton without assistance attracted the highest rating of 3.9 ± 1.22 .

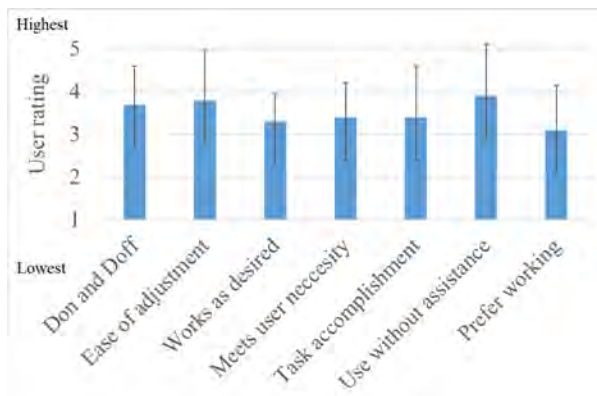


Figure 4. Ease of use

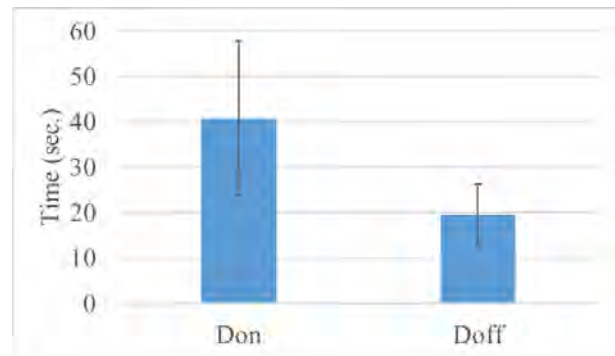


Figure 5. Donning and Doffing

4.3 Ease of Learning

The ease of learning questionnaire included one negative question (i.e., on prior knowledge - Figure 6) and five positive questions. A low rating of 1.8 ± 1.17 for a negative question suggests that the participants did not require any previous knowledge to effectively learn the exoskeleton's functioning. Overall, for the positive questions, the subjects provided a high rating of 4.2 ± 0.21 suggesting good ease of learning. When asked whether the participants' could assemble, adjust and check the fit, a high rating was recorded. The ability of the participants to remember the procedure was high as well (4.3 ± 1.01). Similar to ease of use, the highest rating (4.5 ± 0.69) was recorded for the participants' ability to use the exoskeleton without assistance.

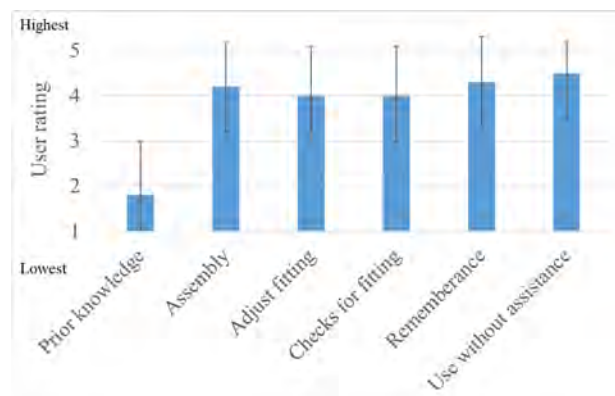


Figure 6. Ease of learning

4.4 Comfort

The participants were asked whether the exoskeleton restricted their movement (2.6 ± 1.43) as well as if the exoskeleton interfered with the work environment (2.2 ± 0.87). Both negative questions received a low to moderate rating, inferring that the exoskeleton posed minimal interference with the rebar task. The participants'

satisfaction with the exoskeleton for rebar tasks received a moderate rating (3.2 ± 1.32).

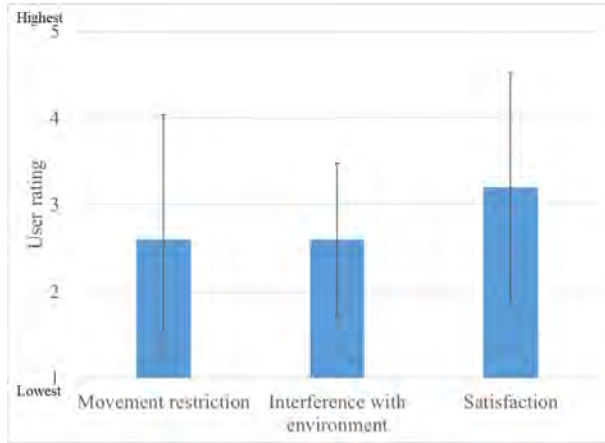


Figure 7. Comfort

4.5 Overall Usability

The overall usability included positive and negative questions. The rating for positive questions (Figure 8) was highest for ease of learning (4.2 ± 0.21), followed by ease of use (3.5 ± 0.29), and comfort (3.2 ± 1.32). Overall, moderate to high rating was registered across all the three positive questions. The negative questions (Figure 9) received a low to moderate rating. The higher rating for positive questions and lower ratings for negative questions suggest good usability of the exoskeleton for rebar work.

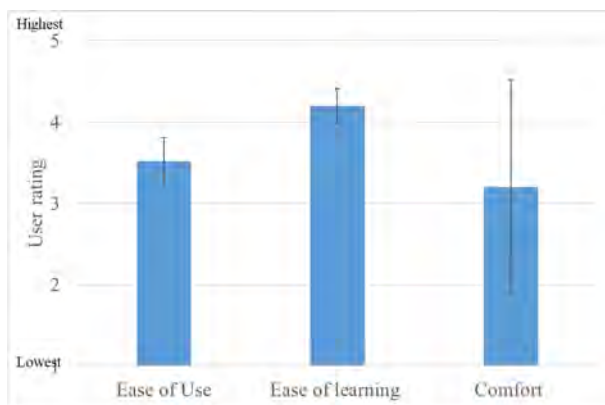


Figure 8. Overall rating for positive questions

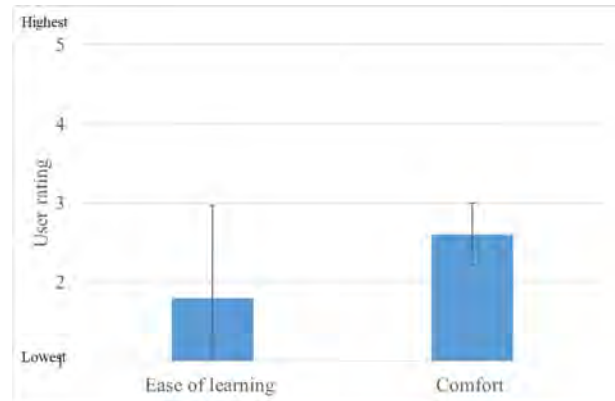


Figure 9. Overall rating for negative questions

4.6 Level of Perceived Discomfort

Table 1 presents the summary of the ANOVA (Mean, F-Value, P-Value, and effect sizes (η^2)) performed across the different body parts and exoskeleton conditions (see table 1). P-values with '*' have a confidence level < 0.05 .

Table 1. Perceived level of discomfort of participants with and without the exoskeleton during rebar task

Outcome Measure		Body Parts	Experimental conditions	B x E
LOD	Mean	3.85	0	1.48
	F-value	11.8	0	3.01
	P-value	2.08E-09*	1	0.0079*
	η^2	0.246	0	0.037

Note: E = Experimental Condition, B = Body parts

The ANOVA results (Table 1) suggest a significant difference ($p < 0.05$) was observed in the perceived discomfort across different body parts as well as body parts and experimental conditions with an effect size of 0.246 and 0.037 respectively. Discomfort level increased in the neck (22.73%), chest (90.91%) and thighs (45.45%) while using the exoskeleton. The perceived discomfort decreased for the low back (118.18%) and lower legs (54.55%) while performing the rebar tasks with the exoskeleton. The upper arm, shoulder and hand/wrist did not have a significant impact between the two experimental conditions.

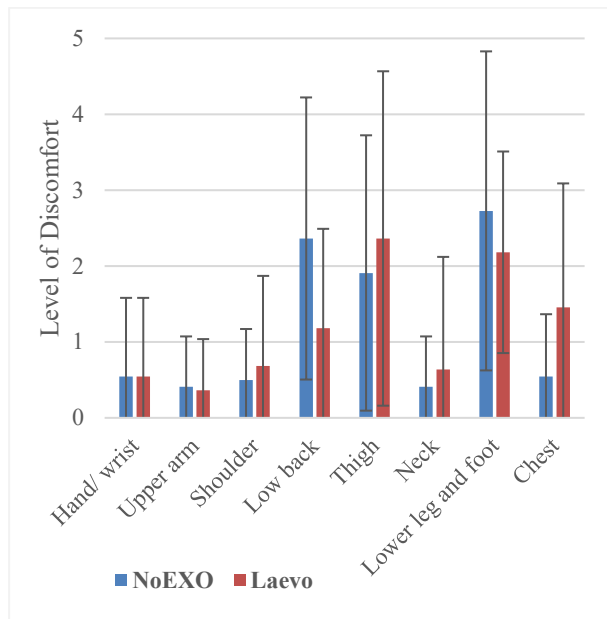


Figure 10. Level of Perceived Discomfort

5 Discussion and Conclusion

The need to improve construction health and safety and reduce work-related musculoskeletal disorders has ushered the adoption of exoskeletons into the construction industry. With the growing interest in exoskeletons and its potential to mitigate at least 60% of construction WMSDs [1], there is a need to assess the usability of exoskeleton for construction activities. This study investigates the potentials of Laevo back-support exoskeleton for mitigating WMSDs by assessing users' perceptions of the ease of use, ease of learning, comfort, and perceived level of discomfort during simulated rebar tasks.

Overall, the study revealed that the exoskeleton was easy to use. Participants could don and doff the exoskeleton without assistance in less than a minute and could easily adjust it to their desired comfort level. This may suggest a low impact of the exoskeleton on increasing rebar task completion time.

However, the participants were moderately satisfied with the ability of the exoskeleton to support rebar tasks and moderately preferred to work with the exoskeleton. This may be influenced by their moderate ratings of the exoskeleton to meet their demands during rebar tasks which can impact the willingness to adopt the exoskeleton.

The assessment of ease-of-learning reveals that it is easy to learn the use of Laevo exoskeleton for rebar tasks. The functionality of the exoskeleton was not complicated, and easy to remember. This may suggest that it can appeal to the construction industry with workers of different ethnicity and educational levels.

Despite the prospects of the exoskeleton, assessing the comfort of users is important for usability. The exoskeleton did not restrict participants' movement and did not interfere with the work environment. This is important as construction activities are dynamic and requires workers to assume awkward postures while working under aggressive environmental conditions. The comfort of working with the exoskeleton was further investigated by assessing users' perceived level of discomfort across different body parts. The findings revealed that the exoskeleton reduced discomfort felt in the lower back and lower leg. This supports similar studies [2, 3] where the use of exoskeletons reduced discomfort at the back and suggests the potential of the exoskeleton for reducing low back injuries during rebar tasks. However, unintended consequences such as increase in the discomfort in the neck, thigh and shoulder and a significant increase in the discomfort at the chest were reported in this study. This may imply the need for improvements to the design of the exoskeleton to suit the dynamic nature of construction activities.

There are some limitations of this study that may impact the generalizability of the results. The study was a laboratory simulation where participants were students and not construction workers and interaction of the exoskeleton with real site conditions is unknown. Hence, the usability assessment of the exoskeleton with experienced construction rebar workers will be explored in future studies. In addition, a high variability across the perceived discomfort levels was reported which shows that these results are highly subjective. Hence, objective measures such as measuring the muscle activity using electromyography sensors, identifying unsafe postures using inertial measurement unit, and assessing discomfort across different body parts are required to further validate these findings.

Acknowledgement

This material is based upon work partly supported by the Construction and Infrastructure Research Affiliates Program (CIRAP).

References

- [1] 1. Cheng, T. and J. Teizer, Real-time resource location data collection and visualization technology for construction safety and activity monitoring applications. *Automation in construction*, 2013. **34**: p. 3-15.
- [2] 2. Statistics, U.S.B.o.L.a. Nonfatal cases involving days away from work. 2020 2/27/2022]; Available from: <https://data.bls.gov/pdq/SurveyOutputServlet>.
- [3] 3. Lee, H., et al., Detecting excessive load-carrying tasks using a deep learning network with a

- Gramian Angular Field. Automation in Construction, 2020. **120**: p. 103390.
- [4] 4. Inyang, N., et al., Ergonomic analysis and the need for its integration for planning and assessing construction tasks. Journal of Construction Engineering and Management, 2012. **138**(12): p. 1370-1376.
- [5] 5. Ogunseiju, O., et al., Subjective Evaluation of Passive Back-Support Exoskeleton for Flooring Work. EPiC Series in Built Environment, 2021. **2**: p. 10-17.
- [6] 6. Ofori-Bah, C.O., US firms performance during recessions: a comparative case study. 2020.
- [7] 7. Taylor Moore, J., et al., Construction workers' reasons for not reporting work-related injuries: an exploratory study. International journal of occupational safety and ergonomics, 2013. **19**(1): p. 97-105.
- [8] 8. Yeau, K.Y. and H. Sezen, Load-rating procedures and performance evaluation of metal culverts. Journal of Bridge Engineering, 2012. **17**(1): p. 71-80.
- [9] 9. Forde, M.S. and B. Buchholz, Task content and physical ergonomic risk factors in construction ironwork. International Journal of Industrial Ergonomics, 2004. **34**(4): p. 319-333.
- [10] 10. Albers, J.T. and S.D. Hudock, Biomechanical assessment of three rebar tying techniques. International Journal of Occupational Safety and Ergonomics, 2007. **13**(3): p. 279-289.
- [11] 11. Yang, K., C.R. Ahn, and H. Kim, Deep learning-based classification of work-related physical load levels in construction. Advanced Engineering Informatics, 2020. **45**: p. 101104.
- [12] 12. Wang, D., F. Dai, and X. Ning, Risk assessment of work-related musculoskeletal disorders in construction: state-of-the-art review. Journal of Construction Engineering and management, 2015. **141**(6): p. 04015008.
- [13] 13. Cho, Y.K., et al. A robotic wearable exoskeleton for construction worker's safety and health. in ASCE construction research congress. 2018.
- [14] 14. Gonsalves, N., et al., Influence of a Back-Support Exoskeleton on Physical Demands of Rebar Work. EPiC Series in Built Environment, 2021. **2**: p. 1-9.
- [15] 15. Zhu, Z., A. Dutta, and F. Dai, Exoskeletons for manual material handling—A review and implication for construction applications. Automation in Construction, 2021. **122**: p. 103493.
- [16] 16. Kim, S., et al., Potential of exoskeleton technologies to enhance safety, health, and performance in construction: Industry perspectives and future research directions. IISE Transactions on Occupational Ergonomics and Human Factors, 2019. **7**(3-4): p. 185-191.
- [17] 17. Kim, S., et al., Assessing the potential for “undesired” effects of passive back-support exoskeleton use during a simulated manual assembly task: Muscle activity, posture, balance, discomfort, and usability. Applied Ergonomics, 2020. **89**: p. 103194.
- [18] 18. Alemi, M.M., et al., Effects of two passive back-support exoskeletons on muscle activity, energy expenditure, and subjective assessments during repetitive lifting. Human factors, 2020. **62**(3): p. 458-474.

Industry Perspectives of the Potential of Wearable Robot for Pipe Installation Work

N. J. Gonsalves, M. Khalid, A. Akinniyi and A. Akanmu

Virginia Polytechnic Institute and State University, USA

E-mail: gonsnihar@vt.edu, khalidm21@vt.edu, abiola@vt.edu, abiola@vt.edu

Abstract –

The physically demanding nature of construction work exposes workers to ergonomic risks resulting back-related musculoskeletal disorders. Back injuries amongst pipe installers increased by 2.3 times in the last year. Back support exoskeletons are emerging as a potential intervention to address back injuries. Without willingness of construction workers to use back-support exoskeletons, the intervention will not be successful in the construction industry. This paper presents the perception of pipe installers regarding the suitability of a commercially available back-support exoskeleton for pipe installation work. Fourteen pipe installers performed their regular work task with a passive exoskeleton during which they provided their experience with the wearable technology. The results indicate that the benefits of the exoskeleton, barriers to the use of the exoskeleton and modifications to the design of the exoskeleton. Participants perceived health benefits in terms of reduced back stress and recommended employing the exoskeleton for outdoor manual labor activities. There were concerns about the use of the exoskeleton in confined spaces and the compatibility of the exoskeleton with on-site safety provisions. Integration of safety harness with the wearable robot was identified to be an essential modification. The findings showcase willingness amongst the pipe installers to adopt exoskeletons. This study contributes to the existing literature on the suitability of passive back-support exoskeleton in construction.

Keywords –

Wearable robot; User assessment; Pipe installation; Exoskeletons

1 Introduction

The construction industry is a labour-intensive sector with high risk of injuries. Each year, federal agencies such as the United States Department of Labor record non-fatal injuries in the construction industry. Non-fatal injuries are sometimes triggered by the physically

demanding nature of construction work [1], which exposes workers to awkward work postures (e.g., twisting, reaching, pulling, lifting and bending) resulting in work-related musculoskeletal disorders (WMSDs) [2].

WMSDs account for 33% of all workplace injuries [1]. In 2020, the United States Bureau of Labor and Statistics [3] reported a WMSD incidence rate of 40.6 musculoskeletal disorders (MSD) per 10,000 full time workers (FTE) which is approximately 1.7 times higher than the average of all the industries. This condition is more severe amongst pipe layers and fitters whose WMSD incident rate is 1.4 times the rate of other construction trades [4]. In fact, the rate of WMSDs amongst pipe layers tripled between 2018 and 2019 [4]. Pipe layers spend a significant amount of time in back-bending postures, when performing pipe installation tasks, that can impact the muscles, nerves, discs, and ligaments in the back. This causes injuries or disorders in the back, which accounts for 43% of all the affected body parts in the construction industry [5]. Evidence suggests that back injuries could cause permanent impairment leading to early retirement of the workforce[6]. Besides the health risks, WMSDs have significant financial consequences and has resulted in work absenteeism in the construction industry [7]. Moreover, WMSDs are one of the leading causes of lost productivity among construction workers [8]. Thus, the health and safety performance of the workforce has a direct impact on the profitability of construction projects [9].

Wearable robots, also referred to as exoskeletons, are increasingly being recognized as a promising ergonomic solution for preventing WMSDs. Wearable robots are designed to reinforce the wearer's performance by augmenting key body parts (e.g., back and shoulder) or the full body. Wearable robots or exoskeletons can be classified as 'active' or 'passive'. An active exoskeleton uses actuators to support the wearer's effort and stimulates the joints, while passive exoskeletons use springs to store energy from the wearer's motion to provide ergonomic support [10]. Unlike active exoskeletons, passive exoskeletons are lighter and more cost-effective. Considering that the back is the body part that is most affected by WMSDs during pipe installation,

back-support exoskeletons could be a potential solution to reduce the physical demands and fatigue experienced by pipe workers, thus improving their safety, health, and performance [11].

As evidence continues to evolve regarding the suitability of back-support exoskeletons in the construction industry [12], there is a need to evaluate user acceptance amongst construction workers. Lack of user acceptance has long been an impediment to the successful deployment of technology in the workplace [13]. Perspectives of end-users could help identify workers' willingness to use the technology, task-specific applications, and facilitators and barriers to advancing back-support exoskeletons in the construction industry.

Existing studies [14, 15] on exoskeletons are mostly laboratory-based studies, which do not reflect the nature of interactions between exoskeletons and the work environment. Construction sites are often characterized as harsh environments with somewhat unsuitable conditions such as dusty and muddy surroundings, hot and cold weather conditions, and confined spaces [16]. Under these conditions, the use of back-support exoskeletons could have unintended consequences such as discomfort to the body parts, device failure and incompatibility with personal protective equipment [5]. Insights into how commercially available exoskeletons can be deployed in such environments could be beneficial for designers to adapt designs of existing exoskeletons to suit construction work. This is significant, as the commercially available exoskeletons are not designed specifically for use in the construction industry [5].

Therefore, the study aims to understand the perceptions of potential beneficiaries of exoskeleton (e.g., pipe layers) regarding the suitability of a commercially available back support exoskeleton for construction work. User perception is captured in terms of the benefits of back-support exoskeletons, barriers to the use of back-support exoskeleton for pipe installation work, and modifications necessary to adapt back-support exoskeletons to pipe installation work.

2 Background

Over the years, there has been several attempts to address the issue of WMSDs in the construction industry. These efforts have ranged from proactive to more reactive measures (e.g., using exoskeletons). The reactive approach includes educating workers about the risk associated with their work so that they can self-manage or control their exposure. For example, developed training manuals and programs to educate workers on how to perform manual material handling tasks in safe postures. To provide workers with opportunities to practice safe work procedures, immersive training environments (e.g., virtual reality) have been proposed

[16]. Yan, Li [17] also developed a framework that uses sensing technologies (e.g. inertial measurement units) to track workers movements while on the job and alerts them on unsafe work habits so that they can control their exposures. The alerts could serve as a distraction and affect workers' productivity.

Commercially available passive back-support exoskeletons, such as BackX and Laevo, are being recognized as a preventive approach to WMSDs. BackX has been demonstrated, via laboratory studies, as being more suitable for work involving back bending and repetitive lifting [18-20]. Till date, there is scarce evidence on how back-support exoskeletons (e.g., BackX) can benefit construction workers. In particular, little is known of how exoskeletons would benefit potential end-users such as pipe layers. It is possible that existing design configurations would need to be modified to suit construction work. However, these can only be determined by deploying the exoskeleton in the field and obtaining feedback from construction workers.

3 Methodology

3.1 Participants

Fourteen pipe layers, installing sewage and water pipelines in Northern Virginia in the United States, volunteered to participate in this study. All the participants signed the informed consent form approved by the Virginia Tech Information Review Board (#IRB 19-1180). None of the workers reported any muscle related injuries, which could influence their ability to carry out daily tasks or could potentially affect their perception of the exoskeleton. The demographics of the participants are as shown in Table 1.

Table 1. Participant's demographics

Demographic Parameters	Mean	SD	Max.	Min.
Age (yrs.)	37.63	10.19	55	22
Height (ft.)	5' 7"	3.95"	6' 11"	5'
Weight (kgs.)	84.12	10.97	108.86	63.5
BMI (kg/m ²)	27.96	3.95	38.7	20.7
Exp. (yrs.)	12.92	7.37	35	2
Note: SD = Standard Deviation, Exp. = Experience, yrs. = years, Max. = Maximum, Min. = Minimum.				

3.2 Wearable Robot

BackX™ S, a commercially available passive back-support exoskeleton was used in this study. The choice of BackX™ S for this study was guided by the benefits documented in Kazerooni, Tung [21]. The exoskeleton is

designed to reduce back strain when performing activities that require bending, stooping, or reaching. BackX™ S weighs 3.4kg and can sustain a load of up to 13kg. BackX™ S, shown in Figure 1, comprises of a frame and a harness. The frame houses a torque generator (i.e., the activation point), a chest-plate, and thigh and leg straps. The harness is made up of a chest pad, a hip belt, and shoulder straps, which are all attached to the body via the frame. The exoskeleton has numerous fixture configurations that may be customized to fit the user's body structure.

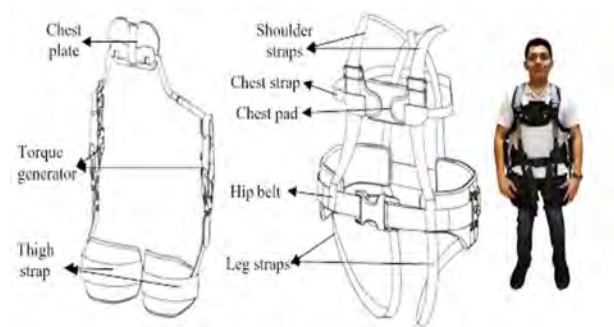


Figure 1. BackX exoskeleton (a) frame (left), (b) harness (middle) and (c) complete exoskeleton (right)

3.3 Experimental Design

Prior to commencing the study, the study procedure was explained to the participants. Subsequently, each participant signed the informed consent form. Thereafter, the back-support exoskeleton was fitted on the participants (Figure 2) considering their height, chest width and waist size as per the guidelines provided by the manufacturers. The functioning of the exoskeleton was explained to the participants and they had the opportunity to explore its operations until they were confident to deploy it. Subsequently, the participants performed their regular pipe installation tasks while wearing the exoskeleton. The tasks were performed for four hours. The tasks involved climbing, squatting, kneeling, bending, cutting pipe, carrying heavy tools, working in trench boxes and confined spaces, working with fall protection, pouring inverts, levelling and shovelling. While performing the tasks, the participants were prompted to provide verbal feedback regarding their experience (e.g., body discomfort and interference with work) with the back-support exoskeleton. At the end of the task, the participants completed a questionnaire designed to capture their perceptions regarding the usability of the exoskeleton, their comfort with the exoskeleton, impact on their performance when using the exoskeleton and how safe they were while using the exoskeleton. The questionnaire also contained open-

ended questions aimed at further capturing the participants' subjective feedback regarding the pros and cons of the back-support exoskeleton, context for the use of the exoskeleton in construction and any modifications that could be made to the design of the exoskeleton to make it more suitable for construction work. Given the length restrictions of this paper, only the subjective feedback to the open-ended questions are reported in this paper.



Figure 2. Pipe layer wearing BackX

3.4 Data Analysis

The participants' responses to the probing questions (obtained while performing work) and the open-ended questions (obtained after the task) were recorded by the investigators. Transcripts of the responses were imported into NVivo 11, and analysed using Thematic analysis [22]. NVivo is a qualitative analysis software which is commonly used by researchers to analyse qualitative data from interviews, surveys, and focus groups [23]. The data were coded using various themes that emerged from the responses. Thereafter, using the inductive coding process, similar or related codes were clustered to form meaningful themes which were identified as categories as shown in Figure 3 and presented in the next section. To confirm the validity and reliability of the findings, an inter-coder reliability testing was conducted using Cohen-kappa coefficient. Cohen-kappa is a commonly used method to measure the level of agreement between coders, where the value of the coefficient ranges from 0 to 1, with 0 being no agreement and 1 being perfect agreement [24]. The assessment showed a percentage agreement of 78% between two coders. A Cohen-kappa coefficient of 0.62 was obtained, showing substantial agreement.

4 Results

Three categories of the themes were extracted from the responses (Figure 3). These include benefits of the back-support exoskeletons, barriers to implementation of back-support exoskeletons on construction projects, and modifications to the design of the back-support exoskeleton.



Figure 3. Categories and sub-categories from participants' responses

4.1 Benefits of Exoskeletons

85% of the participants perceived the exoskeleton to be beneficial for construction related tasks and described specific tasks where the exoskeleton could be most useful. The participants suggested tasks such as shovelling, levelling, working in inverts, forward bending, walking on uneven surfaces, working outdoors and applying lubricant for pipe joints: *"While shovelling and applying the pipe lubricant, it was helpful"*, *"I would use it for invert"*, *"While walking on an uneven surface, on a pile of dirt, I did not have any problems and no imbalances out of the ordinary. Also, while walking uphill the exoskeleton helped me. While picking loads straight up, it is useful"*. About two-thirds of the responders perceived health benefits as a major outcome of using a back-support exoskeleton. According to one of the respondents: *"good device for people with back injuries as it provides good support."* Another participant reported, *"Very beneficial as it provides support to the thighs and chest which helps me reduce the stress from the back"*. One of the participants felt like the exoskeleton helps in keeping the back straight while working: *"felt like the exoskeleton also helps in keeping the back straight which is very useful.....What I like most is the ability to feel no strain on back and knees*

while working". The participants also emphasized some general benefits of the design of the back-support exoskeleton e.g., lightweight, comfortable chest and thigh support: *"it is not heavy, rather lighter than the fall protection that we use"*, *"I like the chest and the thigh support"*. Furthermore, the participants also provided some suggestions for the use of exoskeletons *"the exoskeleton can be very helpful when we use it in an outdoor environment"*. Some participants also felt *"While using the ladder it is better to switch the support system to off"*. Of all the participants who perceived the back-support exoskeleton to be beneficial, 38% suggested potential construction tasks where the use of exoskeletons would be beneficial, 35% suggested health benefits of the exoskeleton, 15% suggested ways by which the exoskeleton can be operated for effective and 12% suggested design benefits (Figure 4).

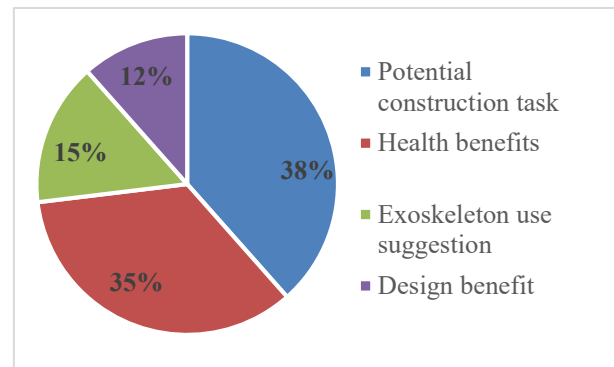


Figure 4. Benefits of back-support exoskeleton

4.2 Barriers to Implementation of Exoskeletons

Overall, 75% of the participants identified potential barriers for the implementation of back support exoskeletons. Some of these respondents had concerns about the design of the exoskeleton and perceived these as barriers to implementing the technology for pipe-laying tasks. A participant mentioned, *"the metal rods caused problems for my under arms."* Another participant reported, *"the chest pad makes it very hard to go down"*. Similarly, some of the participants expressed concerns about the discomfort from the device: *"With the exoskeleton, it is not easy to bend and work...The exoskeleton puts pressure on my body parts especially chest and hips...it causes discomfort to my hips."* The participants also concluded that wearing an exoskeleton with a safety harness would be challenging: *"it does not work with the fall protection harness and I sweat much more while working."* Some of the participants were also concerned about pressure imposed on their chests by the pads, which caused their heartbeats to race and made

them feel exhausted: “When working hard, due to the pressure my heart pumps faster which makes me tired.” On the other hand, trenches and manholes are the most common areas that pipe layers work, and the participants reported that their mobility was restricted while working in such tight spaces: “I do not think it is suitable for pipe work as we work in tight spaces”. With a variety of exemplary benefits, the exoskeleton also introduces several safety issues: “if anyone falls while wearing the exoskeleton, it can hurt them because of the metal parts”. From these implementation barriers, design barriers (27%) and discomfort (27%) were identified as the most significant, followed by work environment (20%), safety (17%), work preference (7%), and weather barriers (3%) (Figure 5).

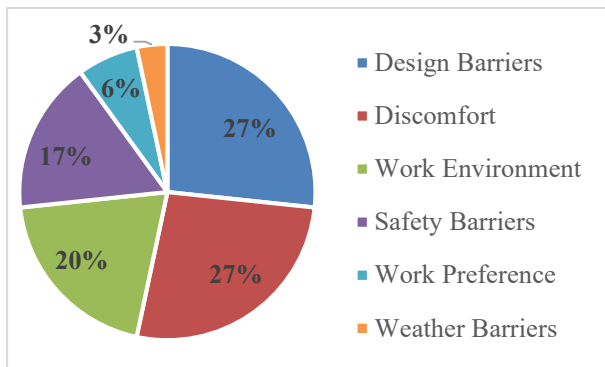


Figure 5. Implementation barriers

4.3 Suggestions to Modification of Exoskeleton

Although the exoskeleton is expected to provide considerable support to workers during pipe installation work, numerous modifications were suggested to make the device more suitable for construction related tasks.

71% of the participants suggested modifications to the exoskeleton. For example, some participants suggested the integration of the safety harness with the back-support exoskeleton: “If there was a design with in-built fall protection, it would be great.” Another important add-on feature for the exoskeleton was suggested by a participant: “If the harness had any spots to carry my tools it would be very useful... have straps for my tape and communication device etc.”. One of the interviewees recommended a weather adaptability function for the device: “If we can change the colour to white then maybe it would be much better”. Another suggestion is, “If the torso was closer to the body and not coming out, it would be better”. One participant felt that having pressure points on the back body part would be beneficial: “It would have been better if the pressure was on the back”. Accordingly, the integration of the safety harness with the back-support exoskeleton was the most

suggested modification (40%), followed by reducing the pressure from the chest pad (10%), repositioning of the metal torso (10%), and back support (10%) (Figure 6).

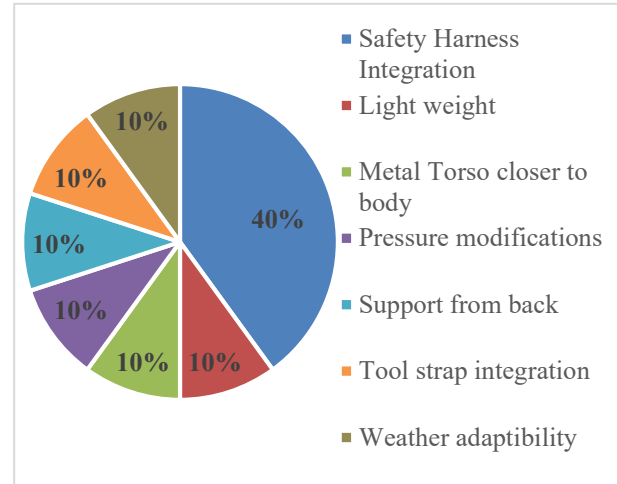


Figure 6. Modifications to back-support exoskeletons

5 Discussion

This study presents a user-assessment of a commercially available back support wearable robot for pipe installation. The result of the assessment includes the benefits of back-support exoskeleton, barriers to the use of back-support exoskeleton and modifications that could be made to back-support exoskeletons to improve their suitability for construction work. The back-support exoskeleton can provide health benefits such as reduction of stress on the back muscles, which is consistent with previous studies [18, 25]. The workers found value in using the back-support exoskeleton for forward bending tasks such as shovelling, levelling and pouring storm-water inverts. This is contrary to Bosch, van Eck [26] where there was an increase in discomfort in the chest when the participants performed work in similar postures. The discomfort was due to pressure from the chest and thigh pad which some participants in this study found beneficial. Also, the workers felt that wearing the back-support exoskeleton could help them work for longer hours which would increase their productivity. This is consistent with the study of Kim, Moore [27] where the authors found that using back support exoskeleton would help workers perform work faster and for longer durations.

Despite the benefits of using a back-support exoskeleton, some barriers were also noticed which should be taken into consideration. Pressure exerted by the metal torso to the chest caused discomfort to the chest while wearing a back-support exoskeleton. This, on the

other hand, does not correlate with the findings of Antwi-Afari, Li [28] which showed that the use of a back support exoskeleton significantly reduced discomfort on the chest. This disparity could be due to the first-time effect of wearing the back-support exoskeleton, and may not suffice after prolonged use.

The workers felt that using both the exoskeleton and tool belt would impact their productivity and ability to work in confined spaces. A key modification suggested to the design of the back-support exoskeleton includes integrating the exoskeleton with workers' tool belt. This could reduce the time to don and doff the exoskeleton. A similar attempt was made by Salvietti, Franco [29] who integrated a passive exoskeleton with a robotic supernumerary finger to improve grasp compensation in Chronic stroke patients. Furthermore, Kim, Moore [27] identified cost as a critical factor for adoption of exoskeleton in construction industry. The integration of tool belt and safety harness with the exoskeleton as suggested by the pipe workers could increase the cost of the device, which could be a potential barrier for adoption of exoskeleton.

6 Conclusion

This study focused on understanding construction workers' perception of the suitability of a commercially available back-support exoskeleton for construction work. The back-support exoskeleton was found to be beneficial for construction tasks and having significant health benefits of reducing back stress. While there were some discomfort experienced from the use of the exoskeleton, most of the participants found it beneficial to supporting the body during pipe installation. Workers could be more willing to adopt the exoskeleton if design can be improved to accommodate existing work wearables such as tool belts and personal protective equipment (e.g., safety vests and harness).

Furthermore, a small sample size was adopted in this study, which is not sufficient to generalize the findings to the entire construction industry. Future work will involve a larger sample size with demographics representative of the construction industry. In addition, the exoskeleton was tested for pipe work, thus the findings might not be adaptable to other types of construction work. In order to promote widespread adoption of exoskeleton to mitigate occurrences of WMSDs, similar studies will need to be carried out to identify suitability of the exoskeleton for supporting other construction trades. Furthermore, this is a short-term field study where the participants used the exoskeleton for four hours. To understand the willingness of construction workers to adopt exoskeletons and evaluate the impact of prolonged use of exoskeletons for construction work, a long-term field study is necessary.

Acknowledgement

This material is based upon work partly supported by the Construction and Infrastructure Research Affiliates Program (CIRAP).

References

1. Li, C. and S. Lee, Computer vision techniques for worker motion analysis to reduce musculoskeletal disorders in construction, in *Computing in Civil Engineering* (2011). 2011. p. 380-387.
2. Umer, W., et al., Low-cost ergonomic intervention for mitigating physical and subjective discomfort during manual rebar tying. *Journal of Construction Engineering and Management*, 2017. **143**(10): p. 04017075.
3. Statistics, U.S.B.o.L.a. Nonfatal cases involving days away from work. 2020 2/27/2022]; Available from: <https://data.bls.gov/pdq/SurveyOutputServlet>.
4. Data, B.o.L.S., 2022.
5. Ogunseiju, O., et al., Subjective Evaluation of Passive Back-Support Exoskeleton for Flooring Work. *EPiC Series in Built Environment*, 2021. **2**: p. 10-17.
6. Taylor Moore, J., et al., Construction workers' reasons for not reporting work-related injuries: an exploratory study. *International journal of occupational safety and ergonomics*, 2013. **19**(1): p. 97-105.
7. Siedl, S.M. and M. Mara, Exoskeleton acceptance and its relationship to self-efficacy enhancement, perceived usefulness, and physical relief: A field study among logistics workers. *Wearable Technologies*, 2021. **2**.
8. Akhavan, R. and A.H. Behzadan, Smartphone-based construction workers' activity recognition and classification. *Automation in Construction*, 2016. **71**: p. 198-209.
9. Yeau, K.Y. and H. Sezen, Load-rating procedures and performance evaluation of metal culverts. *Journal of Bridge Engineering*, 2012. **17**(1): p. 71-80.
10. Wang, D., F. Dai, and X. Ning, Risk assessment of work-related musculoskeletal disorders in construction: state-of-the-art review. *Journal of Construction Engineering and management*, 2015. **141**(6): p. 04015008.
11. Cho, Y.K., et al. A robotic wearable exoskeleton for construction worker's safety and health. in *ASCE construction research congress*. 2018.
12. Gonsalves, N.J., et al., Assessment of a passive wearable robot for reducing low back disorders during rebar work. *Journal of Information*

- Technology in Construction (ITCON), 2021. **26**: p. 936-952.
13. Davis, F.D., User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International journal of man-machine studies*, 1993. **38**(3): p. 475-487.
14. De Bock, S., et al., Passive shoulder exoskeletons: more effective in the lab than in the field? *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2020. **29**: p. 173-183.
15. Luger, T., et al., Using a back exoskeleton during industrial and functional tasks—Effects on muscle activity, posture, performance, usability, and wearer discomfort in a laboratory trial. *Human Factors*, 2021: p. 00187208211007267.
16. Akanmu, A.A., et al., Cyber-physical postural training system for construction workers. *Automation in Construction*, 2020. **117**: p. 103272.
17. Yan, X., et al., Personalized method for self-management of trunk postural ergonomic hazards in construction rebar ironwork. *Advanced Engineering Informatics*, 2018. **37**: p. 31-41.
18. Koopman, A.S., et al., Effects of a passive exoskeleton on the mechanical loading of the low back in static holding tasks. *Journal of biomechanics*, 2019. **83**: p. 97-103.
19. Madinei, S., et al. Assessment of Two Passive Back-Support Exoskeletons in a Simulated Precision Manual Assembly Task. in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 2019. SAGE Publications Sage CA: Los Angeles, CA.
20. Madinei, S., et al., Effects of back-support exoskeleton use on trunk neuromuscular control during repetitive lifting: A dynamical systems analysis. *Journal of Biomechanics*, 2021. **123**: p. 110501.
21. Kazerooni, H., W. Tung, and M. Pillai. Evaluation of trunk-supporting exoskeleton. in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 2019. SAGE Publications Sage CA: Los Angeles, CA.
22. Guest, G., K.M. MacQueen, and E.E. Namey, *Applied thematic analysis*. 2011: sage publications.
23. Welsh, E. Dealing with data: Using NVivo in the qualitative data analysis process. in *Forum qualitative sozialforschung/Forum: qualitative social research*. 2002.
24. Cohen, J., A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 1960. **20**(1): p. 37-46.
25. Hensel, R. and M. Keil, Subjective evaluation of a passive industrial exoskeleton for lower-back support: A field study in the automotive sector. *IIE Transactions on Occupational Ergonomics and Human Factors*, 2019. **7**(3-4): p. 213-221.
26. Bosch, T., et al., The effects of a passive exoskeleton on muscle activity, discomfort and endurance time in forward bending work. *Applied ergonomics*, 2016. **54**: p. 212-217.
27. Kim, S., et al., Potential of exoskeleton technologies to enhance safety, health, and performance in construction: Industry perspectives and future research directions. *IIE Transactions on Occupational Ergonomics and Human Factors*, 2019. **7**(3-4): p. 185-191.
28. Antwi-Afari, M.F., et al., Assessment of a passive exoskeleton system on spinal biomechanics and subjective responses during manual repetitive handling tasks among construction workers. *Safety science*, 2021. **142**: p. 105382.
29. Salvietti, G., et al., Integration of a Passive Exoskeleton and a Robotic Supernumerary Finger for Grasping Compensation in Chronic Stroke Patients: The SoftPro Wearable System. *Frontiers in Robotics and AI*, 2021. **8**.

Development of Framework for Highway Lawn Condition Monitoring using UAV Images

Y. Kim^a, S. Kim^a, Y. Yajima^b, J. Irizarry and^c, and Y.K. Cho^a

^aSchool of Civil and Environmental Engineering, Georgia Institute of Technology, USA.

^bInstitute for Robotics and Intelligent Machines, Georgia Institute of Technology, USA.

^cSchool of Building Construction, Georgia Institute of Technology, USA.

E-mail: ykim858@gatech.edu, skim3310@gatech.edu, yajima@gatech.edu,
javier.irizarry@gatech.edu, yong.cho@ce.gatech.edu,

Abstract –

Planning, monitoring, and maintenance of highway assets is an essential, long-term operation for successful civil infrastructure management. These monitoring and maintenance activities are usually carried out manually, suffering from time-consuming, costly, potentially dangerous tasks. The advancements in Unmanned Aerial Vehicles (UAVs) and computer vision technologies have demonstrated the potential to enable automation of the monitoring workflows. Existing UAS-based approaches are used for various management; however, there was no study to examine the feasibility of aerial image-based computer vision algorithms for the purpose of lawn condition monitoring. This study aims to provide periodic and easy-to-use UAV technology for civil infrastructure maintenance. We developed the comprehensive framework from UAS data collection to build a deep learning model suitable to distinguish areas of interest with vague boundaries robustly, process the outputs into geo-database, and visualize them through a Geographic Information Systems (GIS) platform. The outcome of the proposed framework displays the overall mowing quality in the highway environment in an intuitive way to support decision-making in the management.

Keywords – Civil infrastructure management; Computer-vision; UAVs; Deep Learning; Lawn mowing condition assessment

1 Introduction

Roadways are one of the most vital infrastructures necessary for public transportation and industrial operations. To secure the robustness of roadway infrastructure, the roadway infrastructure condition (e.g., road condition monitoring, lawn condition monitoring) is monitored and assessed for maintenance needs. These inspection, maintenance, and post-maintenance

inspection are often carried out manually with visual inspection by trained inspectors. Thus, it is often time-consuming, laborious, and expensive. One of such tasks is lawn mowing control on the roadway infrastructure. Typically, the goal of lawn mowing maintenance personnel is to verify if the mowing performance carried out by a third-party contractor is acceptable, identify any poorly mowed regions that require additional mowing, and bared regions that require maintenance. Thus, such lawn mowing control requires significant time and resources due to visual observation and manual measurement to verify the mowing performance. Failure to do so may cause traffic safety issues because tall grass can obscure drivers' vision and may not see traffic control devices or pedestrians [1].

The recent advancement of computer vision algorithms and the prevalent use of UAVs offers a great potential for automation in highway landscape maintenance practices. In particular, UAVs equipped with RGB cameras provide an easy-to-use and inexpensive platform to obtain lawn condition information of the field readily. UAVs can capture images of areas where maintenance work was performed, including areas of difficulty (e.g., steep slope). Maintenance personnel can monitor the lawn mowing conditions more frequently and easily by analyzing these aerial images.

To fully utilize the computer vision-based monitoring method, the following limitations should be addressed: 1) to deal with the variability of lawn condition images due to the range of seasonal lawn condition and the different resolutions of aerial images captured at different heights and angles, and 2) the analyzed output needs to be directly mapped onto the actual physical 3D coordinates. To overcome these limitations, we propose a comprehensive, data-driven automated mowing condition monitoring framework integrated with a user interface that visualizes the outcome in a user-friendly manner. We developed a computer vision-based lawn condition monitoring system that adopts fully

convolutional neural networks to perform pixel-wise classification of the aerial images. We examined the feasibility of utilizing computer-vision technologies in the mowing condition maintenance work and evaluated the proposed system with actual test cases. The outcomes of this study are particularly beneficial for practical applications and support the decision-making process.

2 Related Work

2.1 Civil Infrastructure Monitoring

Monitoring and maintenance of civil infrastructure are paramount and have been a significant interest in society. It is crucial to promptly detect and identify the location of possible areas requiring necessary maintenance because the failure to do so can lead to traffic congestion, driver discomfort, and potential safety and operational problems. Traditionally trained personnel carried out onsite visual observation [2]. These practices are non-trivial tasks requiring significant resources, and they can be subjective and error-prone as the outcome is highly dependent on the expertise of the onsite personnel. Acknowledging the limited amount of resources in contrast to the vast volume of areas to be monitored, the recent advancements in visual data collection hardware such as UAVs as well as computer vision techniques to analyze the collected visual data have demonstrated its potential opportunities to automate civil infrastructure monitoring processes [3]. Several examples of UAV-based inspection applications are agricultural crop and weed monitoring [4–6], road inspection applications [7], and construction and infrastructure monitoring [8].

Despite the increasing applications of UAV and deep learning in the civil engineering domain, UAV application has not yet been widely researched in the field of highway mowing and maintenance services. Lawn mowing control on the roadway infrastructure requires significant time and resources for prompt maintenance. However, lawn mowing control poses several challenges. Firstly, it requires the person to distinguish areas of interest with vague boundaries granularly. Also, the processed information needs to be transferred to spatial context information to capture the physical location of interest areas.

2.2 UAV Image Segmentation and Classification

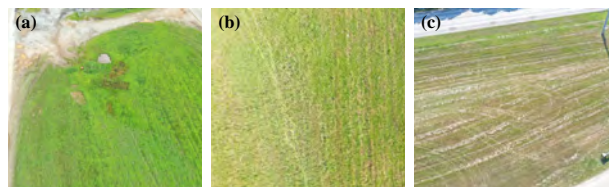
With the prevalent use of UAVs, UAVs gained growing attention for various monitoring applications because they can cover large areas with high-resolution imagery in an inexpensive way. To this end, image segmentation using aerial images captured by UAVs is

also a widely studied topic. For example, in the agricultural areas, UAV images were used to identify weeds from the grass areas [4]. In civil engineering domain, aerial and satellite images were used to assess infrastructure damage in natural disaster areas [9]. In addition, aerial images coupled with machine learning techniques were used to map the rural environment, classifying them into roads, vineyards, asphalt, and roofs [10].

Although there have been studies that demonstrated the capacity of the UAV image segmentation method, there are several challenges to be addressed to deploy in the lawn mowing condition monitoring application. First of all, there is a lack of annotated datasets to develop a segmentation model. Secondly, there is a high intra-variability within the grass conditions based on the seasonal and weather factors (e.g., lighting conditions) and camera factors (e.g., image resolution). To address these limitations, this study aims to develop a framework that can collect and annotate datasets and develop a robust segmentation algorithm for the variance in the collected dataset.

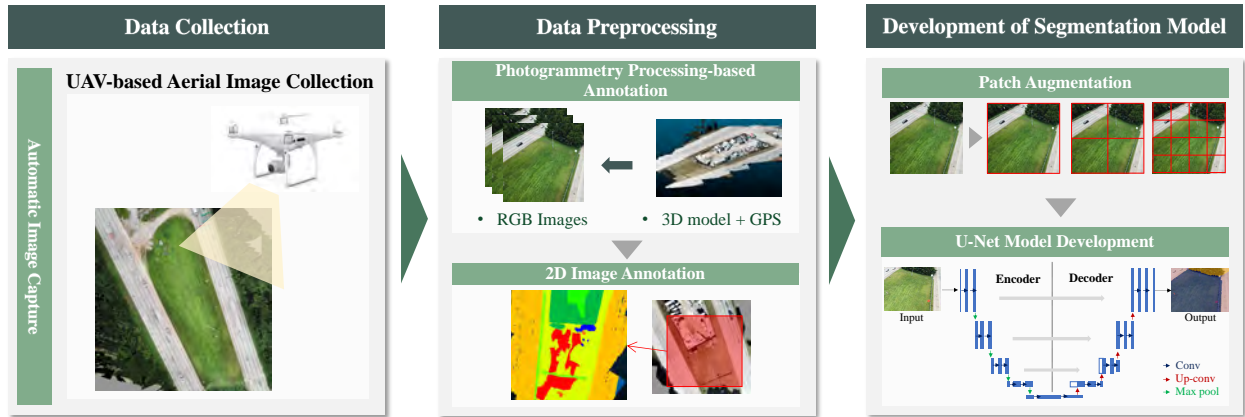
3 Methodology

This study proposes an overall lawn condition classification framework leveraging UAVs equipped with camera sensors. Our proposed framework aims to collect a large volume of dataset for lawn mowing condition monitoring and develop a segmentation algorithm to classify the vegetation conditions into three classes of maintenance personnel's interests: mowed, unmowed, and bared spot, as illustrated in Figure 1. We defined a *mowed* condition as regions where grass height is less than 6 inches, an *un-mowed* condition as regions where grass height is over 6 inches, and a *bare spot* as regions where no grass is detected.



[Fig 1] Examples of lawn conditions: (a) bare spot, (b) un-mowed, and (c) mowed lawn condition

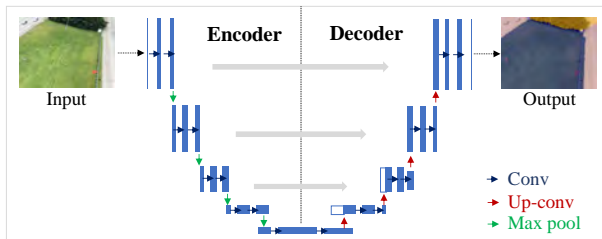
We developed a pixel-wise classification that assigns one of the lawn condition labels to each input image pixel. To automate the annotation process, we labeled the dataset using the 3D semantic segmentation method that converts 3D points into 2D images [11]. The overall framework is depicted in Figure 2.



[Fig 2] Overview of the proposed framework

3.1 Network Architecture

This section describes how our overall framework utilized the image processing algorithms to distinguish the various grass conditions. We adopt U-Net model [12], which is one of the most widely-used network architecture for semantic image segmentation. U-Net has an encoder-decoder structure with fully convolutional network, as illustrated in Figure 3. The first component of the architecture, the encoder, consists of alternating convolution and pooling operations, allowing the model to progressively increase the number of feature maps by downsampling feature maps. The decoder part of the model then upsamples the feature map and semantically projects the features learned by the encoder onto the pixel space to obtain classification results. Overall, this structure of contracting and expansive operations enables to capture of the localized segmentation of the input image and propagates the feature information to successive layers with higher resolutions. Due to this robustness to precise segmentation, this study utilized this architecture to capture fine details of the images.



[Fig 3] U-Net architecture

3.2 Data Collection and Preprocessing

3.2.1 Aerial Image Collection

For data collection, a DJI Marvic Pro and DJI Marvic 2 Pro with 4K cameras and a GPS sensor were deployed. Using a public open mobile application, DJI Go 4 and PIX4Dcapture, a licensed drone pilot set up the flying path and selected the boundaries of the vegetation areas to be collected. The drones automatically captured the images at different elevations and angles based on the flight parameter input, including flight speed, altitude, and camera angle.

3.2.2 Image Annotation Process

As the image collected from UAV inherently contains a significant area of overlap, a 2D-3D co-labeling method [11] that can label the same area only once was adopted to generate image annotations efficiently. First, the collected aerial images were processed through the Structure-from-Motion (SfM) and converted into 3D point cloud data. With respect to the manual 3D annotation, semantic labels were assigned to the original drone images by utilizing the camera projection equation. Since each image's camera intrinsic and extrinsic parameters were known after the SfM step, each 3D point can be associated with a 2D pixel. This approach offers an efficient annotation method because it can produce labels of all aerial images with a one-time manual annotation process.

3.2.3 Data Augmentation

We applied a patch augmentation strategy to make the classification model robust to the variability of vegetation conditions and Ground Sampling Distance (GSD). It processed an $n \times n$ grid over the original image and its corresponding ground truth. The image is then

divided into small pieces, which are treated individually, as depicted in Figure 4. Since the input size of deep learning architecture is much smaller than that of the raw drone images, it can be expected to increase of GSD range in the training dataset.



[Fig 4] Patch-based data augmentation

4 Experiment

4.1 Data Collection

We applied our method to several highway locations in Atlanta, Georgia. Figure 5 shows the drone data collected at actual highway environments: the I-675 Highway that is maintained by the Georgia Department of Transportation (GDOT). The total area of the location is approximately 1.742 (acre), and the length along the primary axis is 150 (m). The data collections were regularly performed to monitor a wide range of grass conditions at different mowing conditions and seasons, as illustrated in Table 1. The test environment includes roads, vegetation, trees, and other objects.

The flight parameters of the UAV consist of flight elevation between 30 to 50 (m) from the ground, GSD of 0.99 – 1.64 (cm/pixels), and camera angles of 45, 60, and 90 degrees. An individual combination of the flight parameters was tested to find the optimistic performance of assessing mowing quality. A trained pilot with an observer controlled the drone's flight and monitored the process to take control of the automated flight if needed.

To acquire ground truth of lawn mowing condition, visual inspection and tape measurements were manually conducted at several locations. Based on the 15cm threshold of grass height, the ground truth of indicating mowed and un-mowed areas was generated for each data collection. These ground truths are used to validate the proposed annotation process.



[Fig 5] Test location under different field conditions collected on (1) Jul 22 and (b) Nov 11, 2021.

4.2 Development of grass condition assessment algorithm using image data

A customized dataset of over 700 aerial images was used to train the proposed U-Net, and the corresponding labeled segmented images. We employed a total of 4 classes, namely mowed, un-mowed, bare spot, and miscellaneous classes, which include non-grass regions such as roads, guardrails, signages, and trees. To improve the model's performance, we experimented with different the GSD of drone images and determined the final input image size. That is, after downsampling to the deep learning input size, the GSD of 20~25cm/pixels, and can be prepared by dividing each original image into four split images in our flight parameters.

Finally, the dataset was divided for training, validation, and testing. The validation dataset was used to tune the hyperparameters of the model that maximize the model accuracy and minimize the sparse categorical cross-entropy loss function. The validation loss is monitored to ensure that there is no overfitting or underfitting during the training process.

4.3 Experimental Evaluation

For evaluation, we produced pixel-wise classification results for input images. To quantitatively validate the performance of our proposed classification framework, we utilize a metric of accuracy, precision, recall, and F1 score, which is the dominant evaluation criteria for the image classification task.

5 Results

5.1 Experiment Results

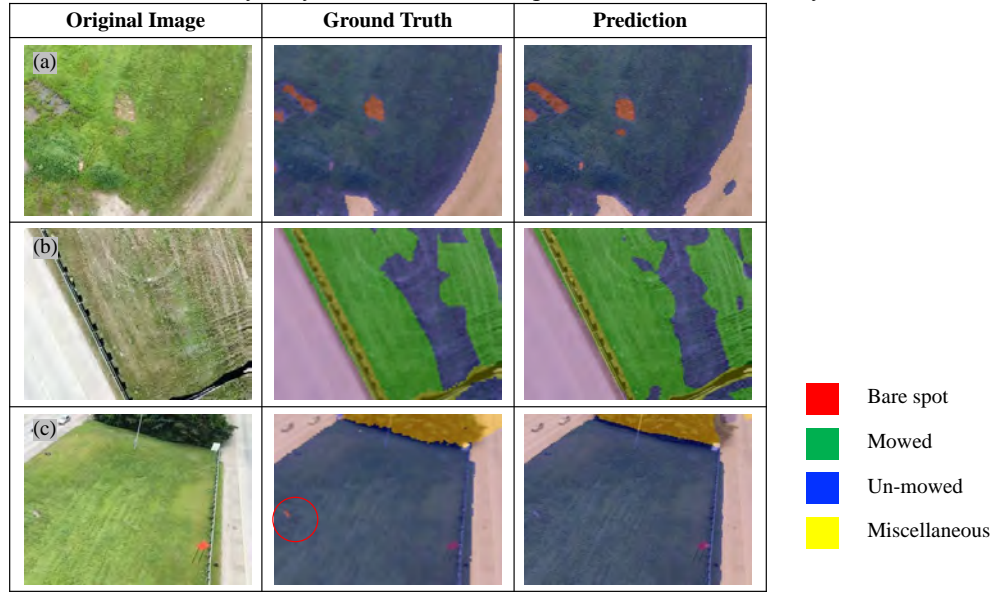
We analyzed the performance of our proposed framework. Figure 6 shows the examples of classification results. The green, blue, and red color legends indicate the lawn conditions of mowed, un-mowed, and bare spots, respectively.

We tested our model in different datasets of grass conditions (Table 1) to check the robustness of our model in terms of handling intra-variability. The overall accuracy was 87.5%, indicating that our model can produce acceptable outcomes across different grass conditions, except for data collected on Nov 11, which had precision and recall values lower than 80%. We believe that this low performance is attributed to the distinctive characteristics of this dataset due to the new construction of building in the test location as illustrated in Figure 5. This resulted in drastic changes in the features of the images.

In addition, we evaluated our model in terms of different classes. Table 2 shows the accuracy results for each class. Compared to other classes, bare spot scores

the lowest accuracy by a large margin. This is due to class imbalance, the lack of a bare spot in the whole dataset. Figure 6 (c) illustrates the relatively very small area of

bare spot compared to other grass condition classes. Therefore, the failure of spotting a small bare spot impacted the overall accuracy a lot.



[Fig 6] Image segmentation results

Table 1. Overall Performance of the Proposed Framework

Collection Date	May 14, 2021	July 22, 2021	July 30, 2021	November 11, 2021
Accuracy	0.872	0.826	0.891	0.912
Precision	0.876	0.768	0.815	0.411
Recall	0.876	0.730	0.874	0.401
F-1	0.811	0.729	0.823	0.405

Table 2. Performance of Lawn Mowing Condition Class

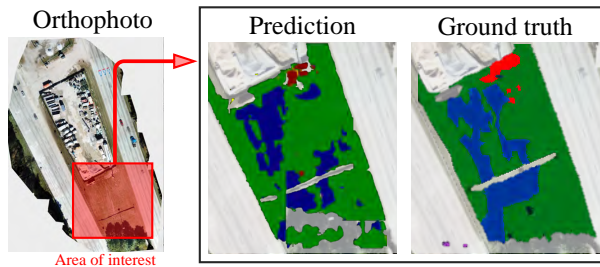
	Mowed	Un-mowed	Bare	Misc
Accuracy	0.989	0.994	0.995	.985
Precision	0.945	0.888	.473	.766
Recall	0.959	0.908	.497	.988
F-1	0.946	0.895	.468	.863

5.2 Geo-Referencing and Digitization

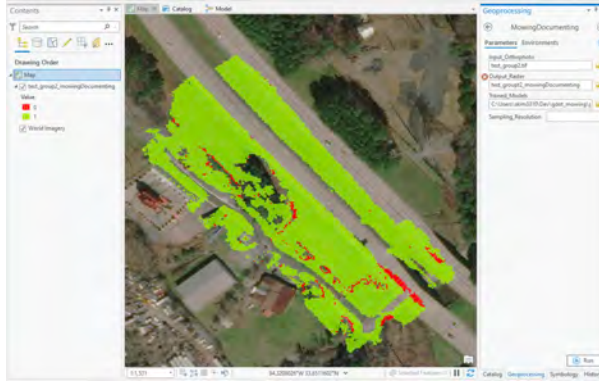
In addition, we mapped the image-level classification results onto the geo-referenced highway site to provide a more user-friendly, informative results. Each orthophoto was produced by stitching and smoothening whole drone images into the whole site was deployed.

To do so, an orthophoto of the highway site was firstly split into multiple patches of the interpretable size in the trained model. Then, each patch was inferred and reverted into the original position while preserving coordinate reference systems. Fig 7 shows an output of

being geo-referenced and digitized classification map in a raster data format. This enables lawn monitoring managers to visualize mowing quality and make decision upon Geographic Information Systems platform. Fig 8 shows the final output in the GIS platform.



[Fig 7] Result with Geo-referencing and Digitization



[Fig 8] Implementation on the GIS platform

6 Discussion

The results demonstrated the potential benefits of our proposed system, there are several limitations to overcome for practical issues for real-world application. First, for practical purposes, orthophoto provide more useful information than a single drone image because it provides the overview with the location information. However, the accuracy of orthophoto classification is lower compared to that of the original aerial image. This is because generating orthophoto causes some texture and color information loss while multiple images are smoothed out and switched together. As a result, orthophoto suffers from low-resolution, incompleteness, lost texture information.

Secondly, the capacity of our proposed model is limited in terms of the highway assets coverage. Several vertical highway assets such as poles are not well detected.

In the future works, these above-mentioned issues can be mitigated by the integration of 3D model. Aerial images from tilted views will be utilized to generate 3D highway scene, and the 3D model will be overlapped with 2D images to further train the classifier.

7 Conclusion

In this study, we showed the feasibility of aerial image-based computer vision algorithms for the purpose

of lawn condition monitoring for highway assets. Our proposed framework expanded the application of computer-vision and drone technology. With the user-friendly GIS-based visualization, it will support the lawn monitoring personnel to easily verify the mowing performance of the contractors and better manage the highway assets condition.

Acknowledgements

The work reported herein was supported by the National Science Foundation (Award #: OIA-2040735) and the Georgia Department of Transportation (GDOT) project RP 20-09/T.O. 2014-99. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF or GDOT.

References

- [1] R.W. Eck, H.W. McGee, United States. Federal Highway Administration. Office of Safety, Vegetation Control for Safety: A Guide for Local Highway and Street Maintenance Personnel: Revised August 2008, 2008.
<https://rosap.nhtl.bts.gov/view/dot/42594> (accessed February 21, 2022).
- [2] K.L. Rens, T.J. Wipf, F.W. Klaiber, Review of Nondestructive Evaluation Techniques of Civil Infrastructure, *J. Perform. Constr. Facil.* 11 (1997) 152–160. [https://doi.org/10.1061/\(ASCE\)0887-3828\(1997\)11:4\(152\)](https://doi.org/10.1061/(ASCE)0887-3828(1997)11:4(152)).
- [3] B.F. Spencer, V. Hoskere, Y. Narazaki, Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring, *Engineering*. 5 (2019) 199–222.
<https://doi.org/10.1016/j.eng.2018.11.030>.
- [4] F. López-Granados, J. Torres-Sánchez, A.-I. De Castro, A. Serrano-Pérez, F.-J. Mesas-Carrascosa, J.-M. Peña, Object-based early monitoring of a grass weed in a grass crop using high resolution UAV imagery, *Agron. Sustain. Dev.* 36 (2016) 67.
<https://doi.org/10.1007/s13593-016-0405-7>.
- [5] A. Brilhador, M. Gutoski, L.T. Hattori, A. de Souza Inácio, A.E. Lazzaretti, H.S. Lopes, Classification of Weeds and Crops at the Pixel-Level Using Convolutional Neural Networks and Data Augmentation, in: 2019 IEEE Lat. Am. Conf. Comput. Intell. -CCI, 2019: pp. 1–6. <https://doi.org/10.1109/LA-CCI47412.2019.9037044>.
- [6] A.I. De Castro, J. Torres-Sánchez, J.M. Peña, F.M. Jiménez-Brenes, O. Csillik, F. López-Granados, An Automatic Random Forest-OBIA Algorithm for Early Weed Mapping between and within Crop Rows

Using UAV Imagery, *Remote Sens.* 10 (2018) 285.
<https://doi.org/10.3390/rs10020285>.

[7] T. Sun, Z. Di, P. Che, C. Liu, Y. Wang, Leveraging Crowdsourced GPS Data for Road Extraction from Aerial Imagery, *Proc. IEEE CVF Conf. Comput. Vis. Pattern Recognit.* (2019).

[8] S. Kim, J. Irizarry, Knowledge-Based Considerations for Developing UAS Operational Procedures on Infrastructure and Construction Task Environments, (2020) 268–277.

<https://doi.org/10.1061/9780784482865.029>.

[9] Y. Pi, N.D. Nath, A.H. Behzadan, Detection and Semantic Segmentation of Disaster Damage in UAV Footage, *J. Comput. Civ. Eng.* 35 (2021) 04020063. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000947](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000947).

[10] S. Bhatnagar, L. Gill, B. Ghosh, Drone Image Segmentation Using Machine and Deep Learning for Mapping Raised Bog Vegetation Communities, *Remote Sens.* 12 (2020) 2602.

<https://doi.org/10.3390/rs12162602>.

[11] Y. Yajima, M. Kahoush, S. Kim, J. Chen, J. Park, S. Kangisser, J. Irizarry, Y. Cho, AI-driven 3D Point Cloud-Based Highway Infrastructure Monitoring System using UAV, in: 2021.

[12] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, *Int. Conf. Med. Image Comput. Comput. Assist. Interv.* (2015).

YAKE-Guided LDA approach for automatic classification of construction safety reports

H. Gadekar^a and N. Bugalia^a

^a Department of Civil Engineering, Indian Institute of Technology Madras, India
E-mail: hrishikesh.gadekar89@gmail.com, nikhilbugalia@gmail.com

Abstract –

Identifying efficient processes for classifying text-based safety reports using Machine-Learning (ML) is an essential area of research. However, much of the previous work on the topic relies on supervised learning approaches, which are often manually intensive and require large volumes of pre-labeled data. To achieve reduced requirements for human intervention during the classification process, the current study tests the applicability and validity of a state-of-the-art unsupervised learning approach, i.e., Yet Another Keyword Extractor (YAKE) integrated with Guided Latent Dirichlet allocation (GLDA). The current study is the first known application of the approach for the construction sector. Web-based, readily accessible information is used to develop a domain corpus. The keywords obtained from the domain corpus using YAKE are seeded in GLDA to classify nearly 13,000 safety reports from two different datasets in 4 commonly used category labels. The study demonstrates that moderate to high classification performance is achievable through the YAKE-GLDA approach. A high F1 score of 0.82 for the Personal-protective equipment category and a total F1 score of 0.62 is achievable. Furthermore, the same domain corpus helps achieve good classification performance across different datasets, highlighting the generalizability of the YAKE-GLDA approach. However, results from novel sensitivity analysis show a non-generalizable trend for sensitivity to hyperparameters. Hence attention is warranted for potential consistency issues facing the approach. The preliminary results demonstrate outstanding potential for the YAKE-GLDA approach for wide-ranging adoption in the construction industry. However, future work should also focus on more granular classification labels applications and improving classification efficiency.

Keywords –

Unsupervised machine learning; Construction safety; Text classification; Topic modeling

1 Introduction

Despite a significant improvement over the years, the construction sector, compared to other sectors, continues to perform poorly for issues relating to Occupational Health and Safety (OHS) [1]. To improve safety performance, one of the central ideas is to collect large volumes of safety observations (SOs), such as accidents, injuries, and near-miss reports, and utilize these reports to enable organizational learning [2].

However, literature has also highlighted construction organizations' challenges in sustaining safety reporting and organizational learning efforts, often due to their resource intensiveness [2]. For example, the safety observations in the construction sector are often unstructured textual data and of poor quality requiring extensive manual efforts to process such information [2]. Therefore, for solving practical issues faced by the construction organizations, identifying efficient processes for the classification of the text-based information (such as the SOs) using Machine-Learning (ML) and text-mining-based approaches continues to be an essential area of research [3,4].

Across domains, including construction, a significant proportion of literature focussing on ML-based classification of SOs continues to rely on supervised ML approaches [4]. Many studies have demonstrated the high classification efficiency of such supervised approaches [3,4]. However, lack of generalizability, the necessity of a large quantity of pre-labeled data, and significant manual inputs during ML-based analysis continue to be a limitation for the broader application of such approaches in actual practice [3,4]. On the other hand, literature exploring unsupervised and semi-supervised approaches is relatively scarce [4,5]. In principle, such approaches can reduce the requirement of human intervention [6].

The key motivation for this paper is to achieve the reduced requirement of human intervention during the classification process while also achieving good classification performance. Consequently, the objective of the current study is to test the applicability and validity of a recently developed unsupervised learning approach,

i.e., Yet Another Keyword Extractor (YAKE) integrated with Guided Latent Dirichlet allocation (GLDA), for classifying construction SOs. The YAKE-GLDA approach is recent and claimed to be a domain-independent approach that has been shown to achieve good classification accuracy with reduced manual efforts [7]. To the best of the authors' knowledge, the current study is the first-ever attempt at the YAKE-GLDA approach for the construction sector [5]. The study makes essential contributions to evaluating the potential of an ML approach in automating SO labeling for the construction sector from an unstructured corpus with minimal manual intervention. Unlike previous studies, the current study also presents the results from sensitivity analysis and hence contributes to state-of-the-art literature on YAKE-GLDA.

The study is structured as follows. Section 2 provides an overview of the literature and identifies the essential gaps where the study contributes. Section 3 describes the integrated YAKE-GLDA method and the analytical methodology adopted in the current study to classify the SOs in construction from two different data sources. Results have been summarized in section 4, followed by discussions in section 5. Conclusions have been summarized in section 6.

2 Literature Review

For the construction sector, literature focusing on analyzing textual SOs using ML techniques using unsupervised learning approaches has been relatively scarce but has been growing [4]. Based on a state-of-the-art literature review on the topic, three broad unsupervised learning approaches are essentially used. These are (a) *Associated Rule Mining* (ARM) approaches, (b) *Text-mining* approaches, and (c) *Clustering* techniques [4].

The primary purpose of ARM approaches is to find the associations or relationships (called rules) among the input variables toward the defined outcomes [4]. ARM's functionality has also been extended toward text classification [8]. However, many challenges facing ARM approaches are that typically large quantities of rules containing many parameters get generated, which are difficult to comprehend for human interpretations [9]. Due to such issues, their applications in classification remain limited. On the other hand, the unsupervised *Text-Mining* approaches have also been used as pre-processing steps [4,10]. They are highly efficient in converting the typically unstructured data available in construction SOs to structured data leading to very high classification performance [10]. However, the overall process is highly resource-intensive, requiring intensive manual intervention and domain expertise to formulate and validate rules [3]. There is also a possibility that the rules

thus developed are specific to one type of data source and not generalizable to other data sources, even within the construction sector. Hence, ARM and Text-mining approaches are not ideally suited for reducing the manual intervention.

On the other hand, *Clustering* is a frequently used unsupervised learning approach useful in grouping the data into different clusters or topics having a substantial similarity among the members belonging to each topic [5]. LDA and K-means clustering are frequently used clustering approaches [4,5]. Clustering approaches are often fully automated, thereby reducing the necessity of manual inputs to a great extent. Studies focusing on construction SOs have also shown the value-addition of various clustering approaches in obtaining meaningful topics relevant to safety [6].

However, there are significant limitations with conventional clustering approaches for classification tasks. Since the clustering process is fully automated, the optimal number of clusters may not always match the user requirements of classification labeling. Furthermore, in many cases, the outputs of clustering approaches are not comprehensible to human decision-makers [7]. The clustering process may not always be guided through the relationship patterns commonly understood in a specific domain, such as construction [7]. Several recent studies have attempted to improve upon the limitations of the traditional clustering process, mainly by seeding keywords while initiating the clustering process. Through such a seeding process, the topics generated by the clustering process are human interpretable, and their number can be controlled [11]. For example, GLDA is a technique that allows users to seed classification categories using domain-specific keywords as an improvement over the conventional LDA approach [12].

However, extracting domain-specific keywords to be used as seeds could still be challenging. The conventional approach of relying on domain experts to identify and assign keywords is expensive, time-consuming, and error-prone. Hence, several recent studies have also explored automated approaches for keyword extraction. These techniques often require statistics-based features such as frequency-of-word, distance-of-word, and structural features [7]. Many unsupervised keywords extracting techniques, such as Term Frequency-Inverse Document Frequency (tf-idf) [13], PageRank method [14], and Rapid Automatic Keyword Extraction (RAKE) [15], are efficient in keyword extraction based on the statistical features described above. The most recent keyword extraction approach is YAKE [16], which has been shown to perform significantly better than tf-idf and RAKE across many standard datasets [16,17]. YAKE also does not require linguistic information and thus can be used for any language [16,17], extending its generalizability.

The first known application of an integrated YAKE and GLDA unsupervised approach demonstrated its potential in classifying SOs for Aviation and Chemical industry [7]. For the construction sector, the most recent studies have only adopted LDA-based approaches using the tf-idf keyword extractor [13]. Most advanced keyword extractors such as YAKE and RAKE are yet to be explored for their applications in the analysis of construction SOs[5]. This is a significant research gap that the current study aims to fulfill. Unlike in the Aviation and the Process industry, safety reporting systems are often not well-established and mature in the construction sector. Hence, the proven validity and applicability of such an approach with reduced manual intervention for the construction sector may significantly impact the industry-wide adoption of ML approaches.

Furthermore, little is known about the sensitivity of classification performance for the YAKE-GLDA framework, as the early-stage studies have focused only on demonstrating the application [7]. Understanding such sensitivity is crucial to exploring any analytical approach to understanding its true potential. The current study also aims to address this gap in the literature.

3 Methodology

The overall analytical process adopted in the current study has been summarized in Figure 1. The process entails three main steps (1) Building a domain corpus for each of the four categories and extraction of domain-specific keywords using YAKE, (2) Pre-processing the input datasets containing SOs, and (3) Classification of the SOs using the GLDA approach. Subsequent sections provide details on each of the three steps.

3.1 Domain-specific keywords using YAKE

To induce domain knowledge to the GLDA classification model, a set of construction domain-specific keywords are automatically extracted using YAKE from the domain corpora of each of the target categories. A total of 4 target categories are identified based on the category labels available in the primary dataset utilized in the study. More information on the dataset has been described later (see Table 1). A domain corpus is a collection of text drawn from sources containing information particular to the domain. Various online literature sources, including journal papers, research articles, and web resources, have been utilized to assemble four separate domain corpora comprising 31,788 words. A subset of the used resources is mentioned in Table 1.

The domain corpus is pre-processed prior to the extraction of keywords. The same pre-processing steps are also applied to the input datasets. The main pre-processing steps performed are (1) Lowercasing, (2)

Punctuation and numbers removal, (3) Spelling correction using spell check library in the programming language python, (4) Tokenization and stop words removal, (5) Stemming and lemmatization [3,18].

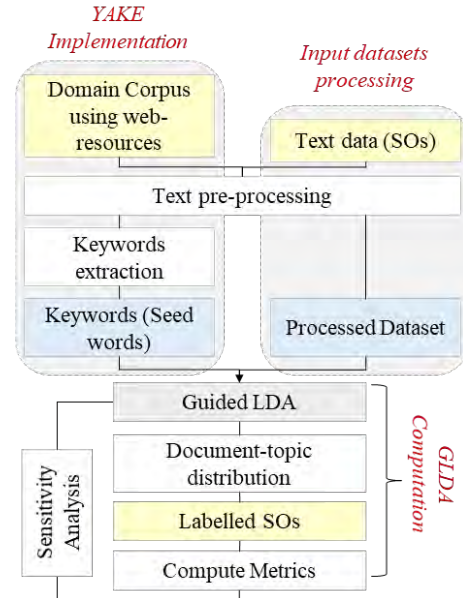


Figure 1. The analytical process of the study

Keywords are a set of content words representing a specific topic. As mentioned earlier, the YAKE method extracts domain-specific keywords from the pre-processed domain corpora. YAKE uses statistical text features extracted from single documents to select relevant keywords. As a document can contain multiple keywords, they are ranked based on their Significance Score $SS(k)$, assigned by the YAKE algorithm. For the computation of $SS(k)$, a set of five features, i.e., Casing (cs), Position (p), Relatedness (r), Frequency (f), and Occurrence (o), is calculated. For brevity, detailed mathematical formulations for these features are not included in the current study, but more information can be found elsewhere [7,16]. The output of YAKE is an ordered list of keywords, ranked based on increasing $SS(k)$, where the smaller the value of $SS(k)$, the more critical the keyword. A set of top 20 keywords for each of the four categories is used for seeding purposes.

3.2 Input datasets and their processing

SOs from two different sources are used in the current study. These datasets represent the diversity of the SOs available in the construction industry.

Dataset 1 is the primary dataset used to demonstrate the validity of the YAKE-GLDA approach for the construction safety domain. *Dataset 1* contains worker-reported near-miss safety observations from a large-scale construction site on a natural gas plant in Kuwait. As per the prevalent reporting practice at this site, the focus is to

promote reporting from the front-end workers as much as possible, rather than obtaining SOs only from safety supervisors [2]. The front-end worker reported data faced several quality-related issues, such as a high proportion of misspelled words, poor sentence structuring, and generally smaller description [18]. Because of such data-quality issues, *Dataset 1* represents the large quantity of usually poor quality SO data generated across the globe. The workers write a brief description of the SO and provide a categorization of the SO into four categories, namely Personal-protective equipment (A), Compliance to safe work (B), Equipment or tools (C), and Housekeeping (D).

Table 1. Overview of Dataset 1

Category labels (% of total data) *	Subcategory label examples	A subset of references used for developing domain corpus
A (21.62)	Ear, eyes, face, hand protection; Harness; respiratory protection	Link 1 , Link 2 , Link 3
B (49.62)	Electrical, Excavation, Fire safety; working at height; Traffic	Link 5 , Link 6 , Link 7 , Link 8
C (08.14)	Equipment usage and selection; color coding; authorization; tags; maintenance	Link 9 , Link 10 , Link 11 ,
D (20.62)	Cleaning; hazardous material management; waste segregation and disposal;	Link 12 , Link 13
*Dataset 1 contains a total of 12490 observations		

On the other hand, *Dataset 2* was used to test the generalizability aspects of the YAKE-GLDA approach. The domain keywords obtained for analyzing *Dataset 1* were also used to analyze *Dataset 2*. Another essential category of SOs prominently prevalent in the construction industry, i.e., textual narratives containing descriptions of injuries/fatalities at construction sites often stored in well-managed databases by safety professionals, is included in *Dataset 2*. Compared to the worker-reported data, the data reported by safety professionals is better in quality and contains longer descriptions [18].

This study utilizes a sample of publicly available fatality/injury narratives provided by the USA's Occupational Safety and Health Administration (OSHA). Goh and Ubeynarayana [3] have used 1000 observations

from the OSHA database and have labeled them into 11 classification categories. *Dataset 2* used in the current study is a subset of the 1000 labeled observations provided in [3]. To test the generalizability of the YAKE-GLDA approach, harmonization of classification labels between the two datasets has been implemented. Hence, the thirteen label categories from *Dataset 2* were mapped with the four broader category labels available from *Dataset 1* (see Table 2). The mapping was also confirmed by reading the detailed description by the two authors experienced in the construction sector. However, not all 11 labels could be readily mapped with the four labels of *Dataset 1*. The extended description in *Dataset 2* included information on multiple potential causes of the accidents/injuries, which could be mapped to different causes of category labels in *Dataset 1* [3]. Hence, only mappings where a precise one-on-one mapping could be obtained have been included in the analysis to avoid inducing errors in the mapping process. As a result, 823 observations out of 1000 available were included in *Dataset 2*.

Table 2. Label Harmonization between datasets 2 and 1

Labels in <i>Dataset 2</i>	Count	Proposed Labels as per <i>Dataset 1</i>
Traffic	63	B (45.4%)
Falls	236	
Fire and Explosion	47	
Electrocute	108	C (32.3%)
Collapse of object	212	
Caught in between	68	
Struck by falling objects	43	
Exposure to chemicals	29	D (4.6%)
Exposure to extreme temperature	17	
Others	43	-- (17.7%)
Struck by moving objects	134	

3.3 Classification of SOs with GLDA seeded using YAKE keywords

The probabilistic model underlying GLDA assumes that document sets can be divided into latent topics, and each topic is made up of different words [7]. It uses Dirichlet distributions in the form of document-topic distribution and topic-word distribution and identifies the topic or category a particular document belongs to, using an iterative procedure. GLDA model uses the domain-specific keywords extracted using YAKE for seeding purposes so that the words are not randomly assigned to a topic during initialization and the topics generated are human interpretable. For seeding, a non-zero weightage

is assigned to the domain-specific keywords during the initialization of GLDA. The *Seed Confidence* (SC) parameter of GLDA can control the weightage assigned to these seed keywords, ranging between 0 and 1. The overall implementation of GLDA is consistent with the previous work [12], which provides details on the involved mathematical formulations. GLDA then uses the YAKE generated seed keywords and the main SOs from the input dataset to generate document-topic distribution. The document-topic distribution provides the probabilities of a document belonging to each of the four categories. The SOs are labeled with the category showing the maximum probability. The classification performance is then evaluated using the commonly adopted F1 score metric [3,7], where the GLDA predicted category labels are compared with those present in the original datasets.

4 Results

4.1 Top-Keywords from YAKE

Table 3 presents an overview of the top-10 (based on estimated weights) keywords extracted for each article. Overall, the top keywords provide an intuitive validity of the keyword extraction process using YAKE. For example, keywords such as "glove," "ppe," "wear," and "protect" represent the category label A, i.e., PPE.

Table 3. Top-10 keywords obtained using YAKE for each category label

A	B	C	D
protection	safety	equipment	waste
glove	fire	conductor	material
ppe	equipment	part	construction
protective	risk	tool	recycling
wear	site	employee	demolition
protect	electrical	material	management
hazard	worker	ladder	project
equipment	ladder	volt	building
type	construction	metal	separate
safety	excavation	expose	product

4.2 F1 scores for Dataset 1

Table 4 summarizes the optimal F1 scores obtained for different category labels and the weighted F1 score. Optimal parameters obtained through a qualitative sensitivity analysis approach are also summarized in Table 4. The optimal F1 scores represent the maximum of 152 runs representing 2, 4, and 19 variations of hyperparameters "Alpha," "Iterations," and "SC." *Alpha* is the Dirichlet parameter for distribution over topics, while *Iterations* is the number of sampling iterations done by the LDA algorithm before convergence. The

optimal total F1 score of 0.62 is obtained. The results also indicate a considerable variation in F1 scores for individual categories. The best performing category is "A," with an optimal F1 score of 0.801. The poorest performing category is "C," with an F1 score as low as 0.37 (see Table 4).

Table 4. F1 scores for Dataset 1

Category	F1 Score	Optimal Hyperparameters*
A	0.801	(0.01, 5000, 0.65)
B	0.56	(0.02, 4000, 0.98)
C	0.37	(0.02, 5000, 0.30)
D	0.69	(0.01, 5000, 0.55)
Total Score**	0.62	(0.02, 4000, 0.98)

*Optimal Hyperparameters are represented using the following notation – ("Alpha," "Iterations," "SC")

** For the maximum Total F1 score, the F1 score for each of the categories are – A (0.796), B (0.56), C (0.34), and D (0.69)

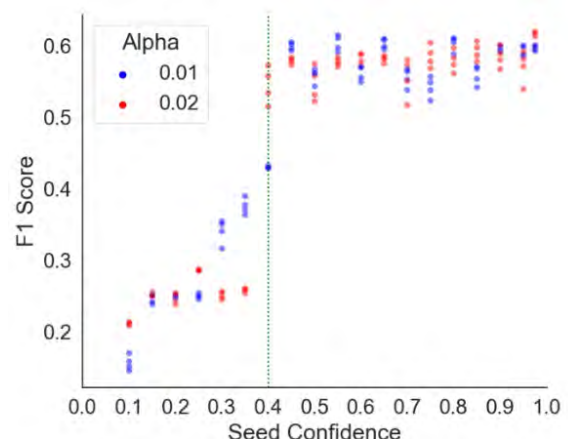


Figure 2. Total F1 Score sensitivity with "SC" and 'Alpha' parameters for Dataset 1

Figures 2 and 3 also summarize the results from sensitivity analysis for the total F1 score and individual category labels. In both figures, four distinct points corresponding to a specific alpha and SC show the different results in variations in the parameter "Iterations." Overall, parameters "Alpha" and "Iterations" do not have a high impact on the F1 score. On the other hand, very high sensitivity in the F1 score around a critical value of the hyperparameter "SC" is obtained. Results from Figures 2 and 3 also suggest that the optimal parameters are different for each category.

4.3 F1 scores for Dataset 2

Table 5 summarizes the optimal F1 scores obtained

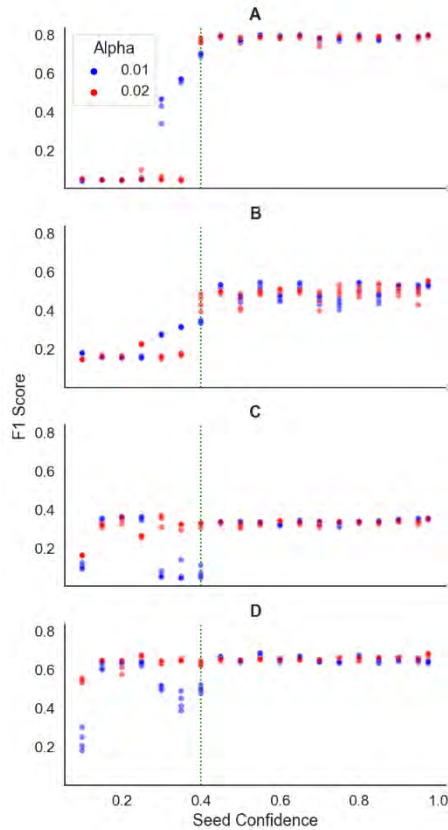


Figure 3. Category wise F1 Score sensitivity with "SC" and 'Alpha' parameters for *Dataset 1*

for different category labels and the weighted F1 score for *Dataset 2*. The optimal total F1 score of 0.48 is obtained. A considerable variation in F1 scores for individual categories is also observed. The best performing category is B, with an optimal F1 score of 0.62. The poorest performing category is D.

Table 5. F1 scores for *Dataset 2*

Category	F1 Score	Optimal Hyperparameters
B	0.62	(0.01, 5000, 0.30)
C	0.56	(0.01, 5000, 0.95)
D	0.24	(0.02, 2000, 0.60)
Total Score**	0.48	(0.01, 5000, 0.30)

** For the maximum Total F1 score, the F1 score for each of the categories are –B (0.62), C (0.33), and D (0.04)

5 Discussions

5.1 Added value of the YAKE-GLDA approach

Overall, a moderate to high classification performance of (F1 score of 0.62 for Total data and 0.8

for category A) has been obtained for the near-miss dataset representative of conditions at the real construction site using the YAKE-GLDA approach. Such classification performance is at par with the previously known application of YAKE-GLDA in the aviation industry [7]. Even for the construction industry, such classification performance is at par with previously reported classification performances obtained using various supervised learning approaches [3]. However, it is also important to note that very high levels of classification performance have also been reported in previous studies utilizing various supervised approaches [5]. Hence, there is still significant scope for further improving the classification performance using the YAKE-GLDA framework. However, the YAKE-GLDA approach has a significant advantage in reducing the need for manual work compared to supervised approaches. The automated keyword extraction process is highly efficient as quality information on various category labels can easily be obtained from domain corpora created using commonly available resources on the web. Furthermore, the GLDA does not require any pre-labeled dataset for learning and classification tasks. In contrast, a requirement for large quantities of the pre-labeled dataset is one of the most significant limitations for the practical implementation of the supervised approaches on construction sites [4,18].

Furthermore, a comparison of results obtained from *Dataset 1* and *Dataset 2* suggests a high degree of generalizability for the domain corpus to obtain classification performance across different types of datasets within the same industry. Typically, ML models trained for one set of datasets perform poorly on classification tasks for similar datasets from different sources [19]. Despite significant differences in the two datasets, domain corpus developed for one application can achieve high classification performance for another dataset, e.g., for categories B and C, where higher classification performance has been obtained than *Dataset 1*. Hence, such preliminary results of the YAKE-GLDA suggest a high potential for rapid and broader implementation of the framework across different construction sites.

5.2 Classification performance and ideas for its improvement

One of the significant aspects affecting the classification performance in the YAKE-GLDA approach is related to the characteristics of the input data in the construction sector. Classification labels typically used in the construction industry are rarely mutually exclusive, creating challenges for ML approaches to classify the observations in a single category [3]. The same is also observed in the current study. Table 6 shows each topic's top words as directly generated by the

GLDA's topic-word distribution based on *Dataset 1*. As highlighted in red in Table 6, many of the top words between categories "B" and "C" are common. Such commonality in the topic words can influence the GLDA's classification accuracy. For example, for the most frequent category in *Dataset 1*, i.e., "B," about 55% of the observations were classified incorrectly. About 50% of these incorrect classifications were classified in category "C." Many ideas on managing challenges related to characteristics of the input data can be implemented in the future studies. GLDA's ability to classify a single observation in multiple-categories should be explored. Furthermore, the unstructured input data need to be converted to structured data as much as possible. Hence, dense vector representation for the text, such as word embeddings, should be explored with GLDA to enhance the performance of the classification tasks[20].

Table 6. Top words for each category for *Dataset 1* as obtained through topic-word distribution in GLDA.

A	B	C	D
circulation	circulation	circulation	our
playful	kind	listed	kind
multiple	workingacceing	underground	underground
listed	towerlight	instruments	instruments
ice	companion	playful	note
earthing	clapboard	earthing	re
underground	toward	molten	inspection
barrication	underground	end	containment
goal	earmuf	electricityhot	fund
instruments	instruments	ice	onsiteafter
seat	earthing	earmuf	barricadding
child	playful	gloves	clapboard
blink	barricadding	barrication	molten
random	double	age	prepare
sss	smoking	kind	tables
workingacceing	fund	dressing	communicate
peak	our	crushed	greenfield
molten	foundation	pope	pepsi
mu	trans	cabin	collar
messy	begin	insulatedwhich	reddy

On the other hand, the domain corpus's quality and comprehensiveness are another significant aspects affecting the classification performance in the YAKE-GLDA approach [7]. Even in the current study, information for some of the subcategories for category C, such as the information related to color-coding of equipment, could not be readily obtained. Hence, the classification performance for category C for *Dataset 1* is inferior (F1 score of 0.37). Whereas, for *Dataset 2*, in which equipment color coding-related factors were not available, the classification performance for category C is significantly high. Hence, a focus of subsequent study could also be to enrich the domain corpus [7].

5.3 Study limitations

The current study is the first to explore the applicability of YAKE-GLDA approach for analyzing the construction SOs. Overall, several promising results have been obtained. However, there are significant limitations of the work requiring improvement.

The current study relies on qualitative methods to assure the validity of the domain corpus generation and data harmonization process. Even though the safety-related experience of authors has helped in the process, more rigorous validations relying on inputs from multiple domain experts are necessary.

The analysis in this study has been focused on classification tasks centered at somewhat broader category levels. However, construction organizations could benefit from tracking trends for other refined categories. In such conditions, the value addition of the YAKE-GLDA approach should also be demonstrated for classification performance at the micro subcategory levels. Such detailed categories are also available in *Dataset 1* and could be explored in future studies.

The study's novel sensitivity analysis results also highlight a lack of generalizable trend in F1 scores with varying SC. Figures 2 and 3 indicate a steep improvement in the F1 score around a critical value of SC and a relatively invariable trend afterward. However, the optimal hyperparameter combination is also different for each category. Such a lack of generalizable trends in sensitivity to hyperparameters indicates potential consistency issues for the approach and may restrict its applicability for new unlabeled data. The results also underscore the importance of conducting a sensitivity analysis to obtain optimal values of the parameters to be used in all subsequent applications. Furthermore, more research is deemed necessary to fully understand the sensitivity of the approach on more parameters that could guide its applications to different construction sites.

6 Conclusions

The current study is the first application of an unsupervised YAKE-GLDA approach for the fully automatic classification of SOs in construction. The process reduces the necessity of manual intervention significantly, provides a moderate classification performance (F1 score of 0.81 for ppe category), and is potentially generalizable to different data sources related to safety in the construction sector. On the other hand, previously unexplored sensitivity to hyperparameters reveals a non-generalizable trend affecting YAKE-GLDA's new application to an unlabeled dataset. Hence, future research is vital for assuring the approach's applicability to various construction sites. Efforts are also necessary for improving the classification performance of the approach. Finally, the study's limitations should

also be addressed to make an objective assessment of the applicability of YAKE-GLDA approach for efficient analysis of SOs in construction.

References

- [1] P. Manu, F. Emuze, T.A. Saurin, B.H.W. Hadikusumo, *Construction Health and Safety in Developing Countries*, Routledge, 2019.
- [2] N. Bugalia, Y. Maemura, K. Ozawa, A system dynamics model for near-miss reporting in complex systems, *Saf. Sci.*, 142:105368, 2021. <https://doi.org/10.1016/j.ssci.2021.105368>.
- [3] Y.M. Goh, C.U. Ubeynarayana, Construction accident narrative classification: An evaluation of text mining techniques, *Accid. Anal. Prev.*, 108: 122-130, 2017. <https://doi.org/10.1016/j.aap.2017.08.026>.
- [4] S. Sarkar, J. Maiti, Machine learning in occupational accident analysis: a review using science mapping approach with citation network analysis, *Saf. Sci.* 131:104900, 2020. <https://doi.org/10.1016/j.ssci.2020.104900>.
- [5] S. Baek, W. Jung, S.H. Han, A critical review of text-based research in construction: Data source, analysis method, and implications, *Autom. Constr.* 132:103915, 2021. <https://doi.org/10.1016/j.autcon.2021.103915>.
- [6] A. Chokor, H. Naganathan, W.K. Chong, M. El Asmar, Analyzing Arizona OSHA injury reports using unsupervised machine learning, *Procedia Eng.* 145, 1588–1593, 2016.
- [7] A. Ahadh, G.V. Binish, R. Srinivasan, Text mining of accident reports using semi-supervised keyword extraction and topic modeling, *Process Saf. Environ. Prot.* 155: 455-465, 2021. <https://doi.org/10.1016/j.psep.2021.09.022>.
- [8] C.-W. Cheng, C.-C. Lin, S.-S. Leu, Use of association rules to explore cause–effect relationships in occupational accidents in the Taiwan construction industry, *Saf. Sci.*, 48: 436-444, 2010. <https://doi.org/10.1016/j.ssci.2009.12.005>.
- [9] M.N. Moreno, S. Segre, V.F. López, Association Rules: Problems, solutions and new applications, *Actas Del III Taller Nac. Minería Datos y Aprendizaje*, Tamida, 317–323, 2005.
- [10] A.J.-P. Tixier, M.R. Hallowell, B. Rajagopalan, D. Bowman, Automated content analysis for construction safety: A natural language processing system to extract precursors and outcomes from unstructured injury reports, *Autom. Constr.*, 62:45-56, 2016. <https://doi.org/10.1016/j.autcon.2015.11.001>.
- [11] J. Jagarlamudi, H. Daumé III, R. Udupa, Incorporating lexical priors into topic models, in: *Proc. 13th Conf. Eur. Chapter Assoc. Comput. Linguist.*, 204-213, 2012.
- [12] S. Zhou, P. Kan, Q. Huang, J. Silbernagel, A guided latent Dirichlet allocation approach to investigate real-time latent topics of Twitter data during Hurricane Laura, *J. Inf. Sci.*, 2021. <https://doi.org/10.1177/01655515211007724>.
- [13] Y. Suh, Sectoral patterns of accident process for occupational safety using narrative texts of OSHA database, *Saf. Sci.*, 142:105363, 2021. <https://doi.org/10.1016/j.ssci.2021.105363>.
- [14] R. Wang, W. Liu, C. McDonald, Using word embeddings to enhance keyword identification for scientific publications, in: *Australas. Database Conf.*, Springer, 257-268, 2015.
- [15] S.J. Rose, W.E. Cowley, V.L. Crow, N.O. Cramer, Rapid automatic keyword extraction for information retrieval and analysis, 2012.
- [16] R. Campos, V. Mangaravite, A. Pasquali, A. Jorge, C. Nunes, A. Jatowt, YAKE! Keyword extraction from single documents using multiple local features, *Inf. Sci. (Ny)*. 509:257-289, 2020. <https://doi.org/10.1016/j.ins.2019.09.013>.
- [17] N. Giarelis, N. Kanakaris, N. Karacapilidis, A Comparative Assessment of State-Of-The-Art Methods for Multilingual Unsupervised Keyphrase Extraction, in: *IFIP Int. Conf. Artif. Intell. Appl. Innov.*, Springer, 635–645, 2021.
- [18] J. Kedia, T. Vurukuti, N. Bugalia, A. Mahalingam, Classification of safety observation reports from a construction site: An evaluation of text mining approaches, in: *PMI Res. Acad. Virtual Conf.* 50–66, Indian Institute of Technology Bombay, Mumbai, 2021.
- [19] L. D'hooge, T. Wauters, B. Volckaert, F. De Turck, Inter-dataset generalization strength of supervised machine learning methods for intrusion detection, *J. Inf. Secur. Appl.* 54:102564, 2020. <https://doi.org/10.1016/j.jisa.2020.102564>.
- [20] F. Zhang, A hybrid structured deep neural network with Word2Vec for construction accident causes classification, *Int. J. Constr. Manag.* (2019) 1–21. <https://doi.org/10.1080/15623599.2019.1683692>.

Automated Construction Contract Summarization Using Natural Language Processing and Deep Learning

X. Xue¹, Y. Hou¹, and J. Zhang^{1*}

¹ School of Construction Management Technology, Purdue University, USA

E-mail: xue39@purdue.edu, hou144@purdue.edu, (*corresponding author) zhan3062@purdue.edu

Abstract –

The interpretation of construction contracts is crucial to the management and success of a project. Correct and accurate interpretation could support the smooth construction of high-quality built assets. Misunderstanding and omissions may lead to costly rework and delay. One main challenge of construction contract interpretation lies in the length of construction contracts. Therefore, the demand for reducing the length of such documents while keeping their main information elements emerges. To address this research need, the authors proposed to use natural language processing (NLP) and deep learning technology to summarize construction contracts (i.e., text summarization). There are many deep learning models available and developed for text summarization. However, their performance on construction contracts is to be tested. To address this gap, the authors proposed a new merit-based evaluation method to evaluate the performance of three deep learning models on text summarization of construction contracts, which were reported the state-of-the-art performance on text summarization tasks in general English corpus. The proposed method evaluated selected models from three aspects: information completeness, information correctness, and human readability. The Distilbart model, which scored 5.23, 4.82, and 5.05 in these three aspects, respectively, outperformed the other models in all three aspects.

Keywords –

Deep Learning; Automated Construction Contract Summarization; Text Summarization; Natural Language Processing; Artificial Intelligence; Construction Management

1 Introduction

Construction contract is a critical type of construction document that details the terms agreed upon by all involved stakeholders [1]. In general, construction contracts specify critical provisions such as payment

schedules, construction costs, and completion dates. Additionally, construction contracts also specify how disputes should be resolved when raised, and other procedural agreements. To ensure that all necessary information is included, and ambiguity is avoided, building contracts typically try to cover every aspect that might potentially be predicted [2]. Therefore, construction contracts may easily become too long for human readers to digest easily.

The length of construction contracts places a high cognitive burden on their human readers and increases the time required to understand and process the contractual information. As a result, reducing the length of such contracts while maintaining their main idea emerges as an urgent research need. The main idea that a body of text conveys can be split into many information elements. Text summarization is the process of creating a condensed version of the body of text by retaining critical information elements and removing uncritical ones [3]. In recent years, many transformer-based deep learning models reported the state-of-the-art performance on the task of text summarization. However, because they were mostly tested on datasets of general English corpus, their performance on domain-specific texts such as construction contracts is not clear.

Evaluating the performance of text summarization is a challenging task [4]. The majority of text summarization metrics were concerned with determining the similarity between automatically generated summaries and some target summaries [5]. These metrics assume that the greater the similarity is, the better the summarizations are and, consequently, the better the model performance is. The target summaries were typically generated by experts and demonstrated what a good summary of the entire text should look like. One benefit of using such metrics is that the measurement can be fully automated. Their shortcomings, on the other hand, are also significant. Certain automated metrics, for example, precision, recall, and f1-score at the word level, do not take into account the sequential order of words in the summaries [5]. The same set of words arranged in a different sequence may have a significantly different meaning or (in the extreme case) make no sense at all.

There were also metrics that take word sequence into account, such as the BiLingual Evaluation Understudy (BLEU) [6] and the Metric for Evaluation of Translation with Explicit ORdering (METEOR) [7]. But they are known to award summaries that make little/no sense or are difficult to read for humans.

To address the above-mentioned problem, the authors proposed a new merit-based text summarization evaluation method that focuses on human perception of the summarization results. Three transformer-based deep learning models were then evaluated using the proposed method in this research: (1) Distilbart [8], (2) Pegasus [9], and (3) Bidirectional and Auto-Regressive Transformer (BART) [10]. These evaluated models were selected because they have been reported with the state-of-the-art performance on many different natural language processing (NLP) tasks, such as machine translation, question and answering, text summarization, and named entity recognition (NER). However, their performance on the summarization of construction contracts has not been tested to the best of the authors' knowledge. This research provides an initial evaluation of the performance of transformer models on the summarization of construction contracts.

2 Background

2.1 Natural Language Processing

Natural language processing (NLP) has a wide range of applications in the architecture, engineering, and construction (AEC) domain. Research on NLP in the AEC domain mostly focused on classifying or extracting information from construction documents (e.g., construction contracts, and building codes) for future processing by human or machine [11]. For example, Caldas and Soibelman [32] used machine learning and NLP in the classification of construction management documents. Zhang and El-Gohary [12] combined semantic NLP-based information extraction with automated reasoning to accomplish automated reasoning with building code requirements to support compliance checking. Xue et al. [13] proposed a semi-automated method to extract regulatory information from tables in building codes. Zhang and El-Gohary [14] developed a deep learning-based information extraction system for extracting regulatory requirements to support automated building code compliance checking. Xue and Zhang [15] increased the accuracy of part-of-speech tagging of building codes by an error-fixing method. Li and Cai [16] utilized NLP in the processing of infrastructure requirements to support compliance checking of underground utility lines. Le and Jeong [17] used NLP techniques, such as Word2Vec, to classify semantic relation between terminologies in transportation asset

manuals. Dimyadi et al. [33] developed a table-based NLP algorithm to convert building codes from normative text to computable rules. Song et al. [34] used deep learning method to convert Korea building codes to predict-argument structure. Al-Qady and Kandil [36] leveraged semantic parsing to extract relations between concepts in construction contraction clauses.

2.2 Text Summarization

The main goal of text summarization is to preserve the main information elements of a body of text while reducing its length [3]. Text summarization systems allow users to obtain the main information elements of documents without having to read the full text. In general, there are single-document systems [18] that summarize one document at a time, and multi-document systems that summarize multiple documents into one summary [19]. Based on the approach of summarization, text summarization systems can be classified into extractive systems that extract important sentences from the documents [20], and abstractive systems that aim to generate a summary by reducing unimportant information from the original documents by processing computerized representations of text (e.g., embedding vector) [21].

Literature review suggested that summarization of contractual clauses is challenging because of the difference between legal document and general document and the limited amount of available training dataset. For example, Manor and Li [37] prepared a dataset of contract documents and evaluated existing text summarization methods on it. They concluded that the summarization of contract documents is challenging because the gap between legal documents and training texts of existing summarization methods is substantial. Elnaggar et al. [38] launched a dataset of legal documents for the tasks of translation, summarization, and classification. They utilized transfer learning (leveraging pre-trained models) and multi-task learning (jointly training one model in multiple tasks) to reach the state-of-the-art performance on these three tasks on a dataset of a few thousand sentences. However, their research cannot be applied directly to this study because their research focused on legal documents in Europe, which is drastically different from construction contracts in the United States.

2.3 Transformer Models

Transformer models belong to one type of deep learning model that frequently achieved the state-of-the-art performances in various NLP tasks in recent years [22]. The core of a transformer model is the attention mechanism, which transforms one sequence to another sequence using an encoder-decoder structure [23]. The

recurrent neural network (RNN) is known not to perform well at capturing long-term dependency. The attention mechanism effectively solved this shortcoming by learning the importance of each token and/or correlations between each token in the source sequence and each token in the target sequence during the training phase. However, the computational complexity of calculating every possible pair of tokens in each sequence is high. Local attention mechanisms (i.e., a subtype of attention mechanism) emerged as a remedy by only calculating attention between surrounding words [24, 25]. Self-attention is another subtype of attention mechanism where the source sequence and the target sequence are the same [26]. Self-attention has been proven to be effective in tasks such as machine translation, named entity recognition, and text summarization. Transformer models are usually pre-trained on self-supervised tasks, such as predicting masked tokens or predicting the next token given all the previous tokens in a sentence [27, 28].

Transformer models are constructed in an encoder-decoder fashion. The encoder receives the input sequence (i.e., text) in the form of embedding vectors. The encoder converted the input vectors to a new vector called “internal state” after processing them. The decoder is then informed of the internal state and generates the outputs (i.e., summary). Typically, transformers are trained on semi-supervised auto-regressive tasks, such as predicting the next word given previous words and predicting masked words given their surrounding words. However, the detail of each transformer varies. BART's encoder was trained to predict masked words, whereas its decoder was trained to predict next word given previous words [10]. Pegasus was trained to generate gap sentences between sentences [9]. The distillbart model is nearly identical to the BART model in most aspects except that it is pre-trained using knowledge distillation, which is a technique that reduces the size of deep learning models by using a larger deep learning model to train a smaller deep learning model. While the larger deep learning model is trained to generate predictions, the smaller deep learning model is trained to match the predictions of the larger model [8]. An early form of knowledge distillation uses data labeled by the larger model (instead of human annotators) to train the smaller model. Knowledge distillation in its current form focuses on matching the behavior of the output layer of two models. The smaller model is trained jointly to make the same predictions and assign the same probability of all predicted classes as those by the larger model [48].

2.4 Challenges

The evaluation of text summarization has been a challenge. Most automated evaluation methods would compare the summarization of a specific model to a gold standard of summarization and generate a score based on

predefined metrics. The more similar the machine-generated summarization is to the gold standard, the higher the score will be. This type of metrics is based on two underlying assumptions both of which may not be very robust. First, it assumes text summarization has a ground truth, and the closer a summarization is to the ground truth, the better the summarization is. However, a text may have multiple good summaries based on different information organizations and expressions. In addition, it is hard for machines to measure how human readers would understand a summarization. Second, it assumes automated text summarization evaluation means n-gram similarity between machine-generated summarization and the gold standard. While the problem of the first assumption may be solved by generating multiple gold standards and averaging scores from all gold standard versions, the cost of such evaluation could increase significantly. The second assumption is also questionable because the meaning of a sentence is very sensitive to the sequence of words, and high word/token-level similarity does not guarantee similar meaning. A slight shift in the sequence of words may completely change the meaning of a sentence.

3 Methodology

It is common to pre-process textual data before it was fed into deep learning models. To analyze the impact of pre-processing on the summarization of construction contracts, the authors compared performance of deep learning models without and with pre-processing. In the first setting, the textual data was fed into deep learning models directly without pre-processing. The inputs to deep learning models are therefore unaltered construction contract text. In the second setting, inputs went through the following pre-processing steps: (1) tokenization (i.e., break down strings of sentences into lists of words and punctuations), (2) lowercasing (i.e., convert all characters into lower case), (3) stop word removal (remove common words that do not carry context-specific information, such as “a” “an” “the”), (4) removal of number, punctuation, and underscores (i.e., fill-in-the-blank space from construction contract template), and (5) lemmatization (reduce inflectional forms of words to their base forms).

The authors proposed a merit-based evaluation of the performance of the state-of-the-art deep learning transformer models on the summarization of construction contracts in this research. The merit-based evaluation has three steps. First, construction contract clauses that are suitable for the summarization text are collected. Clauses that do not need to be summarized (too short) or cannot be summarized properly (contain too much blank space and miss too much information) are removed in the step. In the second step, transformer models generate

summaries of the selected clauses. In the third step, human reviewers manually evaluate the transformer models by evaluating the summaries generated by the models.

Manual evaluation of transformer models was used to avoid the shortfalls associated with automated evaluation. Manual evaluation by domain experts, as an alternative to automated evaluation, can have the following two advantages. First, manual evaluation ensures that a summarization with high score is indeed easy to understand for human readers. Second, manual evaluations can effectively assess the meanings carried by the summarizations. The proposed evaluation quantifies transformer performance on three dimensions: information completeness, information correctness, and human readability. Information completeness is a metric that measures if a summary retains the key information elements from the original text. Information correctness is a metric that measures how accurate a summary is. A good summary should accurately represent the original full text. Human readability refers to the ease with which the summary can be understood. Models' performance in each category is rated on a scale of 0 to 10. The higher the score, the better the performance is. Domain experts are allowed to rate performance according to their understanding of construction contract based on their construction domain knowledge. They are also required to maintain a high-level objectivity and consistency across the evaluation.

4 Experiment

For the purpose of this research, the authors collected publicly available and free-to-download construction contracts or contract templates online. In total, nine construction contracts or contract templates were collected [31, 39-46]. The collected construction contracts were cleaned to fit the needs of the research. Some paragraphs were removed because they were either too short or contained too much blank space. A dataset that contains ninety-two paragraphs of construction contracts or contract templates was prepared. After the dataset was created, a careful evaluation and extensive literature review were conducted to select transformer candidates for the evaluation. Three transformers were selected: (1) Distilbart [8], (2) Pegasus [9], and (3) BART [10] (Table 1). The selected models all have achieved the state-of-the-art performance or near state-of-the-art performance on the task of text summarization in large open datasets of general English corpus [8, 9, 10]. Because transformer models were usually published in different configurations, one configuration of each model was selected based on the popularity of the configuration. For Distilbart, the selected configuration was distilbart-cnn-12-6. For BART, the selected configuration was

bart-large-cnn. For Pegasus, the selected configuration was pegasus-xsum. These configurations were meaningful expressions. For example, the "cnn" in the distilbart-cnn-12-6 means convolutional neural network is used. The 12 refers to 12 layers of encoders. The "large" in bart-large-cnn means this configuration has a large number of parameters. The "xsum" in the pegasus-xsum means the configuration was trained on the xsum dataset [29]. For each model, the configuration with the largest number of downloads in Huggingface Transformer model hub [30] was selected. The selected models weren't altered or fine-tuned in any way for the evaluation.

Table 1. Selected Models

Model Name	Configuration
BART	distilbart-cnn-12-6
Pegasus	pegasus-xsum
Distilbart	bart-large-cnn

Each selected transformer model generated a summary of each paragraph in the dataset with and without preprocessing of the textual data. The summaries were then evaluated by the authors using the merit-based evaluated method. All models were run on a desktop computer with an RTX 3090 graphic card and a Ryzen 3950x CPU.

5 Result

The performance of transformer models with and without pre-processing was summarized in Tables 2, 3 and 4.

Table 2. Model Performance with Pre-processing

	BART	Pegasus	Distilbart	Average
Completeness	4.62	2.12	4.70	3.81
Correctness	4.59	2.16	4.57	3.77
Human Readability	4.62	2.51	4.45	3.86
Overall	4.61	2.26	4.57	3.81

Table 3. Model Performance without Pre-processing

	BART	Pegasus	Distilbart	Average
Completeness	4.74	4.98	5.23	4.98
Correctness	4.72	4.6	4.82	4.71
Human Readability	4.66	4.82	5.05	4.84
Overall	4.71	4.8	5.03	4.76

Table 4. Model Performance Comparison

	With Pre-processing	Without Pre-processing
Completeness	3.81	4.98
Correctness	3.77	4.71
Human Readability	3.86	4.84
Overall	3.81	4.76

The performance of deep learning models was better when inputs were not pre-processed rather than when they were pre-processed. When the inputs were not pre-processed, the Distilbart model outperformed the other two transformer models in all three aspects. This shows the Distilbart model could be a good choice for the summarization of construction contracts. For future refinement on the text summarization of construction contracts, the Distilbart model should be a good starting point. It is very likely that the fine tuning of the model will lead to better performance, which is suggested in literature [38]. Example summarizations by each model are provided in Table 5.

Table 5: Example Summary without Pre-Processing

Deep Learning Model	Original Text	Summary
BART	“Owner shall have a right to inspect the Work at any time and request that Contractor promptly correct any Work that is defective or does not conform to the Contract Documents. If required, the Work shall be inspected and certified by the appropriate state or local agency or health officer at each necessary stage.” [31]	Owner shall have a right to inspect the Work at any time and request that Contractor promptly correct any Work that is defective
Pegasus		The owner of the property where the Work is to be performed has the right to inspect the Work at any time.
Distilbart		Owner shall have a right to inspect the Work at any time and request that Contractor promptly correct any Work that is defective

The performance was not optimal, as no model

achieved a score greater than six. The Distilbart model had the highest score among the three, making it a good choice for starting contract summarization for now. However, it is strongly recommended that a model dedicated to the construction contract text summarization task be developed (i.e., trained on construction contracts corpus). The task of fine-tuning transformer models has been identified as a high-priority area of research. Among all evaluation criteria, the information correctness criterion had the lowest average, indicating that increasing the information accuracy of the summarization is a pressing need.

6 Discussion

The Distilbart model outperformed the other two models in the evaluation. The authors attributed its success to its training strategy and the process of knowledge distillation. The BART model employed a decoder that had been trained to generate the next word given previous words in a sentence, implying that it is capable of generating new texts (summary). Its encoder is trained to predict masked words, enhancing the model’s robustness. The Distilbart model inherited the BART model’s advantages. Additionally, the knowledge distillation process enhanced the model’s generalizability (i.e., ability to generate good summary).

The gap in the performance of deep learning models with and without pre-processing could be attributed to two reasons. First, deep learning models had difficulty in generating complete sentences. It is likely that stop words, although did not carry a lot of useful information themselves, were necessary for the transformer models to generate complete sentences. Second, the Pegasus model generated a lot of summaries, such as, “All photographs courtesy of AFP, EPA, Getty Images, and Reuters” and “All photographs are copyrighted.” that had no correlation with the core of the original text. Therefore, more comprehensive evaluation is necessary before putting deep learning models into practice in construction contract text summarization. The behaviours of deep learning models currently are not guaranteed to perform as users may expect. Based on our experimental results, the performance of deep learning model decreased when the construction contract input is pre-processed. Therefore, to increase the accuracy of summarization, pre-processing that was typically used in other NLP tasks may need to be avoided.

7 Contributions to the Body of Knowledge

This research makes three distinct contributions to the body of knowledge. First, the authors proposed an evaluation method for text summarization based on merit in three different dimensions: information completeness,

information correctness, and human readability. The proposed method ensures that the summary's evaluated score corresponds to its human perception. By involving humans in the evaluation process, the method avoids awarding high scores to summaries with poor human perception. Second, a recommended deep learning model was identified for contract summarization by comparing the performance of three different but popular models: BART, Pegasus, and Distilbart. The Distilbart model outperformed other models on all three criteria, making it a good starting point for the task of summarizing construction contracts. Last but not least, it was found that in contrast to typical NLP tasks, pre-processing does not necessarily help in the construction contract summarization using deep learning models.

8 Limitations and Future Work

The following limitations are acknowledged. First, the evaluation is composed of three components. While these provided a relatively holistic evaluation comparing to the state of the art, additional factors can be included to provide a more detailed and comprehensive evaluation, such as, grammar, structure, and coherence [35]. Second, the proposed evaluation method necessitates manual efforts, which could be costly and time-consuming. Future research could investigate how to make the evaluation more efficient. In addition, human subjectivity may influence the outcome of an evaluation. Certain factors, such as personal preference and educational background, may influence how certain summaries are rated. Future research should investigate how to minimize the impact of human subjectivity on such evaluation.

9 Conclusion

Existing research that leverages NLP in the construction domain mostly directly adopted benchmarks and metrics from other domains, the suitability of which is seldom asked. In this paper, the authors proposed a new merit-based evaluation method for text summarization. The proposed method was used in evaluating the performance of three deep learning transformer models on the text summarization of construction contracts or contract templates. The evaluation was conducted manually by rating performance from three aspects: information completeness, information correctness, and human readability. Three models were selected based on their popularity: Distilbart, Pegasus, and BART. The Distilbart model outperformed the other two models in all three aspects. The performance of deep learning models was found to be better when inputs were not pre-processed than when the inputs were pre-processed. As the direct application of such deep learning model on the

domain specific construction contracts did not achieve very high scores, the authors suggest that the fine-tuning of Distilbart model and/or retraining it based on domain corpus should be in the future direction of text summarization of construction contracts.

10 References

- [1] Chan E. E. Nik-Bakht M. and Han S. H. Sources of ambiguity in construction contract documents, reflected by litigation in supreme court cases. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, 13(4): 04521031, 2021.
- [2] Bagherian-Marandi N. and Akbarzadeh M.R. Two-layered fuzzy logic-based model for predicting court decisions in construction contract disputes. *Artificial Intelligence and Law*, 29(4): 453-484, 2021.
- [3] Abualigah L. Bashabsheh M. Alabool Q. H. and Shehab M. Text summarization: a brief review. *Recent Advances in NLP: The Case of Arabic Language*, 1-15, 2020.
- [4] Gambhir M. and Gupta V. Recent automatic text summarization techniques: a survey. *Artificial Intelligence Review*, 47(1): 1-66, 2017.
- [5] Ermakova L. Cossu J. V. and Mothe J. A survey on evaluation of summarization methods. *Information Processing & Management*, 56(5): 1794-1814, 2019.
- [6] Papineni K. Roukos S. Ward T. and Zhu W.J. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311-318, Philadelphia, Pennsylvania, 2002.
- [7] Banerjee S. and Lavie A. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In *Proceedings of the ACL workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization*, 2005, pp. 65-72.
- [8] Shleifer S. and Rush A. M. Pre-trained summarization distillation. *arXiv preprint. arXiv:2010.13002*, 2020.
- [9] Zhang J. Zhao Y. Saleh M. and Liu P. Pegasus: Pre-training with extracted gap-sentences for abstractive summarization. *arXiv preprint. arXiv:1912.08777*, 2020.
- [10] Lewis M. Liu Y. Goyal N. Ghazvininejad M. Mohamed A. Levy O. Stoyanov V. and Zettlemoyer L. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv preprint. arXiv:1910.13461*, 2019.

- [11] Wu C. Li X. Guo Y. Wang J. Ren Z. Wang M. and Yang Z. Natural language processing for smart construction: current status and future directions. *Automation in Construction*, 134:104059, 2022.
- [12] Zhang J. and El-Gohary N. M. Integrating semantic NLP and logic reasoning into a unified system for fully-automated code checking. *Automation in Construction*, 73:45-57, 2017.
- [13] Xue X. Wu J. and Zhang J. Semiautomated generation of logic rules for tabular information in building codes to support automated code compliance checking. *Journal of Computing in Civil Engineering*, 36(1):04021033, 2022.
- [14] Zhang R. and El-Gohary N. A deep neural network-based method for deep information extraction using transfer learning strategies to support automated compliance checking. *Automation in Construction*, 132:103834, 2021.
- [15] Xue X. and Zhang J. Building codes part-of-speech tagging performance improvement by error-driven transformational rules. *Journal of Computing in Civil Engineering*, 34(5): 04020035, 2020.
- [16] Li S. Cai H. and Kamat Vineet R. Integrating natural language processing and spatial reasoning for utility compliance checking. *Journal of Construction Engineering and Management*, 142(12):04016074, 2016.
- [17] Le T. and Jeong H. NLP-based approach to semantic classification of heterogeneous transportation asset data terminology. *Journal of Computing in Civil Engineering*, 31(6):04017057, 2017.
- [18] Christian H. Agus M. P. and Suhartono D. Single document automatic text summarization using term frequency-inverse document frequency (TF-IDF). *ComTech: Computer, Mathematics and Engineering Applications*, 7(4):285-294, 2016.
- [19] Uçkan T. and Karıcı A. Extractive multi-document text summarization based on graph independent sets. *Egyptian Informatics Journal*, 21(3):145-157, 2020.
- [20] Ferreira R. Cabral L. Lins R. Silva G. Fred Freitas Cavalcanti G. Lima R. Simske S. and Favaro L. Assessing sentence scoring techniques for extractive text summarization. *Expert systems with applications*, 40(14):5755-5764, 2013.
- [21] Song S. Huang H. and Ruan T. Abstractive text summarization using LSTM-CNN based deep learning. *Multimedia Tools and Applications*, 78(1): 857-875, 2019.
- [22] Kalyan K. S. Rajasekharan A. and Sangeetha S. Ammus: A survey of transformer-based pretrained models in natural language processing. *arXiv preprint*, arXiv:2108.05542, 2021.
- [23] Vaswani A. Shazeer N. Parmar N. Uszkoreit J. Jones L. Gomez A. N. Kaiser L. and Polosukhin I. Attention is all you need. In *Proceedings of Advances in Neural Information Processing Systems*, pages. 5998-6008, Long Beach, California, 2017.
- [24] Tjandra A. Sakti S. and Nakamura S. Local Monotonic attention mechanism for end-To-end speech and language processing. *arXiv preprint*, arXiv:1705.08091, 2017.
- [25] Li L. Tang S. Deng L. Zhang Y. and Tian Q. Image Caption with global-local attention. In *Proceedings of the Thirty-first Association for the Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence*, pages 4133-4139, San Francisco, California, 2017.
- [26] Shaw P. Uszkoreit J. and Vaswani A. Self-attention with relative position representations. *arXiv preprint*, arXiv:1803.02155, 2018.
- [27] Devlin J. Chang M. W. Lee K. and Toutanova K. Bert: pre-training of deep bidirectional transformers for language Understanding. *arXiv preprint*, arXiv:1810.04805, 2018.
- [28] Brown T. B. Mann B. Ryder N. Subbiah M. Kaplan J. Dhariwal P. Neelakantan A. Shyam P. Sastry G. Askell A. Agarwal S. Herbert-Voss A. Krueger G. Henighan T. Child R. Ramesh A. Ziegler D. M. Wu J. Winter C. Hesse C. Chen M. Sigler E. Litwin M. Gray S. Chess B. Clark J. Berner C. McCandlish S. Radford A. Sutskever I. and Amodei D. Language models are few-shot learners. *arXiv preprint*, arXiv:2005.14165, 2020.
- [29] Narayan S. Cohen S. B. and Lapata M. Don't give me the details, just the summary! Topic-aware convolutional neural networks for extreme summarization. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages,1797-1807, Brussels, Belgium, 2018.
- [30] Huggingface.co. Model hub. On-line: <https://huggingface.co/models>, Accessed: 07/01/2022.
- [31] County of Montrose. Construction contract. On-line: <https://www.montrosecounty.net/DocumentCenter/View/823/Sample-Construction-Contract>, Accessed: 17/01/2022.
- [32] Caldas C.H. and Soibelman L. Automating hierarchical document classification for construction management information systems. *Automation in Construction*, 12(4), 395-406, 2003.
- [33] Dimyadi J. Fernando S. Davies K. and Amor R. Computerising the New Zealand building code for automated compliance audit. In *Proceeding of New Zealand Built Environment Research Symposium*, pages 39-46, Auckland, New Zealand, 2020.
- [34] Song J. Lee J.K. Choi J. and Kim I. Deep learning-based extraction of Predicate-Argument Structure

- (PAS) in building design rule sentences. *Journal of Computational Design and Engineering*, 7(5):563-576, 2020.
- [35] Allahyari M. Pouriyeh S. Assefi M. Safaei S. Trippe, E.D. Gutierrez J.B. and Kochut K. Text summarization techniques: a brief survey. *arXiv preprint*, arXiv:1707.02268, 2017.
- [36] Al-Qady. M. and Kandil. A. Concept relation extraction from construction documents using natural language processing. *Journal of Construction Engineering and Management*, 136(3): 294-302, 2017.
- [37] Manor. L. and Li. J. J. Plain English summarization of contracts. *arXiv preprint*, arXiv:1906.00424, 2019.
- [38] Elnaggar A. Gebendorfer C. Glaser I. Matthes F. Multi-task deep learning for legal document translation, summarization and multi-label classification. In *Proceedings of the 2018 Artificial Intelligence and Cloud Computing Conference*, pages 9-15, New York, New York, 2018.
- [39] Zhuang. F. Qi. Z. Duan. K. Xi. D. Zhu. Y. Zhu. H. Xiong. H. and He. Q. A comprehensive survey on transfer learning. In *Proceedings of the Institute of Electrical and Electronics Engineers (IEEE)*, pages 43-76, Piscataway, New Jersey, 2020.
- [40] Legaltemplates. The party hiring the general contractor to complete construction on his or her property. what is a construction agreement. On-line: <https://legaltemplates.net/form/lt/construction-contract-agreement>, Accessed: 08/02/2022.
- [41] Eforms. Construction contract template. On-line: <https://eforms.com/employment/independent-contractor/construction>, Accessed: 08/02/2022
- [42] Signaturely. Construction contract agreement template. On-line: <https://signaturely.com/contracts/construction-contract-agreement-template>, Accessed: 08/02/2022.
- [43] BuildingAdvisor. Model construction contract for homeowners. On-line: https://buildingadvisor.com/project-management/contracts/model-construction-contract_1, Accessed: 08/02/2022.
- [44] Washington LLC. Imperium Grays Harbor LLC Construction Agreement Cost Plus Fee. On-line: <https://www.sec.gov/Archives/edgar/data/1381697/000119312507121562/dex107.htm>, Accessed: 08/02/2022.
- [45] City of Colorado Springs. Sample contract template. On-line: https://coloradosprings.gov/sites/default/files/const_ruction_sample_contract_template_version_100316.pdf, Accessed: 08/02/2022.
- [46] Florida Contractors. The construction agreement. On-line: <http://floridacontractors.net/downloads/Residential%20Contract%201.pdf>, Accessed: 08/02/2022.
- [47] The Board of County Commissioners of Escambia County, Florida. Standard construction contract documents. On-line: https://myescambia.com/sites/myescambia.com/files/A_AE.pdf, Accessed: 08/02/2022.
- [48] Hinton G. Vinyals O. and Dean J. Distilling the knowledge in a neural network. *arXiv preprint. arXiv:1503.02531* 2.7 (2015).

Construction Progress Monitoring and Reporting using Computer Vision Techniques – A Review

D. Shamsollahi^a, O. Moselhi^b and K. Khorasani^c

^aDepartment of Building, Civil and Environmental Engineering, Concordia University, Canada

^bCentre for Innovation in Construction and Infrastructure Engineering and Management (CICIEM), Department of Building, Civil and Environmental Engineering, Concordia University, Canada

^cDepartment of Electrical and Computer Engineering, Concordia University, Canada

E-mail: D_shams@encs.concordia.ca, Moselhi@encs.concordia.ca,
kash@ece.concordia.ca

Abstract –

In complex and dynamic construction sites, efficient progress monitoring and reporting play an important role in preventing schedule delays and cost overruns. Such reporting requires detailed and accurate records from job sites to help project managers in comparing project's current state to its as-planned state for enabling timely corrective actions. Currently used manual traditional progress reporting is time-consuming, costly, labor-intensive and can be inaccurate. Recently, there has been a shift toward utilization of digital images and computer vision techniques to overcome the above stated limitations and automate the tasks of progress reporting.

This paper investigates current practices and research works to provide an overview of the main computer vision techniques, tools and applications for progress monitoring and reporting using digital images. The paper also provides current achievements, challenges, and future research direction in this domain. According to the findings it can be concluded that computer vision is a promising path for improving the performance of automated progress monitoring systems in construction.

Keywords –

Progress Monitoring; Computer Vision; Object Detection; 3D Scene Reconstruction

1 Introduction

Monitoring and progress reporting are key management functions in delivering construction projects with least delays and cost overruns. In progress reporting, the current state of the project is compared to its as-planned state in order to identify unfavourable

performance early and apply corrective actions in a timely manner [1], [2]. Construction progress monitoring and reporting encompass data collecting, storing, reviewing, reporting, and representing the findings in an efficient manner [1].

It is also crucial to do progress monitoring in a repeated manner in complex and dynamic construction sites to address potential problems such as unpredicted costs derived from improperly performed work, reworking, lack of coordination, inaccurately allocating resources, etc. [3], [4]. However, current progress monitoring and reporting in construction is a challenging task for project managers due to inefficient manual systems which are time consuming, error-prone and labor-intensive. In addition to manual reporting and updating schedules based on drawings, schedules, and paper surveys, these systems are dependent on experts and their visual inspections for data collection [5].

To address the above stated issues, it is crucial to have an automated system to collect and analyse data accurately, visualize and report the findings in an interpretable format for different responsible parties [5]. Recently, new technologies and methods have been introduced and applied in the construction industry to automate the processes of monitoring and progress reporting. These processes are: (a) data collection of captured as-built/as-is scenes, (b) data analysis of collected data, (c) progress estimation by comparing as-built and as-planned information, (d) visualization of the results [4] as shown in Figure 1.

In recent years, low-priced and high-resolution digital cameras with high-capacity memories enabled construction companies to make use of these cameras in data capturing of construction operations on sites. Digital cameras can produce a large number of images and videos on a daily basis from as-built scenes containing lots of useful and detailed information [5], [6]. However, due to different challenges in image analysis tools such

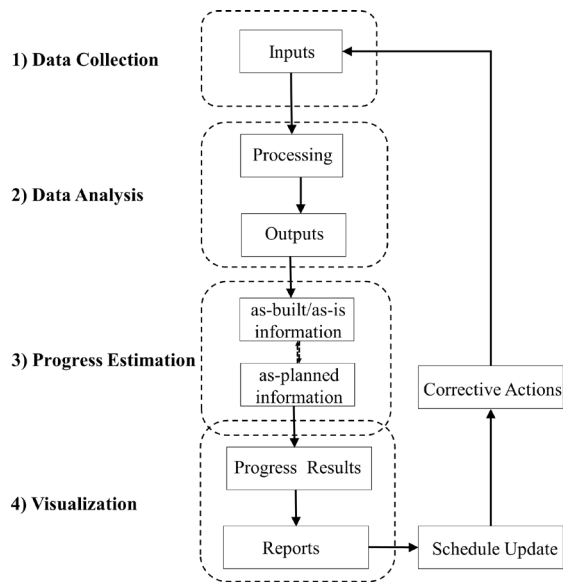


Figure 1 Construction Progress Monitoring and Reporting Process

as computation time, accuracy and cost, the images are analysed manually only for documentation and data recording purposes. As a result, only a small portion of this information were utilized and the rest became useless [6]–[8]. With developments in hardware platforms and algorithms, computer vision (CV) technology improved significantly [9]. Computer vision is a branch of artificial intelligence that uses computers to obtain high-level understanding from visual data, like human visual systems. During the last decade, computer vision has attracted many researchers due to its wide range applications in enhancing automation in construction. It can be applied for different project management purposes such as safety monitoring, quality control, productivity analysis and progress monitoring [10].

In progress monitoring systems which use computer vision techniques, the collected data are analyzed through different algorithms to have an understanding of the project's current state, then the results are compared with as-planned state to find progress deviations. Computer vision techniques can be categorized into 3D scene reconstruction, object tracking, object detection and image segmentation [8].

In recent years, the number of research studies related to automated progress monitoring and computer vision techniques in the construction industry has increased. Kopsida et al. [4] provided an in-depth review of automated progress monitoring steps and their related state-of-the-art technologies and methods. Patel et al. [11] explored recent developments, existing challenges, and future works for progress monitoring in the construction industry. Ekanayake et al. [12] reviewed different research studies regarding computer vision techniques for indoor progress monitoring.

The objective of this research is to provide a systematic review regarding recent developments in automated progress monitoring using digital images and computer vision techniques. Also, their applications in construction sites, related challenges and future works in this domain were discussed.

2 Research Methodology

This research adopted a systematic literature review to obtain findings and developments related to image-based progress monitoring using computer vision techniques. The literature was collected through search engine Google Scholar and academic databases such as ScienceDirect, ASCE, Springer, and IEEE Xplore. The keywords used in this study include: “construction progress monitoring”, “digital images”, “image processing”, “computer vision”, “object detection”, “instance segmentation”, “3D scene reconstruction”, “as-built 3D reconstruction”, “deep learning”, “building information modeling”. Then the related articles from peer-reviewed journals and conference proceedings were selected through analysing abstracts and titles. A total of 58 papers including 42 journal articles and 16 conference proceedings all written in the English language were selected and reviewed. The selected papers were stored by publication years, publication sources and the first author's country. As shown in Figure 2, these papers were published over a period from 2007 to 2021 and are categorized into “object detection and segmentation techniques”, “3D reconstruction” and “review papers”. The year 2021 has the highest number of published articles related to the topic.

Figure 3, shows country-based distribution of the selected articles, the United States has the highest number of publications, then the United Kingdom is the second. Figure 4, shows the number of publication sources. From 58 selected papers, the Automation in Construction journal has the highest number of published articles related to progress monitoring and computer vision techniques and after that is the Journal of Computing in Civil Engineering.

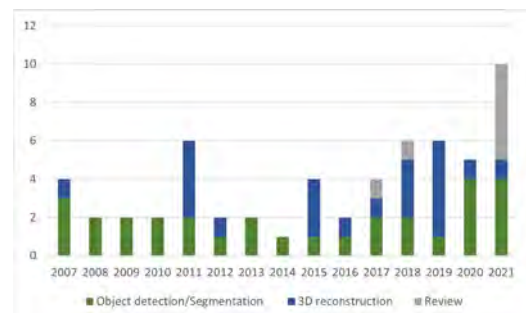


Figure 2. Chronological distribution of the selected articles

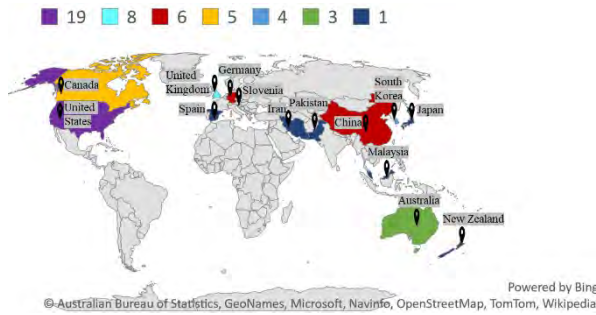


Figure 3. Country-based distribution of the selected articles.

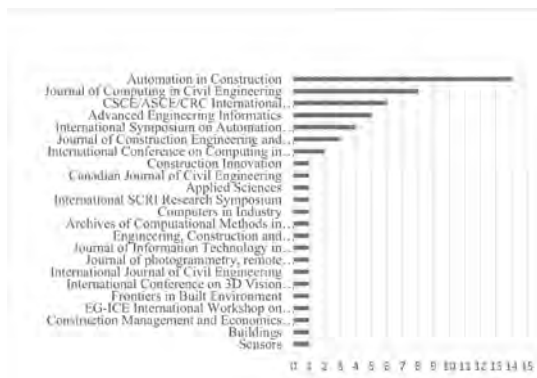


Figure 4. Number of publication and their sources.

3 Results and Discussion

An in-depth literature review reveals that three major computer vision techniques are used for construction progress monitoring and reporting using digital images, namely (i) 3D scene reconstruction, (ii) object detection, (iii) image segmentation. In this section, recent algorithms and technologies in these areas are discussed and the challenges are highlighted. In addition, the integration of building information modeling (BIM) with computer vision techniques are described subsequently in part 3.3.

3.1 3D Scene Reconstruction

In this technique, 3D representations (mesh models, point clouds and geometric models) are generated from one or multiple images taken from construction sites [13], [14] as shown in Figure 5. These 3D representations contain critical information pertinent to the current state of the project, which can then be compared to the as-planned state to track and report the project progress. For this purpose, collected data from cameras (monocular, stereo, video, panoramic, and RGB-Depth) is required to

generate the point cloud models [13], [15].

Computer vision techniques and algorithms for generating 3D scene reconstruction are different due to the characteristics of the input images. The input images are categorized into single and multiple images [16]. Single images can be taken using regular cameras or RGB-Depth cameras such as Azure Kinect. The Azure Kinect camera can easily create as-built 3D scenes using streams of depth and color images. However, for creating 3D reconstruction scenes using regular cameras, there is a need to calculate the depth of pixels in the images using computer vision techniques [17]. Eder et al. [18] developed and trained convolutional neural networks with a dataset containing RGB-D images to predict depth estimation of a single 360° image of an indoor scene that provides all information for creating the 3D as-is model.

The multiple images are divided into (i) multi-perspective 2D images and (ii) video sequences. In general, 3D scene reconstruction using multiple images have fewer challenges and is studied more frequently in the literature since they are more accurate with a higher level of detail compared to approaches using single images. In multi-perspective 2D images, several images with different perspectives of the objects are taken and create 3D scene representations based on parameters and poses of cameras [16], [17]. Fathi et al. [19] proposed a framework to create 3D point clouds using two calibrated cameras. The feature points captured from two video frames were detected using the Speeded Up Robust Features (SURF) algorithm. Automatic point matching between two video frames using Euclidean distance was applied and the outliers were removed using the RANdom SAmple Consensus (RANSAC) algorithm. Then triangulation was used to estimate spatial coordinates of the points in the frames and create point clouds of the construction objects on site.



Figure 5. 3D reconstruction from captured images by Omar et al. [54].

Video sequences can be utilized as an input for computer vision techniques to reconstruct 3D scenes. One of these techniques is the structure from motion (SfM) method that uses the shared information between consecutive frames by repeatedly extracting and matching features between two images, filtering outliers, and estimating poses of images and point clouds through image registration and triangulation [16], [17], [20]. In the study done by Golparvar-Fard [1] unordered daily photographs were used to reconstruct the as-built environment by using the SfM technique. Creating 3D point cloud models and as-built data, enabled project management team to visualize the project's current state through different viewpoints.

3.2 Object Detection

Due to numerous construction activities which use a wide variety of resources including materials, equipment, and workers, it is important to identify which resources are in the scene and which ones are involved in performing the task of interest [21]. Object detection is used to identify tracked building components automatically on site from the captured images and videos. This technique facilitates analysis of tracked activities and material allocation to support progress monitoring and reporting [22]. Object detection is a computer vision task that performs both classification and localization. Meaning that it classifies the objects in the captured image into pre-defined categories and predicts the location of each object in the image as shown in Figure 6.b [21], [23], [24].

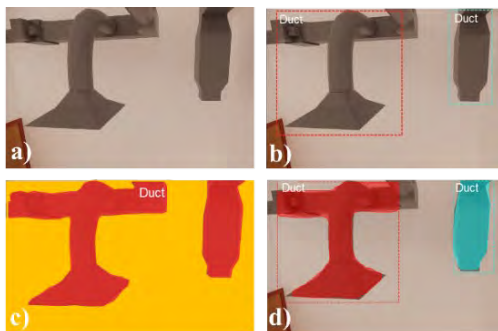


Figure 6. a) Original image b) Object detection result
c) Semantic segmentation result d) Instance segmentation result

In the early stages of object detection, many researchers used traditional (feature-based) algorithms which are essentially performed in a step-by-step process, requiring a specific model for each task. In these algorithms, the image features are extracted using feature descriptors such as Scale Invariant Feature Transform (SIFT), Local Binary Pattern (LBP), Binary Robust Independent Elementary Features (BRIEF) and

Histogram of Gradients (HOG). Next, these feature descriptors are combined with machine learning classifiers such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN), naive Bayes classifiers and neural networks for classification tasks [21], [23], [25].

Hamledari et al. [26] utilized different feature extraction and classification techniques such as color space selection, thresholding, edge extraction, and support vector machine to detect partitioned elements such as drywall, insulation, studs, and electrical outlets to report their actual state. Hui et al. [27] proposed a framework to detect and localize bricks in video frames using image processing techniques such as color thresholding and edge detection, then estimated the number of the bricks on the building façade automatically to improve performance of progress monitoring. These traditional methods have limitations in model generalization for detection purposes since they are based on hand-crafted features and require significant expertise effort for feature selection and extraction [28], [29].

In recent years, deep learning algorithms consisting of neural networks with many hidden layers, have provided solutions with better performance and reduced human involvement. This is achieved by introducing end to end learning process, which means that for completing feature learning, classification and regression tasks only a dataset of annotated images or video frames is required [28], [30], [31].

Many research studies have been developed using deep learning algorithms to monitor construction operations on site to support progress monitoring and reporting. For example, Hou et al. [6] trained Deeply Supervised Object Detector (DSOD) deep learning algorithm and detected building components including columns and beams automatically. Martinez et al. [32] proposed a framework to track progress of construction tasks automatically in offsite jobsites. In this research, Faster Region-based Convolutional Neural Network (Faster R-CNN) is applied to detect and classify the construction resources that are utilized in each task. Also Pour Rahimian et al. [33] developed a framework using computer vision techniques for building elements identification, integrating BIM and virtual reality to provide as-built information.

3.3 Image Segmentation

In image segmentation, which is also named as pixel-level classification, a digital image is separated into different meaningful regions to find how the interest objects are displayed in the image [9]. Image segmentation can be divided into semantic segmentation and instance segmentation [8]. Semantic segmentation refers to assigning a class label to each pixel in the image as shown in Figure 6.c [34]. In instance segmentation, detection and segmentation are joined in one model,

making detected objects distinguishable by pixel-wise masks. Compared to semantic segmentation, objects here from the same class, can be distinguished as separate instances as shown in Figure 6.d [9], [23], [35]. Through segmenting objects of interest by predicted masks, shape and size of objects in the image can be identified. In addition, the object boundaries can be extracted, providing spatial information for further geometry analysis, localization, and tracking which can assist progress monitoring systems [36]–[38]. Wang et al [38] developed an integrated framework using different computer vision tasks including instance segmentation and object tracking to monitor the progress of precast walls. Mask Region-based Convolutional Neural Network (Mask R-CNN) were utilized for detecting and segmenting the walls and DeepSORT for tracking walls through consecutive frames. Shamsollahi et al. [39] trained Mask-RCNN using transfer learning and data augmentation techniques to detect and segment HVAC ducts automatically for progress reporting the status of completed work on site. Table 1. Shows related research papers using computer vision techniques for progress monitoring and reporting.

3.4 BIM and Computer Vision Techniques

Through using BIM, construction entities can create and visualize 3D models, record, manage and analyze construction information. It also facilitates communication and collaboration among users [40]–[43]. 4D simulation is one of the most commonly used BIM methods for progress monitoring, which allows project managers to visualize and compare as-planned and as-built information through semantically enriched 3D models that are linked with project schedules [44]–[46]. As a result, the number of studies utilizing computer vision techniques and 4D BIM for automated progress monitoring in construction sites has increased in recent decades [47]–[49]. For example, Kropp et al. [50] utilized 4D BIM to find information related to the objects associated with specified activities as well as motion information to do a simple 2D classification. To evaluate the model, computer vision tasks including HOG features and SVM classifiers were applied to recognize heating devices in indoor construction site from image frames.

Also, other integrated systems using BIM and computer vision techniques have the potential to support progress monitoring systems. Ying and Lee [36] Developed an automated framework that creates as-is BIM elements using images taken from construction sites. Mask R-CNN, a deep learning-based object recognition algorithm were applied to detect and segment Walls, Doors, and Lifts from images. After the segmentation task, the masks boundaries of detected objects were extracted to generate surface geometries and construct IFC building objects.

Cloud-based BIM technology is another BIM development that provides opportunities for users to have access to project progress information in real-time. It is also a cost-effective collaboration tool that enables project entities to share and exchange necessary information through devices such as tablets and smartphones in different locations. This allows decision-makers to track the progress, organize construction schedules and apply early corrective actions [51]–[53]. Deng et al. [43] developed a method using computer vision and BIM to automatically measure and visualize the progress status of tiles. Computer vision techniques including local binary patterns (LBPs) and SVM classifier were used to identify tiles and the improved edge detection algorithm was applied to extract boundaries of the installed tiles from digital images. By using camera calibration and BIM model information the real tile area was calculated and the results were transferred to the BIM cloud platform for progress visualization.

Table 1. Applied computer vision techniques for progress monitoring and reporting.

Technique	Objectives	Ref.
3D scene reconstruction,	Creating as-built 3D model to visualize the project's current state.	[5,18] [49] [54]–[59]
Object detection,	Detecting construction components to acquire progress information.	[6,27,32] [33,36] [38,39] [60,61]
Image segmentation	Recognizing regions of objects in the image for spatial analysis and track the progress.	[26,33] [36], [38], [39]

4 Research Challenges

Despite many advances in implementation of computer vision tasks for progress monitoring, still some limitations are existed that need to be discussed. Some 3D scene reconstruction related issues are (1) lack of automation level in all the required steps for creating as-built models such as data collection and removing outliers. This increases, the operation time, and errors in the models [14], [15]; (2) Lighting conditions, occlusions, and cluttered backgrounds are unavoidable in construction sites which make 3D representations incomplete and noisy [13], [62]; (3) limited operating spaces in indoor environments [13]; (4) incapability of existed techniques in reconstructing of building elements with complicated geometric shapes (cylindrical, spherical, L-shaped, etc.) which mostly are in indoor environments [17]. The main challenges currently encountered in object detection and segmentation in the

construction industry are (1) lack of available open-source datasets of building elements for training models in a timely manner; (2) existed models are not generalized yet to detect different building elements through one model; (3) Lighting conditions, occlusions experienced on jobsites, camera movements, and limited view range have a negative effect on detection and segmentation performance [6], [26], [38], [61], [63].

5 Conclusion and Future Work

This research study described the importance of automated progress monitoring and introduced research on computer vision techniques. Next, a systematic review of articles published in journals and conference proceedings was carried out focusing on recent applications of digital images and computer vision techniques for construction progress monitoring and reporting. Also, integrated frameworks using computer vision tasks and BIM were discussed. The challenges of these techniques including 3D scene reconstruction, object detection and image segmentation in construction industry were identified and summarized. In future, more research study is required to improve algorithms in terms of accuracy and speed for generating the point clouds and object recognition purposes. Computer vision tasks can be linked with BIM automatically to create as-built models. Also, integration of computer vision techniques with tracking technologies such as radio frequency identification (RFID) and ultra-wide band (UWB) need to be investigated to provide more information about the current state of the project and generalize their applications on complex construction sites [12], [13], [15].

References

- [1] Golparvar-Fard M., Bohn J., Teizer J., Savarese S. and Peña-Mora F. Evaluation of image-based modeling and laser scanning accuracy for emerging automated performance monitoring techniques. *Automation in Construction*, 20(8): 1143-1155, 2011.
- [2] Moselhi O., Bardareh H. and Zhu Z. Automated data acquisition in construction with remote sensing technologies. *Applied Sciences*, 10(8):2846, 2020.
- [3] Yates J. K. and Epstein A. Avoiding and minimizing construction delay claim disputes in relational contracting. *Journal of Professional Issues in Engineering Education and Practice*, 132(2): 168-179, 2006.
- [4] Kopsida M., Brilakis I. and Vela P. A Review of Automated Construction Progress and Inspection Methods. In *Proceedings of the 32nd CIB W78 Conference*, pages 421-431, Eindhoven, The Netherlands, 2015.
- [5] Golparvar-Fard M., Peña-Mora F. and Savarese S. Automated Progress Monitoring Using Unordered Daily Construction Photographs and IFC-Based Building Information Models. *Journal of Computing in Civil Engineering*, 29(1): 04014025 , 2015.
- [6] Hou X., Zeng Y., and Xue J. Detecting Structural Components of Building Engineering Based on Deep-Learning Method. *Journal of Construction Engineering and Management*, 146(2):04019097, 2020.
- [7] Nieto A., Vilarino D.L. and Sánchez V.B. Towards the optimal hardware architecture for computer vision. *Machine Vision-Applications and Systems*, IntechOpen, London, 2012.
- [8] Paneru S. and Jeelani I. Computer vision applications in construction: Current state, opportunities & challenges. *Automation in Construction*, 132:103940, 2021.
- [9] Feng X., Jiang Y., Yang X., Du M. and Li X., Computer vision algorithms and hardware implementations: A survey. *Integration*, 69:309-320, 2019.
- [10] Xu S., Wang J., Shou W., Ngo T., Sadick A. M., and Wang X. Computer Vision Techniques in Construction: A Critical Review. *Archives of Computational Methods in Engineering*, 28(5):3383-3397, 2021.
- [11] Patel T., Guo B. H. W. and Zou Y. A scientometric review of construction progress monitoring studies. *Engineering, Construction and Architectural Management*, 2021.
- [12] Ekanayake B., Wong J. K. W., Fini A. A. F. and Smith P. Computer vision-based interior construction progress monitoring: A literature review and future research direction. *Automation in Construction*, 127:103705, 2021.
- [13] Xue J., Hou X. and Zeng Y., Review of image-based 3d reconstruction of building for automated construction progress monitoring. *Applied Sciences*, 11(17):7840, 2021.
- [14] Lu Q. and Lee S. Image-Based Technologies for Constructing As-Is Building Information Models for Existing Buildings. *Journal of Computing in Civil Engineering*, 31(4):04017005, 2017.
- [15] Ma Z. and Liu S. A review of 3D reconstruction techniques in civil engineering and their applications. *Advanced Engineering Informatics*, 37:163-174, 2018.
- [16] Ham H., Wesley J. and Hendra H. Computer vision based 3D reconstruction : A review. *International Journal of Electrical and Computer Engineering*, 9(4): 2394, 2019.
- [17] Kang Z., Yang J., Yang Z. and Cheng S. A review

- of techniques for 3D reconstruction of indoor environments. *ISPRS International Journal of Geo-Information*, 9(5):330, 2020.
- [18] Eder M., Moulon P., and Guan L. Pano Popups: Indoor 3D Reconstruction with a Plane-Aware Network. In *Proceedings of the 2019 International Conference on 3D Vision (3DV)*, Pages 76-84, Quebec City, Canada, 2019.
- [19] Fathi H., Brilakis I. and Vela P. Automated 3D structure inference of civil infrastructure using a stereo camera set. In *Proceedings of the Computing in Civil Engineering (2011)*, pages 118-125, Miami, Florida, 2011.
- [20] Jiang S., Jiang C. and Jiang W. Efficient structure from motion for large-scale UAV images: A review and a comparison of SfM tools. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167: 230-251, 2020.
- [21] Seo J., Han S., Lee S. and Kim H. Computer vision techniques for construction safety and health monitoring. *Advanced Engineering Informatics*. 29(2): 239-251, 2015.
- [22] Lin J.J. and Golparvar-Fard M., Construction progress monitoring using cyber-physical systems, *Cyber-Physical Systems in the Built Environment*, pages: 63-87, Springer, 2020.
- [23] Wu X., Sahoo D. and Hoi S. C. H. Recent advances in deep learning for object detection. *Neurocomputing*, 396:39-64, 2020.
- [24] Athira M. v. and Khan D. M. Recent Trends on Object Detection and Image Classification: A Review. In *Proceedings of the 2020 International Conference on Computational Performance Evaluation (ComPE)*, pages 427-435, Shillong, India, 2020.
- [25] Murphy K., Torralba A., Eaton D. and Freeman W. Object Detection and Localization Using Local and Global Features, *Toward category-level object recognition*, Springer, Berlin, Heidelberg, 2006.
- [26] Hamledari H., McCabe B. and Davari S. Automated computer vision-based detection of components of under-construction indoor partitions. *Automation in Construction*, 74:78-94, 2017.
- [27] Hui L., Park M.-W. and Brilakis I. Automated Brick Counting for Façade Construction Progress Estimation. *Journal of Computing in Civil Engineering*, 29(6):04014091, 2015.
- [28] Kim H., Kim H., Hong Y. W. and Byun H. Detecting Construction Equipment Using a Region-Based Fully Convolutional Network and Transfer Learning. *Journal of Computing in Civil Engineering*, 32(2):04017082, 2018.
- [29] Nath N. D. and Behzadan A. H. Deep Learning Models for Content-Based Retrieval of Construction Visual Data, In *Proceedings of the Computing in Civil Engineering*, pages 66-73, Atlanta, Georgia, 2019.
- [30] Wang J., Ma Y., Zhang L., Gao R. X. and Wu D. Deep learning for smart manufacturing: Methods and applications. *Journal of Manufacturing Systems*, 48:144-156, 2018.
- [31] O'Mahony N., Campbell S., Carvalho A., Harapanahalli S., Hernandez G.V., Krpalkova L., Riordan D., Walsh J. Deep Learning vs . Traditional Computer Vision. In *proceedings of the 2019 Computer Vision Conference (CVC)*, pages 128–144, Las Vegas, USA, 2019.
- [32] Martinez P., Barkokebas B., Hamzeh F., Al-Hussein M. and Ahmad R., A vision-based approach for automatic progress tracking of floor paneling in offsite construction facilities. *Automation in Construction*, 125:103620, 2021.
- [33] Pour Rahimian F., Seyedzadeh S., Oliver S., Rodriguez S. and Dawood N. On-demand monitoring of construction projects through a game-like hybrid application of BIM and machine learning. *Automation in Construction*, 110:103012, 2020.
- [34] Hao S., Zhou Y. and Guo Y. A Brief Survey on Semantic Segmentation with Deep Learning. *Neurocomputing*, 406:302-321, 2020.
- [35] Liu X., Deng Z. and Yang Y. Recent progress in semantic image segmentation. *Artificial Intelligence Review*, 52(2):1089-1106, 2019.
- [36] Ying H. Q. and Lee S. A mask R-CNN based approach to automatically construct As-is IFC BIM objects from digital images. In *Proceedings of the International Symposium on Automation and Robotics in Construction*, pages 761-771, Banff Alberta, Canada, 2019.
- [37] Kang K. S., Cho Y. W., Jin K. H., Kim Y.B. and Ryu H. G. Application of one-stage instance segmentation with weather conditions in surveillance cameras at construction sites. *Automation in Construction*, 133:104034., 2022.
- [38] Wang Z., Zhang Q., Yang B., Wu T., Lei K., Zhang B. and Fang T. Vision-Based Framework for Automatic Progress Monitoring of Precast Walls by Using Surveillance Videos during the Construction Phase. *Journal of Computing in Civil Engineering*, 35(1):04020056, 2021.
- [39] Shamsollahi D., Moselhi O. and Khorasani K. A Timely Object Recognition Method for Construction using the Mask R-CNN Architecture. . In *Proceedings of the International Symposium on Automation and Robotics in Construction*, pages 372-378, Dubai, UAE, 2021.
- [40] Wang X. and Love P. E. D. BIM + AR: Onsite information sharing and communication via advanced visualization. In *Proceedings of the 2012*

- IEEE 16th International Conference on Computer Supported Cooperative Work in Design (CSCWD), pages 850-855, Wuhan, China, 2012.
- [41] Oh M., Lee J., Hong S. W. and Jeong Y. Integrated system for BIM-based collaborative design. *Automation in Construction*, 58:196-206, 2015.
- [42] Alizadeh Salehi S. and Yitmen İ. Modeling and analysis of the impact of BIM-based field data capturing technologies on automated construction progress monitoring. *International Journal of Civil Engineering*, 16(12): 1669-1685, 2018.
- [43] Deng H., Hong H., Luo D., Deng Y. and Su C. Automatic Indoor Construction Process Monitoring for Tiles Based on BIM and Computer Vision. *Journal of Construction Engineering and Management*, 146(1):04019095, 2020.
- [44] Campagna-Wilson J. and Botton C. Challenges Related to 4D BIM Simulation in the Construction Industry. In *Proceedings of the International Conference on Cooperative Design, Visualization and Engineering*, pages 270-278, Bangkok, Thailand, 2020.
- [45] Alizadehsalehi S. and Yitmen I. The Impact of Field Data Capturing Technologies on Automated Construction Project Progress Monitoring. *Procedia Engineering*, 161:97-103, 2016.
- [46] Braun A., Tuttas S., Borrmann A. and Stilla U. A concept for automated construction progress monitoring using BIM-based geometric constraints and photogrammetric point clouds. *Journal of Information Technology in Construction*, 20, 2015.
- [47] Tuttas S., Braun A., Borrmann A., and Stilla U. Acquisition and Consecutive Registration of Photogrammetric Point Clouds for Construction Progress Monitoring Using a 4D BIM. *PFG—journal of photogrammetry, remote sensing and geoinformation science*, 85(1):3-15, 2017.
- [48] Kropp C., Koch C. and König M. Interior construction state recognition with 4D BIM registered image sequences. *Automation in Construction*, 86:11-32, 2018.
- [49] Han K. K. and Golparvar-Fard M. Appearance-based material classification for monitoring of operation-level construction progress using 4D BIM and site photologs. *Automation in Construction*, 53:44-57, 2015.
- [50] Kropp C., König M. and Koch C. Object recognition in BIM registered videos for indoor progress monitoring. In *Proceedings of the EG-ICE International Workshop on Intelligent Computing in Engineering*, Vienna, Austria, 2013.
- [51] Matthews J., Love P. E. D., Heinemann S., Chandler R., Rumsey C. and Olatunji O. Real time progress management: Re-engineering processes for cloud-based BIM in construction. *Automation in Construction*, 58:38-47, 2015.
- [52] Wong J., Wang X., Li H., Chan G. and Li H. A review of cloud-based bim technology in the construction sector. *Journal of Information Technology in Construction*, 19:281-91. 2014.
- [53] Afsari K., Eastman C. M., and Shelden D. R. Cloud-based BIM data transmission: Current status and challenges. In *Proceedings of the ISARC. International Symposium on Automation and Robotics in Construction*, pages 1073-1080, Auburn, USA, 2016.
- [54] Omar H., Mahdjoubi L. and Kheder G. Towards an automated photogrammetry-based approach for monitoring and controlling construction site activities, *Computers in Industry*, 98:172-182, 2018.
- [55] Brilakis I., Fathi H., and Rashidi A. Progressive 3D reconstruction of infrastructure with videogrammetry, *Automation in Construction*, 20(7): 884-895, 2011.
- [56] Ahmed M., Haas C. T., and Haas R., Using digital photogrammetry for pipe-works progress tracking, *Canadian Journal of Civil Engineering*, 39(9):1062-1071, 2012.
- [57] Han K., Degol J., and Golparvar-Fard M., Geometry- and Appearance-Based Reasoning of Construction Progress Monitoring, *Journal of Construction Engineering and Management*, 144(2): 04017110, 2018.
- [58] Mahami H., Nasirzadeh F., Ahmadabadian A. H., and Nahavandi S. Automated progress controlling and monitoring using daily site images and building information modelling. *Buildings*, 9(3):70, 2019.
- [59] Asadi K., Ramshankar H., Noghabaei M. and Han K. Real-Time Image Localization and Registration with BIM Using Perspective Alignment for Indoor Monitoring of Construction. *Journal of Computing in Civil Engineering*, 33(5):04019031, 2019.
- [60] Roh S., Aziz Z., and Peña-Mora F. An object-based 3D walk-through model for interior construction progress monitoring. *Automation in Construction*, 20(1):66-75, 2011.
- [61] Nath N. D. and Behzadan A. H. Deep Convolutional Networks for Construction Object Detection Under Different Visual Conditions. *Frontiers in Built Environment*, 6:97, 2020.
- [62] Han X. F., Laga H. and Bennamoun M. Image-based 3d object reconstruction: State-of-the-art and trends in the deep learning era. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(5): 1578-1604, 2021.
- [63] Zhou Y., Guo H., Ma L., Zhang Z. and Skitmore M. Image-based onsite object recognition for automatic crane lifting tasks. *Automation in Construction*, 123:103527, 2021.

BIM, Twin and Between: Conceptual Engineering Approach to Formalize Digital Twins in Construction

F.R. Correa^a

^aEscola Politecnica, University of Sao Paulo, Brazil

E-mail: fabiano.correa@usp.br

Abstract –

Construction is looking closely as Industry 4.0 paradigm (I4.0) transforms many processes in Manufacturing. Technologies associated with I4.0 were born out of the ever-present necessity of automation (and integration) on the shop floor, as well as of better management in product lifecycle with computer-aided software. Although indeed relevant outside Manufacturing, it is yet not clear how to transpose and apply some I4.0 technologies – for instance, Digital Twins (DTs) – outside the context of a factory or of complex, one-of-a-kind products. Many researchers and software companies from Architecture, Engineering, and Construction (AEC) and Facilities Management (FM) sectors are already working with an ill-defined concept of a DT, and some difficulties had arisen in dissociating it from Building Information Modeling (BIM). Without a direct counterpart outside AEC/FM, the practice of BIM could be likened to Product Lifecycle Management (PLM). Aiming at creating a clearer picture of where, when, and what should be a DT on AEC/FM, through literature review, clustering of common terms found, and a conceptual engineering approach, the present work develops a concept of DT, and layout differences between it and BIM model. It is advocated that DTs should be more about “functional models” for simulations than “product models” for information visualization and organization. As a result, DTs could be used to predict “behavior”, and thus enhance and transform management, operations, and maintenance practices. In doing so, there is yet a set of challenges that need to be addressed before one could create and employ properly Digital Twins in Construction.

Keywords –

BIM; Digital Twin; Simulation; Industry 4.0

1 Introduction

Construction is looking closely as Industry 4.0 paradigm (I4.0) transforms processes in Manufacturing,

following the digitalization (the use of digital technologies) phenomena which are underway in every aspect of our life in society.

Although indeed relevant outside Manufacturing, technologies associated with I4.0 were born out of the ever-present necessity of automation and integration on the shop floor, as well as of better management in product lifecycle with computer-aided software.

In a way, I4.0 development is an evolution of the quest for a Computer-Integrated Manufacturing (CIM), which Construction had also shared in the past, with its analogous Computer-Integrated Construction (CIC) [1] – a digital representation of the factory or of the manufacturing/construction process to control automated machines, robots, and even cyber-physical systems. Beyond that, there is the operation of the manufactured item / built environment that could be leveraged likewise with I4.0 technologies – although it could only make sense in complex, or automated products or systems.

Regardless its importance, it is yet not clear how to transpose and apply some I4.0 technologies – for instance, Digital Twins (DTs) – outside the original context in which it was first conceived, in factories or for operation of mechatronic systems.

DTs are already being explored by many researchers and software companies from Architecture, Engineering, and Construction (AEC) and Facilities Management (FM) sectors [2]. However, due to an ill-defined concept of DT, some difficulties had arisen in dissociating it from Building Information Modeling (BIM). BIM had been the backbone of all recent advances in digitalization throughout the entire lifecycle of the built environment.

BIM (and Virtual Design and Construction (VDC) [3]) practices have not a direct counterpart in other industries but could be likened to Product Lifecycle Management (PLM) [4][5], as encountered in automotive and aerospace industries, for example. Marketing around the idea of BIM made researchers of other fields, for example from the naval industry [6], to dedicate some attention to its particularities, but in reality, it is not an entirely different technology from

CAD/CAE/CAM or PLM practices.

Aiming at creating a clearer picture of where, when, and what should be a DT for AEC/FM, through literature review, clustering of common terms found, and a conceptual engineering approach, the present work develops a concept of DT, and layout differences between it and BIM models.

Although BIM models could be a framework for all information (geometric, material, and so on) about assets, it is advocated in this paper that Digital Twins, at its core, should be more about “functional models” for simulations [7], which could be used to predict “behavior”, and thus enhance and transform management, operations, and maintenance practices.

As one example, a DT composed of a set of models to simulate ageing in infrastructure, and employing historic data collected from structural health monitoring system to fine tune the models, could not only provide a safer approach to preventive maintenance, but also both retro feed information for better and optimized design (data-driven approach) as well as help in wisely management of scarce resources to the most needed infrastructure assets in a city, or country. Would such DT be possible to build and work with our technologies and knowledge?

So, this work aims to demonstrate that BIM and Digital Twin are different practices and are implemented in distinct systems. In this way, a new set of challenges and research questions arises, which need to be addressed before one could create and employ properly Digital Twins in Construction.

2 Understanding BIM

BIM could be viewed as part of a tradition in creating, organizing, and visualizing information inside computer systems about the design, construction, and operation of a product.

That tradition started with CAD (Computer-Aided Design) systems, which were developed to increase productivity and decrease time-to-market in engineering new products [8]. From a software to “draw lines” and produce technical documentations, to feature-based 3D solid modeling, not only drawings but the entire design process was transformed with the use of CAD – and later of CAM (Computer-Aided Manufacturing) and CAE (Computer-Aided Engineering) systems to deal with fabrication, simulation, and its intertwined relationship with design.

At the same time, came the recognition that CAD files, and all engineering knowledge of one enterprise, could be somehow managed by computer systems as well: there was the necessity to create a *framework* to link all data produced.

Figure 1 illustrates the evolution of such systems,

from the use and development of CAD Data Management to the PLM of today, for managing all the data related to a product or all products of an enterprise, from initial ideas to design, engineering, manufacturing, operation, maintenance, and disposal.

In the Architecture, Engineering, and Construction (AEC) sector, a similar trend started to be conceived at the end of the 70's: “... the design of a computer system useful for storing and manipulating design information at a detail allowing design, construction, and operational analysis” [9].

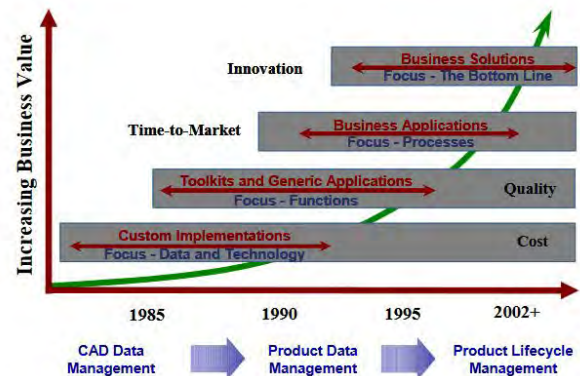


Figure 1. The Evolution of PLM [8]

Without the right technology to put idea into practice at the time [10], it was only years later that BIM authoring software started to appear in the market and being adopted by architects – and it was first likened as the evolution of CAD (Computer-Aided Design) programs. But from the outset, BIM had been conceived as a practice, and its associated software tools, that occurs throughout the lifecycle of the built environment.

Although sometimes BIM was used as acronym of Building Information *Model* [11], opting for *Modeling* emphasizes the *process* of using building information models, as is the case in PLM practices.

Thus, BIM could be viewed as “a socio-technical system that ultimately involves broad process changes in design, construction, and facility management” [4]. The object-oriented nature of BIM 3D modeling, and the highly fragmented platform of software tools for different disciplines in the project of a building, for example, is among the particularities which distinguishes BIM from other similar systems – and the inherent “multi-dimensional nature of the BIM domain” [13].

All the existent workflows with BIM models throughout the lifecycle of the building, potentially bring to attention the problem of interoperability, which now is addressed by the IFC (Industry Foundation Classes) open data schema.

The marketing around such a ‘disruptive’ solution to Construction known problems, leaded people from other segments to consider using BIM, instead of considering PLM for their business [6]. However, BIM and PLM are similar approaches to the same problem [5][4].

As Digital Twins were first thought in the context of PLM [14], there is this potential ill-defined concept of DT which likens it to the BIM model itself. Nowadays, at least two different development paths connect BIM with Digital Twins: Digital Twin of products – operation and Digital Twin of Factories – manufacturing, production.

Firstly, for BIM practitioners, a new trend started to be relevant: how to manage information around old buildings and infrastructure using BIM software tools, if all the design process were done without a digital model.

The use of laser scanning technology – as the service to employ it decreased considerably – and photogrammetry – as the quality of the calculated geometry increased – lead to a trend that many started to call “creating the Digital Twin” or a high fidelity, and digital, model of the built environment. It has been used for facility management, and even Heritage BIM (HBIM) [15].

Secondly, even if 4D BIM models are already used to manage and control production onsite (and there is a natural synergy between BIM, Lean management practices, and I4.0 technologies), the lack of development, and the difficulty in the representation of construction processes (IFC schema has entities representing processes inside the Construction Industry), had been lead to underachievement with regard to the foreseen potential of BIM – in fact, BIM models are mostly Product models [16], lacking a due representation for processes information.

In trying to differentiate DT from BIM, an important fact is that although the idea of a “[...] Digital Twin started off relatively sparse as a CAD description [...], in recent years it is becoming *actionable*” [14]. The meaning of the term actionable is given as such: “We can now simulate physical forces on this [CAD] object over time in order to determine its behavior” [14].

BIM models being analogous (although it contains semantic information) to CAD objects, one could argue if the *simulation* should be addressed *inside* the idea of BIM, or if it should be a synergetic action alongside it.

3 The role of Twins in the industry

As already mentioned, the *original idea* behind what now is referred to as the ‘Digital Twin’ was first presented by Dr. Grieves [14] in the context of a PLM course. With the advancement of the digital technologies, researchers started to evaluate the possibility of ‘exchanging atoms for bytes’ [17] –

instead of building and relaying in prototypes, decision making in design would be based on computer models and digital simulation.

The *term* Digital Twin appeared for the first time in the context of spaceship development inside NASA with the following meaning: “The digital twin concept is an approach to enable a suite of comprehensive multidisciplinary physics-based models that represent all of the physical materials, processes, and products, and ultimately incorporating these capabilities in the production and operation of spacecraft” [18].

One use for DTs would be to overcome current practices in product development aiming for *performance*. Current practices “[...] are largely based on statistical distributions of material properties, heuristic design philosophies, physical testing and assumed similitude between testing and operational conditions” [19]. Part of that change would be to focus on data-driven approaches, but also on different kind of models, which could be likened to the DTs.

It is important to recapture the necessity of a Twin as first experienced in reality [14]. Few controverse do exist around when the concept of a twin emerged in the industry, and it could be associated with the Apollo 13 Mission. Throughout many different project series to achieve know-how in space travel, NASA developed different physical simulators (15 were in use at the time of the Apollo 13 mission), so it could use them to prepare all team personnel and do diagnostics and test solutions to problems reported from the equipment traveling in space, which are inaccessible to the engineers on Earth. In the advent of problems detected with the flight of Apollo 13, NASA scientists and engineers had to rely on their physical simulators, and on data sent from the real Apollo 13, to test procedures and bring back the astronauts alive. What NASA needed was a solution to predict how systems respond to the environment, to the physics of the real word using computers.

Today, there are also initiatives in building Digital Twins of factories, and its shop floor, where the focus is production optimization, to study different scenarios with new machines, and robots, new shop-floor layouts, without stopping production – and only stop it for a few hours to make changes.

4 Between BIM and Twin

As is happening in the AEC-FM context with BIM, inside the PLM context there are also difficulties in separating and understanding the relationship between PLM and DTs. Among the differences pointed out in [20], Table 1 summarizes the *dynamic* aspects of accessing both systems, considering system-based data structures for DTs.

What could be highlighted from Table 1 is that DTs, differently from PLM system, could be inherently connected to IoT (to receive real-time data) and to Big Data Analytics (for processing), which make it more *dynamic*, with the possibility of working with data streaming.

Table 1. Comparison between PLM and Digital Twins (Adapted from [20])

Criteria	PLM	Digital Twin
Data exchange	Mainly document exchange; Collaborative access to the documents;	Microservices; Via the IoT-platform;
Real-time data usage	Storage; Simulation within the linked models; Hardly possible in document-centered structure;	Machine learning within IoT-platform; Cross-impact analysis within system model; Integration of Big Data algorithms;

Although many elements are expected to be present in a Digital Twin, one should ponder the questions:

- If what constitutes a Digital Twin depends on the existence of a real-time link (data flow) between the real object and its digital representation, it is impossible to have a Digital Twin in the Design and Planning Phase, where there is not yet a real object – in some areas, in these phases it should be called the Digital Thread;
- Timeseries data being important, for what reason it should be used, if not to simulate what is happening with the real object, and to gather knowledge that helps to understand better that specific built environment against theories, knowledge, and experience?;
- What is the use of a high-fidelity 3D model? Just for documentation or traditional management (BIM-HBIM)? In simulation scenarios, would this high-fidelity geometry really be used, or a more abstract model should be used to be computationally tractable? Think of the simplified models for simulating building energy;

All the raised concerns are somehow tied to the idea or ability to simulate a more realistic scenario based on real geometry or real data, or both. Also, it depends on the purpose of the simulation, and in the capacity of a computer to really simulate and predict behavior of the system based on the involved variables.

There is emphasis on the link between the real and the virtual spacecraft, but the data could only be

leveraged by means of the existence of models that relates input to behavior. For what purpose one needs “a suite of comprehensive multidisciplinary physics-based models” if not to simulate it on the computer, and predict its behavior in each scenario?

Recently, an increased number of research have been dealing with DTs in Construction. The differences of BIM and Digital Twin were addressed in the literature [21]. But the results of a systematic review of such a theme (a cross-learning between entirely different industries) inside a given community (i.e., AEC-FM) will largely depend on how the first authors interpreted and conceived a framework for, and a definition of, DT to be used in Construction –the conclusions could lead to propagating a biased definition.

4.1 Digital Twin in Construction

The methodology employed in this research consisted in doing a literature review, analysis of the relevant articles which helped in finding not only frequent terms and concepts associated with BIM and Digital Twins and using those concepts to build the concept of a Digital Twin following a Conceptual Engineering approach.

A search in the Scopus database with the combined term (“BIM” OR “Building Information Model”) AND “Digital Twin” returned 258 results, in February 2022 – the first result from 2017, and some from 2022. Two of the most cited articles provide disparate “visions” over the definition of a Digital Twin for Construction: 1) Focusing on the entire lifecycle of the built environment, authors see the Digital Twin as one evolution of BIM [22]; 2) Focusing on the operation phase, DT are BIM models for predictive behavior based on data from the real asset [23].

Performing a Semantic Analysis in the text of the abstracts of all the 258 results, considering the words found in each text, forming features of one and two words to characterize the topics discussed in each article, and then performing a Clusterization with KMeans, provided one analysis tool to investigate which concepts are commonly associated with DTs, and how they are associated with BIM and its predicted evolution.

Looking into the results of Literature Review, the relationship between BIM and DTs, as seen by the majority of researchers in the AEC-FM field, could be largely divided between two approaches.

As an example of the first proposition [21], BIM models should evolve to become the *de facto* DT for Construction. Between BIM and Digital Twin (a transition or evolution from one to another), there would be 5 steps: ‘BIM+Simulation’ (considering that most common use of simulation with BIM models), ‘BIM+Sensors’, ‘BIM+AI’, and finally Digital Twin. A

highlighted point here is the importance of simulation in this approach: without it, there is no DT.

The other proposition recognizes that DTs is a different thing from BIM, but both could be employed together for a different use. One example was presented by [24], with their Digital Twin Construction (DTC) emphasizing the act of building (production on a construction site) with a Digital Twin.

Also, looking into both propositions, one could identify at least two different types of DTs: one for the operation of the asset itself; and other for the production onsite (or even offsite). Both twins could be associated with data from the real twin obtained by a set of sensors.

4.2 Conceptual Engineering Digital Twins for Construction

Following a Conceptual Engineering approach, which “aims to revise rather than describe our concepts” [25], and is a more general approach than Conceptual Analysis, a proposition on understanding DTs for Construction is made.

However biased, let’s consider the role of simulation in the digitization process, and how it enables the integration (it blends more naturally around) of other I4.0 technologies such as Internet of Things (IoT), Cloud and Edge Computing, Big Data Analytics, Artificial Intelligence, and Cyber-physical Systems (CPS).

This consideration, while true to the first application of the terminology or idea of a (digital) twin presented in the literature, it emphasizes the role of simulation, rather than emphasizing the real time link with the real counterpart (which is needed), as the core of the technology – while using it to highlight the challenges in implementation for buildings and construction sites.

In that way, the meaning of the Digital Twin technology was analyzed to elaborate a definition that is useful for the particularities of the Construction sector. During reasoning, it was considered one definition to encompass at least two different applications of Digital Twin for Construction: DTs of the production on the construction site; and DTs of the built environment itself, be it a building or infrastructure.

Figure 2 illustrates the reasoning behind the analysis. In the term ‘Digital Twin’ there is the two terms: ‘Digital’ which means that the DT exists only inside a computer; and ‘Twin’, meaning that there is a real counterpart, and there is a link or relationship between the real and the digital. Following the reasoning, by ‘Digital’, there is the data, or the ‘Representation’ or ‘Model’, and there is a program that uses that data to do something – the use of a Digital Twin. The link or relationship is in fact a ‘Data Flow’ from the real twin to the program that runs based on the ‘Model’ and the ‘Data Flow’.

Delving further the concepts, the ‘Model’ could be a ‘Product Model’, as BIM models could be understood to be [16], and more importantly for this context, the ‘Functional Models’ or ‘Dynamic Models’ – how the current state of model should evolve on the computer with time and data that comes from the real twin. So, with this reasoning, the ‘Program’ is a ‘Simulation’ or a routine to ‘Control’ the real twin based on mathematical representation and a real time data, making a feedback loop.

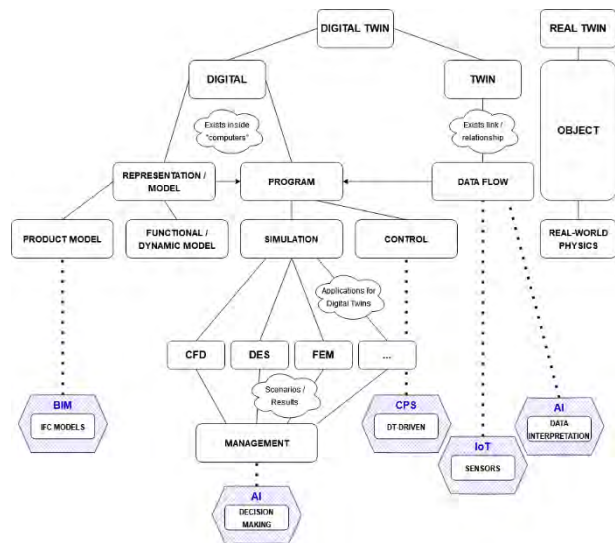


Figure 2. Concept Engineering for Digital Twin in Construction.

In the end, the simulation allows the ‘Management’ and / or ‘control’ of the real twin. So, the Digital Twin should be used to manage or control a real system and is comprised of a product model and at least on functional model that allows the state of the model to change based on external data and a representation of its inner work. In that way, although it is possible to have general functional models, they could be specific based on parameters and real data streaming.

In that definition of a Digital Twin for Construction (Figure 2), it is also illustrated where other technologies associated to the Industry 4.0 concept could be linked with it.

Two points that makes the task to understand the digital twin of the built environment: 1) when one conceives the digital twin of a factory, or the digital twin of a turbine, it is implicit that dynamic models of the robots in the factory and of the manufacturing processes with its transportation and feeding mechanisms, and the mechanical behavior of the turbine do exists; 2) to have sensor information about one thing without having a numerical model in which that data could be used as input, relegates to just data analytics or

business intelligence – the only connection possible with the building model is done by human interpretation of the data.

4.2.1 Implications of the proposed concept for DTs

1. The nature of the digital models

As data structure of both PLM and DTs are based on “multiple partial interlinked models” [20], as also is the case in BIM, one import question to be addressed is how to simplify development of DT models, and how to integrate those models in a specific simulation.

If DTs are the combination of functional models and a simulation engine, and the object to be simulated depends on different physical phenomena that works at the same time in the object, one could investigate the state-of-the-art of simulating multi-body, multi-physics, and multiscale simulations in other fields.

As for advanced fabrication methods, there would even concern about multi-dimensional modelling of materials to understand properties of components or parts designed for additive manufacture.

Also, hierarchical models should be considered, as for example, in the case of building simulations which could become part of cities simulations – as the case of DT applied to urban planning and city infrastructure management [26].

One implication of that discussion is that in many applications, the model used to simulate some physical phenomena, it is an abstraction or simplification of the real object – take in consideration the use of bond graphs for simulation [27]. Thus, point cloud-derived BIM models should be less than a desire as the representation for a DT application, being better for visualization and documentation – which are already BIM uses, and not DT uses. Current building simulations use a simplification of the 3D building geometry. Fidelity or High Precision or Accuracy in geometric form are sometimes against limited computation resources in doing simulations.

What need to be clear in deciding the model representation for a DT, is that it is directly associated to its intended use: for example, one thing is a model of the fabrication process – a Digital Twin of the shop floor for prefabricated timber wall panels, or a Digital Twin of onsite operations – and other thing is for the Operation of the Product itself – a Digital Twin of a Building, or Bridge.

2. Applications of DT in the lifecycle of the asset

From all the pioneered works on DTs, it is possible to list three main applications of DTs in Construction:

- The most obvious is to have a DT of a building, facility, or infrastructure for operation, management, and maintenance. The physical

representation could be derived from BIM models, or acquired by reality capture techniques, and simulations could be very specific to a given end, such as optimizing operation to save energy, or to predict structural failure, and would involve Computational Fluid Dynamics (CFD) and Finite Element Methods (FEM);

- One of the most important and difficult to implement would be the DT for management and control of production on construction sites, due to the nature of the activities, and the difficulties of tracking people, material, equipment, and building components on site. Also, the nature of the simulations which involves teams of workers, and many dynamic and unpredictable aspects should prove expensive to provide a meaningful result. Probably, simulation would be based on Discrete Event Simulation (DES);
- The most advanced use would be in automated activities in construction, where digital fabrication is possible, and the entire process of automation and control could be based on DT models. In this scenario, a Digital Twin-driven Cyber-Physical Systems could be employed, such as 3d-printing houses;

4.3 Challenges

Despite the lack of consensus around what is a Digital Twin, what few works addressed is what are the differences between new DT and “old” BIM models and practices. Also, what are the challenges in terms of development and implement of DT in Construction? As one could argue, there exists few if any applied DT worth of mentioning.

As is advocated in the previous section, the main difference could be established as the simulation core capabilities of DTs, against the management and visualization of information of BIM. Although some kind of simulation is already applied in design phases, simulation in the context of DT have a different nature, and a different proposition.

When one starts to consider simulation possibilities in Construction, what could be highlighted is the large presence of manual labor in production, and of artisanal practices on the field – even if it is more common some industrialized practices for prefabricated houses and apartments all around the world.

Few published articles addressed the questions about simulation of construction processes or simulation on the context of building operation and infrastructure maintenance.

4.3.1 People and Education

The first challenge is who will model, and how DTs will be modelled, if the professionals of the sector are

not well versed in that kind of technology? Modeling for simulation, and the know-how to deal with IoT systems and data flow is yet beyond the common knowledge received at academic institutions.

4.3.2 DTs Uses

One reason for that is the main challenge ahead to the way to adopt Digital Twins in Construction: how to simulate something that largely depends on human labor? How to predict the performance of each system that comprises a building, i.e., structure, envelope, etc...

4.3.3 Automatic Onsite Monitoring

Another challenge is how to improve and make widespread onsite monitoring, how to integrate with more predictable manufacturing outputs of prefabricated building components, and how to integrate logistics and the supply chain?

4.3.4 Digital Fabrication and Automated Construction

As the industrialization process becomes the norm in Construction, and the lack of construction works become a reality in more countries, further automated processes, mainly in production, would increase the need of a simulation tool both to test different approaches, as well as to control the process in real time, and to respond to deviations of the plans accordingly.

5 Conclusions

Digital Twins are core components of the Fourth Industrial Revolution based on the digitalization phenomena. It could play an important role on the global efforts to promote productivity and competitiveness in Construction Industry.

Although current practices of BIM could evolve and encompass the functionalities and workflows of the use of a Digital Twin, it seems more naturally and less complex that both work in a synergetic manner towards the end of management.

As core BIM practice is around product models, it will be interesting to enhance the simulations that could be made possible with adding functional models and data acquired from sensors to use it in the operation and maintenance of infrastructure and in Building Management Systems (BMs).

Four challenges were enumerated based on the proposed approach to DTs. These challenges should be addressed so that BIM models of today could work alongside the Digital Twins of the near future.

Future research should address how current simulations related to buildings, infrastructure, and onsite construction could be leveraged to work with enriched BIM models, and with data streams from

sensors. Not only better management and maintenance practices could be achieved, but also feedback for optimizing design and planning for new projects.

References

- [1] M. Fischer, M. Betts, M. Hannus, Y. Yamazaki, J. Laitinen, Goals, dimensions, and approaches for computer integrated construction, in: *Management of Information Technology for Construction*, 1993. <https://itc.scix.net/paper/w78-1993-2-421> (accessed February 22, 2022).
- [2] C. Newman, D. Edwards, I. Martek, J. Lai, W.D. Thwala, I. Rillie, Industry 4.0 deployment in the construction industry: a bibliometric literature review and UK-based case study, *Smart and Sustainable Built Environment*. (2020). <https://doi.org/10.1108/SASBE-02-2020-0016>.
- [3] J. Kunz, M. Fischer, *Virtual Design and Construction: Themes, Case Studies and Implementation Suggestions*, 2012. http://www.stanford.edu/group/CIFE/online_publications/WP097.pdf.
- [4] S. Aram, C. Eastman, Integration of PLM Solutions and BIM Systems for the AEC Industry, in: *International Symposium on Automation and Robotics in Construction (ISARC)*, 2013. <https://doi.org/10.22260/ISARC2013/0115>.
- [5] J.R. Jupp, Cross industry learning: a comparative study of product lifecycle management and building information modelling, *International Journal of Product Lifecycle Management*. 9 (2016) 258–284. <https://doi.org/10.1504/IJPLM.2016.080502>.
- [6] R. Luming, V. Singh, Comparing BIM in Construction with 3D Modeling in Shipbuilding Industries: Is the Grass Greener on the Other Side?, *IFIP Advances in Information and Communication Technology*. AICT-467 (2015) 193–202. https://doi.org/10.1007/978-3-319-33111-9_18.
- [7] S. Boschert, R. Rosen, Digital Twin—The Simulation Aspect, *Mechatronic Futures: Challenges and Solutions for Mechatronic Systems and Their Designers*. (2016) 59–74. https://doi.org/10.1007/978-3-319-32156-1_5.
- [8] CIMdata, *PDM to PLM: Growth of an Industry*, 2003.
- [9] ERIC - ED113833 - An Outline of the Building Description System. Research Report No. 50., 1974-Sep, (n.d.). <https://eric.ed.gov/?id=ED113833> (accessed February 22, 2022).

- [10] C.M. Eastman, *Building Product Models: Computer Environments, Supporting Design and Construction*, CRC Press, 1999.
- [11] T. Hartmann, H. van Meerveld, N. Vosseveld, A. Adriaanse, Aligning building information model tools and construction management methods, *Automation in Construction*. 22 (2012) 605–613. <https://doi.org/10.1016/J.AUTCON.2011.12.011>.
- [12] C.M. Eastman, P. Teicholz, R. Sacks, K. Liston, *BIM Handbook: a guide to building information modeling for owners, managers, designers, engineers, and contractors*, John Wiley & Sons, Inc., 2011.
- [13] B. Succar, Building Information Modelling Maturity Matrix, in: *Handbook of Research on Building Information Modeling and Construction Informatics: Concepts and Technologies*, 2010: pp. 65–103. <https://doi.org/10.4018/978-1-60566-928-1.CH004>.
- [14] M. Grieves, J. Vickers, Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems, *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches*. (2017) 85–113. https://doi.org/10.1007/978-3-319-38756-7_4.
- [15] J. Moyano, I. Gil-Arizón, J.E. Nieto-Julián, D. Marín-García, Analysis and management of structural deformations through parametric models and HBIM workflow in architectural heritage, *Journal of Building Engineering*. 45 (2022). <https://doi.org/10.1016/J.JOBE.2021.103274>.
- [16] W. Pan, B. Ilhan, T. Bock, Process Information Modelling (PIM) Concept for On-site Construction Management: Hong Kong Case, *Periodica Polytechnica Architecture*. 49 (2018) 165–175. <https://doi.org/10.3311/PPAR.12691>.
- [17] Dr.M. Grieves, *Virtually Perfect - Driving both lean and innovative products through Product Lifecycle Management*, Space Coast Press, 2011.
- [18] NASA Technology Roadmaps - TA 12 Materials, Structures, Mechanical Systems, and Manufacturing, (2015). <https://www.nasa.gov/offices/oct/home/taxonomy> (accessed February 17, 2022).
- [19] E.H. Glaessgen, D.S. Stargel, The digital twin paradigm for future NASA and U.S. Air force vehicles, *Collection of Technical Papers - AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*. (2012). <https://doi.org/10.2514/6.2012-1818>.
- [20] D. Adamenko, S. Kunnen, A. Nagarajah, Digital Twin and Product Lifecycle Management: What Is the Difference?, *IFIP Advances in Information and Communication Technology*. 594 (2020) 150–162. https://doi.org/10.1007/978-3-030-62807-9_13.
- [21] G. Editors, K. Ruikar, K. Kotecha, S. Sandbhor, A. Thomas, M. Deng, C.C. Menassa, A. Professor, V.R. Kamat, From BIM to digital twins: a systematic review of the evolution of intelligent building representations in the AEC-FM industry, *ITcon Vol. 26, Special Issue Next Generation ICT - How Distant Is Ubiquitous Computing?*, Pg. 58–83, <http://www.itcon.org/2021/5.26> (2021) 58–83. <https://doi.org/10.36680/J.ITCON.2021.005>.
- [22] C. Boje, A. Guerriero, S. Kubicki, Y. Rezgui, Towards a semantic Construction Digital Twin: Directions for future research, *Automation in Construction*. 114 (2020). <https://doi.org/10.1016/J.AUTCON.2020.103179>.
- [23] S.H. Khajavi, N.H. Motlagh, A. Jaribion, L.C. Werner, J. Holmstrom, Digital Twin: Vision, benefits, boundaries, and creation for buildings, *IEEE Access*. 7 (2019) 147406–147419. <https://doi.org/10.1109/ACCESS.2019.2946515>.
- [24] R. Sacks, I. Brilakis, E. Pikas, H.S. Xie, M. Girolami, Construction with digital twin information systems, *Data-Centric Engineering*. 1 (2020). <https://doi.org/10.1017/DCE.2020.16>.
- [25] S. Koch, Engineering what? On concepts in conceptual engineering, *Synthese* 2020 199:1. 199 (2020) 1955–1975. <https://doi.org/10.1007/S11229-020-02868-W>.
- [26] R. Al-Sehrawy, B. Kumar, R. Watson, A digital twin uses classification system for urban planning & city infrastructure management, *ITcon Vol. 26, Special Issue Construction 4.0: Established and Emerging Digital Technologies within the Construction Industry (ConVR 2020)*, Pg. 832–862, <http://www.itcon.org/2021/45.26> (2021) 832–862. <https://doi.org/10.36680/J.ITCON.2021.045>.
- [27] P.J. Gawthrop, M.I. Wallace, D.J. Wagg, Bond-graph based substructuring of dynamical systems, *Earthquake Engineering & Structural Dynamics*. 34 (2005) 687–703. <https://doi.org/10.1002/EQE.450>.

Inspection Data Exchange and Visualization for Building Maintenance using AR-enabled BIM

J. Park^a, S. Chang^b, H. Lee^c, and Y.K. Cho^a

^aSchool of Civil and Environmental Engineering, Georgia Institute of Technology, USA

^bSchool of Construction Management Technology, Purdue University, USA

^cConstruction Technology Department, BRICON LAB, South Korea

E-mail: jpark711@gatech.edu, chang776@purdue.edu, hyeongi84@gmail.com, yong.cho@ce.gatech.edu

Abstract

After the COVID-19 outbreak, a new concept of building maintenance (BM) systems is needed because current approaches highly rely on physical contact between workers, engineers, and managers. It imposes health and safety risks as increasing concerns about infections and spreads. This adds burdens to take unavoidable close contact and health risks to building owners, occupants, workforce, and society at large. In this respect, a new BM system was developed that enables reliable virtual communication and reduces BM response times by filling gaps between users and building managers. The proposed system is based on a concept of a cyber-physical system (CPS) using augmented reality (AR) and building information modeling (BIM) to promote non-contact building management. In this system, AR plays an important role in inspecting and visualizing defects in the real world, and the detected defect information is stored and managed by cloud-based BIM in cyberspace. This paper focuses on data visualization and management in the CPS-based non-contact building management system. A cloud-based database and mobile application are developed for data management purposes. In addition, this paper presents BIM data exchange and visualization in AR applications. Target image-based localization and tracking in BIM are also tested. The test results showed that the model alignment and localization accuracy are reliable for building maintenance works. Using the new BM mechanism, we expect that the related workers, building owners, and occupants will experience a reliable building maintenance process based on CPS-based information exchange from both users and facility managers while maintaining social distance.

Keywords –

BIM; AR; CPS; Building Maintenance System

1 Introduction

Although the accelerated number of vaccinations reduces the risk of infection, People's lives are still very different from before the Pandemic. After the outbreak of the tragic virus, many people want to minimize physical contact with others, and this trend may continue even after the COVID ends. For this reason, many health care and medical fields have been developing non-contact health care systems using cyber-physical systems (CPS) [1–3]. For example, Shah et al. presented a non-contact medical CPS framework with wireless sensing systems to monitor patients' physical activities [2], and Al-hababi et al. proposed a non-contact sensing network for post-surgery monitoring with artificial intelligence (AI) [3]. However, the study on non-contact systems for building maintenance and construction management is still lacking. Especially after the pandemic era, the need for the non-contact system is emphasized more because people spend more time indoors and are reluctant to physically contact others for maintenance and repair tasks.

Typical building maintenance is processed by the following steps: 1) reporting maintenance issues to facility managers or operators, 2) hosting a repairperson or contractor, 2) communicating face-to-face about the issue at hand, and 4) repairing defects and checking. Traditionally, physical contact between the occupants and maintenance workers is inevitable. Recently, some apartments have introduced a web-based maintenance system, in which the residents need to report the defects on the management website, and building managers carry out the maintenance work based on the resident's reports. In this way, the maintenance technicians can repair the reported defects without face-to-face instruction. However, repair request information provided by building occupants may be incomplete and inaccurate. For the correct maintenance work without the guidance of occupants, the occupants should fill in highly detailed

information such as the type, location, and size of building objects where the defects are found. It can be troublesome, time-consuming work for nonprofessionals. Moreover, in most cases, the repairpersons leave a hand-written report for their maintenance jobs, which also makes it difficult for occupants to understand what is done. To address the challenges in the current BM process, this research intends to devise a contact-free process for building maintenance and repair by enabling non-contact communication and information transition between occupants and building managers in a cloud-based building information model (BIM)

BIM is an intelligent digital representation of a building or other infrastructure entity such as a road [4], bridges [5], and tunnels [6]. BIMs represent the digital revolution's most important contribution to the architecture/engineering/construction (AEC) industry. BIM can account for all components of a building, including walls, floors, ceilings, columns, stairs, railings, doors, windows, wiring, plumbing, and HVAC systems [7–10]. It can store semantic information that describes each component's appearance, dimensions, weight, material, thermal performance, and other physical properties [11]. BIM supports decisions [12] and facilitates convergent collaboration [13] throughout a building's life cycle. During the operation and building maintenance phase, BIM can help owners monitor, maintain, renovate, and repair a building [14], assure its security [15], optimize its safety [16], and track its assets [14]. Therefore, by offering a single source of building data and streamlining data exchange, implementing cloud-based BIM for non-contact BM systems can benefit a broad population of architects, engineers, contractors, facility operators and managers, and building occupants.

With the advantages of BIM, we propose a novel CPS using augmented reality (AR) enabled-BIM. The main objective of this system is to achieve non-contact building maintenance in preparation for post-COVID-19 by providing a way of managing inspection data in cloud-based BIM and visualizing inspection data with AR application. In the proposed system, the building occupants can upload maintenance requests simply with a smartphone app instead of a traditional burdensome reporting process. This research is driven by the hypothesis that the physical contact between the BM stakeholders can be reduced by enabling seamless communications between cyber and physical components while preserving individuals' privacy. The proposed CPS system automatically computes the user's location with visual-inertial odometry and identifies the type of objects and defects with deep learning technology. By doing that, this system automatically obtains all necessary information for building defects maintenance, and the collected information is uploaded to cloud-based

BIM. By modeling an interoperable schema, the defect information is mapped to BIM, and the updated BIM is sent to the repairperson or facility managers. Among the proposed CPS component, this paper mainly focused on building inspection data management using cloud-based BIM and visualization using a mobile AR application. The following sections describe the way of exchanging and visualizing the building inspection data in CPS.

2 Related Works

Collaborative methods for BIM-AR technology have been employed to design, construction, and maintenance processes. By combining fuzzy multiple-criteria decision-making (MCDM), researchers can identify maintenance priorities among essential building information detected from BIM-AR applications [17]. However, BIM only provides static and predefined building data and information [18, 19] and may not contain target objects detected from AR. To update new building components, several ways of creating a new library of reusable parametric objects have been studied [20]. Parametric modeling using a scripting language and procedural modeling techniques using shape grammars can be combined and used for generating as-built BIM models. While procedural modeling enables automatic object generation and scalable geometric representation etc., manual methods are still required to create detailed elements [20]. Specifically, since the maintenance data is divorced from the BIM dataset [21], as-designed building information models do not contain an up-to-date status of buildings. For example, Construction operations building information exchange (COBie), a subset of the Industry Foundation Classes (IFC) schema, has been developed to collect data that can be used during operation and maintenance (O&M) stages [22]. However, COBie is non-geometrical data, so building models need to contain a high level of information, including installed conditions [22]. Although the COBie exports building elements data with their interdependent relationships, there are still missing links, and manual updates are necessary to maintain the recent building status [23]. COBie data exchange can also induce errors in the data transfer process [24]. Therefore, this research addresses current challenges of data exchange and updates in BM and intends to advance fundamental understanding of the feedback-loop formed during BM stages across building users, facility managers, and digitized building data and information. Another key challenge of cloud-BIM data transmissions is the lack of standardization [25]. The lack of cloud BIM-specific standards impedes BIM interoperability between data generators and feedback providers.

As discussed above, three major challenges need to be overcome before applying cloud-BIM and AR

technologies to building management: 1) disconnections of BIM and AR, 2) one-way data transfer only from BIM to AR or other systems, and 3) lack of cloud standards for BIM data exchange and AR visualization. In this study, therefore, we present a way of exchanging and managing inspection data between AR and BIM and visualizing the information in AR applications.

3 Methodology

3.1 Human-in-the-Loop CPS (HiLCPS)

This research applied CPS concepts for building maintenance. One of the fundamental challenges in enabling field applications of CPS is the difficulty of digitalization and virtualization of physical data acquired in the real world. In the proposed system, the inspection data acquired by building occupants are automatically entered into the BIM model in cyberspace, and this information is then communicated back to maintenance workers in the real world. However, the existing BIM format does not have a data schema to digitalize and manage the damage information acquired from the AR application. Another distinct but related issue is the high dimensionality and the complexity of the physical world, with noisy, poorly collected, and uninterpretable data, which can cause difficulties in data communication between the physical world and cyber systems. In the proposed CPS system, the inspection or maintenance information collected from the physical world via smartphone should be mapped to the exact location in the 3D cyberspace.

One way to tackle these challenges is to create an artificial tunnel that connects cyberspace and the physical world [26]. In this pathway, the sensing data in the physical world is digitalized according to a predefined data schema, and the digitalized information is mapped to BIMs through the cyber environment. The updated BIM securely exports specific data for maintenance works via the artificial tunnel, and the exported information is visualized in the maintenance workers' AR app. For this purpose, this research aims to develop a human-in-the-loop CPS (HiLCPS) for non-contact building or building maintenance after the COVID-19 era. Figure 1 shows the HiLCPS concept of the proposed non-contact building maintenance system. In this loop system, the inspection information is generated by occupants, and maintenance workers perform repair works based on the given inspection information. The maintenance statement written by workers is delivered through the loop system, and then the occupants confirm if the maintenance is done right or not.

3.2 Data exchange in cloud-BIM using fiducial markers

The major challenge of updating cloud-BIM data is a lack of an open-source framework [27] and a break of cloud models from the original building dataset [21]. To address the challenge, we created a two-way information model that can bridge between cloud-BIM and AR-detected datasets by developing an interoperable schema. This task intends to enable the interoperability between AR and Cloud BIM for mutual communications between

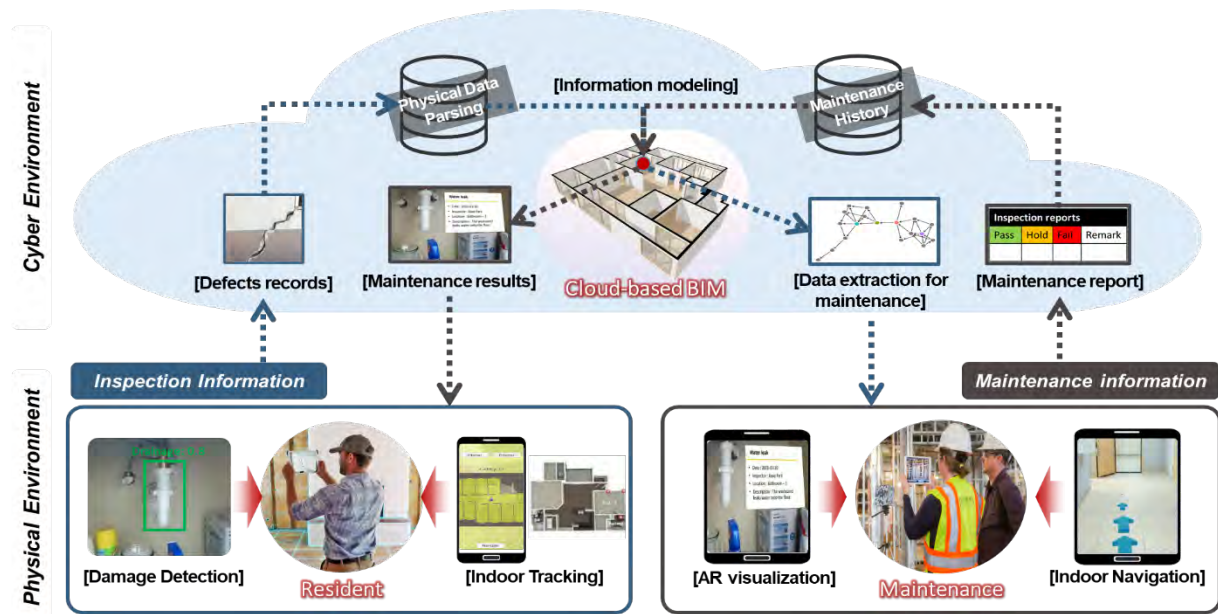


Figure 1. HiLCPS concept of the non-contact building maintenance system

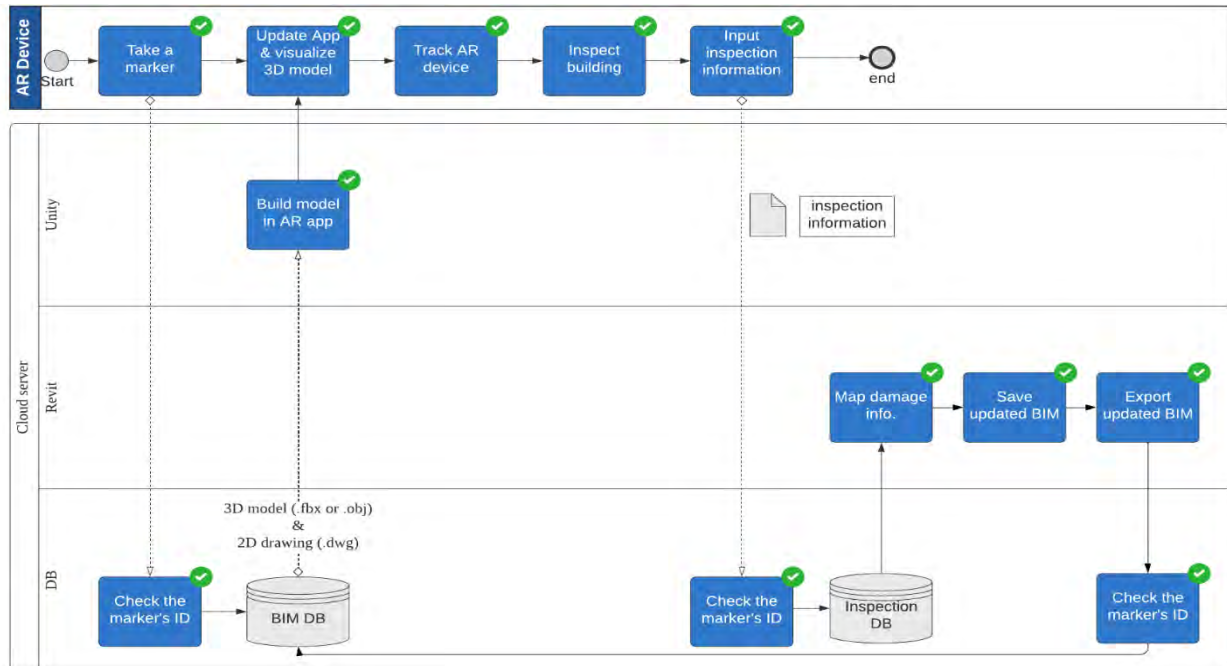


Figure 2. Data exchange model between the AR application and cloud-based BIM

multiple users and the proposed cyber-physical system and data exchange between two different systems of AR applications and Cloud BIM platform. This research team compared database schema, identified data loss, and built the integration by assembling building components in the original building models, cloud BIM, AR applications. Figure 2 shows the data exchange diagram. The integrated data schema can identify differences in the database in Cloud BIM platforms and AR applications, manage the effects of user inputs, and highlight and discover optimal information flow. The feasibility of the interoperable schema can be tested by modeling an integration task using DYNAMO, a combined environment of visual programming and textual programming. This research team also developed a DYNAMO script to access the AR database in BIM environments based on the integrated schema and verified the operational feasibility of the schema. The overall HiLCPS using the interoperable schema was tested its functions through a case study at an actual building. The outcome of integrating database schema can automate workflows of human inputs and building model updates and ultimately facilitate interactive communications of HiLCPS.

3.3 Indoor Localization and Model Alignment using Visual-Inertial Odometry (VIO)

For the fine localization and tracking of AR devices in GPS-denied environments, the AR camera's pose and location should be computed in real-time based on visual-inertial odometry (VIO). However, the VIO-based

localization cannot express the absolute coordinate values with a single monocular camera on a mobile device. For this purpose, we developed a BIM-assisted depth estimation and localization method to compute the AR device's pose and to output absolute coordinates in a map. The BIM models can be imported by a fiducial marker, called AprilTag, attached to each maintenance unit (e.g., a room and an office). The AprilTags can also return a single pose at 6-degree of freedom (DOF) relative to the camera frame of reference. By doing so, the initial location and posture of the AR device can be located at a corresponding point in the BIM model. The BIMs imported to the AR scene assist in estimating initial location and global depth. The AR app then continuously estimates the camera's trajectory using VIO. The estimated 6-DOF is used for two functions of the application to be developed: 1) estimating the exact location of detected damages and 2) navigating maintenance workers to the location of the defects, as shown in Figure 3.

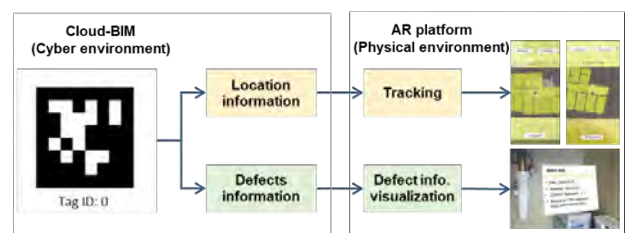


Figure 3. Localization and mapping in AR application using fiducial markers.

4 Case Study

4.1 Test Environment

To test the proposed HiLCPS-based building maintenance system, we conducted a case study in a one-story office building, as shown in Figure 4. The building was retrofitted in 2020, and there are a total of 14 rooms on both sides along the hallway. Through the case study, the performance of the data exchange between the AR application and cloud-based BIM using visual marker was tested. In addition, we also validated the accuracy of the localization and model alignment using VIO. A BIM was generated from the 2D drawings for building retrofit.



Figure 4. The outside and inside of the office building for test

4.2 Development Environment for the Building Maintenance System

For this test, we used an Apriltag for model alignment and multiple visual targets for localization. An Android smartphone with an operating system of 9.0 Pie was used during the test. The development environments for the database and AR application are:

- Spring Boot (4.11.0)
- MySQL (8.0.25.0)
- Unity (2020.3.13f1)

The file format for the 3D models imported to the AR application was Filmbox (.fbx), which is converted from Revit (.rvt) file through Dynamo. Two models, internal and external models, were generated from BIM, as shown in Figure 5. Tag ID #(1) assigned to the testing models.

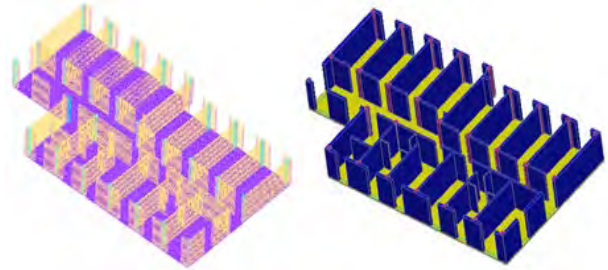


Figure 5. Internal 3D model (left) and external 3D model (right)

4.3 Test Result

4.3.1 Model Alignment

The accuracy of the model alignment was tested by measuring the discrepancies of real structure and visualized 3D model in the AR application. The 3D models are imported to the AR application by taking the Apriltag. Once the models are imported, the AR application visualizes the interior or exterior structure on the real-world scene. The AR application then tracks the position and pose of the mobile device in real-time with VIO. Therefore, the accuracy of model alignment and localization can be decreased as the distance from the fiducial marker increases. As described in Figure 6, The tests were performed on two walls 2 and 5 meters away from the fiducial marker. The measured alignment errors are described in Figure 7 and Table 1.

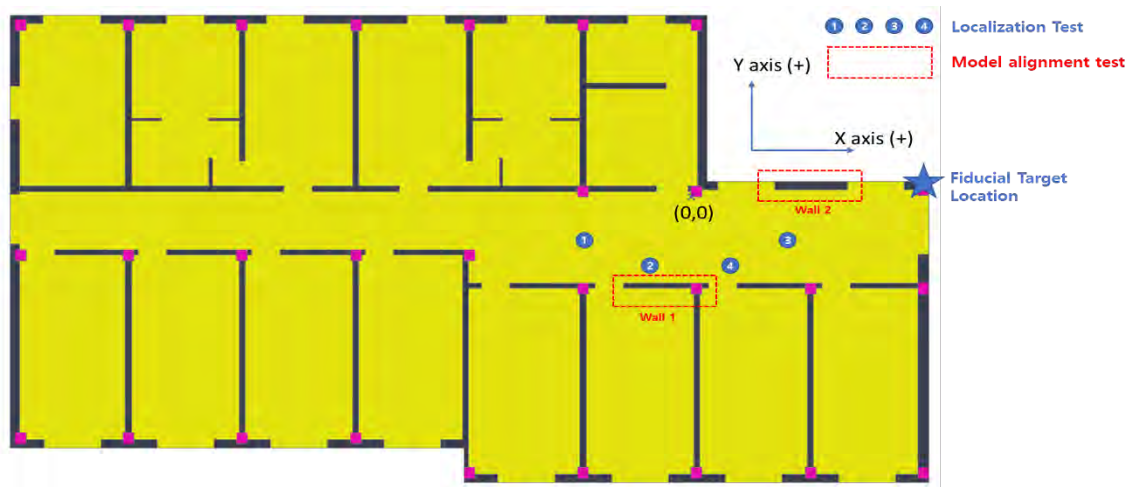


Figure 6. The positions for the localization test and model alignment test

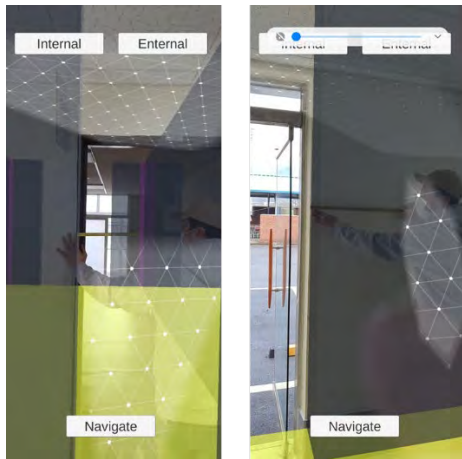


Figure 7. The model alignment test results: Wall 1 (left) and Wall 2 (right)

Table 1. Measured model alignment errors

Target	Error (mm)	Error per meter from marker
Wall 1	12	2.4
Wall 2	10	3.3

4.3.2 Localization and Tracking

The accuracy of the localization using VIO was evaluated by measuring the position errors at four positions, as described in Figure 6. The XY coordinates indicated in the application are the coordinate system of BIM, and metric units are used. The ground truth positions were measured by a tape measure and a laser ranger. The results are demonstrated in Figure 8 and Table 2.

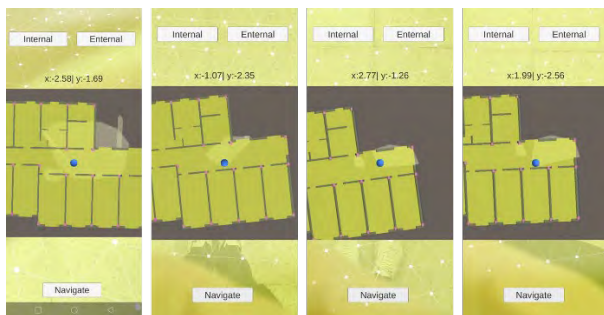


Figure 8. The localization test results at the test position 1, 2, 3, and 4

Table 2. Positioning errors

Target	Δx (m)	Δy (m)	Δxy (m)
Position 1	0.11	0.09	0.14
Position 2	0.10	0.12	0.16
Position 3	0.08	0.05	0.09
Position 4	0.10	0.15	0.18

4.3.3 Data Exchange

The mutual communications and data exchange between the AR application and the BIM database was carried out with a client application, as shown in Figure 9. The client app receives inspection and maintenance information from the mobile AR apps, formats it, and sends it to a web-based database. At this time, a predefined fiducial marker was used for model synchronization. Most of the inspection information, including the location of the observed damage, inspection date, and photos, is automatically entered into the client app. Figure 10 depicts the updated 3D assets in Unity with the inspection information collected from the AR application.

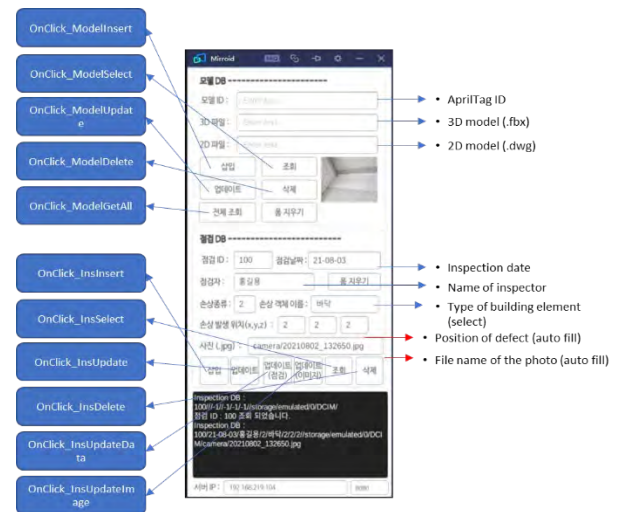


Figure 9. A client application for Inspection data management

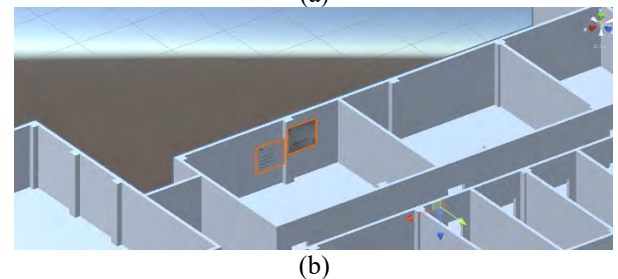


Figure 10. The inspection data management in web database (a) and 3D mapping on BIM

4.3.4 Data Visualization

The inspection information was managed by a web database, and the information could be mapped on BIM in real-time. The inspection information is then formatted in the Dynamo script as a text asset in the Unity application so that it can be visualized in the AR application. Figure 11 is a screen capture of the AR application visualizing the inspection information. For this test, we intentionally entered damage information into the inspection database. As shown in the figure below, the AR application visualized the damage information as a plain text box on the screen when the mobile phone camera was aimed at the location of the damage.

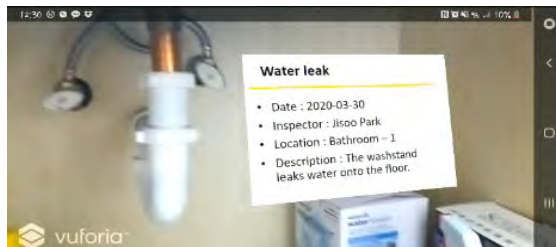


Figure 11. Screen capture of the AR application showing the inspection information

5 Discussion

This research presented a building inspection data management and visualization method using AR-enabled BIM. To validate the presented methods, we conducted a case study at an actual office building. The test results showed that the AR application could visualize the 3D models in real-world view within 1.2 cm when the distance from the fiducial marker was 5 m. In addition, the positioning error was less than 2 cm, which is reliable for inspection data localization and user tracking. Finally, we also tested if the data exchange between BIM and AR applications works well through the data visualization test. We confirmed that the AR application was able to visualize the inspection information with a virtual text box at the correct location.

Although our test results were positive and reasonable, we also found some technical problems on model alignment and localization. The AR application was not able to track the rapid movement and motion change of the AR camera precisely. Hence, the user's sudden movement and motion change caused significant errors in localization and model alignment. To address this problem, we recommend using multiple fiducial markers. Since our AR application was designed to update the 6-DOF whenever the predefined AprilTag is displayed on the screen, the localization and model alignment error would be easily adjustable by employing multiple

fiducial markers.

6 Conclusion

This research has three key scientific goals: (1) broadening understanding of Human-in-the-Loop Cyber-Physical Systems (HiLCPs), (2) advancing fundamental knowledge on Cloud-BIM data interoperability, and (3) developing an integrative network of AR-enabled BIM. Through the practical intervention of two user groups, residents and facility managers, the cyber-physical system improves its reliability and usability. Also, to fill the gaps between as-built building information model (BIM) data and operation and maintenance information updates, a semantic data model employed in cloud BIM was developed. This system can be utilized to support housing management decisions by providing integrated and timely data.

We expect the research findings will innovate current building maintenance practices through reliable virtual operation, maintenance, and repair of buildings. By facilitating near real-time communications, the proposed method can detect various building operation issues promptly and reliably, including miscellaneous repair, major structural risks, or occupants' health and safety. In addition, we expect the non-contact building maintenance system can improve the security from infection of occupants in the COVID-19 era by eliminating physical contact between the occupants and maintenance workers.

Acknowledgments

This work was supported by the Starting growth Technological R&D Program(S2872724) funded by the Ministry of SMEs and Startups(MSS, Korea).

References

- [1] Shah, S.A., Fan, D., Ren, A., Zhao, N., Yang, X. and Tanoli, S.A.K. Seizure episodes detection via smart medical sensing system. *Journal of Ambient Intelligence and Humanized Computing*, 11(11):4363-4375, 2020.
- [2] Al-hababi, M.A.M., Khan, M.B., Al-Turjman, F., Zhao, N. and Yang, X. Non-contact sensing testbed for post-surgery monitoring by exploiting artificial-intelligence. *Applied Sciences*, 10(14): 4886, 2020.
- [3] Jirgl, M., Bradac, Z. and Fiedler, P. Human-in-the-loop issue in context of the cyber-physical systems. *IFAC-PapersOnLine*, 51(6):225-230, 2018.
- [4] Chong, H.Y., Lopez, R., Wang, J., Wang, X. and Zhao, Z. Comparative analysis on the adoption and use of BIM in road infrastructure projects. *Journal of Management in Engineering*, 32(6):05016021,

- 2016.
- [5] Pipinato, A. and Modena, C. Structural analysis and fatigue reliability assessment of the Paderno bridge. *Practice Periodical on Structural Design and Construction*, 15(2):109-124, 2010.
- [6] Sharafat, A., Khan, M.S., Latif, K. and Seo, J. BIM-based tunnel information modeling framework for visualization, management, and simulation of drill-and-blast tunneling projects. *Journal of Computing in Civil Engineering*, 35(2):04020068, 2021.
- [7] Khosakitchalert, C., Yabuki, N. and Fukuda, T. Improving the accuracy of BIM-based quantity takeoff for compound elements. *Automation in Construction*, 106:102891, 2019.
- [8] Choi, J., Choi, J. and Kim, I. Development of BIM-based evacuation regulation checking system for high-rise and complex buildings. *Automation in construction*, 46:38-49, 2014.
- [9] Hossain, M.A., Abbott, E.L., Chua, D.K., Nguyen, T.Q. and Goh, Y.M. Design-for-safety knowledge library for BIM-integrated safety risk reviews. *Automation in Construction*, 94:290-302, 2018.
- [10] Esfahani, M.E., Eray, E., Chuo, S., Sharif, M.M. and Haas, C. Using Scan-to-BIM Techniques to Find Optimal Modeling Effort; A Methodology for Adaptive Reuse Projects. In *Proceedings of the 36th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 772-779, Banff, Canada, 2019.
- [11] Eastman, C.M., Eastman, C., Teicholz, P., Sacks, R. and Liston, K. *BIM handbook: A guide to building information modeling for owners, managers, designers, engineers and contractors*. John Wiley & Sons, New Jersey, USA, 2011.
- [12] Akbarieh, A., Jayasinghe, L.B., Waldmann, D. and Teferle, F.N. BIM-based end-of-lifecycle decision making and digital deconstruction: Literature review. *Sustainability*, 12(7):2670, 2020.
- [13] Piroozfar, P., Farr, E.R., Zadeh, A.H., Inacio, S.T., Kilgallon, S. and Jin, R. Facilitating building information modelling (BIM) using integrated project delivery (IPD): A UK perspective. *Journal of Building Engineering*, 26:100907, 2019.
- [14] Becerik-Gerber, B., Jazizadeh, F., Li, N. and Calis, G. Application areas and data requirements for BIM-enabled facilities management. *Journal of construction engineering and management*, 138(3): 431-442, 2012.
- [15] Zhang, X. Automatic underground space security monitoring based on BIM. *Computer Communications*, 157:85-91, 2020.
- [16] Wetzel, E.M. and Thabet, W.Y. The use of a BIM-based framework to support safe facility management processes. *Automation in Construction*, 60:12-24, 2015.
- [17] Wang, T.K. and Piao, Y. Development of BIM-AR-based facility risk assessment and maintenance system. *Journal of Performance of Constructed Facilities*, 33(6):04019068, 2019.
- [18] Park, J., Chen, J. and Cho, Y.K. Point cloud information modeling (PCIM): An innovative framework for as-is information modeling of construction sites. In *Construction Research Congress 2020: Computer Applications*, pages 1319-1326, Tempe, USA, 2020.
- [19] Park, J. and Cho, Y.K. Point Cloud Information Modeling: Deep Learning-Based Automated Information Modeling Framework for Point Cloud Data. *Journal of Construction Engineering and Management*, 148(2):04021191, 2022.
- [20] Dore, C. and Murphy, M. Semi-automatic generation of as-built BIM façade geometry from laser and image data. *Journal of Information Technology in Construction (ITcon)*, 19(2):20-46, 2014.
- [21] Garbett, J., Hartley, T. and Heesom, D. A multi-user collaborative BIM-AR system to support design and construction. *Automation in Construction*, 122:103487, 2021.
- [22] Kumar, V. and Teo, A.L.E. Development of a rule-based system to enhance the data consistency and usability of COBie datasheets. *Journal of Computational Design and Engineering*, 8(1):343-361, 2021.
- [23] Kumar, V. and Lin, E.T.A., Conceptualizing "COBieEvaluator": A rule based system for tracking asset changes using COBie datasheets. *Engineering, Construction and Architectural Management*. 27:1093–1118, 2020.
- [24] Demirdöğen, G., Işık, Z. and Arayıcı, Y. Facility management information taxonomy framework for queries in healthcare buildings. *Journal of Building Engineering*, 44:102510, 2021.
- [25] Afsari, K., Eastman, C.M. and Shelden, D.R. Cloud-based BIM data transmission: current status and challenges. In *Proceedings of the 33rd International Symposium on Automation and Robotics in Construction (ISARC)*, pages 1073-1080, Auburn, USA, 2016.
- [26] Price, L.C., Chen, J., Park, J. and Cho, Y.K. Multisensor-driven real-time crane monitoring system for blind lift operations: Lessons learned from a case study. *Automation in Construction*, 124:103552, 2021.
- [27] Logothetis, S., Karachaliou, E., Valari, E. and Stylianidis, E. Open source cloud-based technologies for BIM. In *Proceedings of the The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 607-614, Riva del Garda, Italy, 2018.

Comparison of TLS and Photogrammetric 3D Data Acquisition Techniques: Considerations for Developing Countries

E. Mengiste ^a, S. A. Prieto ^a, B. García de Soto ^a

^aS.M.A.R.T. Construction Research Group, Division of Engineering, New York University Abu Dhabi (NYUAD), Experimental Research Building, Saadiyat Island, P.O. Box 129188, Abu Dhabi, United Arab Emirates
eyob.mengiste@nyu.edu, samuel.prieto@nyu.edu, garcia.de.soto@nyu.edu

Abstract –

With the use of digital data, in conjunction with technologies such as BIM and digital twin, construction professionals have the ability to monitor the progress efficiently and perform detailed quality assessments. The rapid development of multiple data acquisition technologies, such as terrestrial laser scanning (TLS) and photogrammetry techniques, has allowed their broad integration in mostly middle to large-scale construction projects worldwide. However, this type of technology is often not accessible to small-scale contractors, especially in developing countries where high upfront costs and lack of skilled workers might be limiting factors. This study compares the economic, quality, feasibility, and value of TLS versus Photogrammetric data acquisition methods; i.e., the scope is not to discuss the benefits of using 3D digitization techniques but to consider if small-scale contractors can still take advantage of said benefits by using affordable technologies. Moreover, the impact of achieving complete automation by employing robotic platforms with TLS systems is also considered. A small case study is used to illustrate the quality and economic comparisons.

Keywords –

Construction 4.0; Terrestrial Laser Scanning; Photogrammetry; BIM; Construction Robots

1 Introduction

Lack of proper management and control in the construction is one reason for delay and cost overruns in developing countries [1]. In pursuit of improving construction project control and enhancing productivity, researchers have proposed progress monitoring approaches based on digitization. Cutting-edge technologies used in construction sites are associated with high infrastructural requirements and cost of equipment, heavy computational power requirements

and skilled/trained manpower [2]. These elements are typically lacking in construction companies in developing countries [3].

Data acquisition technologies are widely applied in the construction environment. However, except for 2D imagery [4] and basic terrestrial laser scanning (TLS) [5] devices, the use of such technologies makes it privative for low-budget construction companies due to the special infrastructural requirements and limitations that may arise. Other technologies, such as depth cameras, do not add any significant value to be considered in this study. For those reasons, this study compares point clouds generated using photogrammetry (2D imagery) and TLS to identify how low-cost applications can lead to high-effective outcomes regarding the integration of technology in developing countries. Other technologies, such as depth/stereo cameras, fall between the ones chosen but are not widely used in the field.

2 Comparison criteria

The comparison between the applicability of reality capture technologies in this study is conducted based on the following dimensions.

Economical: The initial and operational costs incurred in collecting and processing the data are compared. This includes the procurement of the initial device as well as the cost associated with the sufficient computational power required for the deployment and successful application.

Quality: Quality of the point cloud data that is obtained. It is measured in terms of point cloud density, noise volume, and lack of points due to occlusions.

Feasibility: Feasibility is measured in terms of ease of application on the construction site. Considering the specific characteristics of a construction site, the level of hindrance to deployment is measured.

Value added: Any improvement in construction automation needs to add value compared to regular manual operation. The comparison considers the level of detail required for a construction operation and the level

of information obtained through applying the alternative data collection and processing scheme. This parameter is subjective and is defined based on specific project characteristics.

3 Case Study

In order to test the two approaches, a small-scale construction site in the NYUAD campus (Saadiyat Island, United Arab Emirates) was chosen. Being on campus, the construction site was convenient to the authors, and it included many real conditions of a construction site (e.g., highly cluttered areas with temporary structures and storage of construction materials), making it a great space to illustrate this study.

3.1 Data acquisition and processing

The data acquisition was performed in four ways: (1) TLS mounted on a tripod: two different scan positions where the scanner was manually moved between them. (2) Autonomous Robotic System (ARS): autonomous robotic platform equipped with the TLS. Same scanning locations as the previous approach. (3) Static photogrammetry: a digital single-lens reflex (DSLR) camera mounted on a tripod, taking pictures in the same scanning locations as the two previous approaches. (4) Free-roam photogrammetry: handheld DSLR camera, data capturing as the camera moves around, with free tilt, rotation angle, height, and orientation

The scanning positions were the same for all (1 through 4). The only difference between 1 and 2 is how the scanner was moved between said scanning positions. In the case of the TLS mounted on a tripod (1), the scanner was moved manually by an operator. In the case of the TLS on board the ARS (2), the navigation was performed autonomously by the ARS, completely removing the human interaction in the process.

For the case of static photogrammetry (3), a DSLR camera mounted on a tripod was used to take two pictures every three seconds while rotating 360 degrees on the vertical angle. The same routine of image data collection is repeated five times at the same spot by varying the camera tilt angle approximately from -90° to 90° (measured from the horizontal plane) with 45° increments. This static photogrammetry is conducted at two different locations where the locations are selected manually to maximize the visibility of the whole area. For the given area, a total of 263 images were collected.

In the free-roam photogrammetry (4), data was collected roaming around the construction site, manually locating views and planes to reduce the clutter spots. The camera was set to take a picture every 2 seconds. The elevation of the camera was also manually adjusted based on the height of nearby objects. For the given experimental area, a total of 845 images were collected.

4 Results and discussion

4.1 Economical comparison

The economic representation of each technology examined in this case study is presented in Table 1. Table 1 presents the economic factors in terms of initial investment to buy the equipment and the operating expenses required to process the data and convert it to a complete representation of the construction site. In general, photogrammetric methods have relatively affordable initial expenses. Although the experiment is conducted using a high-end DSLR camera, photogrammetry can be achieved with affordable devices such as smartphone cameras. On the other hand, TLS requires relatively expensive tools. To achieve complete automation by mounting the TLS device on an autonomous robot, the initial cost can be doubled.

The initial investment is not the only factor that should be considered, but also the computational capacity required to process the raw data and the skilled/unskilled labor needed to provide a tangible result in any of the approaches. In this regard, the conducted case study revealed that although photogrammetric methods only require a basic camera and operation skills for a relatively shorter period, the data processing could be a lengthy task and requires a specific set of photogrammetric skills. This generally causes a relatively higher amount of man-hours and computational equipment utilization cost.

Photogrammetric methods could appear viable options for short-living and low-budget construction projects. However, the decision must consider multiple factors that could be influenced by applying this particular technology.

4.2 Quality comparison

Quality is defined based on multiple parameters, including noise ratio, point cloud density, and point cloud coverage of a given space. The results of the quality measure are presented in Table 2. Attribute-wise, all methods provide semantic information in terms of geometry (X, Y, Z), point normal (N_x , N_y , N_z) and color (RGB). However, the 3D laser scanner comes with an additional reflectance feature (R), which expands the applicability of a specific method to a wider domain.

The percentage of noise from the point cloud data is computed as a ratio of the removed data during multiple layers of filtering to the total pre-filtered point clouds. As a result, photogrammetry-based methods had less noise with 14.6% and 13.76% for the static and free-roaming camera-based photogrammetry, respectively. The percentage of noise values for the TLS-based methods were 30.6% and 20.91% for static and ARS mounted TLS, respectively. Although the noise volume appears smaller in photogrammetric methods, this can be attributed to the low density and short span reaching

abilities these specific methods have. Both filtering layers applied are statistics-based, which depends on the dispersity and semantic inputs of the neighboring points.

As the samples are less dense, outliers could fit in the model and mislead the filtering process.

Table 1: Economic comparison

Approach /Technology	Investment		Operation requirements			
	Data collection equipment	Processing	Data collection		data processing	
			Skills	Man-hours	Skills	Man-hours
(1) Static TLS	\$19,000	\$8,000	Use of TLS	12.52 min	CAD, 3D data processing	5 min
(2) TLS in ARS	\$39,000	\$8,000	Use of TLS / Robotic skills	17.70 min	CAD, 3D data processing	5 min
(3) Static photogrammetry	\$1,000	\$8,000	None	5.25 min	Photogrammetry engineer	22 min
(4) Free-roaming photogrammetry	\$1,000	\$8,000	None	6.17 min	Photogrammetry engineer	67min

Moreover, the static TLS resulted in 10% more noise when compared with the ARS mounted TLS. This can be attributed to the ability of the mobile system to take a closer scan of objects that are reflective and cluttered.

A point cloud's volume is measured based on the volumetric space enclosed within the set of neighboring points. The search of points enclosed around a given point is conducted using a 0.077 m search radius consistently among all data sets. The results obtained are summarized as the mean and standard deviation of a density histogram (Table 2). Comparatively, among the

four methods, the TLS mounted on the ARS provided the densest results but were very close to the results obtained from the static TLS. However, density can only represent the compactness of points around a given scan. For instance, through a visual inspection, static photogrammetry has clear openings where limited objects are reconstructed. It is challenging to measure the magnitude of occlusion through the point cloud density. Therefore, the point cloud population in rasterized cells is computed to observe the relative completeness of points in a given space.

Table 2. Quality comparison

Data source	Attributes	Point count		Point Cloud Density		Per-cell population ratio
		Pre- filtering (1x10 ⁶)	Post - filtering (SOR and MLS) (1x10 ⁶)	Mean (1x10 ⁹)	STDEV (1x10 ⁹)	
TLS	*	59.23	41.09	2.21	1.35	260.03
ARS	*	54.27	42.92	2.23	1.54	264.57
Static camera	**	2.10	1.79	0.24	0.23	42.44
Free-roaming camera	**	27.06	23.34	0.67	0.83	131.82

* Geometry (X, Y, Z), Point Normal (N_x, N_y, N_z), Color (RGB), and Reflectance (R) ** Geometry, Point Normal, and Color

The per-cell population ratio represents the average distribution of points in equal-sized cells with a magnitude of 0.05. Table 2 presents ARS mounted TLS as the most complete while the static camera-based photogrammetry is the least complete. The mobile aspect of the robotic system has helped improve the completeness of the TLS-based system. Similarly, the free-roaming photogrammetry improved the static camera photogrammetry by more than 3-fold.

4.3 Feasibility comparison

The experimental site consists of flat planes, clutter, narrow paths, and hidden areas (e.g., underneath

scaffoldings and equipment). Tripod-mounted methods (static camera photogrammetry and static TLS) share blind spot occurrences around the hidden areas or behind solid objects. On the other hand, the free-roaming camera and ARS mounted TLS have the advantage of mobility to address hidden zones to reduce occlusion.

Although the ARS can navigate through the room with a clear understanding of obstacles and obstructions, narrow paths and clutter could hinder the system from accessing certain areas. In a similar observation, unlike the TLS methods, photogrammetric methods heavily depend on natural illumination. Inter-reflection between shiny objects and less illuminated areas within the

experimental area were some of the primary causes of noise.

4.4 Value added to the project

The technologies in this case study show that complete automation can only be achieved through the TLS mounted ARS approach. This approach has proved efficient in construction sites in various instances [5]. However, this method commonly requires trained workers to collect and process raw information. It has proved to be fast and provides relatively complete point clouds with the highest density value. The method comparatively results in the highest standard data, but for a higher cost. This method can be proposed for large projects with higher contract amounts. In most cases, large construction sites produce a relatively large amount of waste and are easy victims of mismanagement. Manual operations could result in delayed data with little to no room to produce decisions before resources are wasted. On the other hand, small-scale and low-budget construction projects need to perform a cost-benefit analysis before choosing automated data acquisition methods. One of the important aspects of the examined technologies in this case study is the ability to create a simplified way of visualizing the construction site in 3D. This helps contractors, owners, and consultants devise construction decisions, safety measures and many more, which are usually time-taking activities.

The applicability (i.e., benefits) of the four comparison criterion for the methods evaluated in this case study is summarized in Figure 1. Applicability in the y-axis is a comprehensive measure composed of the four comparison criterion used in this study. Given the availability of digital cameras (economy and feasibility), photogrammetry is the most applicable method for small constructions. However, the quality of the data and the value that the processed data adds to the project varies between static and free-roaming photogrammetry. As the complexity of the project increases, the size of image data and the requirement of processing capacity (economy) increases. Based on the quality of data, the ease of obtaining the results (feasibility) and the value it adds to the project, TLS-based methods appear to be relatively applicable. Given the complexity of the construction, blind spots could be avoided, and data registration could be error-free if it is conducted with the help of an ARS.

5 Conclusions and future work

TLS and photogrammetry are used indistinctly of the applicability or suitability of the situation in which they are implemented but based on the availability of the required hardware/technology to the user. However, small-scale and low-budget contractors, especially in developing countries, often have to be specific regarding hardware requirements due to cost, availability or lack of

qualified personnel. This study looks at the development of point clouds acquired with TLS and photogrammetry and serves as a base to objectively compare the feasibility and usability of each technology.

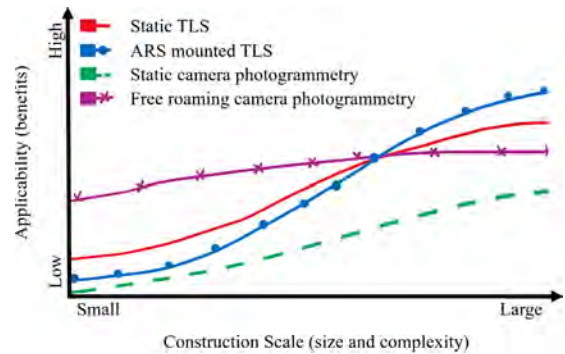


Figure 1: Comparative benefit of technologies investigated vs. the scale of construction project

The results are based on a small case study, which is not the best scenario to accurately represent the benefits of some of the technologies used. The study concludes that automated data acquisition provides reliable control with relatively low budgets for large and complicated construction projects. However, image-based methods could be more applicable for smaller construction projects, given their mild initial cost and skill requirement. Nevertheless, the value added to the process must be evaluated relative to the conventional methods.

References

- [1] Frimpong Y, Oluwoye J, Crawford L. Causes of delay and cost overruns in construction of groundwater projects in a developing countries; Ghana as a case study. *International Journal of Project Management*. 2003 Jul 1;21(5):321-6.
- [2] Kochovski P, Stankovski V. Supporting smart construction with dependable edge computing infrastructures and applications. *Automation in Construction*. 2018 Jan 1; 85:182-92.
- [3] Adepoju OO, Aigbavboa CO. Assessing knowledge and skills gap for construction 4.0 in a developing economy. *Journal of Public Affairs*. 2021 Aug; 21(3):e2264.
- [4] Mostafa K, Hegazy T. Review of image-based analysis and applications in construction. *Automation in Construction*. 2021 Feb 1; 122:103516.
- [5] Wu C, Yuan Y, Tang Y, Tian B. Application of Terrestrial Laser Scanning (TLS) in the Architecture, Engineering and Construction (AEC) Industry. *Sensors*. 2021 Dec 30;22(1):265.

Mud Dauber: Prototype of the Mobile Gantry Architecture

P.J. Staritz^a, J.C. McClurg^a, C.S. Miller^a, M.J. Schlenker^a, S.A. Wozniak^a

^aDepartment of Physics and Engineering, Taylor University, USA
Email: peterstaritz@taylor.edu, josiah_mcclurg@taylor.edu, caleb_miller5@taylor.edu, moriah.schlenker@gmail.com, shanewoz@ymail.com

Abstract

The Mobile Gantry is a robot architecture in which a rail-less gantry rolls directly on the unprepared planetary surface. Mud Dauber is a prototype of this architecture developed to enable testing of the concept and understand the capabilities and limitations of the architecture. This paper describes the major subsystems developed and tested in Mud Dauber: coarse positioning, fine positioning, localization, print head and control.

Keywords

3D Printing; Sulfur Concrete; Mobile Gantry; Robot Architecture; Mars

1 Introduction

Creating a self-sustaining colony on Mars requires the ability to autonomously construct habitable structures using locally available (in-situ) resources. 3D printing is an excellent solution to this problem, and in recent years research into print materials, print techniques, robotic autonomy, and building architectures have examined many aspects of this approach. A growing body of research into robotic 3D-printed construction has primarily explored three robotic architectures: fixed base radial arm, mobile robot and gantry.

In the *fixed base radial arm architecture* [1] a long robotic arm is used to manipulate the print head while the base of the arm remains fixed. Concepts of the type often envision the radial arm mounted to a mobile base to allow the system to move to other locations *between* prints. Because each structure is printed from a single location, the size of the structure is directly related to the size of the robot. The limited reach of this architecture means that the footprint of the printed structures are generally measured in meters and usually self-contained. For example, a winner of the NASA Habitat Challenge, envisions pod-like structures isolated from one another that require pressurized suits when traversing to other habitats. [2]

In the *mobile robot architecture* [3, 4] robots capable of maneuvering during the build process employ shorter, more agile arms to print as they maneuver in and around the structure. This architecture envisions a beehive of activity in which many robots work together to print the structure. In this way, the mobile robot architecture enables much larger and more capable structures. However, this capability comes at a cost, impacting robot control complexity, introducing constraints on habitat design, and requiring frequent recharging. The control complexity is much greater than the other architectures because robots must be able to maneuver within the structure to print internal walls. As a result, mobility paths must be coordinated between robots, and robot manipulator paths must avoid collision with already printed structure and other robots. In addition, because of these various constraints, robots of this architecture cannot ensure continuous extrusions along the print path. This results in frequent discontinuities in the path, which yields a weaker structure and limits the utility of reinforcing fibers. The requirement that the robots must be able to maneuver within the structure also means that the structure must have internal dimensions (doorways, hallways, etc.) large enough to accommodate the robots that are building it. Finally, this architecture requires the robots to be tetherless in order to avoid power cable tangling and damage from other robots rolling over the cables. The need to be tetherless requires portable power sources, for example rechargeable batteries which must be frequently recharged, limiting robot operational time and increasing overall system complexity.

In the *gantry architecture* [5, 6, 7] a single robot, positioned on rails, maneuvers the print head from above, in a manner very similar to a desktop 3D printer. Printing from above, simplifies control and allows for very long continuous print paths. However, only a single robot can work in the workspace at one time, limiting the speed at which large structures can be built. As with the fixed base architecture, the size of the printed structures is directly related to the size of the robot. Thus, the footprint of the building is bounded by the position and length of the rails and the height of the building is bounded by the height of the gantry. When building sequential structures, this architecture requires

repositioning of the rails. This repositioning process is complex and would likely require additional robots dedicated to the task.

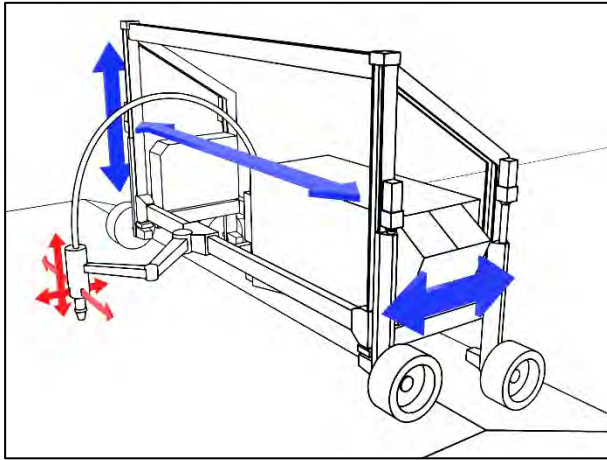


Figure 1. Artist's rendition of the mobile gantry robotic architecture with coarse positioning system (blue) and fine positioning system (red).

This team has developed a hybrid architecture, the *mobile gantry*, which benefits from many of the positive characteristics of the other architectures, eliminates some of the most difficult negative aspects, and introduces new capabilities. The mobile gantry architecture is a gantry without rails. Robots of this architecture maneuver directly on the planetary surface using wheels and can maneuver from print area to print area using their mobility system. During the printing process these robots roll back and forth as though on a set of virtual rails. Figure 1 presents an artist's concept of this architecture. The architecture benefits from simple control and long continuous print paths, enabling the building of structures measuring 10's or even 100's of meters in length. Because robots of this architecture are designed to print while maneuvering on the planetary surface, they are capable of using their printed structures as scaffolding that they can climb using their mobility system. In this way, the mobile gantry architecture is able to build structures taller than the individual robots. Moreover, multiple robots can print on the same structure at the same time, increasing the speed that structures are printed. The paper [8] describes the benefits and tradeoffs of this hybrid architecture in detail. In this work, early prototype results are presented, demonstrating core functionality from the primary subsystems of the Mud Dauber mobile gantry.

2 Mud Dauber Robot

The Mobile Gantry architecture maneuvers directly on the unprepared planetary surface using wheels. To print while moving, these robots isolate the motion of the

print head from the motion of the robot body. Large-scale motion is performed using the *coarse positioning system*, consisting of vehicle mobility (X-axis) and two coarse degrees of freedom (Y & Z Axes). The *fine positioning system* is moved in the workspace by the coarse positioning system and provides precision positioning of the print head, allowing it to follow the print path at the required velocity. (Figure 1) A *localization system* is used to measure the print head position in the world frame. The *print head system* enables heating, mixing, and extrusion of a high-quality sulfur concrete slurry. The *control system* commands and coordinates robot actions.

Mud Dauber (Figure 2) is a prototype mobile gantry designed to enable testing of the robotic architecture concept and understand its capabilities. Designed for laboratory testing, it is 2.4m wide, 1.65m tall and 1.75m in length. Through the development of Mud Dauber, the team is developing insights with regard to mobility systems, fine positioning, localization and print head technology.



Figure 2. Mud Dauber is a prototype testbed of the mobile gantry architecture.

2.1 Coarse Positioning System

Mud Dauber's coarse positioning system consists of the robot's frame, wheels and a leg extension module. This system enables the robot to traverse long straight paths in the test area and decouples the length of the building from the size of the robot.

The wheels move the robot along the primary axis of the system (X-axis, see Figure 2) and a leg extension module enables the robot to navigate over obstacles and uneven terrain (Figure 3). This device changes the length of the right rear leg and allows Mud Dauber to bypass obstacles up to 17.5 cm in height. In future iterations of the coarse positioning system the design will include coarse Y and

Z axis degrees of freedom and a leg extension module on each leg. However, to control the scope and cost of this initial prototype these additional degrees of freedom were not implemented.

The addition of Y and Z degrees of freedom will allow the movement of the fine positioning system in 3 axes and enable the printing of structures up to 1.8m wide and 1.3m tall. The addition of the 3 other leg extension modules will allow the traversal of slopes and rough terrain while maintaining the orientation of the robot during printing.

Initial testing of the Mud Dauber has confirmed basic functionality of the coarse positioning system. The robot traverses the X-axis at a maximum velocity of more than 2.5 cm/s. The leg extension module articulates at 0.8 cm/s, can articulate over obstacles and is capable of lifting loads in excess of 110kg. The robot has also demonstrated the ability to climb inclines of more than 25 degrees.



Figure 3. The leg extension mechanism keeps the robot level when traversing obstacles and slopes.

2.2 Fine Positioning System

The fine positioning system isolates the motion of the print head from the coarse motion of the robot. This allows the mobile gantry robot to print precise paths over very long distances. To demonstrate this capability, a 3-axis fine positioning system was developed. This system is capable of moving the print head within a print volume of 30 cm on a side and allows print head positioning to within ± 1 mm. The system employs an inverted frame that ensures the print head is the lowest component in the print area, preventing collisions with printed structures.

The fine positioning system enables the print head to traverse in all directions at the maximum print velocity of 2.5 cm/s in the world frame regardless of coarse robot motion. This capability allows the robot to position the printing workspace independent of the print head motion, ensuring that the print path remains continuous for the entire printed structure.

Testing of the fine positioning system has confirmed baseline functionality. The X and Y axes traverse at a velocity of more than 5.0 cm/s and with a precision of ± 1 mm. Z axis motion only occurs during the transition from one printed layer to the next and is limited to 0.1cm/s. The fine positioning system is capable of responding to three-axis position and velocity commands from Mud Dauber's control system.

In future work this version of the fine positioning system will be replaced with a 3 degree of freedom radial manipulator arm. This concept is illustrated in Figure 1 and will enable Mud Dauber to reach in front of and beyond the front wheels of the robot. This capability is critical for the demonstration of self-scaffolding whereby the robot prints the walls of the habitable structure so that it can use the walls as scaffolding. Using this approach, future versions of Mud Dauber will be able to print structures taller than the robot itself.

2.3 Localization System

Because the entire mobile gantry is moving through 3D space, the absolute position of the print head must be determined in the world frame. To enable Mud Dauber to print high quality structures, this system must support an update frequency of at least 30Hz and with an accuracy of 1mm. To accomplish this, a precision wireless localization system is being developed.

In this approach, four transmitters are positioned around the outside of the print envelope at fixed locations in the world frame. Receivers are mounted on the print head and robot frame to provide the position and velocity of both the print head and the robot as a whole. The system measures the distances to the fixed bases and calculates the position and velocity of the robot and print head using multilateration.

Progress has been made on an early low-precision version of the distance estimation using off-the-shelf digital transceivers with wired reference clocks separated by a known phase offset. Additionally, the prototype localization system includes a multilateration estimator

running on embedded hardware that is capable of streaming data to Mud Dauber's control system.

In future work, a high-speed wirelessly-synchronized clock reference and a low jitter RF (radio frequency) stack will allow the proposed system to be capable of precise distance measurements at high sample rates.

2.4 Print Head System

Future Martian construction systems will use locally available (in-situ) resources as raw materials. A promising material for 3D printing structures on Mars is sulfur concrete which is made by mixing sulfur and regolith and then heating the mixture above the melting point of sulfur (~113°C). When the slurry cools, the sulfur freezes and acts as a binding agent for the regolith.



Figure 4. The first-generation print head extrudes a mixture of sulfur and sand.

Mud Dauber's print head (Figure 4) is capable of heating, mixing and extruding sulfur concrete. Raw materials are loaded into the mixing hopper (1250 cm³) where they are heated and mixed. An auger extrudes the slurry through a 2.8cm diameter nozzle at a maximum rate of 18 cm³/s. The print head is capable of imparting up to 4500W of heat in two independently controllable zones: mixing chamber and extrusion nozzle. Testing has validated mixing performance, extrudate homogeneity, extrusion rates and the impact of extrusion temperature on slump. (Figure 5) These tests revealed problems with the flow of extrudate from the mixing chamber to the extrusion nozzle which is being addressed in a second generation print head.

In future work the print head will be redesigned to enable continuous extrusion, improved efficiency, and better sensing and closed loop control. The new design will allow for continuous operation in which the print head is

constantly ingesting, mixing, heating and extruding the sulfur concrete. Eliminating the batch-based design enables continuous extrusions and minimizes discontinuities. The print head will change the heating approach to improve the thermal efficiency of the system and decrease overall print head power consumption. The print head will incorporate improved sensing to allow for higher quality extrusions in a wider range of laboratory conditions.



Figure 5 – Test extrusions from the first-generation print head show the variability in the material slump due to changes in extrusion temperature.

2.5 Control System

Control of Mud Dauber is based on an STM32 microcontroller. The controller coordinates robot mobility, terrain compensation, fine positioning, and print head action. The STM32 communicates with a user interface and issues commands to subsystems which employ microcontrollers for closed loop control.

3 Terrestrial Applications

Today, commercial companies [5, 6] are 3D printing terrestrial structures using robots of the gantry and fixed base architectures. These robots have shown great promise in their ability to build useful structures for human use at low cost. But the printing process still includes substantial involvement from humans during setup, operation and finishing of the structures.

The mobile gantry architecture has the potential to further decrease the cost of commercial terrestrial 3D printed structures by eliminating or reducing the effort associated with setting up and moving robots during the print. For example, the mobile gantry architecture can print a series of structures one after another without ever needing to move a set of rails. This architecture can also increase the size of such structures, leading to longer and taller buildings.

The mobile gantry architecture becomes even more valuable when considering military applications. Structures in forward operating areas must be built quickly, often under threat, with minimal existing infrastructure and support. The mobile gantry architecture requires fewer support personnel (reducing risk), is more easily transported (due to the ability to roll out of the transport container) and only requires site clearing and not extensive site preparation. These factors make the mobile gantry an ideal candidate for future 3D printing of military structures.

Applying this architecture in these areas will require continued advances in the coarse positioning, fine positioning, and localization subsystems as well as the integration of print heads that are relevant in terrestrial applications.

4 Conclusion

Mud Dauber is a prototype of the *mobile gantry architecture*. Robots of this architecture roll directly on the planetary surface but move back and forth as though on a set of virtual rails. These robots are simple to control, generate continuous print paths, and can build structures measuring 10's or even 100's of meters in length. A key new capability of this architecture is the ability to use the robot's printed structures as scaffolding upon which the robot can climb to print taller structures. This capability allows the printing of structures taller than the robots themselves. These capabilities open the door to the construction of complex and capable structures that are taller and longer than those generated by robots of competing architectures and similar size.

Mud Dauber is an early prototype of this architecture, designed to perform 3D printing of test structures while demonstrating key capabilities. This paper detailed the five major subsystems developed for the prototype and reported on early validation testing of these subsystems. Mud Dauber is a first step in the development of more capable 3D printers for construction on Earth, Mars and beyond.

4 Acknowledgement

This research was performed under funding from the Indiana Space Grant Consortium (INSGC) at Taylor University. Prototype development was performed as part of engineering design classes and as such was supported by a large team of undergraduate students. Special thanks to: H. Childs, J. Conejero, C. Deckard, N.

Eshuis, C. Gardner, M. Jacques, K. King, C. Lehrian, L. Mason, J. Meleski, N. Reisler, J. Richey, M. Shearer, N. Streitmatter, R. Cartwright, and S. Dalcher.

5 References

- [1] "Meet Frank and his family" Apis Cor, 2022. Online: <https://www.apis-cor.com/technology>, Accessed: 2/20/2022
- [2] "Architecture on Mars" AI Spacefactory, 2022. Online: <https://www.aispacefactory.com/marsha>
- [3] Yuan X., Zhang J., Zahiri B., and Khoshnevis B. "Performance of Sulfur Concrete in Planetary Applications of Contour Crafting." Additive Manufacturing Conference, Jan. 2016.
- [4] Howe A. S. *et al.* "Modular Additive Construction Using Native Materials." *Earth and Space 2014*, St. Louis, Missouri, Jun. 2015, pp. 301–312.
- [5] Cesaretti G., Dini E., De Kestelier X., Colla V., and Pambaguian L. "Building components for an outpost on the Lunar soil by means of a novel 3D printing technology." *Acta Astronautica*, vol. 93, pp. 430–450, Jan. 2014.
- [6] Khorramshahi M. R. and Mokhtari A. "Automatic Construction by Contour Crafting Technology." *Emerg Sci J*, vol. 1, no. 1, p. 28, Jul. 2017.
- [7] "Meet ICON's next generation Vulcan construction system." <https://www.iconbuild.com/vulcan> Accessed: 4/24/2022
- [8] Staritz P.J., Miller, C.S., and McClurg, J.C. "The Mobile Gantry: A Robotic Architecture for 3D Printing Structures on Mars" Available EngrXiv.org

A model driven method for crack detection in robotic inspection

E. Pellegrino¹ and T. Stathaki¹

¹Department of Electrical and Electronic Engineering, Imperial College London, UK
e.pellegrino16@imperial.ac.uk

Abstract

Nowadays civil infrastructure is exposed to several challenges such as daily vehicular traffic and extreme weather conditions. It is well known that these may determine structural deterioration and damages, which can even cause catastrophic collapses related to significant socio-economic losses. For this reason automatic inspection and maintenance are expected to play a decisive role in the future. With the objective of quality assessment, cracks on large civil structures have to be identified and monitored continuously. Due to the availability of cheap devices, techniques based on image processing have been gaining in popularity, but they require an accurate analysis of a huge amount of data. Moreover, fracture detection in images remains a challenging task due to high sensitivity to noise and environmental light conditions. This paper proposes a mathematical method for detecting cracks in images along with a parallel implementation on heterogeneous High Performance Computing (HPC) architectures aiming both at automatizing the whole process and at reducing its execution time.

Keywords

Machine Vision; Smart Infrastructure; High Performance Computing

1 Introduction

Infrastructures can be exposed to different loading conditions, recurrent ones due to vehicular traffic and extraordinary ones caused by earthquakes, wind and strong rain. The consequently induced stresses may determine structural deterioration and damage, which can even cause catastrophic collapses related to significant socio-economic losses [1]. For this reason increasing the level of automation and maintenance is object of research interests. As matter of fact, several authors have pointed out that current visual inspections, which highly rely on a human subjective and error-prone empirical evaluation [2], can be enhanced by robotic/automatic assisted operations [3], [4]. Actually, actions performed by inspectors require often a long time to examine large areas and specialized equipment such as large under bridge units, heavy trucks etc. In most cases those expensive solutions may cause high logistical efforts and even interfere with normal

operational conditions.

Recent works address the problem of the automation of inspection and maintenance tasks based on robotic systems [5]. Existing automatic or robotic systems based on ground or aerial solutions have been proposed for inspection of dangerous sites or those difficult to access, but at the present state-of-the-art, human-based procedures are not yet completely substituted. Examples of ground systems used for inspection are wheeled robots [6] and legged robots [7]. In case of inspection of vertical surfaces, wall-climbing robots were developed using magnetic devices [8] or using vacuum suction techniques [9]. Recently, unmanned aerial vehicles (UAVs), equipped with high definition photo and video cameras to get high-quality data, have shown a great potentialities in inspection applications [10, 11, 12]. Although they can significantly enhance infrastructure inspections and provide data to feed digital twin models [13], robotic systems are not yet fully embraced. The main reason is that algorithms are computationally intensive and require a significant computing capacity on hardware platforms that are subjected to strict limits on size, weight and power. For these reasons energy-efficient heterogeneous HPC architectures equipped with many-core processing units have been considered in order to speed up mathematical computations inherent to signal and image processing, motion planning etc [14].

Most of the infrastructure and civil structures are made by concrete, steel and masonry, which are prone to cracks due to deterioration of reinforcements. Crack information (e.g., the number of cracks and crack width and length) represents the current structural health indicators, which can be used for the proper maintenance to improve the structural safety [15]. Nowadays damages in buildings and bridges can be easily captured using a commercial digital camera and consequently analyzed and classified by image processing algorithms, but the detection of fractures is still challenging in image processing. The main reasons are that they have a complex topology, a thickness similar to the image resolution and are easily corrupted by noise [16].

Amongst the most widely used techniques there are those based on color detection and neural networks. In [17] a comparative analysis is proposed among different color spaces to evaluate the performance of color image

segmentation using an automatic object image extraction technique. In [18] an Red Green Blue (RGB) based image processing technique was proposed for rapid and automated crack detection. Recently, an algorithm based on the Convolutional Neural Network (CNN) was considered in [19] to detect concrete cracks without calculating the defect features [20], [21]. Furthermore, a modified architecture of Fusion Convolutional Neural Network to handle the multilevel convolutional features of the sub-images is proposed and employed in [22] for crack identification in steel box girders containing complicated disturbing background and handwriting. Even though these techniques allow fast processing they are not suitable for real-world applications because they are affected by a significant false-positive rate due to a high sensitivity to environmental light conditions and noise [3].

In this paper we propose a mathematical method for detecting cracks in images that is more robust against all these factors. Aiming both at automatizing the whole process and at reducing its execution time a parallel implementation on heterogeneous HPC architectures is provided.

2 Mathematical model and parallel implementation

Variational methods have addressed successfully problems such as image segmentation and edge detection. They proposed a minimizer of a global energy as a solution. A first example is described by Mumford and Shah (MS) in their paper [23] where they proposed a first order functional, whose minimization determines an approximation of the image by means of a piecewise smooth function and detects edges as singularities in the image intensity. However, this model is not suitable for cracks because they do not represent singularities in the intensity function, but in its gradient instead. For this reason, we propose a second order variational model based on the Blake-Zissermann (BZ) functional [24]. This was introduced with the aim of overcoming some limitations of the MS approach, such as the over segmentation and the lack in detecting gradient discontinuities. Being the original formulation not suitable for numerical treatment, we focused on a different approach that is based on the approximation proposed by Ambrosio and Tortorelli (AT) for the MS functional [25, 26]. In their model, they replaced the unknown discontinuity set by an auxiliary function which smoothly approximates its indicator function. In our case two auxiliary functions are introduced as indicators of both intensity discontinuity and gradient discontinuity sets. As numerical minimization algorithm we chose an “inexact” block-coordinate descent scheme (BCD) to address the heterogeneous computing environment. In order to process very large images a tiling scheme is adopted: the minimizer is assembled by merging together local solutions restricted

to portion of images.

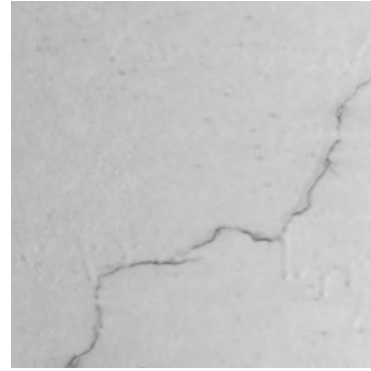


Figure 1. Crack on a concrete wall

3 Results

We tried our method on images of cracks taken in tunnels in Greece (Fig. 1) aiming at reconstructing the whole structure avoiding the effect of the noise and the environmental conditions (i.e. light conditions). We compared the results with the state-of-art technique based on mathematical morphology [16].



Figure 2. Our method

While in Fig. 3 the structure is broken in several points, our reconstruction is closer to the original one (Fig. 2). In order to reduce the execution time and to provide a automatic procedure we tested a sequential implementation with a parallel one based on the OpenMP framework that implements a strategy for collaboratively executing a program on an environment composed by devices of different types (aka heterogeneous architectures). Both the versions were executed on a High Performance Computing (HPC) cluster equipped with x86-64 processors, running a CentOS 7.6 operating system. The table 3 shows the execution times for Fig. 1. Overall, the parallel version is significantly more efficient with respect to the sequential one.

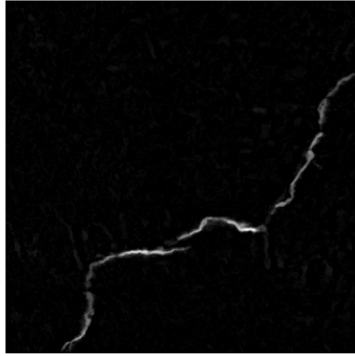


Figure 3. State-of-art based on mathematical morphology

algorithm	time (s)
sequential	13.184868
parallel 24 cores	1.056701
parallel 48 cores	0.5915374

Table 1. Run time comparisons for a single image

4 Conclusions

In this paper we proposed an automatic procedure for detecting cracks in images. This is based on a variational method and its parallel implementation on heterogeneous HPC architectures. We got promising results for both the quality of the reconstruction and its execution time. As future plan, we would like to test our procedure on real world scenarios in order to understand if it could be used as useful tool for assessing civil structures. In order to reduce further the run time we plan to test the execution on several domain specific hardware accelerators. Moreover, we would like to address different structures for detecting new kind of damages.

References

- [1] Gian Michele Calvi, Matteo Moratti, and Gerard J. O'Reilly. Once upon a time in Italy the tale of the morandi bridge. *Structural Engineering International*, 29(2):198–217, 2019.
- [2] H. Kim, Sung-Han Sim, and S. Cho. Unmanned aerial vehicle (uav)-powered concrete crack detection based on digital image processing. In *6th International Conference on Advances in Experimental Structural Engineering*. 11th International Work-
- shop on Advanced Smart Materials and Smart Structures Technology, 2015.
- [3] Eftychios Protopapadakis et al. Automatic crack detection for tunnel inspection using deep learning and heuristic image post-processing. *Applied Intelligence*, pages 1–14, 2019.
- [4] Dongho Kang and Young-Jin Cha. Autonomous uavs for structural health monitoring using deep learning and an ultrasonic beacon system with geo-tagging. *Computer-Aided Civil and Infrastructure Engineering*, 33(10):885–902, 2018.
- [5] Ronny Salim Lim, Hung Manh La, and Weihua Sheng. A robotic crack inspection and mapping system for bridge deck maintenance. *IEEE Transactions on Automation Science and Engineering*, 11(2):367–378, 2014.
- [6] Shigeo Hirose and Hiroshi Tsutsumitake. Disk rover: A wall-climbing robot using permanent. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. Vol. 3. IEEE, 1992.
- [7] Claudio Semini, Nikos G Tsagarakis, Emanuele Guglielmino, Michele Focchi, Ferdinando Cannella, and Darwin G Caldwell. Design of hyq—a hydraulically and electrically actuated quadruped robot. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 225(6):831–849, 2011.
- [8] Lin Guo, Kevin Rogers, and Robin Kirkham. A climbing robot with continuous motion. In *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*. IEEE, 1994.
- [9] J. Savall, A. Avello, and L. Briones. Two compact robots for remote inspection of hazardous areas in nuclear power plants. In Michigan Detroit, editor, *Proceedings of IEEE international conference on robotics and automation*, 1999.
- [10] Norman Hallermann and Guido Morgenthal. Visual inspection strategies for large bridges using unmanned aerial vehicles (uav). In International Conference on Bridge Maintenance, editor, *Proc. of 7th IABMAS*. Safety and Management, 2014.
- [11] Fabio Ruggiero, Vincenzo Lippiello, and Anibal Ollero. Aerial manipulation: A literature review. *IEEE Robotics and Automation Letters*, 3(3):1957–1964, 2018.
- [12] Barrie Dams, Sina Sareh, Ketao Zhang, Paul Shepherd, Mirko Kovac, and Richard J Ball. Aerial

- additive building manufacturing: three-dimensional printing of polymer structures using drones. *Proceedings of the Institution of Civil Engineer*, 173(1): 3–14, 2020.
- [13] Rafael Sacks, Ioannis Brilakis, and Mark Girolami. Construction with digital twin information systems. *Data-Centric Engineering*, 1, 2020.
- [14] Joseph Ortiz, Mark Pupilli, Stefan Leutenegger, and Andrew J. Davison. Bundle adjustment on a graph processor. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.
- [15] Yufei Liu et al. Automated assessment of cracks on concrete surfaces using adaptive digital image processing. *Smart Structures and Systems*, 14(4): 719–741, 2014.
- [16] Hugues Talbot. *Oriented patterns in image analysis*. PhD thesis, Université Paris Est, December 2013.
- [17] Dina Khattab et al. Color image segmentation based on different color space models using automatic grabcut. *The Scientific World Journal*, 2014, 2014.
- [18] HweeKwon Jung, ChangWon Lee, and Gyuhae Park. Fast and non-invasive surface crack detection of press panels using image processing. *Procedia Engineering*, 188:72–79, 2017.
- [19] Young-Jin Cha and Wooram Choi. Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 32(5):361–378, 2017.
- [20] Young-Jin Cha et al. Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types. *Computer-Aided Civil and Infrastructure Engineering*, 33(9): 731–747, 2018.
- [21] Yi-zhou Lin, Zhen-hua Nie, and Hong-wei Ma. Structural damage detection with automatic feature-extraction through deep learning. *Computer-Aided Civil and Infrastructure Engineering*, 32(12):1025–1046, 2017.
- [22] Yang Xu et al. Surface fatigue crack identification in steel box girder of bridges by a deep fusion convolutional neural network based on consumer-grade camera images. *Structural Health Monitoring*, 18 (3):653–674, 2019.
- [23] David Mumford and Jayant Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on pure and applied mathematics*, 42(5):577–685, 1989.
- [24] Andrew Blake and Andrew Zisserman. *Visual reconstruction*. MIT press, 1987.
- [25] L. Ambrosio-VM. Tortorelli. Approximation of functionals depending on jumps by elliptic functionals via gamma-convergence. *Comm. Pure Appl. Math*, 43: 999–1036, 1990.
- [26] Luigi Ambrosio and Vincenzo Tortorelli. On the approximation of free discontinuity problems. *∴*, pages 105–123, 1992.

A Case Study: Conception of digitalizing prefabrication processes in the construction industry

Z. Cai¹, M.B. Tarfaoui², S. Kessler³ and J. Fottner⁴

¹School of Engineering and Design, Technical University of Munich, Germany

²Department of Informatics, Technical University of Munich, Germany

³School of Engineering and Design, Technical University of Munich, Germany

⁴School of Engineering and Design, Technical University of Munich, Germany

zhen.cai@tum.de, mohamed.tarfaoui@tum.de, stephan.kessler@tum.de, j.fottner@tum.de

Abstract -

During the last few years, the importance of prefabrication and its digitalization has grown significantly in the construction industry. In this context, a German construction company intended to extend an existing warehouse application into a system that additionally monitors production time to identify the bottleneck. In this case study, a concept of digitalizing the prefabrication was proposed with the components of hardware, middleware and software. In terms of hardware components, an RFID system was tested in the production plant and a feasibility study of a Bluetooth system was done. In the software concept, a new framework was designed and implemented with new functions, e.g. introducing a resolver layer to reduce controller workload. Based on this concept, the company is able to utilize data from both production and warehouse.

Keywords -

Prefabrication; Modular Housing; Digitalization; RFID; Bluetooth

1 Introduction

In the construction industry, prefabrication is widely recognized as an important method to enhance industrialization progress [1]. Prefabrication enables an off-site production of construction elements, from simple construction components to complex building modules [2]. Not only can the prefabrication improve the level of automation, but also reduce energy consumption and environmental pollution [3]. In this case study, the authors have cooperated with the German company maxmodul, which plays a leading role in the prefabricated house section with its solid construction.

In many manufacturing industries, the production processes are highly digitalized and automated by systems such as Manufacturing Execution System (MES) and Warehouse Management System (WMS) [4, 5]. In comparison, few construction companies have applied those systems in their prefabrication plants because of the lack

of research and practical applications in the construction industry [6]. RIB iTwo MES [7] is one of the very few MES software in the construction industry but only supports limited processes. To overcome the gap, we propose in this paper a concept to digitalize the prefabrication processes based on a case study in company maxmodul. In the next section, we explain the current situation, and requirements from maxmodul, which leads to the designed concept. After that, both hardware and software are implemented and the test results are summarized. Finally, the results are evaluated and future work is proposed.

2 Initial Situation and User Requirements

2.1 Initial Situation

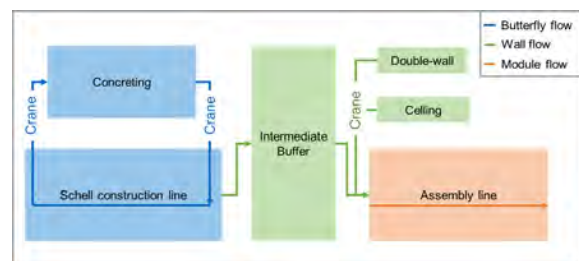


Figure 1. Material flow in the prefabrication plant maxmodul[8]

In the **production floor**, a material flow simulation was conducted by Fischer et al. [8]. The production follows the principle of mixed-model Assembly. The modular housing elements go firstly through the concreting stations and then enter the buffer areas, waiting for assembly. Two elements (double-wall and ceiling) are produced in separate stations and transported with cranes, as shown in Figure 1.

After the assembly line, the products (modules) are stored temporarily in an outdoor **block warehouse**. A portal crane is responsible for the transportation of modules. On the portal crane, an RFID reader and a GPS device

are installed. On each module, an RFID tag with the module number is attached. The reader on the portal crane records the number of the module and the GPS data every time it changes the position of the modules. The data are displayed and stored in internally developed software, which has some technical flaws due to its outdated architecture.

2.2 User Requirements

After interviewing the users of maxmodul company, the most urgent requirements are selected as follows:

- Digitalizing the production time: as like in any mixed-model assembly, the bottleneck of the production in maxmodul shifts from one station to another depending on their production time. The digitalization of the production time is essential for the company to identify the bottleneck and take measures in early time to avoid production stops.
- Refactoring and extending the software: the main goal is to refactor the software to a stage with robustness and efficiency. In addition, the output of the system focuses on the graphical display of the database entries in both table form and on the maps. Also, the communication with the connected MySQL Database should be optimised.

3 System Architecture

Based on the interviews with six users (two production manager, three workers, one IT expert), a system architecture consisting of the following components is proposed, also shown in the Figure 2:

- Hardware components: the data reader and a gateway ensure the sending information from the physical system to the intranet/internet.
- Middleware: it is a centralised repository that collects data from various sources. It allows persistent data storage and fulfils complex data processing.
- Software components: it provides the users with different functionalities, such as monitoring and planning tools.

Within this architecture, different digitalization scenarios can be realized. However, the scope of implementation in the next chapter is limited to the hardware and software components, reflecting the user requirements.



Figure 2. System architecture based on a digital twin concept from Altexsoft [9]

4 Implementation and Results

4.1 Hardware Components

4.1.1 System concept and installation

To record the production time, two systems were tested in this case study, as shown in Figure 3.:



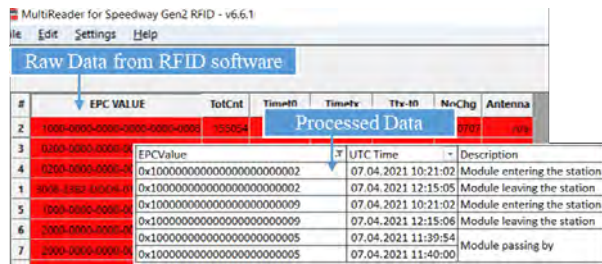
Figure 3. Installation of the two hardware systems

- RFID system: the RFID tags (Omni-ID Exo 750) were tagged on the product, and the RFID antenna (Impinj) stood nearby. Through the antenna, the time of a product entering and leaving the production station was recorded.
- Bluetooth system: the Bluetooth beacons were attached to the scaffold of each station. A Bluetooth reader was attached to the crane. Each time the crane entered or left the area of one working station, the time was recorded. The system of MotionMiners was used in this paper.

Due to the COVID constraints, the Bluetooth experiments were carried out as a feasibility study in the testing facility at the Technical University of Munich.

4.1.2 Test Results

Because of its high equipment cost, in this case study, only the double-wall station was tracked with the RFID system to test the production time as an example. The



MultiReader for Speedway Gen2 RFID - v6.6.1

Raw Data from RFID software

#	EPC VALUE	TotCnt	TimeIn	TimeOut	Thx-In	NoChg	Antenna
2	1000-0000-0000-0000-0000	333004					
3	0260-0000-0000-0000-0000						
4	0260-0000-0000-0000-0000						
1	0000-1000-0000-0000-0000						
5	1000-0000-0000-0000-0000						
6	0000-0000-0000-0000-0000						
7	0000-0000-0000-0000-0000						

Processed Data

#	EPCValue	UTC Time	Description
3	0x10000000000000000000000000000002	07.04.2021 10:21:02	Module entering the station
4	0x10000000000000000000000000000002	07.04.2021 12:15:05	Module leaving the station
5	0x10000000000000000000000000000009	07.04.2021 10:21:02	Module entering the station
6	0x10000000000000000000000000000009	07.04.2021 12:15:06	Module leaving the station
7	0x10000000000000000000000000000005	07.04.2021 11:39:54	Module passing by
7	0x10000000000000000000000000000005	07.04.2021 11:40:00	Module passing by

Figure 4. Data samples from RFID system

tag number and timestamp of the double-walls production were recorded, as shown in Figure 4. During the two weeks of testing in the production plant, ten modules were recorded. After processing the data, e.g. filtering the false recorded data, four production time data points are considered plausible by comparing the statement from workers.

For Bluetooth systems, the reader on the crane was able to recognise the RSSI and timestamped Bluetooth signal from the beacons attached to the scaffold. Moreover, the data was automatically transmitted and pre-filtered for outliers. The feasibility of the system has therefore been proven and field tests will be planned at the production plant to obtain real data.

4.2 Software Components

4.2.1 Software Design

The goal from the software aspect is to refactor the framework of the self-developed software and bring it up to date while extending the system for the production floor.

To develop the software components, a detailed requirements analysis is conducted with the users mentioned above. As a result, the requirements are divided into functional and nonfunctional requirements as follows:

- Functional requirements: items table view, warehouse map view, process map view.
- Nonfunctional requirements: data redundancy, portability, usability.

The requirements analysis is visualised by a top-level UML use case model in Figure 5. The red use cases are developed in this case study, whereas the blue use cases already exist in the old software.

To meet the **functional requirements**, the software framework Symfony is used. It consists of a PHP web application framework and a collection of PHP components and libraries. The architectural pattern is the model-view-controller, which splits a software application into three principal logical parts: the model, the view and the controller part. Each of those parts is designed to deal with implementational aspects of the system.

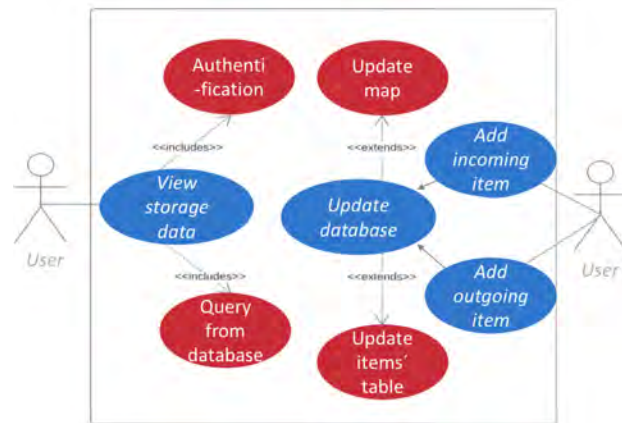


Figure 5. Use cases for the maxmodul application

Furthermore, Twig is used as a template engine and Doctrine is used for object-relational mapping and enabling complex queries to use the object-oriented paradigm of a selected programming language. Among its main features, the ability to query databases using DQL (an object-oriented SQL dialect) is crucial.

To avoid dependency issues and fulfil the **nonfunctional requirements** of portability, the Docker container is used in this project. Docker container creates a ready configured and isolated environment so that any user can run the application easily. In this project, three images are constructed in the container:

- MYSQL: with a mounted volume for data persistence.
- PHP-FPM: with a mounted volume for the application's code.
- NGINX: with mounted volumes for configurations, logs and share mounted volume with PHP-FPM for the application's assets.

4.2.2 Software Implementation

Based on the proposed software design, the initial software is divided into four major artifacts:

- Core: it contains all the business logic, data models and event publishers and subscribers.
- Web: it contains UI elements, their controllers and resolvers, e.g. the map service view or item view.
- API: it contains exposed RESTFUL controllers, API resolvers and object serializers that are still in development.
- Common: it contains parent classes and reusable components for all the elements e.g. controllers,

repositories and resolvers, which reduces the redundancy of codes.

The issues of logic duplication in controllers from the old software are solved with the new structure. Another important change is the introduction of a resolver layer. It handles CRUD and data fetching/providing operations and reduces the workload of controllers. This was initially carried out by the controller in the former software.

4.2.3 Test Results

Since PHP is an interpreted language, it's unable to see the bugs until the program is executed. For this reason, PHPStan is integrated into the project for static code analysis. PHPStan focuses on finding errors in the code without actually running it. It catches whole classes of bugs before writing tests for the code and detects mistakes in the source code, e.g. calling an undefined method or passing a wrong number of arguments to a function. Also, PHP CS is integrated to detect and fix the violations of pre-defined coding standards in the existing code.

After the refactoring process, the software is extended respectively to the requirements mentioned in 4.2. Concerning the functional requirements, the three requirements in Figure 5 have been satisfied, as the maxmodul application now offers all the required tables and map views. Concerning the nonfunctional requirements, the performance and usability criteria have been satisfied since the application displays all the required information.

5 Discussion and Outlook

In this case study, a concept of digitalization in a prefabrication plant is proposed. From the hardware aspect, RFID and Bluetooth systems are tested. The advantage of the RFID system is that it can be easily integrated but the shortage is the high equipment cost in comparison to the Bluetooth system. In addition, the Bluetooth system enables automatic data transmission and pre-processing. For this reason, the Bluetooth system is preferred and future tests will be based on this system.

From the software aspect, the refactored software is fully functional and has met the user requirements. It is handed over to the company for further test. However, other functions, like APIs for data exchange, should be integrated into future work. On the hardware side, the Bluetooth system should be tested for its performance in the company in the future.

References

- [1] Bing Qi, Mohamad Razkenari, Aaron Costin, Charles Kibert, and Meiqing Fu. A systematic review of emerging technologies in industrialized construction. *Journal of Building Engineering*, 39:102265, 2021. ISSN 23527102. doi:10.1016/j.jobbe.2021.102265.
- [2] Nassim Sebaibi and Mohamed Boutouil. Reducing energy consumption of prefabricated building elements and lowering the environmental impact of concrete. *Engineering Structures*, 213:110594, 2020. ISSN 01410296. doi:10.1016/j.engstruct.2020.110594.
- [3] Weisheng Lu, Wendy M.W. Lee, Fan Xue, and Jinying Xu. Revisiting the effects of prefabrication on construction waste minimization: A quantitative study using bigger data. *Resources, Conservation and Recycling*, 2021.
- [4] Hans-Hermann Wiendahl: Andreas Kluth and Rolf Kipp. Marktspiegel business software mes – fertigungssteuerung 2019/2020, 2020. URL <https://www.ipa.fraunhofer.de/de/Publikationen/studien/studie-mes-marktspiegel-20192020.html>.
- [5] R. Caridade, T. Pereira, L. Pinto Ferreira, and F.J.G. Silva. Analysis and optimisation of a logistic warehouse in the automotive industry. *Procedia Manufacturing*, 13:1096–1103, 2017. ISSN 23519789. doi:10.1016/j.promfg.2017.09.170.
- [6] Yanwu Xiao and Jyoti Bhola. Design and optimization of prefabricated building system based on bim technology. *International Journal of System Assurance Engineering and Management*, 2021. ISSN 0975-6809. doi:10.1007/s13198-021-01288-4.
- [7] RIB iTWO MES. itwo smart production: Intelligente softwarelösungen für die vorfertigung und modulares bauen, 2022. URL <https://www.rib-saa.com/de/>.
- [8] Anne Fischer, Jorge Rodriguez Llorens, Zhen Cai, Michael Wilke, Stephan Kessler, and Johannes Fottnier. Implementation of a digital twin framework in the modular housing industry. *28th IEEE ICE/ITMC & 31st IAMOT Conference IEEE*, 2022.
- [9] Altexsoft. Digital twins: Components, use cases, and implementation tips., 2021.

On the Bar Installation Order for the Automated Fabrication of Rebar Cages

M. Momeni¹, J. Relefors², L. Pettersson³, A.V. Papadopoulos¹, and T. Nolte¹

¹Mälardalen University, Västerås, Sweden

²Solving Robotics Sweden AB, Västerås, Sweden

³Skanska Sweden AB, Stockholm, Sweden

mahdi.momeni@mdu.se

Abstract -

Robotics automation is a promising solution for the fabrication of structures made out of reinforced concrete. The reinforcement is often installed directly in the form and bar-by-bar. Using bigger pre-fabricated units (cages) may be beneficial for saving construction time and better labor safety. In this paper, we focus on the problem of automating the generation of a plan for the installation of rebars, given the digital twin of a desired reinforcement cage design, and of its basic components. More specifically, the plan describes the assembling order for the rebars such that (i) it is possible to fabricate the reinforcement cage by the robots, and (ii) the end product is the final reinforcement cage, ready for installation in the form for the concrete structure. In this paper, we propose an algorithm to automatically compute a feasible installation order for a generic rebar cage. The feasibility of the generated order is also case studied and simulated on a simplified rebar cage under the given assumptions.

Keywords -

Robotics for construction; Rebar installation; Automation

1 Introduction

The process of installing reinforcement bar-by-bar in reinforced concrete structures is time-consuming and potentially harmful to the workers' safety. Pre-fabricating the reinforcement in bigger units (often referred to as "cages") and placing them in the form may help save construction time. Automatic fabrication of rebar cages can improve the construction time even further, and introduce other benefits, e.g., improved safety of construction workers.

One of the major challenges in automating rebar cage fabrication is that they are high mix and low volume. Automation cannot rely on repeating the same motions over and over, but a higher degree of flexibility is required. This challenge can be divided into two major challenges. First, the adopted industrial robots require accurate motion planning that cannot be easily reused between one cage and another. Second, the ordering sequence of one cage can significantly differ from another one. This calls for novel algorithmic approaches that can handle such flexibility.

In [1], we explored how the first challenge can be addressed, by automatically generating motions for a gantry-robot setup with 27 degrees-of-freedom (DOFs), shown



Figure 1. Gantry robot system.

in Fig. 1. However, the installation order was manually determined for the specific cage used in the experimental demonstration, therefore limiting the possibility of reusing the same solution for different types of cages.

In this paper, we focus on the problem of identifying a feasible installation order, given a generic rebar cage. "Installation order" refers to the determination of how and in what sequence the rebars should be placed to assemble the cage. If no such *valid order* exists, or if one cannot be found, the cage is considered not *robotically fabricable*. A valid order should take into consideration the gripping as well as tying poses while ensuring that collision-free paths for placing the rebars exist. Moreover, it should ensure that the rebar cage remains stable during fabrication and at the end of the process. Several valid installation orders may exist but in this context finding one such order is enough.

2 Background

In [2], the authors used Constraint Satisfaction Problem (CSP) to formulate sequence and motion planning problems (SAMP) for spatial extrusion of 3D trusses. The authors present an automated approach to finding an extrusion sequence. The sequence in their work refers to the order of the end-effector's feasible direction for each extrusion direction. In this paper, however, the sequence refers to the order of placing rebars one after another.

In [3], the authors present an automated planning ap-

proach to find a construction sequence and plan robot motion jointly for additive manufacturing. In their work, the construction sequence is defined as an ordering of the elements which are welded or glued together at their ending. Besides, their work deals only with straight elements. In this work, the considered rebars can be connected at any point on the bar and the rebars can have different shapes.

In [4], the authors presented a fabrication process method based on a cooperative assembly approach. They use two robots to cooperatively fabricate a full-scale vault brick-by-brick with identical bricks. The construction sequence is designed taking into account the temporary stability of neighboring brick assembly while maintaining global stability. Their approach is developed based on the principle that the added self-weight should be efficiently transferred to the foundation. In our work, the rebar cage is built using a gantry robot system (shown in Fig. 1). In addition to that, a rebar cage in our problem can contain hundreds of rebars, many of them of different types.

The problem of finding the bar installation order is split into two sub-problems: (i) finding a way to deal with a large number of ordering possibilities, which is a combinatorial problem, and (ii) finding gripping and tying poses, as well as the robots' trajectories. More specifically, we analyze a rebar cage structure to reduce the complexity of the problem by reducing the search space. To this end, we order the rebars based on the geometry of the rebars and the way that the path planning algorithm presented in [1] works. We test this ordering on a simplified cage. The input used is a digital twin of the cage, i.e., a digital version of the cage where all necessary bars are present. More specifically, the digital twin contains the position and geometry of all the bars in the cage. Each bar has a unique ID which is defined using information such as: (i) the bar mark which contains information regarding the bar shape type and a serial number which identifies a group of identical bars; (ii) a serial number identifying an individual bar in a group of identical bars; and (iii) the way the bar should be fixed to other bars, e.g., welding or tying.

The remainder of this paper is organized as follows. Section 3 presents the problem addressed in this paper. Section 4 describes the methodology developed to compute the bar installation order. Section 5 discusses the obtained results, and Section 6 concludes the paper.

3 Problem Statement

The problem of finding a rebar placement sequence from a digital twin can be formulated as follows:

Given the digital twin of a rebar cage and an automation system, find a feasible order in which the bars can be installed by the system, one after another, to fabricate the entire rebar cage.

Finding such a sequence is far from being trivial, due

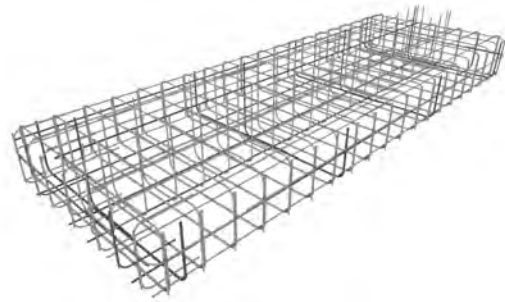


Figure 2. The digital twin of a cage.

to its combinatorial nature [5]. The main focus of this work is to investigate ways of reducing such complexity by exploiting the nature of the problem and the domain knowledge and expertise. In particular, in this paper, we focus on finding an installation order for a single cage while using rebar trajectories that are compliant with the trajectories of the path planning algorithm proposed in [1].

We are then looking for an installation order for a cage with the following constraints and assumptions:

1. The rebars are placed following vertical movements complying with the path planning algorithm in [1].
2. The rebar should remain connected to, at least, one other rebar from either the same or another layer at each installation step, except for the rebar(s) in the first layer which is(are) assumed to rest on the ground and may not connect to another rebar (s).
3. The rebars are stiff, i.e., their geometry corresponds to the geometry given by the digital twin.

In theory, the number of installation orders is equal to the factorial of the number of rebars in each cage. Rebars enumeration to solve the problem is therefore not possible. We are interested in finding one feasible solution and thus, orders evaluation can be stopped as soon as a solution is found. This suggests using a backtracking and Depth-First Search (DFS) [6] algorithm to solve the problem.

To further analyze the structure of the problem, we look at all states that the cage can assume during fabrication, denoting this as the *states of a cage*. A key observation is that the same state of the cage can be part of different installation orders. The history of how the cage came to a state is not important. Saving the computations following a certain state of the cage means that the computations following from that state do not have to be performed again if another order leads to the same state of the cage.

Summarizing the analysis, the problem is well suited for a DFS algorithm. Furthermore, saving any computations made during the search to be re-used is important since the same state of a cage can be visited more than once. Finally,

Algorithm 1: Forward Method Algorithm.

input : DT = Digital Twin
output : R = Placement Sequence

```

1  $Rebar \leftarrow RebarData(DT)$ ; /* Determines
   rebar parameters: ID, type, and
   position w.r.t the system coordinate
   frame */
2  $[s, Z] \leftarrow Layers(Rebar)$ ; /*  $s$ : Number of
   layers,  $Z$ : HLs coordinates, each layer
   includes HLs with the same
    $z$ -coordinates */
3 for  $j \leftarrow 1$  to  $s$  do
   // find all the rebars with at least
   one Horizontal Leg in layer  $j$ 
4    $tmp \leftarrow RebarsInEachLayer(Z(j))$ ;
   // sort rebars based on their "x-",
   "y-" and "z-" coordinates
5    $SortedRebars(j) \leftarrow SortRebars(tmp)$ ;
   /* from left to right */
6 end
7  $R \leftarrow SortedRebars$ 

```

the criteria of complying with the path planning algorithm mean that the DFS should be guided in a direction where blocking yet-to-be-placed rebars are avoided.

4 Methodology

In this section, we present a heuristic to compute the bar installation order that we call the *Forward Method*, presented in Algorithm 1. In this method, rebars are added one by one starting from scratch. This can be compared with the *Backward Method*, presented in [7], which is based on the main principle of disassembling the cage.

As motivated in the previous section we are going to use a depth-first search to find an installation order. The search needs to be guided in a way that increases the likelihood of finding a feasible installation order. This means that the path planning algorithm has to be taken into account.

The planning algorithm is designed to place rebars by moving them straight down, followed by an approach movement. Looking at a cage from above, the horizontal rebar legs have a larger surface area than the vertical rebar legs. We interpret this to mean that horizontal legs in a partially assembled cage are more likely to block the installation of further rebars. This is especially true for horizontal legs which are at the top of the cage.

To use this idea to guide the search algorithm we assign a height coordinate to each rebar, denoted as the Z -coordinate of the rebar, as the height of the top horizontal leg of the rebar. The height is defined with respect to the

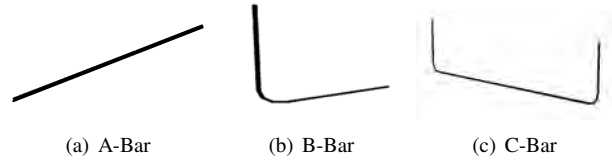


Figure 3. Different rebar types used in this paper.

system coordinate frame which can be arbitrarily placed anywhere on the floor. We then initialize the search algorithm with a list of rebars where the order is from lower to higher Z -coordinate. The list is then used in the search by adding rebars to the cage in the order in which they show up in the list.

There may be, and often is, many rebars with the same Z -coordinate in a cage. We denote collections of rebars with the same Z -coordinate as being in the same *layer*. In a given layer the order in which the rebars should be placed is not guided by the Z -coordinate. This freedom is used by ordering the rebars in a given layer by their distance to the origin of the system coordinate.

The inclusion of an arbitrary origin may seem strange at first. The rationale is that starting from some point in space should maximize the size of free space around the rebars in a partially constructed cage. Note, however, that changing the position of the origin can affect both the installation order and the convergence of the search.

5 Results and Discussion

The pseudo-code of the Forward Method algorithm is explained in Algorithm 1. We used MATLAB[®] to implement the algorithm. To test the algorithm, we have chosen the digital twin of a reinforcement cage, see Fig. 2, which was designed for a bridge structure in Stockholm, Sweden. For the sake of this paper, we simplified the cage to the one with fewer rebars (from 198 bars in the original one down to 44 bars), see Fig. 4. Noting that there might be a loss of generality with any simplification, this simplification is done for the digital twin to be imported to MATLAB[®] for a faster prototyping phase. For the final product, however, the software we are working on is being developed in C++ and will take the original digital twin directly as the input.

The simplified cage contains three types of bars, shown in Fig. 3. For this case study, the algorithm can find a sequence without backtracking in the depth-first search. An order is also produced regardless of the position of the coordinate frame. An example can be found in [8]. There are overall 8 different layers in this case study, as shown in in Fig. 4. which means that the cage will be fabricated in 8 different stages. The required number of robot's translational movement to fabricate the cage, however, is equal to the number of rebars in a cage.

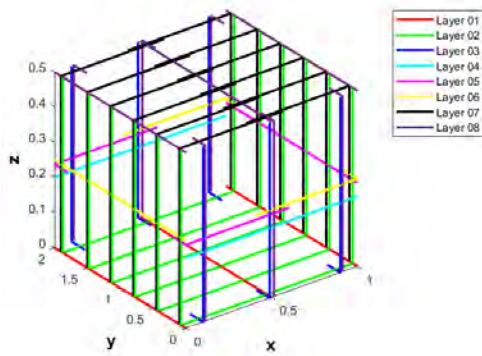


Figure 4. The simplified model of the cage in Fig. 2.



Figure 5. A row of 6 bars of two different types.

Unlike the presented case study, we can find cases where the position of the origin affects the number of branches that need to be explored in the search. An example is given by the collection of rebars shown in Fig. 5 which contains 6 rebars. Depending on whether the origin is set to O1 or O2 the number of branches that need to be explored and backtracking differs.

6 Conclusion

Finding a valid installation order so that the robots can automatically fabricate a reinforcement rebar cage is a challenging task. Especially if we take into account the gantry robot's trajectory to install the rebars as well as all the many different rebar types. We, therefore, tried to split the problem into two subproblems of (i) finding a sequence(s) of rebar placement as an initial starting point, and (ii) verifying that the determined sequences are executable by gantry-robot systems. We addressed the first part of the problem in this paper and presented an algorithm to find a sequence(s) for placing the rebars in a cage.

Theoretically, the problem of finding the bar installation order is combinatorial with respect to the number of rebars in the cage. This paper proposes a heuristic to be able to compute a feasible solution. The executability of the solution has been simulated on a simplified cage. The solution, however, may not necessarily be the fastest possible solution to fabricate the cage as we are not aiming at an optimal solution at this stage but rather a feasible one.

The presented algorithm requires the rebars to be connected in each step of the assembly process. This is a relaxation of the cage's stability requirement in each step. This relaxation will be addressed in future work.

We also need to address the second part of the problem. One idea in that direction is to use the algorithm presented in this work to produce several different installation orders, possibly by moving the origin into different locations, and see if any of these orderings is fully compatible with the path planning used for the gantry-robot system.

Acknowledgment

This work was supported by the Knowledge Foundation with the Automation Region Research Academy (AR-RAY) and SBUF.

References

- [1] M. Momeni, J. Relefors, A. Khatry, L. Pettersson, A.V. Papadopoulos, and T. Nolte. Automated fabrication of reinforcement cages using a robotized production cell. *Automation in Construction*, 133:103990, 2022.
- [2] Y. Huang, C.R. Garrett, and C.T. Mueller. Automated sequence and motion planning for robotic spatial extrusion of 3D trusses. *Contr. Rob.*, 2:15–39, 2018.
- [3] Y. Huang, C.R. Garrett, I. Ting, S. Parascho, and C.T. Mueller. Robotic additive construction of bar structures: unified sequence and motion planning. *Contr. Rob.*, 5:115–130, 2021.
- [4] Stefana Parascho, Isla Xi Han, Samantha Walker, Alessandro Beghini, Edvard P. G. Bruun, and Sigrid Adriaenssens. Robotic vault: a cooperative robotic assembly method for brick vault construction. *Contr. Rob.*, 4:117–126, 2020.
- [5] L. Nägele, A. Hoffmann, A. Schierl, and W. Reif. LegoBot: Automated planning for coordinated multi-robot assembly of LEGO structures. In *IEEE/RSJ Int. Conf. Int. Robots and Syst. (IROS)*, pages 9088–9095, 2020.
- [6] S.J. Russell and P. Norvig. *Artificial intelligence: a modern approach*. Pearson Education Limited, 2016.
- [7] J. Relefors, M. Momeni, L. Pettersson, A.V. Papadopoulos, and T. Nolte. Installation order in automatic fabrication of reinforcement rebar cages. In *IEEE Conf. Emerging Tech. and Factory Aut. (ETFA)*, pages 1–4, 2021.
- [8] Simulation video. <http://www.idt.mdh.se/~aps01/research/robotics/2022-BarInstallationOrder.mp4>. Accessed: 2022-06-28.

A Taxonomy for Connected Autonomous Plant

C. Browne¹, R. Walker², T. Embley³, M. Akhtar⁴, A. Essa⁴, A. Pass⁴, S. Smith²
and A. Wright¹

¹Transport Research Laboratory (TRL), UK

²School of Engineering, University of Edinburgh, UK

³Costain, UK

⁴National Highways, UK

cbrowne@trl.co.uk, r.m.walker@ed.ac.uk

Abstract -

National Highways commissioned the development of a Roadmap for Connected and Autonomous Plant (CAP), which proposed a programme of activities which would aim to deliver the widespread deployment of CAP. A particular milestone activity identified as an early target within the Roadmap was the development of a taxonomy for understanding the capability of construction plant for operating without human involvement. This would provide a unified language to understand how plant can be used to achieve tasks with reduced or no human intervention. This paper presents an overview of the process used in developing a taxonomy to achieve this purpose, including the principles underlying the taxonomy, and the taxonomy itself. This builds on previous automation taxonomy work and applies it to the construction context and is further applied to two examples of autonomous compaction plant. It is concluded that the levels establish a unified language for the capability evaluation of automation of plant. This will support and catalyse the development of technology roadmaps amongst plant and technology manufacturers, enable procurement processes that incentivise the deployment of CAP within construction management, and support innovation practices by providing an understanding of the safety and operational implications of deploying automation on construction sites. It is also identified that the application of this taxonomy is not limited to the Construction environment.

Keywords -

Connected Autonomous Plant (CAP); Taxonomy; Automation; Autonomy; Plant

1 Introduction

The UK construction industry is undergoing transformational change, as it adopts digitised and automated processes to overcome the challenges facing the sector. Due to the importance of heavy machinery (colloquially known as "plant" within the UK) for construction, the use of Connected and Autonomous Plant (CAP) is of particular inter-

est to the industry, with new technologies being applied to a wide range of activities, such as geofencing of plant operation, the use of 3D machine control to meet the design requirements, remote collection of data for both design and as-built, semi-autonomous extraction and movement of materials, and the introduction of offsite and robotic construction methodologies.

However, the UK construction industry has not adopted a unified approach to this transformation, resulting in varying levels of deployment of CAP across sites, and poor information transfer between organisations. For example, the use of continuous compaction control has been a standard industry practice within mainland Europe for over 15 years but has not seen significant adoption in the UK until recent years. However, while some major projects (e.g., HS2) are implementing it, widespread adoption remains some years away [1]. To alleviate this, National Highways commissioned the development of a Roadmap for Connected and Autonomous Plant, [2]. Development of the Roadmap drew on the expertise of over 75 stakeholder organisations, through a series of questionnaires and workshops. This stakeholder engagement identified a number of barriers to the adoption of CAP including: a lack of a legislative framework that permits and facilitates the use of automation; the need for sufficient financial investment with appropriate recognition of the benefits achieved; contractual programmes which do not incentivise the use of CAP; and the difficulties in developing technology and connectivity across the wide range of plant used in the construction sector. To address these barriers the Roadmap proposes a programme of activities across 9 workstreams, which would aim to deliver the widespread deployment of CAP as milestones are achieved.

The Roadmap was jointly launched by National Highways and the Infrastructure Industry Innovation Partnership (i3P) in June 2020. The Roadmap estimates that, if the deployment of CAP within the UK construction sector can replicate the productivity and efficiency benefits that automation achieved in the manufacturing sector, then benefits of £200Bn could be achieved by 2040. However,

this requires that the steps outlined in the Roadmap are rapidly actioned – a delay of 5 years would see the 2040 savings reduced by over 50% as a result of delays in the deployment of technologies and innovations.

A key early milestone identified within the Roadmap is the development of capability levels and taxonomy for classifying the automated capabilities of construction plant. Standards and taxonomies are useful tools for driving innovation and development by specifying a vision of the future and a potential pathway to achieving it [3]. For example, the creation of a taxonomy provides a unified language for the industry to understand how plant developed by Original Equipment Manufacturers (OEMs) and technology retrofitted by third party developers can be used to achieve tasks with reduced or no human intervention. This facilitates the specification of development strategies, contracts, standards, and procurement strategies that can be readily understood across the industry, promoting a unified direction. To this end, National Highways commissioned the development of a taxonomy for the automated capability of plant.

2 Methodology

As automation has been developed and implemented in other industries, there has been the parallel development of supporting taxonomies so that each industry can understand their progress towards full automation. Hence, the first stage of the development process of a taxonomy for the automation of construction plant was an extensive literature review of other industries' taxonomies. This review considered the agriculture, aviation, manufacturing, maritime, military, mining, rail transport, road transport, and space sectors. The underlying principles and assumptions of each taxonomy were examined, and their applicability to other industries considered, so that best practice for the construction industry could be identified.

Following the review, an initial version of the taxonomy was drafted. This draft was subject to project team peer review throughout its development. Following the production of the final draft version, it was then reviewed through a series of workshops held with industry stakeholders from 32 organisations between the 30th of November and 2nd of December 2021. These stakeholders were drawn from all sectors that interact with plant throughout the construction process, from OEMs that design and develop plant, designers who create the designs that plant implement, procurement and contract writers who determine the types of plant used on site, site managers who control the deployment of plant within a construction site, and plant operators who physically use the plant. These stakeholders were drawn from the community which National Highways established during the development of the Roadmap. Stakeholders were invited to participate

through an online questionnaire which was distributed to the community and publicised through social media and word of mouth. Following this, the positive responses were analysed to understand what type of organisations they belonged to (OEM, client, end user, etc.). During the workshops the stakeholders were asked to rate (on a scale 1-5, where a higher score is more useful / easier to understand) the taxonomy on its utility (achieving a mean score of 3.9/5 and a standard deviation of 0.85) and how easy it is to understand (achieving a mean score of 4/5 and a standard deviation of 0.80). The feedback from these workshops was collated and used to refine the taxonomy. The taxonomy was launched in the UK at FutureWorx in March 2022, and is presented in the following sections.

2.1 Literature Review

The full content of the literature review is given in [4], which presents the contents of this paper in greater detail. Here we present a summary of the review, focusing on the taxonomies which are of greatest interest and relevance to the development of a taxonomy for construction plant.

A key early development of a taxonomy for automation was the 10-point scale classifying the automation of under-sea teleoperators, developed by [5]. In subsequent work [6] Parasuraman recognised that this taxonomy focused on decision selection and action implementation, and did not fully describe the capabilities of the human information processing system. It was therefore proposed that the human information processing system could be abstracted to four classes of functions which could each be automated to different degrees: 1. Information Acquisition, 2. Information Analysis, 3. Decision and Action Selection, and 4. Action Implementation. However, no detailed taxonomy was developed to implement this proposal until [7], who used the 4-stage human information processing system as the basis of a taxonomy to describe automation of Air Traffic Control centres. This taxonomy acted as the basis for a simplified version which was presented in [8].

In [9] Clough presented a taxonomy for the operation of Unmanned Aerial Vehicles (UAVs), drawing on a very similar classification of human behaviour to that proposed above. However, this drew on the Observe-Orient-Decide-Act principles that were developed by John R. Boyd for combat principles and are summarised in [10]. Notably, the central tenant of Clough's taxonomy is that "if building machines to replace human capability, they should be understood in the same way we understand human actions".

Perhaps the most widely known taxonomy has been developed by the Society of Automotive Engineers (SAE) for use in describing Connected Autonomous Vehicles (CAVs) operating on the road network[11]. The basis of this taxonomy is conceptually different to other taxonomies, defining a dedicated driving task (DDT) com-

prised of two aspects of object and event detection and response (OEDR), and the lateral and longitudinal control of vehicle motion. The different levels of automation are achieved by incrementally handing control over each aspect of the DDT to an automated system from the human. The SAE taxonomy has two extensions to the DDT for higher levels of automation – the responsibility for safe operation of the CAV and the Operational Design Domain (ODD), which define the constraints in which the CAV can safely drive without human intervention.

2.2 Basis for the Development of the Taxonomy

Construction presents unique challenges that make it particularly challenging for introducing automation. A large proportion of these challenges feature in the complex ODDs in which plant typically operate:

- Construction sites are of variable size and form which are typically not designed for plant to be there (in contrast to the road environment for CAVs).
- Construction sites are exposed to external uncontrollable factors.
- Operations are regularly concerned with modifying the environment in some way.

This is explored more fully in [4]. In addition to these challenges, the range of tasks which construction plant are expected to perform is significantly more complex than other industries and it is more typical for plant to be augmented with after-market systems which modify its performance in some way. As such, we have developed a more granular taxonomy to accommodate these differences.

In developing the taxonomy for the automated capability of plant, three aspects of the above taxonomies were identified and utilised. Firstly, we adopted the same central tenant as Clough and developed the taxonomy on the basis of the 4-stage human information processing loop (hereafter referred to as the Observe-Understand-Decide-Act (OUDA) loop), with some slight modifications to terminology for clarity. We refer to this as a loop because it happens continuously as we (or a machine) interact with our environment to achieve a particular task, see Figure 1. This task can be considered from a strategic level, which may feature a small number of these loops, or from a very detailed, immediate activity level where hundreds of these OUDA loops are occurring in quick succession (or indeed quasi-simultaneously). It is important to note that the 4 stages are an abstraction to aid in understanding the concept and may not be how an automated machine is implemented – some aspects may be combined or hard coded and some aspects may not be explicitly defined. However, the abstraction is conceptually useful for discussing how

different kinds of automation might be implemented, and the automated capabilities different systems might achieve. It is also the case that the validity of subsequent stages is dependant on information available in the previous stages and that subsequent stages cannot rely on more information than is available in the previous stage. The levels are also asymmetrical, meaning capability comparisons cannot be made between equally levelled stages, that is to say, a level 2 understand stage does not equate in automation to a level 2 act stage.

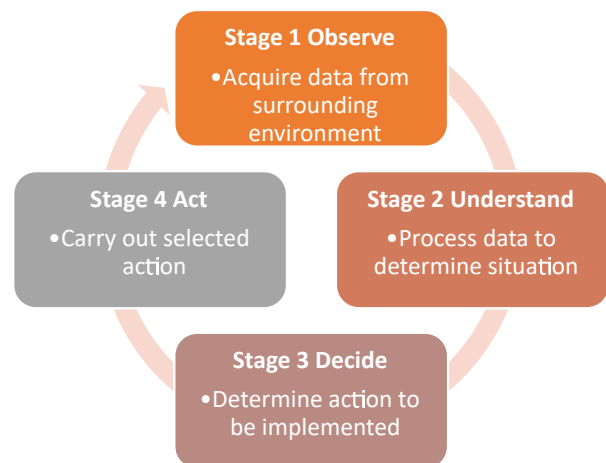


Figure 1. Stages of the human information processing loop [6, 10]

The second area included in the development of the taxonomy is the need for an explicit classification of Fallback and Responsibility. This is of critical importance to the construction industry, where plant is being operated in complex environments with a mix of other vehicles and pedestrian workers. A clear taxonomy for Fallback and Responsibility defines who or what is responsible for the safe operation of the plant and who or what is expected to respond to any unexpected incidents or changes to the operating conditions of the plant. This provides a clear understanding to plant operators and site managers about the safety conditions that need to be established so that plant of particular capability can be deployed on a given site.

A final area, which is important to consider when defining a taxonomy, is the influence of the Operational Design Domain (ODD) on how the taxonomy can be applied. In simple terms, the ODD defines the conditions, both physical and digital, in which the plant can perform at its expected level of automation. It is worth noting that the ODD is a property of both the automating system and the system being automated. That is, the computer controlling the machine has some conditions in which it can safely control said machine, as well as the machine itself having

some conditions in which it can operate, even if it was a human controlling it.

In the SAE taxonomy for CAVs operating on the road network there is no explicit discussion of ODD – it is either limited (i.e., there are some constraints on the system's ability to replicate human driving ability) or it is unlimited (i.e., there are no constraints on the system, and it can replicate or exceed a human's driving ability in all conditions in which the vehicle itself can operate). However, this simplification was felt to be inappropriate for the construction industry due to the highly complex nature of construction sites in comparison to the road network. Hence whilst CAVs are designed to operate in an environment which has been created for their operation, CAP will operate in environment which are not tailored for it. In [4] a more explicit discussion of ODD is included, alongside a detailed consideration of what physical and digital parameters should be included when defining the ODD of a particular piece of plant. The following sections describe the resulting taxonomy for automated plant developed in this work.

3 Taxonomy for the Automated Capability of Plant

3.1 Stage 1: Observe

Observing is the act of acquiring information (Observations) on the current situation from the surrounding environment through various sensing organs / sensors and through any existing communications channels. The taxonomy is presented in Table 1.

Table 1. Participation of human and system for each level of Automation for Stage 1: Observe

Level	Name	Sensor
0	No Automation	Human
1	Partial Automation	Human & System
2	Full Automation	System

3.2 Stage 2: Understand

After the observations on the surrounding environment and current situation are acquired through the Observe stage, this data must be processed to develop an understanding of the situation. To accommodate the different aspects of the Understand stage we consider three components, *Compare*, *Predict*, and *Learn*, defined as:

- *Compare*: Understanding the current state by comparing the Observations to existing values and thresholds.
- *Predict*: Understanding the future state through a pre-defined model against which the Observations are applied.
- *Learn*: Understanding the future state by learning from the outcomes of past Decisions and Actions and applying this.

We include both prediction and learning to provide the potential for operatives and systems to develop their skills. The taxonomy is presented in Table 2.

Table 2. Participation of human and system for each level of Automation for Stage 2: Understand

Level	Name	Compare	Predict	Learn
0	No Automation	Human	Human	Human
1	Automatic Comparison	Human & System	Human	Human
2	Automatic Prediction	Human & System	Human & System	Human
3	Full Automation	System	System	System

3.3 Stage 3: Decide

In the Decide Stage the outcomes of Understanding are used to develop a set of possible actions that could be carried out, and a Decision made on which action to select. Hence Decide contains three components – *Generate*, *Select*, and *Inform* - defined as:

- *Generate*: The creation of a set or list of possible actions based on the understanding of the situation (from the Understand stage).
- *Select*: The choice of one of the actions. This can be an unrestricted selection (pick any option, or even an option not presented as part of the Generation step, i.e., can go off-list - an Open List) or a restricted selection (choose an option that was presented as part of the Generation step, i.e., cannot go off-list - a Closed List).
- *Inform*: Provide awareness to the party responsible for approving and/or implementing the selected action.

It is important to note the generation of a list of actions is unlikely to be explicit in most implementations of an automated system, but it is a useful concept to differentiate the different levels of automation. The taxonomy is presented in Table 3.

Table 3. Participation of human and system for each level of Automation for Stage 3: Decide

Level	Name	Generate	Select	Inform
0	No Automation	Human	Human	Human
1	Open List	System	Human	Human
2	Closed List	System	Human	Human
3	Informed Selection	System	System	Human
4	Full Automation	System	System	System

3.4 Stage 4: Act

The selected Decision must be implemented through Action. To assist in understanding the roles of the human and the system at each level of automation within the Act stage, we consider two components:

- *Implementation*: The party which initiates and carries out the action selected at the Decide stage.
- *Monitoring*: The party which ensures that the action selected at the Decide stage is being carried out as intended.

The taxonomy is presented in Table 4.

Table 4. Participation of human and system for each level of Automation for Stage 4: Act

Level	Name	Implement	Monitor
0	No Automation	Human	None
1	Automated Guidance	Human	System Guides Human
2	Automated Intervention	Human	System Restricts Human
3	Supervised Automation	System	Human Monitors System
4	Full Automation	System	System

3.5 Fallback and Responsibility

The concept of Fallback and Responsibility underpins the information processing chain. We define these as:

- *Fallback*: Who or what ensures that, when the plant suffers from a component failure, encounters an unexpected situation, or leaves its Operational Design Domain, it either continues to operate or fails in a safe manner.
- *Responsibility*: Who or what ensures that the task is being carried out (either manually or automatically) in a safe and proficient manner to the desired quality.

It is possible for each component of the information processing chain to feature different levels of automated fallback and responsibility, which would create a complex classification system. To simplify this, the lowest level of automation of fallback and responsibility can be used to indicate the level of human responsibility expected. To describe the taxonomy for fallback and responsibility, we have considered there to be two aspects of fallback and responsibility which can be automated - *Judgement and Intervention*, defined as:

- *Judgement*: Who or what decides when the plant has entered a situation which is not within the ODD of the plant.
- *Intervention*: Who or what takes over the operation of the plant when it has entered a situation which is not within the ODD of the plant.

The taxonomy is presented in Table 5.

Table 5. Participation of human and system for each level of Automation for Fallback and Responsibility

Level	Name	Judgement	Intervention
0	Human Monitoring	Human	Human
1	Human on Request	System	Human
2	System	System	System

3.6 Combining the Individual Taxonomies

In the above, we have established a taxonomy for each stage of the OUDA loop and for Fallback and Responsibility, (together referred to as OUDA-R for simplicity) for CAP. These have been discussed as separate, standalone classifications such that each could be independently automated. However, when applying these to a particular

piece of plant they would interact, with each stage having implications for future stages.

A simple example is that of undertaking the Observe stage with a level 0 system. Here no observations are made by autonomous components, and hence it is not possible for any meaningful level of autonomy to be applied to the Understand, Decide, and Act stages. This can be generalised to a useful principle for determining if a particular combination of levels is allowed – information collected and used at each stage must be passed to the relevant party for that party to remain involved in the processing loop in subsequent stages. In simpler terms, this means that if a human or automated system is not involved at an earlier stage then this party should not be involved in subsequent stages.

4 Application to Example Plant

We have selected two types of existing technology to demonstrate the potential application of each of the taxonomy components described above. We have chosen to focus on two example compactors, as this is a relatively simple construction activity which primarily involves controlling a plant to operate within a defined area to achieve a target amount of compaction. In contrast to an excavator, the way in which a compactor interacts with its environment is much simpler, as the compactor does not add or remove material. As such, automation within the compaction space is more easily achieved than other construction tasks.

4.1 Robomag



Figure 2. BOMAG Robomag Autonomous Roller

The Bomag Robomag, Figure 2, is an autonomous roller which can be deployed without a manual operator to carry out a pre-defined compaction task which is specified in the design documents uploaded to it. The Robomag operates within a geofenced area, determining how to achieve the

target level of compaction in a safe manner and provides a continuous record of how this task was completed.

Observe Level 2 - Robomag is equipped with GNSS, LIDAR and stereo cameras to allow it to observe its environment in high-fidelity automatically. It also has drum sensors and other proximity sensors for collision avoidance.

Understand Level 2 - Robomag uses observations to understand its environment automatically. For example, the GNSS receiver enables the machine to understand its position within a geofenced site area, the drum sensors provide automated understanding of stiffness.

Decide Level 3 - The machine calculates the most efficient way of completing pre-set tasks within the geofence and begins operating, a manual route override is available prior to tasks beginning.

Act Level 4 - The machine implements the predefined compaction task, without the ability to override these operational inputs. Utilising interpreted data from drum sensors, Robomag automatically adjusts compaction force for consistent stiffness, increasing for soft spots and decreasing for hard spots.

Responsibility Level 2 - The machine is responsible and capable of maintaining its safety without intervention from a human.

4.2 CAT Command for Compaction



Figure 3. CAT CS56B Compactor

The Cat CS56B, Figure 3, is an example of a standard compactor which has been augmented with an automatic system to achieve compaction, the CAT Command for Compaction package. This enables a human operator to define a target area by driving the compactor to each corner of the space, set a target compaction value or number of passes, and record a driving line which the automated system will replicate until the target is achieved.

Observe Level 2 - The CAT CS56B contains all required sensors including dual RTK positioning and radar object detection to allow the vehicle to run in AUTO mode, but requires a human to observe the environment for safety reasons.

Understand Level 2 - The system uses observed parameters to understand all environmental parameters required to complete the compaction task.

Decide Level 3 - The operator inputs an overall task and the system assesses if it is able to run in AUTO mode to achieve this task by deciding what actions are necessary. If the systems are operating as normal, the machine will carry out its task until completion.

Act Level 3 - The machine implements the input compaction task. The difference to Robomag, and hence its lower level, is that this machine is being more closely supervised. The operator must react to system messages to continue in AUTO mode.

Responsibility Level 1 - The system is responsible for maintaining the safety of the machine in AUTO mode; however, it will ask for human intervention where required.

5 Discussion

The experience gained in applying the taxonomy to example plant has shown that there is a need to be clear on whether the levels describe a subsystem, or the system (plant) as a whole. It is possible for a specific subsystem of plant to be very highly automated and hence score highly, but this may not apply to the entire machine. When classifying item of plant it is important to clearly convey what is being assessed for its automated capabilities. When considering the classification of plant holistically, we recommend labelling it as the lowest scoring combination of the considered subsystems as this will overestimate the role of the human operator. This is beneficial in the construction environment where safety is the paramount factor.

A concept related to this is that the taxonomy has been developed for application to specific pieces of plant, but it would be equally possible to consider automation as being a property of the task to be completed. That is, there may be benefit in assessing the degree of human involvement required to successfully implement a particular construction task, no matter the number of machines involved in achieving this.

We also note that in the above examples both of the compactors score relatively highly within the taxonomies. This is partly due to the specific and narrow scope of the operations the compactors are expected to perform. Focusing on specific types of task could enable OEMs to develop highly automated systems, as the expected ODD and task requirements are well understood. In contrast an excavator, for example, undertakes a wide range of activities, which presents a greater challenge for OEMs to

overcome when developing automated systems. Thus, as the use of automated technologies expands on construction sites we might expect the initial applications to be in the more achievable areas.

It is also worth noting that although a system may score highly against the OUDA-R taxonomy this does not imply that it would be a 'good' system for use within a given site. There are two elements to this. Firstly, although a system may be highly automated it may have significant mechanical or operational deficiencies when used in a particular environment that renders the machine unsuitable for deployment. Secondly, a high level of automation may not be the best way of achieving some tasks – automation is best suited for highly repetitive, precise tasks which suit computerised control, whereas human control allows for highly adaptive and responsive operations to be implemented.

Finally, we have discussed above the Roadmap for CAP, and its milestones. The Roadmap suggests that the implementation of CAP could be encouraged through both technical and commercial development. Hence, in addition to supporting a common language across automated technology in CAP, the levels could help to provide a common language for use in contracts that aim to encourage deployment of automated construction techniques. Our examples suggest that this will require care to ensure that contractual requirements are staged to match the technological capabilities to the defined construction activity – the levels provide the tools to support this.

6 Conclusion

This paper has presented the development of a new taxonomy for classifying the automated capability of plant, and the resultant definitions of each level within the taxonomy. The levels have drawn on previous experience in other industries, refined and adapted to meet the needs of autonomy in construction. This has led to a four stage process plus a fallback stage (OUDA-R). An initial peer review of the levels carried out in the stakeholder workshops and feedback received during the launch has suggested that the proposed approach should be both practical and understandable to the industry. To demonstrate application we have applied the taxonomy to two different compaction systems to show how an assessment of automated capability could be carried out, and discussed the challenges associated with such an assessment. These levels can be utilised by industry to greatly improve information transfer when discussing and considering automation of plant. This will support and catalyse the development of technology roadmaps amongst plant and technology manufacturers, enable procurement processes that incentivise the deployment of CAP within construction management, and support innovation practices by providing an understand-

ing of the safety and operational implications of deploying automation on construction sites.

The proposed taxonomy must be applied by industry to provide an understanding of the functional suitability of the taxonomy, so that it can be refined. Industry would also benefit from considering what approaches to automation will be adopted, and if there is a need to develop further taxonomy for the automation of tasks. As identified in the Roadmap [2], there should be no delay to adoption of CAP technologies and supporting work such as presented here, as this may reduce 2040 savings by over 50%.

7 Future Work

Future work for the authors, National Highways, and wider industry is to deploy the taxonomy across a wide-range of construction sites to begin tracking autonomous capability of multiple machines at scale. From this key performance indicators (KPIs) can be created and then CAP expectations and mandates made into contracts to affect these KPIs. Additional work has already begun with standardisation bodies (BSI/ISO) to ensure the taxonomy can be ratified at a larger scale as well as influence other standards. A method of certifying plant against these levels may also be explored to reduce duplication of effort as well as ensure uniformity and fairness in scoring.

Although the taxonomy was developed for use within the construction industry, there is no fundamental aspect which restricts its use to the construction sector. As part of the refinement process, understanding how it applies to other sectors and its utility beyond the construction sector will be explored. This is of particular relevance to industries which also make use of heavy machinery such as mining or agriculture.

Acknowledgements

This work has been completed by TRL and Costain as part of the National Highways Connected and Autonomous Plant Phase 2 project. This forms part of a wider programme of activities that National Highways are implementing to create a digital transformation in the construction, operation, and management of the UK Strategic Road Network. We are grateful to Bhavin Makwana, Francis McKinney, and Tristan Bacon of Zenzic for their discussions and review during the development of this taxonomy, and to the wider industry stakeholders that have given their time and effort in the development and review of the taxonomy.

References

- [1] Mike G. Winter. Continuous compaction control in the UK: history, current state and future prognosis. *Proceedings of the Institution of Civil Engineers - Geotechnical Engineering*, 173(4):348–358, 2020. URL <https://doi.org/10.1680/jgeen.19.00120>.
- [2] National Highways, TRL, and i3P. Connected and autonomous plant roadmap to 2035, 2020.
- [3] Robert H Allen and Ram D Sriram. The role of standards in innovation. *Technological Forecasting and Social Change*, 64(2):171–181, 2000. doi:[https://doi.org/10.1016/S0040-1625\(99\)00104-3](https://doi.org/10.1016/S0040-1625(99)00104-3). URL <https://www.sciencedirect.com/science/article/pii/S0040162599001043>.
- [4] National Highways and TRL. A frame for classifying the capability of connected autonomous plant, 2022.
- [5] Thomas B. Sheridan and William L. Verplank. Human and computer control of undersea teleoperators, 1978.
- [6] R. Parasuraman, T. B. Sheridan, and C. D. Wickens. A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 30(3):286–297, 2000. doi:10.1109/3468.844354.
- [7] Luca Save and Beatrice Feuerberg. Designing human-automation interaction : a new level of automation taxonomy. In *Human Factors: a view from an integrative perspective*, 2013.
- [8] SESAR Joint Undertaking. European ATM master plan, 2020.
- [9] B. T. Clough. Metrics, schmetrics! how the heck do you determine a uav’s autonomy anyway? In *Proceedings of the 2002 Performance Metrics for Intelligent Systems Workshop (PerMIS-02)*, 2002.
- [10] J. R. Boyd. The essence of winning and loosing. On-line: <https://web.archive.org/web/20110324054054/http://www.danford.net/boyd/essence.htm>, 1995.
- [11] Society of Automotive Engineers. Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles, 2014.

Digital Commissioning Processes for the Oil and Gas Sector

D. L. M. Nascimento^a, A. B. Roeder^b, D. Calvetti^c, A.C. Lopez^b, F.R. Gonzalez^d and F.N.M. Araújo^b

^aUniversity of Jaén, Jaén, Spain

^bCERTI Foundation, Florianópolis, Brazil

^cConstruction Institute, CONSTRUCT/Gequaltec, Porto University, Portugal

^dFederal Fluminense University, Industrial Construction Department, Brazil

E-mail: dmattos@ujaen.es, are@certi.org.br, diegocalvetti@fe.up.pt, axl@certi.org.br, fernandorgonza@gmail.com, fna@certi.org.br

Abstract

Oil and Gas (O&G) projects have complex engineering endeavours. A digital information management process is vital to increasing efficiency due to projects' size and complexity. The industrial plants' commissioning is regarding on-site testing to check performance and ensure final compliance of all equipment and systems installed. On-site commissioning procedures are mainly still managed by paper-work-based conventional methods. This work presents a digital solution' architecture developed based on an O&G company commissioning processes mapped. The novelty is assuring through a new integrator system agility and data traceability over digitalisation on-site commissioning procedures. Finally, the commissioning topic has limited scientific applied research, which this paper enhances future works.

Keywords –

O&G; Commissioning; Digitalisation; Processes

1 Introduction

South America has abundant deep-waters oil reserves [1]. Oil and Gas sector exploration projects have substantial budget investments, multiple interfaces, and complex engineering endeavours [2]. Due to the project's size and complexity, the management process is vital to increasing efficiency [2], [3]. The industrial plants commissioning is regarding on-site testing of all equipment and systems installed to check performance and ensure final compliance [4]. Therefore, the digitalisation of the commissioning activities will contribute to high-quality assurance providing digital data traceability [5] increasing performance [6].

Commissioning on-site procedures of inspection and tests are mainly still managed by paper-work-based conventional methods [7], [8]. Lacking interoperability and the information post record over software systems

are persisting issues that inhibit digital process-oriented management [6] making the processes prone to errors [9]. The IFC (Industry Foundation Classes) is a neutral and open data format standardised by the ISO 16739-1 and used for modelling activities and attributes related to an engineering project [10]. Although the use of IFC has increased over time, this format is not widespread in the O&G sector. As a result, there are no IFC applications for facility commissioning information management.

To face this challenge, CERTI Foundation, a Centre of Reference in Innovative Technologies and an O&G (Oil and Gas) Company from South America, started an innovation project to develop a digital solution to deploy commissioning on-site with the integration of existing systems and new mobile applications. The project aims to develop an application interface that can implement inspections and tests commissioning procedures customised by managers to be deployed on-site where the workforce can receive and record the execution of the tasks in near real-time.

This work presents the digital solution' architecture developed based on the commissioning processes mapped. An IDM (Information Delivery Manual) [11] is performed to capture the current business process (as is) integrating all stakeholders. The identification of the status quo allows the understanding of detailed commissioning specifications concerning stakeholders, processes, systems and information. After, together with the O&G commissioning management team a "as to be" process is developed setting the new system requirements. Finally, the novelty is regarding assuring through a new integrator system agility and data traceability over the digitalisation of manual paper-based on-site commissioning procedures. It should be highlighted that the topic has limited scientific applied research, and this research contributes to enhancing future works.

2 Commissioning procedures in the O&G sector

2.1 Background

Once O&G projects assembly is consolidated, and before plant operations, a detailed commissioning process is mandatory [12]. The commissioning process guarantees that the assets will be delivered to users in complete conditions, with all tests and preservation carried out and updated documentation [13]. The commissioning phase is time-consuming and can be associated with health, safety and environmental (HSE) risks [12].

- The O&G company where the project is conducted uses three basic activities over the commissioning process.
- Commissionable tagged items – Any physical component previously identified with a vital function over the plant process.
- Preservation – This is a cyclical process of activities that aim to preserve the best conditions and functioning of the items that constitute the industrial plant.
- Systems Mesh verification – Interconnected set of commissionable items that compose a system mesh often concerning a plant piece of processing (e.g., gas injection system, cooling water system).

While for some authors, activities can be separated into pre-commissioning and commissioning, others consider commissioning a unique phase that falls within the pre-operation and construction phases [14], [15]. However, regardless of the project breakdown structure, the commissioning information management success is connected to integrating and managing information through systems and databases. Most importantly, the software communication is integrated into different languages and needs a fully integrated data exchange. However, software integration is often prone to error, lacking completion and low accuracy [9], demanding interoperable actions.

Poor data recording can significantly threaten project performance [15]. Information management is crucial for the O&G sector and directly impacts project quality assurance [16]. In a non-automated process, paper documentation assigned forms will directly prove that the procedure is done. The information generated during the commissioning process must harmonise and correspond to the project requirements [16]. Often, last moment documentation collection and preparation to deliver to clients impacts projects delay and claims [15]. Also, periodic reports are vital for the commissioning execution as the management team can exchange information about the project progress and identify areas

of concern [15]. Therefore, it is essential to use a system that allows the management of documents that lead the procedures involved in the project. Furthermore, a commissioning system should fully visualise the processes and enable communication between those involved.

2.2 As is commissioning procedures

The O&G commissioning process studied is currently performed in a database software system that creates PDF files to orient on-site activities. The actual database system works with a fixed process, which means that the process has to happen as designed in the first version of the system. It is not allowed to update, change or customise the activities. Most importantly, all the on-site procedures are performed based on printed paper forms. As activities performed on-site are later recorded on the system by human input. The following steps performed by the specific actors are conducted for commissionable items and system mesh verifications:

- Managers: Login the system, input information regarding commissionable items and systems mesh and release the commissioning procedures (items verifications, mechanical inspections and functional tests).
- System Operators: Print the technical inspection procedures and send them to the on-site workers.
- On-site workers: Receives the paper forms with the technical inspection activities, performs them, fills the paper forms with the execution information, and sends the paper to the system operators.
- System Operators: Record the activities information to the system and send the paper forms for approval.
- Managers: Fill out the paper forms with the process approval and send the papers back to the system operators.
- System Operators: Record the approval information on the system.

The preservation is a cyclical and parallel process and is over just when all the commissioning process is completed, it follows the steps below:

- Managers: Login in the system and input information regarding preservation procedures.
- System Operators: Input the information that will release the preservation procedures. Print labels with the procedures and place them in the respective items on-site.
- On-site workers: Find the items with the labels, perform the procedure and fill the paper with the activities information.
- System Operators: Go on-site to collect the items' execution information, record them in the system,

send them to approve and send an email to inform the managers about the approval process.

- Managers: Approve the preservation procedure in the system.

2.3 As to be commissioning procedures

As identified, the main bottleneck was related to manual commissioning activities, instructions, and reports. Based on that, the new process designed targets eliminate all paper-related steps. Deploying the commissioning procedures for digitalisation using a process-oriented architecture increases users' usability by providing a tailor-made tool. That concept will increase quality assurance where the company engineering team is now in charge of modelling the procedure flows and checking the field procedures execution.

As the process is dynamic, it is possible to add as many approval steps as they want to assure that the right information is uploaded. So, if one activity is refused, it will return to the field worker with the appropriate guidelines for correction. And until the evaluator does not approve, the activity does not follow the flow. The new processes facilitated and eliminated users' work. The following steps demonstrate the new process:

- Managers: To log in to the web application to design a commissioning procedure using standardised processes or model new ones.
- On-site workers: To log in to the mobile application and identify the procedures to be conducted. To perform and report the activities execution concerning the items verification, system mesh verification and preservation.
- Managers: Receive the procedures completion alerts and directly in mobile or web applications perform the approval or not.

With the new flow, there is no longer a need for a dedicated System Operator as Managers can quickly check the data uploaded and executed activities. In addition, the managers can adapt the workflow of the procedure through the system to satisfy the new demands. The software will have embedded continuous improvement since it is possible to analyse the new workflows designed and implement improvements over the procedures.

The system interoperability will be made over Industry Foundation Classes (IFC) to assure accuracy in data exchange when necessary and improve the transparency in the process. An IFC commissioning schema was developed and published in an open repository (doi.org/10.5281/zenodo.5786902) where different vendors can access [17]. The IFC commissioning standard in IFC JSON was validated by the buildingSMART converter [17].

Thereby, the commissioning process developed will

have significant improvements, such as:

- Significantly lead-time reduction.
- Data consistency and traceability in near real-time.
- Workforce reallocation to functions that add value to the process.

Figure 1 shows the system use case diagram that helps to describe the relationship between the system, actors, and case uses. The managers (primary actors) can use the web system for configuration, customisation, and validation of the commissioning procedures. The on-site workers (secondary actors) will utilise the mobile to receive the procedures and record the work done.

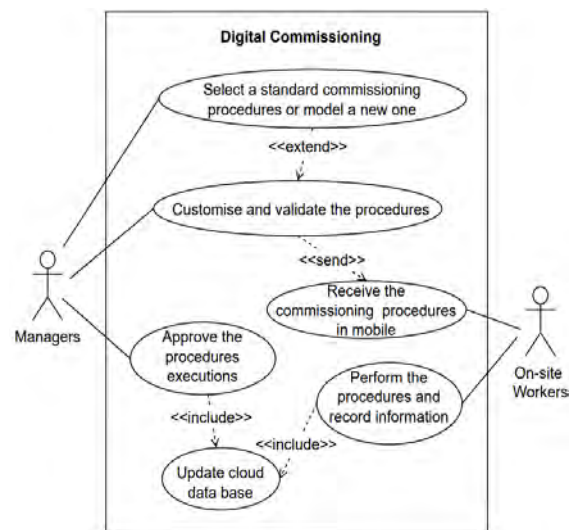


Figure 1. Digital commissioning use case diagram

Figure 2 presents the architecture systems concept. The Digital Commissioning system is a web and mobile application with the same interface, making a two-factor authentication by the ADFS (Active Directory Federation Service).

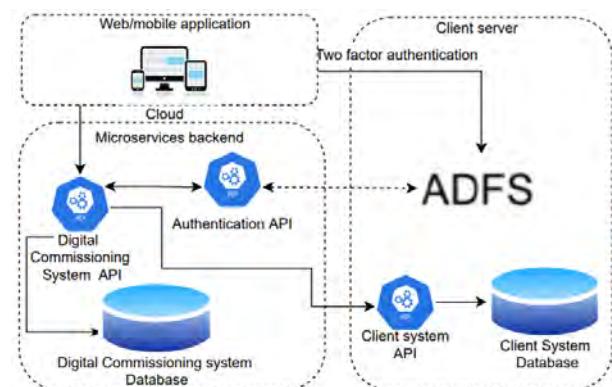


Figure 2. Digital commissioning architecture system concept

The ADFS service enables a secure sharing of identifying information through the client system to call the service backend by an authentication token. The backend is based on microservices in a cloud computing server. It stores the data in an internal database (Digital Commissioning database) while calling the client system API to synchronise the same data in the client system database.

2.4 Remarks and Future Directions

This work summary presents a system that brings a digital solution that automates the commissioning process. The digital management of the information directly affects procedures performance evidence. When the digital solution is fully implemented, the commissioning process will be done quickly, increasing quality assurance and eliminating non-value activities. Finally, bringing easiness to workers' routine with digital data traceability. This work is limited to presenting the commissioning flows only considering the procedures on-site. In future research, this digital solution will add IoT (Internet of Things) and integrated Building Information Modelling (BIM) to develop mixed-reality tools for on-site workers.

References

- [1] IBP, '2019 Report', IBP - Instituto Brasileiro de Petróleo, Gás e Biocombustíveis, 2019. [Online]. Available: <https://www.ibp.org.br/personalizado/uploads/2020/03/relatorio-atividades-2019-ibp-compactado.pdf>
- [2] A. B. Badiru and S. O. Osisanya, *Project Management for the Oil and Gas Industry: A World System Approach*, 0 ed. CRC Press, 2013. doi: 10.1201/b13755.
- [3] Z. Rui *et al.*, 'Development of industry performance metrics for offshore oil and gas project', *Journal of Natural Gas Science and Engineering*, vol. 39, pp. 44–53, Mar. 2017, doi: 10.1016/j.jngse.2017.01.022.
- [4] K. E. Arnold, *Petroleum Engineering Handbook*. Society of Petroleum Engineers, 2017.
- [5] A. Jain, D. A. Vera, and R. Harrison, 'Virtual Commissioning of Modular Automation Systems', *IFAC Proceedings Volumes*, vol. 43, no. 4, pp. 72–77, 2010, doi: 10.3182/20100701-2-PT-4011.00014.
- [6] H. Lu, L. Guo, M. Azimi, and K. Huang, 'Oil and Gas 4.0 era: A systematic review and outlook', *Computers in Industry*, vol. 111, pp. 68–90, Oct. 2019, doi: 10.1016/j.compind.2019.06.007.
- [7] A. Borrmann, M. König, C. Koch, and J. Beetz, 'Building Information Modeling: Why? What? How?', in *Building Information Modeling*, A. Borrmann, M. König, C. Koch, and J. Beetz, Eds. Cham: Springer International Publishing, 2018, pp. 1–24. doi: 10.1007/978-3-319-92862-3_1.
- [8] D. L. de M. Nascimento, O. L. G. Quelhas, M. J. Meiriño, R. G. G. Caiado, S. D. J. Barbosa, and P. Ivson, 'Facility Management Using Digital Obeya Room By Integrating Bim-Lean Approaches – An Empirical Study', *Journal of Civil Engineering and Management*, vol. 24, no. 8, pp. 581–591, Dec. 2018, doi: 10.3846/jcem.2018.5609.
- [9] C. Palmer, Z. Usman, O. Cancigliieri Junior, A. Malucelli, and R. I. M. Young, 'Interoperable manufacturing knowledge systems', *International Journal of Production Research*, vol. 56, no. 8, pp. 2733–2752, Apr. 2018, doi: 10.1080/00207543.2017.1391416.
- [10] NBIMS, 'National Building Information Modeling Standard Version 1 - Part 1: Overview, Principles, and Methodologies', National Institute of Building Sciences, 2007.
- [11] ISO/DIS 29481-1, 'ISO/DIS 29481-1 Building information models - Information delivery manual - Part 1: Methodology and format'. International Organisation for Standardisation, 2014.
- [12] T. R. Wanasinghe *et al.*, 'Digital Twin for the Oil and Gas Industry: Overview, Research Trends, Opportunities, and Challenges', *IEEE Access*, vol. 8, pp. 104175–104197, 2020, doi: 10.1109/ACCESS.2020.2998723.
- [13] ASHRAE Handbook, *Heating, ventilating, and air-conditioning applications*. Inc.: Atlanta, GA, USA., 2011.
- [14] David Horsley, *Process Plant Commissioning: a user guide*. IChemE, 1998.
- [15] T. Bendiksen and G. Young, *Commissioning of Offshore Oil and Gas Projects: The Manager's Handbook*. AuthorHouse, 2005.
- [16] N. N. Samie, *Practical Engineering Management of Offshore Oil and Gas Platforms*. Elsevier, 2016. doi: 10.1016/C2014-0-04721-1.
- [17] D. L. de M. Nascimento *et al.*, *IFC commissioning standard in IFC JSON*. 2021. doi: 10.5281/ZENODO.5787461.

Overview of the State-of-Practice of BIM in the AEC Industry in the United States

H. Nassereddine^a, M.B. Hatoum^a, and A.S. Hanna^b

^aDepartment of Civil Engineering, University of Kentucky, USA

^bDepartment of Civil and Environmental Engineering, University of Wisconsin – Madison, USA

E-mail: hala.nassereddine@uky.edu, mbh.93@uky.edu, and ashanna@wisc.edu

Abstract –

The Architecture, Engineering, and Construction (AEC) industry has undergone a significant and radical transformation in its design and documentation process as it evolved from the days of the drafting board to today's Building Information Modeling (BIM) process. As BIM remains the center of this transformation, it is important to keep both practitioners and academicians updated on the current state-of-adoption of BIM in construction projects. Thus, this paper presents the results of a BIM survey conducted on 125 respondents representing 83 companies located in the United States of America, United Kingdom, Netherlands, and Canada. The types of the targeted companies varied between Owner, Owner's Representative (OR), Architect/Engineer (A/E), General Contractor/Construction Management (GC/CM), Mechanical Contractor, Electrical Contractor, Sheet Metal Contractor, Plumbing Contractor, Fire Protection Contractor, Structural Steel Contractor, and Facility Manager. Findings of the paper elaborate on why companies are using and requiring BIM, why companies are not using and requiring BIM, and how BIM is being used by the different company types across the project lifecycle.

Keywords –

Building Information Modeling; BIM; AEC Industry; Construction Projects; Usage; Practices; Users; Construction 4.0; Construction Project Lifecycle

1 Introduction

The fourth industrial revolution, coined as Industry 4.0, has been reshaping the way work is done in the twenty first century, and the construction industry is no exception to this transformation [1][2]. The mapping of Industry 4.0 onto the Architecture, Engineering, and Construction (AEC) industry has been coined as Construction 4.0 – a radical transformation that is

digitizing and industrializing the AEC industry using technology [3][4]. Major technologies such as augmented reality, robotics, big data, drones, and digital twins are being heavily investigated to increase their use across the construction project lifecycle [5][6][7][8][9].

One major component at the center of Construction 4.0 is Building Information Modeling (BIM). Studies that investigated Construction 4.0 have identified BIM as a major enabler for the construction transformation [10][11]. [4] noted that BIM is at the center of the construction industry transformation since it is the technology that interacts with every other Construction 4.0 technology. BIM has also laid down the foundation for construction firms to adopt a wide variety of technology because it incorporates people, processes, and technology [12].

Given its essential role in advancing the industry, the continuous investigation of BIM in the AEC industry is critical to keep both academicians and practitioners updated on the latest state-of-adoption of BIM in the industry [13]. This paper provides an overview of the state-of-practice of BIM surveying 128 construction practitioners on their rationale for using BIM and their perceptions on the usage and users of BIM across the construction project lifecycle.

2 Background

2.1 The Evolution of Construction Information Technology

Ever since humanity started building structures, there have been accompanying methods of drawing, sketching, and planning of these structures. The two-dimensional (2D) drawings for architectural purposes have been traced back to Ancient Egypt [14] and have evolved over the course of history to keep pace with the advancing complexity and ambition of the built environment.

The most common purpose of 2D construction and architecture drawings is the presentation and visualization of an as-yet unbuilt structure,

communication of the designer(s) intentions, and instructions for later on-site work. The earliest known drawings of this type are Egyptian, as previously stated. The next evolution of construction documentation occurred in middle-ages Europe. During that time, construction was overseen in all aspects by a 'Master Builder' who would plan, manage, and execute a project for an owner or patron. To communicate the particulars of the design to that patron, the master builder would employ scale models [15]. The patrons, usually landed nobility, provided the funding for many of the most iconic structures we know today – the castles and fortresses of feudal Europe. However, the term 'construction documents' as currently used still did not yet exist. The master builder relayed instructions to the workers verbally or through demonstration, rather than disseminating plans and drawings. Many particularly complex aspects of the project were developed as full-scale mockups on site, using real materials.

In the Renaissance, projects grew larger and more complex, and the master builder spent more time off-site working through engineering problems in the 'office'. Eventually, early engineering drawings emerged. They served a twofold purpose – to communicate to experienced craftsmen what should be built, and to show a particular detail or section to the patron(s) for their approval [16]. The consequence of the master builder spending more time off-site was the creation of the superintendent position, as the project still required supervision on-site. Thus, the master builder assumed the new responsibility of coordinating communications between the patron (owner) and the superintendent, while making design changes. As construction continued to grow more complex, the various trades began to specialize – masons, carpenters, joiners, etc.

The Pharaohs of Egypt, the master builders of the middle-ages, the architects of the renaissance, and even constructors today all face a common problem: buildings are three dimensional, but documents are not. Thus, the use of 2D drawings and instructions in a 3D world requires multiple translations – from the initial concept in the designer's head, onto paper, and then into reality. As such, numerous efforts have been made to improve the quality of design drawings. These efforts are motivated by the need to reconcile planned solutions with practical implementations, poor communication between project parties, and inefficient scheduling of construction activities [17]. [18] postulated that the need for teamwork, flexibility, coordination, and communication in construction gave the industry a great potential to integrate Information Technology (IT). Froese has divided the innovations in IT into three eras [19][20]. The first era is comprised of "stand-alone tools that improve specific work tasks – Computer Aided

Design (CAD), Structural Analysis, Estimating, Scheduling – which are all individual programs that each works on a single facet of the construction process". During the early 1980s, CAD became commonplace in architectural work and soon supplanted the drafting board as the most common method of producing drawings. This is because CAD allows for quick replication with a high degree of accuracy. Eventually CAD also supported 3D design, making it a more attractive and efficient option than hand-drafting [21][22]. The second era includes computer-supported communications (i.e. email, web-based messaging), and document management systems. The third era is where construction currently sits – "reconciling the first two eras into a unified platform wherein project teams can collaborate to produce a virtual model of all aspects of the construction project". A major problem faced during the early iterations of CAD was the lack of understanding of relationships between the spatial geometric objects and how these relationships functioned. For example, while drawings communicated that a beam is connected to a column, the number, size and placement of the bolts to connect it would not be communicated [23]. This problem was addressed in more modern iterations of CAD, and the inclusion of this process is commonly known as Building Information Modeling (BIM).

2.2 The Evolution of BIM

The concept of BIM can be traced back to 1962 when Engelbart presented a hypothetical description of computer-based augmentation system [24]. Later, [25] recognized the shortcomings of 2D drawings and developed a computer-based Building Description System (BDS) that arranges and connects the geometric, spatial, and property description of the various elements of a building into an actual 3D building. This system serves as a database that provides a single description of each building element and of its relation to other components in the building and can be used during design, construction, and operation. In addition, if change is needed, designers need to make the change to the element once and the drawings will be automatically updated. This system designed by [25] paved the way for the concept of Building Information Models, a term that was first introduced by [26].

BIM has transformed the traditional paradigm of construction industry from 2D-based drawing information systems to 3D-object based information systems [27][28]. For more than a decade, BIM has been one of the most important innovation means to approach building design holistically, enhance communication and collaboration among key stakeholders, increase productivity, improve the overall quality of the final product, reduce the fragmentation of

the construction industry, and improve its efficiency [29][30]. One of the greatest benefits of BIM is its ability to represent in an accessible way the information needed throughout a project lifecycle, rather than being fragmented [31]. Being a shared knowledge resources through the project lifecycle, BIM centralizes information on a facility and acts as a “reliable basis for decisions” during the facilities’ lifecycle [32]. [33] defined BIM as “a technology that describes an engineering project consisting of intelligent facilities with their own data properties and parameter rules, in which each object’s appearance and its internal components and features can be displayed in the form of three-dimensional figures”.

BIM has been widely hailed as a successful innovation in the construction industry [34], with numerous competing products available on the market today: AutoCAD MEP, Revit® (Autodesk®), BIM 360™ Glue®, Navisworks® (Autodesk®), Sketchup (Trimble®), Synchro Bentley Systems, Graphisoft, and Nemeschek [23][35].

BIM has also evolved from the 3D modelling (object model) to further dimensions such as 4D (time), 5D (cost), and 6D (as-built operations) [36]. This evolution represents added information that is placed in the model and attached to intelligent objects [37].

3 Methodology and Research Questions

To provide an update on the perception of the AEC industry on the use of BIM, a survey was developed and distributed to the industry. Five stakeholders were targeted: Owner, Owner’s Representative (OR), Architect/Engineer (A/E), General Contractor/Construction Management (GC/CM), and specialty contractors (Mechanical Contractor, Electrical Contractor, Plumbing Contractor, i.e., MEP Trades). As a result, 128 responses were captured from all five stakeholders. Descriptive and statistical analyses were then employed to describe, summarize, and analyse the collected data to answer the following research questions:

1. Why are companies not using/requiring BIM?
2. Why are the companies using/requiring BIM?
3. What are the most common BIM practices and how do they vary between company types?
4. What is the usage of BIM for different construction stakeholders and how is this usage perceived by different company types?
5. Where is BIM used across the project lifecycle and how does this usage vary between company types?

4 Analysis

A total of 128 responses were collected from 83 different companies in the AEC industry.

4.1 Data Characteristics

4.1.1 Geographic Distribution

The bulk of the respondents (around 96%) were located in the United States of America (USA). Other respondents were located in Canada, United Kingdom, and Netherlands. Within the USA, most responses were collected from Wisconsin, California, Illinois, and Minnesota. Given that the bulk of the data was from USA, the remaining sections of the paper will encompass the analysis of the 125 USA data points.

4.1.2 Types of Companies

Respondents were asked to identify the type of their company among the following options: Owner, OR, A/E, GC/CM, and MEP Trades. The data was recategorized into the five types as shown in Figure 1.

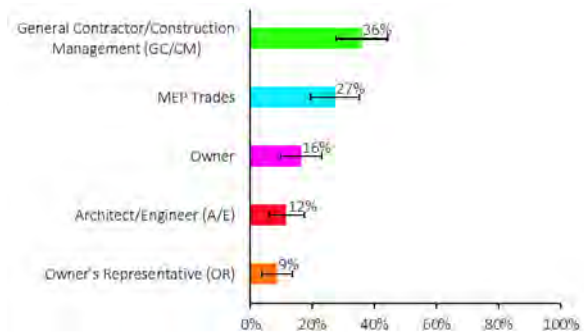


Figure 1. Breakdown of respondents by company type

4.1.3 Respondent Occupation

Respondents were asked to provide their job titles. Their responses were then categorized into one of the following occupations: Technologist, Field, and Top Management. Figure 2 shows the distribution of the respondents based on their occupation.

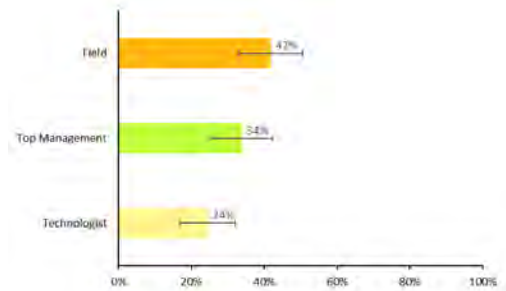


Figure 2. Breakdown of respondents by occupation

4.2 BIM Usage

4.2.1 Companies Distribution

The 125 responses were collected from the employees of 80 companies: 59 companies represent A/E, GC/CM, MEP Trades, or OR (Group 1), and 21 companies represent owners (Group 2). Out of the companies that belong to Group 1, 79% use BIM and the rest (21%) do not use BIM. Among the companies that belong to Group 2, 43% indicated that they require the use of BIM and 57% indicated that they do not. These numbers show that the dominant majority of companies use BIM despite it being only required by a slight majority of Owners.

4.2.2 Respondents Distribution

Out of the 104 respondents who work for companies in Group 1, 87% indicated that their company uses BIM and 13% reported that their company does not. Among the 21 respondents who belong to companies of Group 2, 43% indicated that their organization requires the use of BIM and 57% indicated that they do not.

4.2.3 Reasons for Not Using and Requiring BIM

Respondents were provided with 6 frequently reported reasons for not using BIM, as shown in Figure 3. Out of the 14 respondents of Group 1 who reported that their companies do not use BIM, 64% reported that the high cost of implementation and the fact that BIM is not requested/required by owners are the main reasons for not using BIM. Only 7% of the respondents indicated that social and habitual resistance to changes is a driver for not using BIM. Respondents also had the option to provide other reasons for not using BIM. One respondent indicated that they still use CAD and another reported that their company manages designers who use BIM.

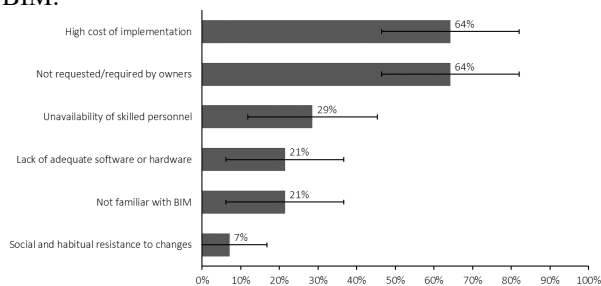


Figure 3. Reasons for not using BIM

Similar to the companies that do not use BIM, owners were given 5 reasons for not using BIM and were asked to select those they relate to them, as shown in Figure 4. Out of the 12 owners who reported that their company does not require the use of BIM, 33% reported that high cost of implementation is the main reason for not requiring BIM.

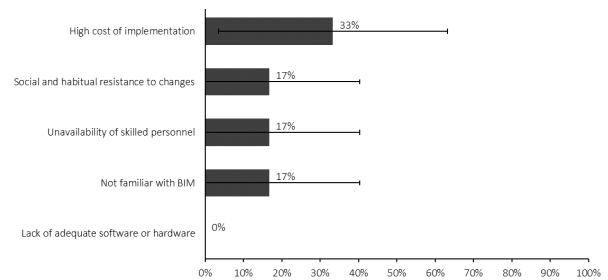


Figure 4. Reasons for not requiring/requesting BIM

Overall, it shows from Figures 3 and 4 that “High Cost of Implementation” is the biggest barrier for the implementation of BIM.

4.2.4 Initial/Ongoing Reasons for Using and Requiring BIM

Respondents who indicated that their companies use BIM were asked to select the factors that impacted their initial decision to use BIM and the factors that impacted their company’s ongoing use of BIM (shown in Figure 5).

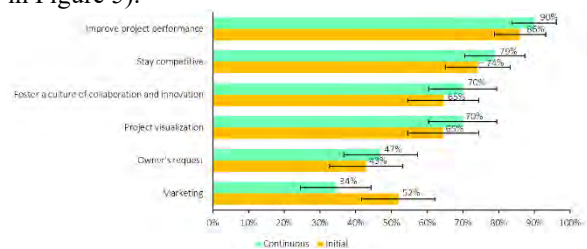


Figure 5. Reasons impacting company’s initial and continuous decision to use BIM

Similarly, respondents who work for Owners were asked to identify the factors that impacted their initial and ongoing decision to require the use of BIM on their projects (shown in Figure 6).

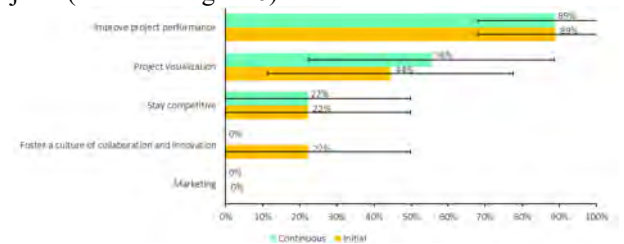


Figure 6. Reasons impacting company’s initial and continuous decision to require/request the use of BIM

It can be noticed from Figure 5 and Figure 6 that most respondents, whether their company uses or requires BIM, indicated that BIM is vital for improving the project performance.

4.2.5 BIM Practices: Overall Industry Usage

Thirteen BIM practices (shown in Figure 7) were identified, and respondents were asked to identify their company's level of use of each practice. As seen in Figure 7, between 20% and 55% of respondents indicated that these practices are not used within their organization, while the rest reported their organization uses these practices. Notably, Clash Detection was reported to be the most used BIM practice with 80% of respondent using it. It was followed by Visualization (79% use it and 21% do not), design collaboration (77% use it and 23% do not) and understanding constructability (77% use it and 23% do not). Environmental Analysis was the least BIM practice used, with only 45% of the respondents using it.

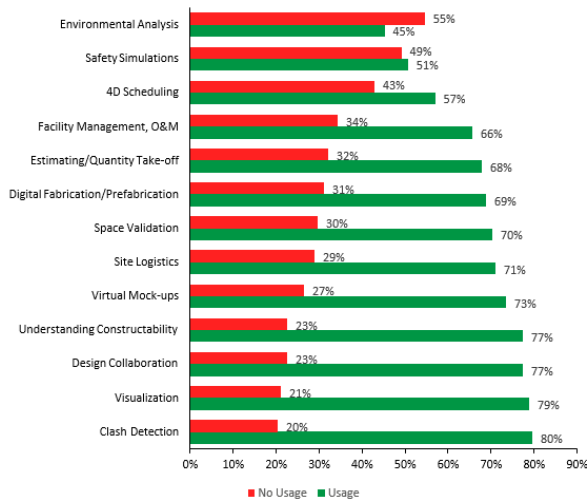


Figure 7. Usage/No Usage distribution of BIM Practices

Respondents who have indicated that their company uses a given BIM practice were subsequently asked to rate their level of usage on a scale from 1 (very low) to 5 (very high). The respondents' overall average level of usage of each BIM practice is displayed in Table 1 indicating that, on average, Clash Detection has the highest level of usage in the industry (4.20) and Safety Simulation has the lowest (2.09). The cluster analysis performed on the 13 BIM practices, showed that Clash Detection, Design Collaboration, Visualization, and Understanding Constructability have, on average, the highest level of usage in the AEC industry.

Table 1 Clustered table of the ranked BIM practices based on their average level of usage

BIM Practices	Overall Average	Clusters
Clash Detection	4.20	Cluster 1
Design Collaboration	4.09	

Visualization	3.78	Cluster 2
Understanding Constructability	3.70	
Space Validation	3.42	
Virtual Mock-ups	3.20	
Digital Fabrication/Prefabrication	3.06	
Site Logistics	2.87	Cluster 3
Estimating/Quantity Take-off	2.84	
4D Scheduling	2.41	
Facility Management, Operation & Maintenance	2.26	
Environmental Analysis	2.17	
Safety Simulations	2.09	

4.2.6 BIM Practices: Usage Per Company Type

The average level of usage of the BIM practice was then broken down by company type and shown as a heatmap in Figure 8. The darker the color the higher the average level of usage of BIM.

BIM Practice	Average Level of Usage				
	A/E	GC/CM	MEP Trades	OR	Owner
Clash Detection	3.64	4.56	4.38	3.14	3.33
Design Collaboration	4.55	4.09	4.13	3.33	3.88
Visualization	4.18	4.05	3.41	3.83	3.33
Understanding Constructability	3.90	3.74	3.77	3.57	3.11
Space Validation	3.78	3.13	3.68	3.67	3.38
Virtual Mock-ups	3.90	3.31	2.70	3.43	3.25
Digital Fabrication/Prefabrication	2.50	3.00	3.45	2.50	2.25
Site Logistics	2.50	3.58	2.04	2.14	2.71
Estimating/Quantity Take-off	2.70	3.18	2.54	2.33	2.57
4D Scheduling	2.25	2.92	1.84	1.50	1.50
Facility Management, O&M	1.80	2.36	2.15	2.17	3.00
Environmental Analysis	2.56	2.03	2.50	2.25	1.67
Safety Simulations	2.33	2.14	2.33	1.80	1.20

Figure 8. Heatmap of the average level of usage of BIM practices by company type

The analysis showed that the average level of usage of the different construction companies is similar for the following BIM practices: Design Collaboration, Visualization, Understanding Constructability, Space Validation, Virtual Mock-ups, Digital Fabrication/Prefabrication, Estimating/Quantity Take-off, Facility Management, O&M, Environmental Analysis, and Safety Simulations.

While for the three remaining BIM Practices, it was shown using the Kruskal-Wallis test and its post-hoc Conover-Iman non-parametric test that for:

- Clash Detection: GC/CM have a significantly higher average level of usage than MEP Trades.
- Site Logistics: GC/CM have a significantly higher average level of usage than OR and Owners, and MEP Trades have a significantly higher average level of usage than OR.
- 4D Scheduling: GC/CM have a significantly higher average level of usage than MEP Trades.

4.2.7 BIM Stakeholders: Overall Industry Perceived Usage

The use of BIM by 10 stakeholders (listed in Table 2) was investigated. More than 50% of respondents reported that, in their opinions, inspectors do not use BIM and 63% indicated that workers also do not use BIM. Conversely, the usage of BIM is reported to be mostly uniform between 70-78% for other stakeholders with project managers reported to be most users of BIM.

Respondents who have indicated that a particular stakeholder uses BIM were asked to rate this stakeholder's perceived level of usage of BIM on a scale from 1 (very low) to 5 (very high). The respondents' overall perceived average level of usage of BIM of each stakeholder is illustrated in Table 2 indicating that, on average, "Architects/Engineers" have the highest perceived level of usage (3.91), followed by "Project Managers" (3.51). "Project Executives" (2.08) and "Inspectors" (1.90) have the lowest perceived level.

Table 2 Clustered table of the ranked stakeholders based on their average perceived level of usage of BIM

Stakeholders	Overall Average	Clusters
Architects/Engineers	3.91	Cluster 1
Project Engineers	3.51	
Project Managers	3.00	Cluster 2
Superintendents	2.75	
Foremen	2.71	
Owners/Owner's Representatives	2.42	Cluster 3
Workers	2.35	
Facility Managers	2.24	
Project Executives	2.08	
Inspectors	1.90	

4.2.8 BIM Stakeholders: Perceived Usage Per Company Type

Each stakeholder's level of usage was then broken down by company type and is illustrated as a heatmap in Figure 9. The darker the color the higher the average level of usage of BIM.

BIM Users	Average Perceived Level of Usage					
	Average	A/E	GC/CM	MEP Trades	OR	Owner
Architects/Engineers	3.91	4.36	4.17	3.52	3.71	3.63
Project Engineers	3.51	3.73	3.57	3.26	3.43	3.88
Project Managers	3.00	2.73	2.98	2.94	3.29	3.50
Superintendents	2.75	1.67	2.76	3.10	2.29	3.00
Foremen	2.71	2.17	2.49	3.19	1.80	3.00
Owners/Owner's representatives	2.42	2.11	2.41	2.17	3.00	3.29
Workers	2.35	2.33	2.13	2.81	1.50	2.40
Facility Managers	2.24	1.89	2.21	2.07	2.43	3.29
Project Executives	2.08	1.89	1.92	2.10	2.60	2.71
Inspectors	1.90	1.75	1.89	1.69	2.20	2.40

Figure 9. Heatmap of stakeholders' reported level of use of BIM by company type

The analysis of this heatmap using the Kruskal-Wallis non-parametric test showed that respondents working for different types of companies have, on average, similar perception of the level of usage of BIM of stakeholders except for the following:

- Architects/Engineers: GC/CM reported a higher level of usage of BIM by Architects/Engineers that what was reported by MEP Trades.
- Superintendents: GC/CM and MEP Trades reported a higher level of usage of BIM by Superintends than what A/E reported.
- Foremen: MEP Trades reported a higher level of usage of BIM by Foremen that what was reported by GC/CM.

4.2.9 BIM Phases: Overall Industry Perceived Usage

The use of BIM was investigated throughout the seven construction project lifecycle phases: Planning, Design, Pre-Construction Planning, Construction, Commissioning, Operation and Maintenance, and Decommissioning. The perceived usage of BIM throughout these phases varied among respondents. Between 20% and 48% indicated that BIM is not used in a particular phase, while the rest reported the use of BIM in a certain phase. Notably, 80% of the respondents believe that BIM is used in the design phase and 48% of the respondents reported that BIM is not used in the Decommissioning.

Respondents who have indicated that BIM is used in a phase were asked to rate the perceived level of usage of BIM on a scale from 1 (very low) to 5 (very high). On average, respondents reported that BIM has a high level of usage in the Design, Pre-Construction Planning, Construction, and Planning phases and a low level of usage in Commissioning, Operation and Maintenance and Decommissioning (Table 3).

Table 3 Clustered table of the ranked project phases based on the average perceived level of usage of BIM in each phase

Phases	Overall Average	Clusters
Design	3.97	Cluster 1
Pre-Construction Planning	3.83	
Construction	3.77	
Planning	3.36	
Commissioning	2.29	Cluster 2
Operation and Maintenance	2.04	
Decommissioning	1.78	

4.2.10 BIM Phases: Average Perceived Usage per Company Type

The perceived level of usage of BIM throughout the lifecycle of a construction project was then broken down by company type as shown in Figure 10. The analysis of this heatmap using Kruskal-Wallis showed that all company types have a similar perception of the level of usage of BIM in each of the seven phases of the lifecycle of a construction project.

Phases	Average Perceived Level of Usage				
	A/E	GC/CM	MEP Trades	OR	Owner
Design	4.55	4.07	3.63	4.00	4.00
Pre-Construction Planning	3.64	3.91	3.91	3.14	4.00
Construction	3.36	3.88	3.94	3.14	3.63
Planning	3.36	3.54	3.13	3.67	3.22
Commissioning	2.55	2.39	2.21	2.00	1.88
Operation and Maintenance	2.10	2.10	1.89	2.14	2.13
Decommissioning	2.00	1.77	1.76	1.83	1.50

Figure 10. Heatmap of the average level of usage of BIM throughout the lifecycle of a construction project by company type

5 Conclusions and Future Works

The results of the questionnaire showed that while 72% of A/E, GC/CM, MEP Trades, and OR use BIM, only 43% of Owners require the use of BIM on their project. Respondents whose companies don't use or require BIM reported that high cost of implementation is the biggest barrier for implementing BIM. On the other hand, respondents whose companies use or require the use of BIM indicated that BIM is vital for improving project performance. Out of 13 identified BIM practices, Clash Detection, Design Collaboration, Visualization, and Understanding Constructability were reported to have, on average, the highest level of usage. Architects/Engineers and Project Managers were reported to be the stakeholders with the highest level of usage of BIM and Owners/Owner's Representatives, Workers, Facility Managers, Project Engineers, and Inspectors to be the stakeholders with the lowest level of usage of BIM, on average. In addition, it was shown that BIM has, on average, a high level of usage in each of the Design, Pre-Construction Planning, Construction, and Planning phases and a lower level of usage in Commissioning, Operation and Maintenance, and Decommissioning phases.

6 References

- [1] Hatoum, M.B., Nasserredine, H., and Badurdeen, F. Reengineering construction processes in the era of construction 4.0: a lean-based framework. In *Proc. 29th Annual Conference of the International Group for Lean Construction (IGLC)*, pages 403–412, Lima, Peru, 2021.
- [2] Whitmore, D., Papadonikolaki, E., Krystallis, I., and Locatelli, G. Are megaprojects ready for the 4th industrial revolution? In: *Proceedings of The Institution of Civil Engineers-Management, Procurement and Law*, pages 1–9, 2020.
- [3] Ammar, A. and Nasserredine, H. Blueprint for construction 4.0 technologies: A bibliometric analysis. *IOP Conference Series: Materials Science and Engineering*, 1218(1):012011, 2022.
- [4] El Jazzar, M., Schranz, C., Urban, H., and Nasserredine, H. Integrating construction 4.0 technologies: A four-layer implementation plan. *Frontiers in Built Environment*, 7 2021.
- [5] Bou Hatoum, M., Piskernik, M., and Nasserredine, H. A holistic framework for the implementation of big data throughout a construction project lifecycle. In *Proceedings of the 37th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 1299–1306, Kitakyushu, Japan, 2020.
- [6] Bou Hatoum, M. and Nasserredine, H. The use of drones in the construction industry: applications and implementation. In *Proceedings of the 39th International Symposium on Automation and Robotics in Construction (ISARC)*, Bogota, Colombia, 2022.
- [7] El Jazzar, M., Piskernik, M., and Nasserredine, H. Digital twin in construction: an empirical analysis. In *EG-ICE 2020 Proceedings: Workshop on Intelligent Computing in Engineering*, pages 501–510, Germany, 2020.
- [8] Hatoum, M.B. and Nasserredine, H. Developing a framework for the implementation of robotics in construction enterprises. In *EG-ICE 2020 Proceedings: Workshop on Intelligent Computing in Engineering*, pages 453–462, Germany, 2020.
- [9] Nasserredine, H., Schranz, C., Bou Hatoum, M., and Urban, H. A comprehensive map for integrating augmented reality during the construction phase. In *Proceedings of the Creative Construction e-Conference 2020*, pages 56–64, Online, 2020.
- [10] Botton, C., Rivest, L., Ghnaya, O., and Chouchen, M. What is at the root of construction 4.0: a systematic review of the recent research effort. *Archives of Computational Methods in Engineering*, 28(4):2331–2350, 2021.
- [11] Maskuriy, R., Selamat, A., Ali, K.N., Maresova, P., and Krejcar, O. Industry 4.0 for the construction industry—how ready is the industry? *Applied Sciences*, 9(14):2819, 2019.
- [12] R. Sacks, Construction tech: innovation founded on lean theory and BIM technology. Online: <https://www.ucl.ac.uk/bartlett/construction/events/2019/feb/construction-tech-innovation-founded->

- lean-theory-and-bim-technology-professor-rafael.
- [13] Nassereddine, H. and Bou Hatoum, M. Understanding and leveraging BIM efforts for electrical contractors. In *Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 963–970, Dubai, UAE, 2021.
- [14] Babič, N.Č. and Rebolj, D. Culture change in construction industry: from 2d toward bim based construction. *Journal of Information Technology in Construction (ITcon)*, 21(6):86–99, 2016.
- [15] Kymmell, W. *Building information modeling: planning and managing construction projects with 4d cad and simulations*. McGraw-Hill Construction, 2008.
- [16] Weisberg, D.E. *The engineering design revolution: the people, companies and computer systems that changed forever the practice of engineering*. Cyon Research, 2008.
- [17] Chi, H.L., Kang, S.C., and Wang, X. Research trends and opportunities of augmented reality applications in architecture, engineering, and construction. *Automation In Construction*, 33:116–122, 2013.
- [18] Ahmad, I.U., Russell, J.S., and Abou-Zeid, A. Information technology (it) and integration in the construction industry. *Construction Management and Economics*, 13(2):163–171, 1995.
- [19] Froese, T.M. Impact of emerging information technology on information management. In *Computing in Civil Engineering (2005)*, pages 1–10, Cancun, Mexico, 2005.
- [20] Froese, T.M. The impact of emerging information technology on project management for construction. *Automation In Construction*, 19(5):531–538, 2010.
- [21] D. Cohn. Evolution of Computer-Aided Design: How we got to where we are, and where are we headed. Online: <https://www.digitalengineering247.com/article/evolution-of-computer-aided-design/>
- [22] D. Cunz and D. Larson. Building Information Modeling. Online: http://www.imageserve.com/naples2013/eunder_construction_12_06.pdf
- [23] Howell, I. and Batcheler, B. Building information modeling two years later—huge potential, some success and several limitations. *The Laiserin Letter*, 22(4):3521–3528, 2005.
- [24] Antunes, R. and Poshdar, M. Envision of an integrated information system for project-driven production in construction. *ArXiv Preprint ArXiv:1807.04966*, 2018.
- [25] Eastman, C., Fisher, D., Lafue, G., Lividini, J., Stoker, D., and Yessios, C. *An outline of the building description system*. Research report no. 50, 1974.
- [26] Van Nederveen, G. and Tolman, F. Modelling multiple views on buildings. *Automation In Construction*, 1(3):215–224, 1992.
- [27] Antunes, R. and Poshdar, M. Envision of an integrated information system for project-driven production in construction. *ArXiv:1807.04966*, 2018.
- [28] BIM Alliance. BIM Alliance on BIM. Online: <https://www.bimalliance.se/vadaer-bim/bim-alliance-om-bim/>
- [29] Schweigkofler, A., Monizza, G.P., Domi, E., Popescu, A., Ratajczak, J., Marcher, C., Riedl, M., and Matt, D. Development of a digital platform based on the integration of augmented reality and bim for the management of information in construction processes. In *Product Lifecycle Management to Support Industry 4.0*, pages 46–55, Cham, 2018.
- [30] Succar, B. Building information modelling framework: a research and delivery foundation for industry stakeholders. *Automation In Construction*, 18(3):357–375, 2009.
- [31] A. Carlsen, O. Elfstrand. *Augmented Construction: Developing a framework for implementing Building Information Modeling through Augmented Reality at construction sites*. Luleå University of Technology, 2018
- [32] Rossini, F.L., Novembri, G., and Fioravanti, A. BIM and agent-based model integration for construction management optimization. In *25th Annual Conference of the International Group for Lean Construction*, pages 111–118, Heraklion, Greece, 2017.
- [33] Li, Junjie and Yang, Hui A research on development of construction industrialization based on bim technology under the background of industry 4.0. *MATEC Web Conf.*, 100:02046, 2017.
- [34] Yeutter, M.R. *The use of building information modeling in the electrical contracting industry*. University of Wisconsin-Madison, USA, 2012.
- [35] JBKnowledge. *Construction Technology Report*. JBKnowledge, 2021. Online: <https://contechreport.com/>
- [36] Smith, P. BIM & the 5d project cost manager. *Procedia-Social And Behavioral Sciences*, 119:475–484, 2014.
- [37] O’Keeffe, S.E. *Synergy of the developed 6D BIM framework and conception of the nD BIM framework and nD BIM process ontology*. The University of Southern Mississippi, USA, 2013.

Machine Learning for Construction Process Control: Challenges and Opportunities

B. Sankaran

Scaled Robotics, Spain

bharath@scaledrobotics.com

Abstract -

One of the primary drivers of inefficiencies plaguing construction projects world over is the lack of efficient process control in the construction process. This paper explores the idea of construction process control through BIM based quality control. It discusses the challenges of practically implementing BIM based quality control on a construction site and the opportunities for innovation in this space using state-of-the-art Machine Learning methods.

Keywords -

BIM; 3D Machine Learning; Laser Scanning; Process Control

1 Introduction

Construction, one of the worlds largest industries is also one of its most wasteful and inefficient [1, 2]. One of the major reasons the industry is unable to curb its waste and inefficiencies is because there is no quick and easy way to compare what is being built (reality) to what was designed (BIM). In industrial manufacturing settings this method of comparison is termed *Manufacturing Process Control* [3] whereas in construction settings it is called *Project Control* [4]. This inability to do process control in construction, limits the ability to reduce waste and inefficiency. This article proposes a paradigm for construction process control (project control) which compares the BIM to the onsite construction process using reality capture and machine learning (ML). In the rest of the paper, the general problem of process control, the challenges associated with the data inputs in construction, and the potential opportunities to deploy ML based solutions are addressed.

2 Process control in construction

The process of building construction is a specialized form of onsite manufacturing where a 3D model is recreated in physical space using machines, material and labour. Though construction bears many similarities to other onsite manufacturing processes like automotive, electronics, etc.; there exist stark differences between these processes and construction that affect the practical implementation of efficient process control.

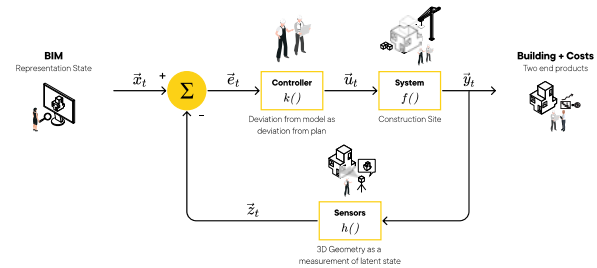


Figure 1: BIM based construction process control

List of Constraints

- Extremely **non linear state evolution** of the building under manufacture due to lack of correlation between states in the state vector. For example, the state (stage of construction) on floor A seldom informs the state on a different floor B within the same building.
- The **lack of correlation** results in different parts of the building having **different rates of evolution**, making the process model extremely complex.
- As the building is also the warehouse, it becomes an **open system for process control**, that is plagued with clutter and other detractors. This results in the building state being only **partially observable**.

These issues make the problem of implementing process control in construction challenging. Section 2.1 will discuss how one can implement process control in construction by measuring variables that can capture the complex latent state space of the construction process that involves material, machines, labour and schedule, without directly measuring them.

2.1 Construction process control via proxies

Since the actual state of the construction process is not directly measurable, one can use the constructed onsite geometry of the building as a proxy for the construction process. It is reasonable to assume that if the physical geometry of the constructed structure is incorrect (not in accordance with the BIM), the underlying process has not

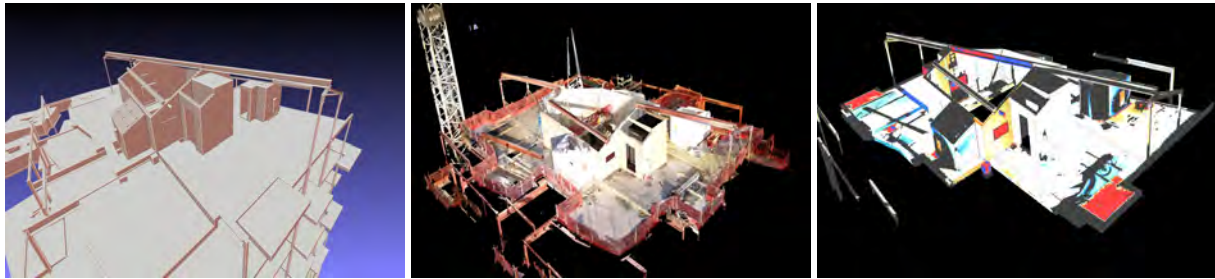


Figure 2: From left to right figure shows the idealized BIM model, the reality captured on the construction site using laser scanning and the point wise divergence between BIM and reality. The laser scans from the construction site are susceptible to both low and high frequency noise.

evolved accurately. Leveraging the BIM, one has access to a 4D signal (3D building geometry evolving over time) of how the ideal construction process should evolve. This can be used as a reference signal for the construction process control problem. For the actual signal of the construction process, one can measure the onsite constructed geometry using laser scanners or similar devices. This process is illustrated in Fig 1.

With this raw information, the physical 3D geometry of the built structure can be reconstructed using structure from motion techniques [5] like SLAM [6]. This 3D geometry once transformed into the reference coordinate space of the BIM, provides a canonical form for comparing the BIM to reality. The divergence between the BIM and reality can be used as the feedback error signal to perform process control on the construction site. This error signal contains both the error in the constructed geometry and the location of this error within the built structure (state space of construction). This error is illustrated in Fig 2.

3 Challenges with the quality of data

This section highlights how input data quality can be a detriment for implementing process control. Both the reference signal (BIM) and the onsite process signal (reality capture) are generally corrupted and noisy.

3.1 Divergence between BIM and reality

Though the difference between the BIM and reality can be used as a representation of divergence in the construction process, this is seldom the case in the real world. BIMs are generally not good reference signals as in practice they are seldom an exact representation of the intended on site geometry [7, 8]. Most practical implementations utilize a LOD 200 or 300 (level of detail) which is often a best case approximation of the constructed geometry. This results in situations where differences between the BIM and reality are not always errors in construction. Some reasons of divergence are:

- Prefabricated elements have a different level of manufacturing detail than their CAD modeled counterpart. This is highlighted via an architectural section drawing in left most figure in Fig: 3, where the constructed element is circular (represented by red lines) but BIM element is a low polygon CAD element.
- Elements during installation may differ entirely from the BIM, or not be represented at all in the model geometry. For example, in the center Fig: 3, a generic sewer placeholder prism is modeled in the BIM that does not match its location in reality.
- Construction elements often suffer *organic* forces and changes, which result in plastic deformations that are not represented in the CAD model. For example in the rightmost figure Fig: 3, non rigid deformations are encountered in concrete slabs due to flexion forces.

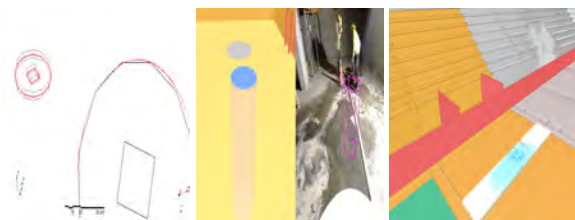


Figure 3: The figures above portray common divergences between the BIM and reality. From left to right the figures show the divergence between BIM and prefabricated elements, the installation differences between BIM and reality, and the divergence due to plastic deformations

3.2 Domain specific low frequency noise

Unlike an industrial manufacturing site, a construction site is also its warehouse. Construction sites, by nature are scenes with a significant amount of clutter, tools, auxiliary props, etc. These onsite artefacts naturally translate to occlusions and spurious measurements in the pointclouds captured by reality captured devices as shown in Fig: 4.

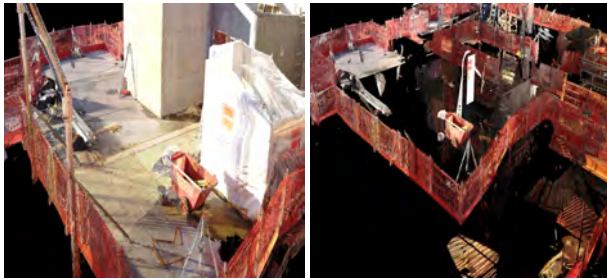


Figure 4: Original pointcloud and the segmented low frequency noise: segmented clutter, props, tools, etc

This unmodeled low frequency noise can easily lower the quality of data. The final sparse signal (pointcloud) after filtering low and high frequency noise is shown in Fig: 5.



Figure 5: Original pointcloud and the filtered signal cloud that can be used to compare to BIM after segmentation

3.3 Into-the-wild sensor acquisition

Another stark contrast between industrial manufacturing settings and a construction site is that, given the open nature of the environment, a construction site is susceptible to uncontrolled events such as bad weather (rain, snow, etc.); bad illumination conditions (Fig: 6) and exposure to materials (glass and reflective surfaces) that cannot be measured with current depth sensors. This degrades the onsite data captured for comparison to the BIM. An example of snow affecting data capture is shown in Fig: 7



Figure 6: Examples of construction sites with bad illumination conditions due to the nature of the environment

4 Challenges with the quality of labels

This section will highlight the issues related to label quality in the input data.

4.1 Lack of reliable ground truth

Building reliable ML models requires access to accurate ground truth (BIM vs Reality). A practical problem that plagues the acquisition of accurate ground truth labels on construction sites is the hierarchical risk management profile within the industry. Asset owners hire general contractors who hire subcontractors to execute on-site work. This incentivizes risk being pushed down the value chain which results in a feedback mechanism where no single entity on a construction site can provide accurate information (ground truth) regarding the state of the construction project. Resulting in very inaccurate mappings between the BIM and reality.

4.2 Lack of consistent semantics

Another source of poor label quality is the lack of reliable and consistent semantics between the BIM and on-site processes. For example, a semantic label such as "wall under construction", can have multiple different geometrical interpretations depending on the type of wall. These poorly defined semantics have distinctly different outcomes when measuring the divergence between the BIM and reality. Moreover domain experts also tend to commonly disagree on the semantics of these divergences.

5 Opportunities for ML in Construction Process Control

Despite the challenges discussed in Sec 3, 4, the problem of process control in construction reveals significant opportunities to employ ML based techniques by exploiting peculiarities in the input data or the problem structure.

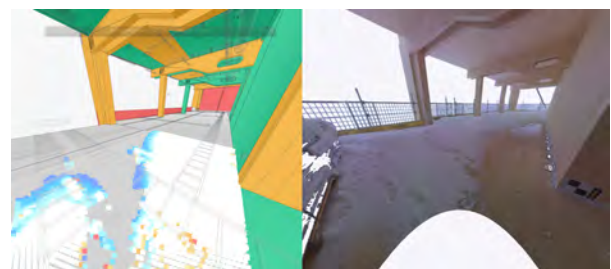


Figure 7: Environmental effects on captured onsite data. The onsite snow interferes with laser returns

5.1 Equivariant learning for construction process control

The lack of consistent semantics discussed in Sec 4.2 are consequences of descriptive characteristics of BIM elements being subject to various transformations such as shift, rotation, scaling etc. This results in the divergence (BIM vs reality) being subject to similar transformations. ML models that are equivariant to group operations can account for such variations. These models can handle variation within the mesh representation of a single element in a BIM or in the entire building graph structure represented by the BIM. At *Scaled Robotics*, we use deep learning models such as PointNet and its extensions [9, 10] to handle equivariance with respect to the point set. We also utilize graph neural networks [11] to model equivariance with regards to the building graph structure.

5.2 Localizing insights for efficient process control

Insights regarding the divergence between BIM and reality also need to be geometrically localized to specific surfaces within BIM elements. As shown in Fig 2, the divergence specific to a BIM element can have its own unique semantic meaning. We extend attention models [12] to work with geometric data to capture these insights.

5.3 Applications for Unsupervised Learning

Construction process data is extremely high dimensional, with non trivial correlations between the inputs. This makes the data domain ripe for the application of unsupervised learning [13] and dimensionality reduction techniques. At *Scaled Robotics*, we utilize both techniques to cluster the divergence to extract specific insights to inform the process control problem.

6 Conclusion

This article introduces the problem of process control in construction and a potential solution to the problem involving BIM to reality comparison using reality capture and machine learning. It also underscores the issues with data and label quality for implementing ML based solutions for construction process control. Finally the article highlights opportunities for exploiting particular threads in ML research that can benefit construction process control.

References

- [1] S. Nagapan and et al. Issues on construction waste: The need for sustainable waste management. *2012 IEEE Colloquium on Humanities, Science and Engineering (CHUSER)*, pages 325–330, 2012.
- [2] S.E. Sapuay. Construction waste – potentials and constraints. *Procedia Environmental Sciences*, 35:714–722, 2016. ISSN 1878-0296. doi:<https://doi.org/10.1016/j.proenv.2016.07.074>. Waste Management for Resource Utilisation.
- [3] National Research Council. *Manufacturing Process Controls for the Industries of the Future*. The National Academies Press, 1998. ISBN 978-0-309-06184-1.
- [4] W.J.D. Pico. *Project Control: Integrating Cost and Schedule in Construction*. RSMears. Wiley, 2013. ISBN 9781118139233.
- [5] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, New York, NY, USA, 2 edition, 2003.
- [6] C. Cadena and et al. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Transactions on Robotics*, 32(6):1309–1332, 2016. doi:[10.1109/TRO.2016.2624754](https://doi.org/10.1109/TRO.2016.2624754).
- [7] M. L. Trani, M. Cassano, D. Todaro, and B. Bossi. Bim level of detail for construction site design. *Procedia Engineering*, 123:581–589, 2015. doi:<https://doi.org/10.1016/j.proeng.2015.10.111>.
- [8] P. et al Uusitalo. Applying level of detail in a bim-based project: An overall process for lean design management. *Buildings*, 9(5), 2019. doi:[10.3390/buildings9050109](https://doi.org/10.3390/buildings9050109).
- [9] C. R. Qi and et al. Pointnet: Deep learning on point sets for 3d classification and segmentation. 2016.
- [10] C. R. Qi, L. Yi, H. Su, and L. J. Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. 2017.
- [11] V. G. Satorras and et al. E(n) equivariant graph neural networks. *CoRR*, abs/2102.09844, 2021. URL <https://arxiv.org/abs/2102.09844>.
- [12] A. Vaswani and et al. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, 2017.
- [13] G. Hinton and T.J. Sejnowski. *Unsupervised Learning: Foundations of Neural Computation*. Computational Neuroscience Series. MIT Press, 1999. ISBN 9780262581684.

5G Wireless Communication for Autonomous Excavation

R. Heikkilä^a & M. Immonen^a & H. Keränen^b & O. Liinamaa^a & E. Piri^c & T. Kolli^a

^aCivil Engineering Research Unit, University of Oulu, Finland

^bSatel Oy, Finland

^cKaitotek Oy, Finland

E-mail: rauno.heikkila@oulu.fi, matti.immonen@oulu.fi; mikko.hiltunen@oulu.fi; tanja.kolli@oulu.fi, olli.liinamaa@oulu.fi, heikki.keranen@satel.com, esa.piri@kaitotek.fi

Abstract

The applicability of wireless 5G network to the information communication needs of autonomous excavations is studied. An autonomous excavator developed by the University of Oulu is used for the experiments in a 5G test network at the Linnanmaa Campus of the University of Oulu. In the experiments, remote and autonomous control methods were used for test drives, in which the operation working was observed and measured. As a reference system, 4G wireless network was used. Remote control worked well in 4G and 5G. In 4G, delay of 1,7 s was observed from joysticks to camera view, while in 5G, the delay was 799 ms. The 5G network delay of 15 ms improvement was measured. Camera and image processing were still the primary reason for delays, approximately about 784 ms. In the 5G test, some jerking movements and steering crashes in autonomous steering were observed. In current 5G networks, uplink speeds are a limiting factor, which requires further attention on future 5G networks and standardization to be better applicable for industry use scenarios.

Latency of the connectivity is an important performance characteristic, even more than throughput. 5G can bring some benefits to autonomous excavation, especially more capacity for real-time information utilization.

Keywords – Wireless network;
Autonomous excavator; 5G

1 Introduction

1.1 Background

5G is the 5th generation of wireless mobile networks, a significant evolution of today's 4G LTE networks. 5G has been designed to meet the very large growth in data

and connectivity of today's modern society, the Internet of Things (IoT) with numerous connected devices, and tomorrow's innovations [1].

On a general level, there are at least three common expectations for the use of the 5G network [2]:

- massive machine to machine communications (i.e., IoT connecting billions of devices without human intervention at a scale not seen before
- ultra-reliable low latency communications
- enhanced mobile broadband (faster data speeds and greater capacity).

Lukau (2014) pointed out that there are many difficult solutions, tasks, challenges, and requirements that need to be faced and successfully resolved in order to heft the mobile telecommunication technology to the next 5G level. [2]

Excavator is one of the most used earth-moving machines with the most hours of use on construction sites. Excavators typically have a lot of different booms, tools, additional equipment, and other systems included or connected to the basic machine frame. The working tasks are quite versatile and challenging to be automated. Most typical infra construction tasks are earthwork cutting and loading. In Finland, based on several interviews with the industrial key players, master class human operators are rare (about 1/15).

Vernersson et. al (2013) have evaluated that in road construction sites and quarries, there are clear improvement potentials with the introduction of wireless communication technologies especially for safety improvements and productivity optimization. In quarries, they suggested that a productivity increase of up to 30% can be achieved assuming a reliable wireless connectivity to minimize waste in the production. They also saw a clear potential in accident avoidance using wireless communication-based warning systems. [3]

Rylander (2021) has studied and experimented wireless short-range Communication Performance in a Quarry Environment with also a focus on earth moving

operations. Rylander [4] points out that there is a clear potential in optimizing site throughput by better coordinating machine movements. Such an optimization requires a site control system with wireless communication to control the speed of each machine. Machines can travel at a more energy-efficient speed with lower fuel consumption instead of moving at maximum speed and then waiting at the destination. The research included the wireless short-range technologies ZigBee, 802.11g, and 802.11p using frequencies of 868 MHz, 2.4 GHz, and 5.9 GHz. In the study, 802.11p reached a range of 1767 meters in Line-of-Sight (LOS) conditions. The 802.11g performed well at the quarry top but gave a poor result in the quarry pit where it lost the most packets out of all standards. The 802.11g delivered an unstable communication in the pit and was sensitive to Non-Line-of-Sight (NLOS) areas and obstacles, and was unable to benefit from reflecting surfaces. The ZigBee RF provided the best results (only two data packet losses were observed, the maximum range of operation was 1753 m) and seemed to be unaffected by NLOS. In his study, Rylander [4] also evaluated the potentiality of fully automated machines and machine fleets and concluded that continuous activity measurements and feedback loops are required in the control process.

In Finland, infra construction machines with automatic machine control systems have been tested and used using commercial operator networks, 3G and 4G wireless networks, and/or using different Wifi network solutions. So far, we have not yet found any experiments with autonomous work machinery in 5G network. It is clear that the higher the level of automation the control of mobile machines are developed, the more speed, capacity, and reliability are required for the wireless data transmission required by the control. [6][7]

1.2 Aim

The aim of the research was to study and experiment the applicability of the wireless 5G test network to the information transfer demands of autonomous excavation work process.

2 Materials and Methods

2.1 Autonomous Excavator

In the study, we used the autonomous excavator developed by the University of Oulu (Smart Excavator, Bobcat E85) [5], which is presented in Figure 1 as well as Figure 2 with omni antennas on the roof of the excavator. In Figure 3 is showed how VR glasses can be utilized to the remote control of the excavator. The

control of the Smart Excavator is based on Open Infra Building Information Models (Open Infra BIM) that provide the needed information for machine control systems. Current types of automatic 3D machine control systems are presented in Table 1: (1) remote (human operates machine remotely), (2) guidance (human operates manually machine using computer user-interface to BIM model); (3) coordinated (human operates the machine and manages the tool blade manually with the help of inverse kinematics), (4) partial automation (human moves the machine and controls some of the blade movements while the system drives automatically the other movements), (4) teach-in (human operator drives a series of movements, which are recorded and can later be automatically repeated), and (5) autonomous (machine can operate without human operator).

Table 1. Levels of Automation for Infra Construction Machinery (University of Oulu).

Level	Name	Description of the activity
0	No automation	Human operates machine
1	Remote control	Human operates remotely machine
2	Guidance	Operator supported, the operator drives manually machine and blade using computer user-interface to BIM model
3	Coordinated	Tip control, the operator moves the machine and manages the tool blade manually with the help of inverse kinematics
4	Partial automation	Controlling, the operator moves the work machine and manages the part of the tool blade manually while the system drives automatically some of the movements
5	Autonomous	Machine can operate without human driver
6	Autonomous machine swarm	Autonomous operation of work machines, interactivity and collaboration of working machines

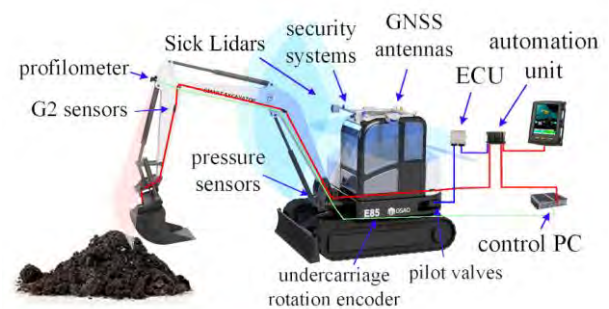


Figure 1. Smart Excavator of the University of Oulu – hardware, sensing, and control systems. [5]



Figure 2. Smart Excavator of the University of Oulu (top), Omni directional antenna on the top of the machine (below).



Figure 3. Remote control with VR glasses from the University of Oulu campus to the test site.

2.2 5G Test Network (5GTN) and equipment

5GTN is an R&D project with several partners in Finland [1]. The aim of 5GTN has been to provide access to the test network and related expertise to interested third parties. In the experiment, the used band was n78 (3500 MHz) with non-standalone (NSA) architecture for connecting excavator to the network. Modem integration to the Smart Excavator for 5GTN connection was developed as bachelor's thesis (Joona 2021). In the experiments, the 5G link was for connecting Smart

Excavator, which was controlled from control cabin by wired (Figure 5).

The UE (User Equipment) used in the experiments was Vehicle Connectivity Unit (VCU) provided by Satel Oy. The UE is a robust machine connectivity unit designed to be used in heavy machinery. The modem supports theoretically DL 2.5 Gbps; UL 600 Mbps and LTE DL Cat 16/ UL Cat 18 speeds.

The test setup required the connectivity between the control PC and Lidar and Camera receivers in the cabin to be connected to the excavator internal network using L2 Ethernet. The solution to connect from the remote to the excavator over Ethernet was to establish an OpenVPN connection from the control cabin to Satel's VCU. The algorithms used were SHA256 for authentication and AES128-CBC for encryption. The OpenVPN connection was also established without encryption or authentication for performance comparison between non-encrypted, non-authenticated, and fully secured connectivity.

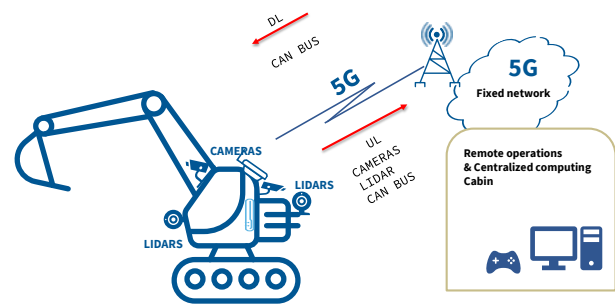


Figure 4. Network Schematics used in the 5G experiment (Satel Oy).

The quality of the application traffic from video streams, LiDARS, and CAN bus over the network connection was measured by Kaitotek's Qosium solution. Qosium is a passive and real-time network Quality of Service (QoS) measurement software. It measures QoS for the real ongoing application traffic over the network path of interest without creating any artificial test traffic to the network.

The measurement setup was a two-point measurement necessary for measuring QoS. We deployed a Qosium measurement agent on the remote controlling side in the end device. In the excavator, the measurement agent was running in an Ethernet bridge device through which all application traffic flows, being deployed right beyond the 5G router device.

2.3 Experiments

Real excavation work was not allowed at all in the Linnanmaa campus area. The excavator and the control booth were transported by truck to the campus. The excavator was operated on the 5G network (Fig 5) using two different automation methods, i.e., 1) remote control and 2) model-based autonomous control. Both automation methods were carried out from the nearby control booth. Plane excavating above the asphalt was selected for the model-based autonomous control. (Fig 6)

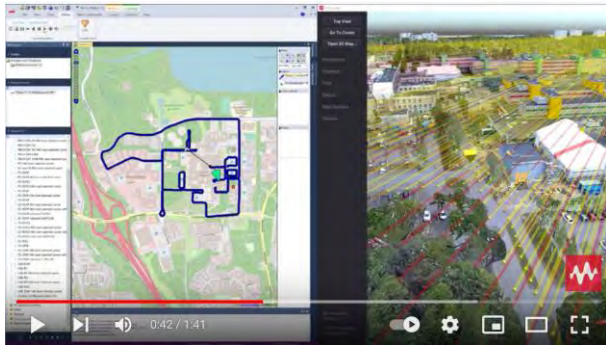


Figure 5. 5G Test Network at Linnanmaa Campus of University of Oulu.

During both control methods, lidar data and camera stream were transferred with the excavator control signals. Lidar data was sent by burst, and the camera stream was continuous. Data transfer rates and delays were measured with Kaitotek's Qosium and Wireshark software, and the results obtained from 4G (reference) and 5G networks were compared. Both control methods were also driven with and without encryption. The functionality of the automatic control was observed during the experiments.



Figure 6. Remote and Autonomous Control Operation.

The measurement solution used enabled us to see the performance of the 5G connection directly to the applications relevant to the remote control. The measurement solution only factors in the connectivity effect for the end-user experience. Thus, for example, video encoding and decoding latencies are not considered in the measurement results.

The statistics collected comprised QoS statistics, including delay, jitter, packet loss, and connection break. In addition, traffic and flow statistics showing, for example, bitrates, packet sizes, and packets per second were also measured. All statistics are one-way, meaning that the results are provided for the traffic originating from the control desk to the excavator and vice versa separately.

3 Results

Remote control worked well for both 4G and 5G connections. We measured, on average, 12 ms connection delays for the control signal traffic, with no significant difference between the access technologies. However, we observed a 1.7 s latency from joystick input to be shown on the camera view in 4G. In 5G, the latency was, on average, 799 ms. Thus, 5G performed much better in delays when there was more traffic over the network. The 5G network delay improvement of 15 ms was measured. Camera and image processing were still the primary reasons for latencies, approximately $799 \text{ ms} - 15 \text{ ms} = 784 \text{ ms}$.

The moving excavator cabin caused the 5G user plane of the employed Non-Standalone (NSA) 5G connection periodically to switch to a 4G network, as visible in Figure 7 (e.g., 5G RSRP values going to zero). Handover to 4G produced increased delays to the excavator control signals. This behavior was most likely due to the azimuth pattern of the used antenna. Using two or more antennas at different orientations could have prevented the observed effect.

In these experiment conditions, 5G throughput capacity measurements with the Iperf traffic generator resulted in only 24 Mbps and 75 Mbps for uplink and downlink, respectively, when no data encryption was used. With encryption, 17 Mbit/s (uplink) and 27 Mbit/s (downlink) were measured. The actual control signal + camera throughput averaged 3Mbit/s and with lidar data peaked to a maximum of 9 Mbit/s. Control signals delays median with and without encryption was about the same (30 ms with encryption and 28ms without). However, there was observed a significant difference with the automation method when using encryption with the loaded lidar data bursts. We observed some jerking movements and a steering crash in automatic steering,

which ended up the safety person pressing an emergency stop. Available bandwidth for 4G anchor frequency was 10 MHz (on B7) and equivalent 5G allocation was 60 MHz.

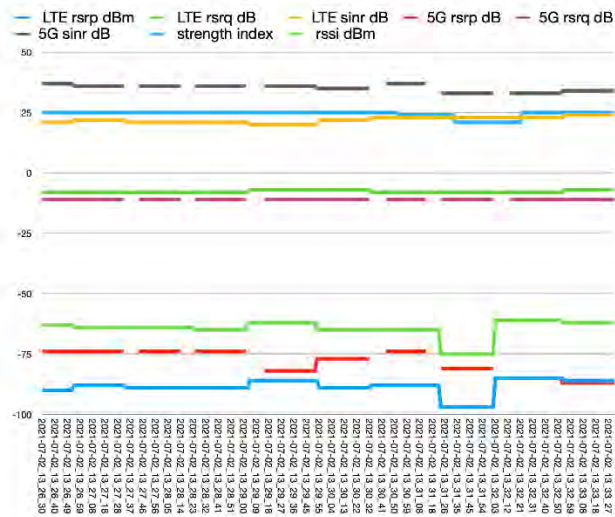


Figure 7. 5G Connectivity Stability (on a vertical scale signal strength, horizontal scale time).

4 Conclusion

Infrastructure construction is typically done on large-scale construction sites, with lengths ranging from kilometers to over 100 km. In addition to work machines, various vehicles, people and unexpected moving objects such as wild animals move on the construction sites. Unexpected work can also happen during the excavation work itself, the truck may drop more material than expected, or the cut ramp may collapse on the excavated cutting floor. As the control of work machines moves to an ever higher level of automation, the requirements of real-time are emphasized for reasons of efficiency and safety. An autonomous, unmanned excavator can basically work with its own sensors and control logic. However, in BIM based control, it must be possible to update the machine control model at any time using the real-time cloud service. To ensure work safety, the work machine is likely to have a continuous real-time wireless connection to the construction site safety system and associated sensors. Job control requires a real-time location of the implement and an accurate snapshot of the work done. Effective cooperation between automatic machines requires a continuous real-time exchange of information between the machines. This enables efficient and secure synchronization of work fleet work processes.

For remote control and automated movement detection, camera sensor and processing delays are

significant in relation to network delays. Encryption processing performance is an important factor, too, especially on higher traffic loads. Encryption and decryption processing delays directly impact the end-user experienced latencies of applications. In current 5G networks, the uplink performance is the limiting factor that affects industrial usage. Most traffic is typically sent to the uplink direction in industrial use scenarios, including, e.g., camera streams. Available bandwidth for experimentation system is also limited compared to commercial networks.

For a remote controlling use scenario with real-time applications, network delays are more relevant to assessing the performance and quality of the network than maximum throughput capacity measurements. 5G can bring some benefits to autonomous excavation, especially more capacity to real-time information utilization. At the moment, 5G, however, needs careful planning and practical testing before moving to operative usage. The best scenario in operative usage is continuous monitoring of connection quality for the mission-critical real-time applications to detect connectivity disruptions before impacting efficiency and safety.

On remote infra construction sites, 5G can still meet challenges for sufficient coverage and especially uplink quality. On some large sites, it is likely that some supporting and backup networks are needed, e.g., WiFi or 4G LTE. In highly automated excavation, it will be essential to consider work management and division between central and edge.

References

- [1] Kanstrén, T. & Mäkelä, J. & Uitto, M. & Apilo, O. & Pouttu, A. & Liinamaa, O. & Destino, G. (2018) Vertical Use Cases in the Finnish 5G Test Network, European Conference on Networks and Communications (EuCNC).
- [2] Lukau, E. (2014) 5G Cellular Network for Machine to Machine Communication. Fraunhofer University, Researchgate.net. DOI: 10.13140/RG.2.1.4193.3520
- [3] Vernersson, S. & Kalpaxidou, E. & Rylander, D. (2013) Evaluation of Wireless Short-Range Communication Performance in a Quarry Environment. IEE International Conference on Connected Vehicles, Las Vegas, USA, pp. 308-313.
- [4] Rylander, D. (2021) Productivity Improvements in Construction Transport Operation through Lean Thinking and Systems of Systems. Mälardalen University Doctoral Dissertation 350. 166 p.
- [5] Heikkilä, R. & Makkonen, T. & Niskanen, I. & Immonen, M. (2019) Development of an Earthmoving Machinery Autonomous Excavator Development Platform. ISARC 2019, The 36th

- International Symposium on Automation and Robotics in Construction, Banff, Alberta, Canada.
- [6] Kilpeläinen, P. & Heikkilä, R. & Parkkila, T. (2007) Automation and Wireless Communication Technologies in Road Rehabilitation. ISARC'2007, The 24th International Symposium on Automation and Robotics in Construction, 19-21 September 2007, Kochi, Kerala, India, pp. 35-40.
- [7] Viljamaa, E. & Kilpeläinen, P. & Pentikäinen, V. & Sarjanoja, E.-M. & Heikkilä, R. (2009) On-line Process Management of Pavement Laying Using Wireless Communication Technologies. ISARC'2009, The 26th International Symposium on Automation and Robotics in Construction, 24-27 June 2009, Austin, Texas, U.S.A., pp. 348-356.

The Use of Drones in the Construction Industry: Applications and Implementation

M.B. Hatoum^a and H. Nassereddine^b

^aDepartment of Civil Engineering, University of Kentucky, USA
E-mail: mbh.93@uky.edu, hala.nassereddine@uky.edu

Abstract –

In a volatile, uncertain, complex, and ambiguous world, Industry 4.0 can be the silver lining of the construction industry which has been facing daunting chronic problems for decades. The transformative impact of Industry 4.0 which has been well documented in the industrial sector has spurred interested among construction researcher to explore the opportunities Industry 4.0 can have on our industry. Among the various Industry 4.0 technologies, drones have found their way into the construction industry and their use has been on the rise. Building on the existing body of knowledge, this paper summarizes the current state of adoption of drones in the construction industry. To achieve the research objective, the various applications of drones throughout the construction project lifecycle are explored, associated benefits, challenges, and costs are identified, and considerations to be accounted for are discussed. The findings summarize the “5W2H” – or the What, When, Where, Why, Who, how, and How Much – of the use of drones in the construction industry.

Keywords –

Construction 4.0; Technology; Drones; Unmanned Aerial Vehicles; UAV; Unmanned Aerial System; UAS; Remotely Piloted Vehicles; RPV; 5W2H

1 Introduction and Background

The fourth industrial revolution, known as Industry 4.0, has been steadily reshaping industries around the world. With the fast advancement in technological capabilities, industries are continuously adopting technologies into their workplaces. The construction industry is no exception. Despite being a late comer, construction is gradually utilizing the different technologies that Industry 4.0 offers, including augmented and virtual reality, robotics, big data, sensors, and 3D printing [1–4]. Another technology gaining momentum is drones – otherwise known as unmanned

aerial vehicles (UAVs), unmanned aerial system (UAS), and remotely piloted vehicles (RPVs).

For the past few years, the use of drones in the construction industry has been on the rise, and the global construction drone market size was valued at \$4.8 billion in 2019 [5]. In the United States of America (USA), the use of drones in the construction industry is controlled by the Federal Aviation Administration (FAA) regulations for the commercial use of UAS published in 2016 [6]. The regulations boosted the use of drones in both vertical and horizontal construction projects for both public and private construction sectors, eventually making drones one of the major innovations in the Every Day Counts program conducted by the Federal Highway Administration [7].

Drones were also helpful in the COVID-19 pandemic. With the move to remote working for project stakeholders and safety regulations like social distancing and decreased number of labor, drones played an important role in providing real-time data and allowing work to continue despite restrictions on worksites [8]. Drones were also used for crown surveillance and enforcing social distancing, broadcasting public announcements, spraying disinfectants, and delivery for medical supplies and essentials [9].

2 Objective and Methodology

Following the rise of drones in the construction industry, numerous efforts have been undertaken to investigate the use of this technology and its impact on construction. Using the “5W2H” method, a technique employed to analyze and characterize in a concise and reliable way the most important information about a matter, this paper builds on the existing body of knowledge where academicians and practitioners publications on drones (*what*) were reviewed to (1) explore the applications of drones through the different phases of the lifecycle of the construction project (*when, who, and where*), (2) identify the benefits and challenges associated with those applications (*why*), (3) provide an overview of the lifecycle cost (*how much*),

and (4) present common practices to implement drones (*how*). The “5W2H” framework of drones in the construction industry is illustrated in Figure 1 to summarize the state of drones for both academicians and practitioners in the post-pandemic world, especially after drones demonstrated great value before and during the COVID-19 pandemic. The 5W2H method was employed to offer a holistic discussion of drones in the construction industry, a type of discussion that has not been presenting in one single paper yet.

To identify relevant research work, Google Scholar was used to identify articles, journal papers, dissertations, and conference proceedings published between 2017 and 2021. Search filters included various keywords to tackle the two layers, such as (“Drones” OR “UAS” OR “UAV” OR “RPV”) AND (“Construction” OR “Construction Industry” OR “Construction Project” OR “Project Lifecycle” OR “Adoption” OR “Implementation” OR “Challenges” OR “Barriers” OR “Benefits” OR “Advantages” Or “Use Case”). The relevant number of papers after scanning titles, abstracts, and the content was 31. In addition to literature, websites for major drone companies that work with construction firms were searched to identify relevant success stories and use-cases, webinars, reports, and publications that can provide insightful observations from the industry perspective. The search resulted in 15 resources which were found relevant to this study.

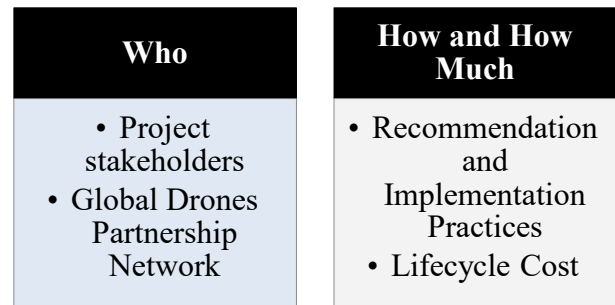
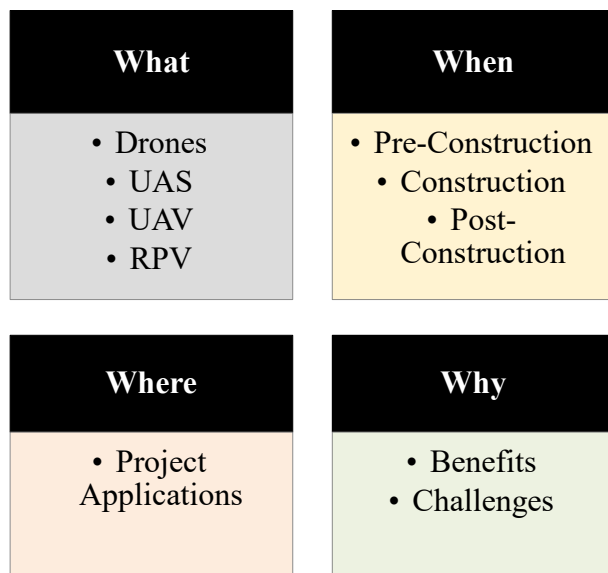


Figure 1. The “5W2H” (what, when, where, why, who, how, and how much) of drones in the construction industry.

3 Applications of Drones Across the Project Lifecycle

Drones can be used across the entire project lifecycle – starting from the early stages of pre-planning and design, passing through the construction stage, to the late stages of asset management, operation, and management. This section elaborates on the major applications for drones across the different phases of a construction project. While there is no single definition of what constitute the lifecycle of a construction project, this paper adopts the following three phases that are commonly referenced in drones related research: pre-construction phase, construction phase, and post-construction phase [10–12].

3.1 Pre-Construction Phase

The pre-construction phase of a construction project includes all stages that are needed before the physical construction starts. During this phase, drones can be used for designing, bidding, and site planning.

In the design stage, the use of drones can provide designers with a good understanding of the project environment. Designers can also benefit from the geometrically accurate 3D scans that drones offer. The 3D scans can be integrated into the modelling software to account for landscape reality when modelling the project [13]. This could help designers and stakeholders visualize the project on its surrounding, and assess how the project could impact its environment from a practical and aesthetic viewpoint [14].

In the bidding stage, drone images can provide contractors with a detailed view of the site location and its surrounding to expose possible risks and hazards that should be accounted for while creating their bid. Moreover, drones can assist in the correction and rendering of terrestrial surveying information to provide geometrically accurate 3D maps for estimating surveying quantities such as earthwork volumes [15]. The accuracy of such measurements can improve the

reliability of the bids and decrease possible estimation errors.

As for site planning, the high accuracy of the drone data allows for an accurate identification of site boundaries and presentation of the construction plot area [16,17]. Construction teams can then use this data to plan different site aspects such as entrance and exit points, crane locations, and material storage areas. This becomes especially important when the project is surrounded with restricted areas and no-go zones such as existing utilities and environmentally protected lands [18].

3.2 Construction Phase

The construction phase of a project includes all the physical construction that happens on the construction site. Drones can serve multiple purposes.

To begin with, drones can be used for safety and health purposes. Safety managers can fly drones over the construction site to collect real-time information on hazardous activities, blind spots from moving equipment, and breaches of site parameters and no-go zones [11,19]. Drones can also carry built-in microphones to allow safety managers to contact workers who are in danger and warn them about potential hazards [20]. Successful attempts were also performed on drones to automatically recognize and warn workers who are not wearing personal protective equipment such as hardhats, and detect opening openings and guardrails that do not conform with the Occupational Safety and Health Administration (OSHA) standards [12,21].

Drones can also be used for resource and logistics management. Drones can provide real-time measurement of the volume of materials on site such as soil and aggregates [16]. Drones can also use real-time location systems to detect, identify, and track the movement of materials and equipment tagged by different sensing technologies [22]. Moreover, drones can serve as assistant tools for unmanned equipment, such as excavators with autonomous intelligent control units, to perform earth moving operations [23].

Another use for drones is to track progress on the construction sites. Drones manoeuvring over the jobsite can display the progress happening in real time for stakeholders whether on-site or away [10]. This creates a timeline of information that can be used to control and validate tasks, provide quality control, compare what is built to what was planned to identify differences between plans and real-time progress, and create 3D as-built models [12,18,19,24].

In addition to 3D models and progress reports, footage gathered by drones can prevent a permanent record of the project that can be referred to at any time. Disputes are very common in the construction industry,

so stakeholders can refer to the recorded drone footage when problems arise for a comprehensive understanding of the situation [25]. Drone footage can also be useful evidence to resolve litigation since they “can clarify or explain oral testimony or documentary narrative in concrete terms” [18,26].

Other uses of drones on the construction project include transporting and installing construction materials like bricks, foam blocks, and steel elements [27–29]; transporting labour tools [23]; illuminating dark areas in the construction site and providing lighting for night works [23]; finishing and painting walls and vertical slopes [30]; monitoring dust and quality of air and detecting environmental violations [23]; inspecting existing structures before repairs or remodelling [12]; and managing demolitions or destruction of structures [31].

3.3 Post-Construction Phase

The post-construction phase of a construction project includes all stages happening after the physical construction ends. During this phase, drones can be used for different mapping, marketing, monitoring, and inspecting purposes.

Drones can be used to map inaccessible and hazardous areas, especially with their ability to enter areas and reach certain heights that can be hard to capture manually or by laser scanners [32]. Thus, the bird’s eye view that drones can provide, when integrated with advanced scanning technology, allows stakeholders to create 3D models of existing structures such as historic landmarks [32]. Moreover, the unique viewpoint that drones provide allow project owners and real estate agents to create unique videos for their constructed projects for marketing purposes, sales, and promotions, especially for commercial projects [33].

Drones can also play a vital role in monitoring and inspecting structures. For example, drones can be used for inspecting and cleaning facades of skyscrapers, detecting corrosion in dam slopes, inspecting hazardous roofs such as those made with slate or clay tile and steep-slope roofs, and identifying roof leaks or electrical hot spots on transformer installations that are not ground-visible to detect inefficiencies in energy usage and develop thermal maps [23]. Moreover, [34] identified several studies that investigated the use of drones in highway infrastructure management, both bridges and roads. For bridges, the use of drones can enhance monitoring and inspecting bridges due to the drones’ ability to manoeuvre above, under, and around bridges without the need of heavy equipment or traffic control. As for roads, drones can enhance the periodic basement of pavement conditions and distress monitoring by reducing the man-hours needed or visual inspection and conducting in-situ tests to measure

different aspects such as distress, unevenness, rutting and cracks [34]. Asphalt 3D printing technologies could also be installed on drones to fill cracks in pavements [35].

Another major use-case for drones is post-disaster reconnaissance. Fires, explosions, and natural disasters such as earthquakes, floods, hurricanes, tornadoes, and landslides can cause unavoidable damage to buildings and infrastructure. Drones can be used to perform comprehensive damage assessment through identifying damage evidence ranging from cracks to complete collapses, and enhance the processes of rescuing, cleaning, rehabilitation, and retrofitting [12,23]. Damage evidence can also be helpful when filing for post-disaster insurance compensations [36].

4 IMPLEMENTATION OF DRONES IN CONSTRUCTION PROJECTS

4.1 Benefits

The different applications of drones across the construction project lifecycle discussed earlier highlight major social, economic, and environmental benefits.

From the social perspective, the ability of the drone to manoeuvre over hazardous areas and dangerous heights for important tasks like inspection or surveying makes the technology a safer alternative than human labour. Land surveyors and inspectors can fly drones over their areas of interest instead of performing their surveying or inspection tasks physically and getting exposed to risks [12]. Drones can also add an extra layer of security to the job site through identifying breaching of parameters or entrances to no-go zones faster [19]. This becomes especially important for site visitors and pedestrians commuting around the jobsite. Moreover, the use of drones for inspecting highway infrastructure such as bridges, roads, and tunnels, can decrease the hardships of traffic control [37].

From an economic perspective, the different use-cases of drones – discussed in the applications section above – show that drones can (1) simplify reporting through capturing and footage in real time for project stakeholders present inside and outside the job site, (2) enhance decision making through remote site access, (3) provide accurate calculation tasks such as earthwork volumes, (4) track progress on construction site with high accuracy and compare it to design to develop as-built statuses, (5) track the movement of materials and resources anywhere on the job site, (6) replace the need for heavy equipment and truck cranes or elevating platforms to inspect hazardous area, and (7) reduce liability through resolving disputes faster [16,18,21,23,32]. All those benefits translate into productive job sites with efficient communication and

collaboration among stakeholders, and significant time and cost savings in a construction project [12]. Moreover, the use of drones allows companies to gain competitive advantage whether when bidding on projects or seeking new clients.

From an environmental perspective, drones do not rely on fossil fuels and are mostly electric-motor driven. Thus, drones release significantly lower levels of carbon dioxide emissions when compared to construction equipment, making them an environmentally friendly alternative for aerial work – i.e. land mapping, photography, and surveying [12]. Further environmental benefits are also perceived through reducing congestion, for drones will decrease the number of human labour and equipment needed to perform critical tasks [28]. Moreover, the use of drones can contribute to sustainable development and construction when monitoring energy projects like pipelines, wind turbines, or solar farms [12].

4.2 Challenges

Despite the different benefits perceived from the use of drones, several challenges remain at the technical, social, and legal standpoints.

From a technical standpoint, technological constraints such as the limits on the battery life can shorten the time of flights and operations [38]. Battery life is affected by factors such as the take-off weight, which in turns affects the types of cameras and sensors that can be used [34]. Another technical constraint is the effect of environmental factors like severe weather conditions and rough terrains; terrains like mountains and hills or underground tunnels and severe weather conditions like extreme temperatures, accumulation of ice, high wind speeds, and heavy rains may hinder signals and connectivity [34,38]. Moreover, drone take-off location and flight paths should be away from large metal objects or heavily reinforced concrete structures since electrical sensors like gyroscope and magnetometer can be affected by communication interference if there are magnetic sources around the drone [39]. Furthermore, even though researchers are studying the ability of drones to perform aerial drone-based construction tasks, material such as bricks or blocks or voxels might require to be altered to meet the drones capabilities [40]. Finally, while drones excel at outdoor environments. Several limitations remain for using drone in indoor environments with the absence of GPS and lack of predefined parameters necessitate more intricate sensing abilities and sophisticated supporting algorithms [38]. To overcome such technical issues, cross-industry cooperation is required between the manufacturer and client to enhance product research and development and improve the capabilities of drones for the construction industry [23].

From a social standpoint, operating drones comes with the risk of collisions with people, objects, and other aerial vehicles [34]. Many reasons can cause the drone to malfunction such as the sudden discharge of the battery, unpredicted weather occurrences like unknown winds aloft, unexpected hazard events like forest fires or earthquakes, loss of command and control link, loss of visual contact with the drone, and loss of navigation control [37]. Even though the probability of such risks may be low, they become more possible when using drones on large scale entails [41]. Noting the labour intensive nature of the construction industry and the critical location of many construction sites in populated areas, drone malfunctions may lead to crashes crash into on-site personnel, pedestrians, vehicles, or structures resulting in bodily injuries and/or property damages [38]. Other social challenges include security and privacy concerns. Drones are equipped with cameras, night vision device, and various sensors that give it the ability to track objects and observe them in different perspectives. Such capabilities can facilitate snooping and capturing unauthorized images of persons or private property, which in turn raises privacy concerns from citizens [38,41]. Moreover, many law enforcement agencies have voiced security concerns on the use of drones since drone operations can be vulnerable for GPS jamming and hacking, making drones prospect for malicious users to conduct cyberterrorism and other unlawful activities [9]. Such social challenges may be maintained by strong security firewall systems and ethical guidelines for drone operations to clearly define the objectives of the flight, the data that should be gathered for analysis, time of data residencies, possible privacy violations, and proper risk and safety mitigation plans and strategies [23].

From a legal standpoint, different aviation rules may restrict the potential of drones. For example, in USA, while the FAA regulations are important to guarantee the safety of people and property, other regulations may cause operation restrictions. Examples of restricting regulations include: the weight of the drone including payload at take-off should not cross 25 kilogram (55 pounds), flying at or under 120 meters (400 feet), flying at or under 160 kilometres per hour (100 miles per hour), staying 8 kilometres (5 miles) away from airports, remaining in Class G airspace, keeping the drone within the visual line of sight of the operator, and yielding the right of way to other aircrafts. It is worth to note that restrictions, aside staying in Class G and the take-off weight of the drone – may be waived by the FAA upon submitting a waiver. It is also worth to note that the restrictions of flying over moving vehicles, flying over people who are not protected by a shield and/or part of the operation, and flying at night were lifted by the FAA as of April 2021 [6]. Other legal challenges come in the

form of liabilities [26]. The absence of common industry standards for risk allocation and mitigation may cause potential liabilities on who might be responsible for the incurred injuries or property damages in the case of a drone malfunction [26]. This becomes especially important because the FAA does not mandate insurance for drone operations, but requires operators to report all drone incidents, which may lead to revoking licenses and jeopardizing the company from using drones in their operations [26]. Thus, insurance companies should develop policies to clearly define and integrate drone usage into the insurance coverage of construction projects and create ideal insurance cost models that account for the accident rates and safety risks for drone operations [23].

4.3 Lifecycle Cost

The lifecycle cost of implementing drones can be broken down into three phases: procurement and acquisition, operations and management, and retirement and disposal [42]. The procurement and acquisition phase is related to the permits needed to buy and operate the drones, and the cost of purchase. Factors that affect the cost of the purchase include the drone's weight (i.e. battery, propellers, motors, controllers, frame, take-off elements), its payload capacity (i.e. gimbles, cameras, sensors), and endurance (i.e. battery service) [42]. Other cost factors include the required software for navigation and flightpath automation, to control, monitor, communicate, and navigate the drone, in addition to gathering, storing, and sharing data [42]. The operation and management phase is associated with incurred costs to operate the drones at regular intervals of time. Examples of operation costs include the lifespan of propellers and batteries, software adaptability, insurance, and operating crew [42]. As for the retirement and disposal phase, it is derived from the recuperated capital at the end of the drone's lifespan [42].

5 Implementation Practices and Recommendations

Several companies have published success stories of the use of drones in the construction industry, whether through webinars, academic papers, or blogs. A number of those resources were reviewed to comprehensively compile a list of common practices needed for implementing drones [12,36–38,43–48,50]. It is worth to note that the practices should be treated as guidelines instead of an official standard for implementing drones.

To begin with, construction companies should understand the need for drones before starting the implementation process. As seen in the applications section, drones have multiple purposes across the entire

lifecycle of a project. Thus, it is important for companies to communicate with project stakeholders and assign one or more champions to define the areas of interests, the type of data needed, the timelines of acquiring and analysing the data, the proper methods to share the data, and the expected deliverables. Another reason for including multiple stakeholders is that one drone flight may serve multiple teams. Thus, understanding the needs of every team helps better understand the use-cases that drones can fulfil. For example, images can be shared with superintendents to note and understand

Since the construction industry is resistive to change, a transition to implement drones in construction processes may not be easy. Thus, it is important for champions to highlight the value of using drones by sending newsletters, conducting webinars, and sharing videos with those who are hesitant to change. Return on Investment (ROI) is also a common practice to show value; champions can perform preliminary analysis by comparing the current incurred costs (such as the current-state labour wages and equipment needed to collect the data and perform analysis) to the drone investment cost (i.e. the Lifecycle cost explained before) and the expected savings (both tangible such as the equipment's replaced by drones, and intangible such as the dollar value of the time savings). Another way to show value is to conduct different pilot test on a small-scale construction product with the help of a third party and gather feedback from involved personnel.

Another important practice is to review the regulations and necessary authorizations to operate drones. While the FAA is responsible to regulate the operations of drones in USA, several states and cities have created their own rules and regulations, with some imposing regulations limitations that are stricter than others. Thus, a company should invest time in understanding the local and federal rules and regulations in their areas of operation.

As discussed in the challenges section, the drone operations are associated with serious safety risks. Thus, it is important for companies to develop risk assessment plans and standardize checklists for drone operators and on-site personnel to use before, during, and after the flight operation. Such practices are also valuable for the involved lawyers and insurance companies.

Finally, the global drone market is increasing and currently involves a growing number of hardware, software, pilot, training, insurance, and distribution partners. With such a strong global partnership network, and with drones successfully exceeding their expectations on construction projects, companies have the support that they need to confidently start their drone initiatives.

6 CONCLUSIONS

This paper recapped the use of drones in the construction industry. Using the 5W2H method, the paper presented the different applications of drones across the lifecycle of a construction project, discussed the benefits of using drones and elaborated on the challenges, and provided an overview of the lifecycle costs. The paper also presented a comprehensive assessment of common practices derived from several successful implementations of drones in the construction industry. Findings in this paper can set a baseline for both construction academicians and practitioners for the state of drones in the industry. Further studies can build on this study by providing industry use-cases and elaborating on other Industry 4.0 technologies that can be integrated with drones to maximize its use.

References

- [1] Bou Hatoum, M., Piskernik, M., and Nassereddine, H. A holistic framework for the implementation of big data throughout a construction project lifecycle. In *Proceedings of the 37th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 1299–1306, Kitakyushu, Japan, 2020.
- [2] El Jazzar, M., Urban, H., Schranz, C., and Nassereddine, H. Construction 4.0: a roadmap to shaping the future of construction. In *Proceedings of the 37th International Symposium on Automation and Robotics in Construction*, pages 1314–1321, Kitakyushu, Japan, 2020.
- [3] Hatoum, M.B. and Nassereddine, H. Developing a framework for the implementation of robotics in construction enterprises. In *EG-ICE 2020 Proceedings: Workshop on Intelligent Computing in Engineering*, pages 453–462, Germany, 2020.
- [4] Nassereddine, H., Schranz, C., Bou Hatoum, M., and Urban, H. A comprehensive map for integrating augmented reality during the construction phase. In *Proceedings of the Creative Construction e-Conference 2020*, pages 56–64, Online, 2020.
- [5] Chinchane A., Singh S. P. and Sumant O. Construction Drone Market Outlook - 2027 (Allied Market Research). Online: <https://www.alliedmarketresearch.com/construction-drone-market-A06247>. Accessed: 6/9/2021.
- [6] [Federal Aviation Administration. Unmanned Aircraft Systems (UAS). Online: <https://www.faa.gov/uas/>. Accessed: 6/9/2021.
- [7] Federal Highway Administration. EDC-5 Innovations (2019-2020). Online:

- https://www.fhwa.dot.gov/innovation/everydaycounts/edc_5/. Accessed: 6/9/2021.
- [8] Carson, K. Digitalization & Automation in the Constructions Trades (Seattle, WA). *Seattle Job Initiative*, 2020.
 - [9] Chamola, V., Hassija, V., Gupta, V., and Guizani, M. A comprehensive review of the covid-19 pandemic and the role of IOT, drones, ai, blockchain, and 5g in managing its impact. *IEEE Access*, 8:90225–90265, 2020.
 - [10] Albeaino, G. and Gheisari, M. Trends, benefits, and barriers of unmanned aerial systems in the construction industry: a survey study in the united states. *Journal Of Information Technology In Construction (ITcon)*, 26:84–111, 2021.
 - [11] Dastgheibifard, S. and Asnafi, M. A review on potential applications of unmanned aerial vehicle for construction industry. *Sustainable Structure And Materials*, 1(2):44–53, 2018.
 - [12] Li, Y. and Liu, C. Applications of multirotor drone technologies in construction management. *International Journal Of Construction Management*, 19(5):401–412, 2018.
 - [13] Efosa, A. Exploring the Role of Building Modeling and Drones in Construction. *UMEA University*, 2019
 - [14] Chiddarwar, K. The Rise of Drones in Construction. Online: <https://www.youtube.com/watch?v=C80UckBJ8nQ>. Accessed: 6/9/2021.
 - [15] Wang, X., Al-Shabbani, Z., Sturgill, R., Kirk, A., and Dadi, G.B. Estimating earthwork volumes through use of unmanned aerial systems. *Transportation Research Record*, 2630(1):1–8, 2017.
 - [16] Saccardo, D. and Langston, C. The Impact of Emerging Technology on the Value of Construction Projects. *Bond University*, 2020.
 - [17] Tugay, A., Zeltser, R., Kolot, M., and Panasiuk, I. Organization of supervision over construction works using uavs and special software. *Science and Innovation*, 15(4):21–28, 2019.
 - [18] Propeller. Drone Data Analytics for Construction. *Propeller*, 2018.
 - [19] Wingtra. Drones in Construction and Infrastructure: Why and How to Use Drones in Construction and Infrastructure. Online: <https://wingtra.com/drone-mapping-applications/drones-in-construction-and-infrastructure/>. Accessed: 6/9/2021.
 - [20] Mosly, I. Applications and issues of unmanned aerial systems in the construction industry. *International Journal Of Construction Engineering And Management*, 6(6):235–239, 2017.
 - [21] Gheisari, M., Rashidi, A., and Esmacili, B. Using unmanned aerial systems for automated fall hazard monitoring. In *Proceedings of ASCE Construction Research Congress*, pages 62–72, 2018.
 - [22] Hubbard, B., Wang, H., Leasure, M., Ropp, T., Lofton, T., Hubbard, S., and Lin, S. Feasibility study of uav use for rfid material tracking on construction sites. In *51st ASC Annual International Conference Proceedings*, 2015.
 - [23] Zhou, Z., Irizarry, J., and Lu, Y. A multidimensional framework for unmanned aerial system applications in construction project management. *Journal Of Management In Engineering*, 34(3):04018004, 2018.
 - [24] Dupont, Q.F., Chua, D.K., Tashrif, A., and Abbott, E.L. Potential applications of uav along the construction's value chain. In *Procedia Engineering*, pages 165–173, 2017.
 - [25] Dukowitz, Z. Drones in Construction: How Drones Are Helping Construction Companies Save Money, Improve Safety Conditions, and Keep Customers Happy. Online: <https://uavcoach.com/drones-in-construction/>. Accessed: 6/9/2021.
 - [26] Loveless, C. Drones in Construction. *University of Mississippi*, 2017.
 - [27] ETH Zurich. Aerial Construction: Building structures with flying machines. Online: <https://idsc.ethz.ch/research-dandrea/research-projects/aerial-construction.html>. Accessed: 6/9/2021.
 - [28] Goessens, S., Mueller, C., and Latteur, P. Feasibility study for drone-based masonry construction of real-scale structures. *Automation In Construction*, 94:458–480, 2018.
 - [29] Paneni, C. Drone Proof of Concept- Australian Patent Application No 2019902083 (Australia: Ausco Modular). Online: <https://vimeo.com/user39979442/review/380421875/f4fcf56b8d>. Accessed: 6/9/2021.
 - [30] Tsuru Robotics. World's First Graffiti Drone (Part 1). Online: <https://diydrones.com/profiles/blogs/first-in-the-world-graffiti-drone-part-1>. Accessed: 6/9/2021.
 - [31] O'Neill, K. Spectacular aerial drone footage gives unique view of controlled hospital demolition. *Mirror (UK)*, Online: <https://www.mirror.co.uk/news/world-news/spectacular-aerial-drone-footage-gives-7118609>. Accessed: 6/9/2021.
 - [32] Luhmann, T., Chizhova, M., and Gorkovchuk, D. Fusion of UAV and terrestrial photogrammetry with laser scanning for 3d reconstruction of historic churches in Georgia. *Drones*, 4(3):53, 2020.
 - [33] Hazem, Z. Drones in Construction – 6 Ways Drones are Driving Construction Innovation

- Drones in Construction Industry. Online: <https://www.planradar.com/us/changes-of-drones-in-the-construction-industry/>. Accessed: 6/9/2021.
- [34] Outay, F., Mengash, H.A., and Adnan, M. Applications of unmanned aerial vehicle (UAV) in road safety, traffic and highway infrastructure management: recent advances and challenges. *Transportation Research Part A: Policy And Practice*, 141:116–129, 2020.
- [35] Eskandari Torbaghan, M., Kaddouh, B., Abdellatif, M., Metje, N., Liu, J., Jackson, R., Rogers, C.D., Chapman, D.N., Fuentes, R., Miodownik, M., and others. Robotic and autonomous systems for road asset management: a position paper. In *Proceedings of The Institution of Civil Engineers-Smart Infrastructure and Construction*, 172(2):83–93, 2020.
- [36] Skycatch, Bechtel, and DJI 2017 Getting Started with Drones in Construction. Online: <https://www.youtube.com/watch?v=WLnouZ8szo4>. Accessed: 6/9/2021.
- [37] Banks, E., Cook, S. J., Fredrick, G., Gill, S., Gray, J. S., Larue, T., Milton, J. L., Tootle, A., Wheeler, P., Snyder, P. R. and Waller, Z. Successful Approaches for the Use of Unmanned Aerial System by Surface Transportation Agencies. *National Cooperative Highway Research Program*, 2018.
- [38] York, D.D., Al-Bayati, A.J., and Al-Shabbani, Z.Y. Potential applications of uav within the construction industry and the challenges limiting implementation. In *Construction Research Congress 2020: Project Management and Controls, Materials, and Contracts*, pages 31–39, 2020.
- [39] Ciampa, E., De Vito, L., and Pecce, M.R. Practical issues on the use of drones for construction inspections. *Journal of Physics: Conference Series*, pages 012016, 2019.
- [40] Melenbrink, N., Werfel, J., and Menges, A. On-site autonomous construction robots: towards unsupervised building. *Automation In Construction*, 119:103312, 2020.
- [41] Kardasz, P., Doskocz, J., Hejduk, M., Wiejkut, P., and Zarzycki, H. Drones and possibilities of their using. *Journal Of Civil & Environmental Engineering*, 6(3):1–7, 2016.
- [42] Eiris, R. and Gheisari, M. Evaluation of small uas acquisition costs for construction applications. In *Proceedings of the Joint Conference on Computing in Construction (JC3)*, pages 931–938, Heraklion, Greece, 2017.
- [43] DJI and Skycatch. DJI Drone Workflows for Construction Webinar. Online: <https://www.youtube.com/watch?v=wieM7eZMGsk>. Accessed: 6/9/2021.
- [44] Drone Deploy. Commercial Drone Industry Trends. *Drone Deploy*, 2018.
- [45] Propeller. How to Start a Drone Program on Your Site. *Propeller*, 2018.
- [46] Skycatch. Robots, Rework, & ROI: Optimizing Construction Operations with Drones. Online: <https://www.youtube.com/watch?v=rGIgSfIFxoI>. Accessed: 6/9/2021.
- [47] Skyward and Skycath. Drones in Construction: Best Practices from Skyward and Skycatch. 2021.
- [48] Skyward. Adding Drones to the Enterprise. *Skyward*, 2021.
- [49] Propeller. The Future of Construction: Three Big Things You Need to Know. *Propeller*, 2021.
- [50] Hatoum, M. B. Nassereddine, H., and Badurdeen, F. Reengineering Construction Processes in the Era of Construction 4.0: A Lean-Based Framework. In *Proc. 29th Annual Conference of the International Group for Lean Construction (IGLC)*, pages 403–412, 2021.

Intelligent BIM-based spatial conflict simulators: A comparison with commercial 4D tools

L. Messi ^{a,b}, B. García de Soto ^b, A. Carbonari ^a, and B. Naticchia ^a

^a Polytechnic University of Marche (UNIVPM), Faculty of Engineering, DICEA Department, Ancona, Italy

^b New York University Abu Dhabi (NYUAD), Division of Engineering, S.M.A.R.T. Construction Research Group, Abu Dhabi, United Arab Emirates

E-mail: l.messi@staff.univpm.it, garcia.de.soto@nyu.edu, alessandro.carbonari@staff.univpm.it,
b.naticchia@staff.univpm.it

Abstract –

In construction projects, the space required for executing each activity is unanimously recognized as a limited but renewable resource, like workers, equipment, and materials. Overloading a given resource space, as demonstrated by statistics, leads to efficiency losses and safety threats. Despite the valuable contributions provided by academics and the construction software industry, a definitive tool for managing and resolving spatial issues is not available yet. In fact, current approaches, generally based on geometric intersection tests between main workspaces in their initial static position, do not account for not-purely-geometric spatial issues (e.g., struck-by hazards, electrical hazards, etc.) and overestimate conflicts affected by unlikely surrounding conditions.

In order to cover these gaps, a workspace management framework integrated into the preparation of the construction schedule is proposed and tested by developing a spatial conflict simulator (i.e., “Enhanced” approach). The simulator implemented using the serious gaming environment Unity3D™ was compared with Synchro Pro (i.e., “Benchmark” approach). Results show a better performance of the “Enhanced” approach, able to extend the range of detected spatial conflicts thanks to physics simulations, and filter unlikely spatial conflicts based on Bayesian inference’s results.

Keywords –

Construction Management; Workspace Scheduling; Spatial Conflicts; BIM; Game Engine; 4D tool

1 Introduction

In construction projects, each activity requires a specific workspace to be executed [1]. A workspace is defined as the suitable occupational volume a crew

and/or equipment occupy during the execution of a certain activity on a predefined geometrical element [2]. As the construction progresses, the space occupied by completed activities is released and reused by other operations [3]. Consequently, the space required by construction activities, i.e., the geometry and the location of workspaces, continuously change over time [4], leading to a sequence of workspaces associated with the project’s activities [4]. As suggested by [5], the space in the construction site must be considered as a limited but renewable resource, similar to workers, equipment, and materials [3]. When the same workspace is occupied simultaneously by two or more activities, a spatial interference occurs, leading to significant problems such as labor safety hazards, construction delay, and loss in productivity. To cite a few statistics, a study related to masonry works has reported that congested workspaces and restricted access cause efficiency losses of up to 65% [6]. In addition to the productivity impacts, a study conducted in the US found that poor workspace planning was the cause of 323 fatalities over a period of 12 years [7]. This trend can be explained by the fact that the dynamic nature of construction activities makes the management of workspaces challenging using conventional planning methods. In [8], the authors asserted that conventional planning methods do not adequately represent and communicate the interference between construction activities and do not consider space constraints in the planning process. As of now, workspace planning is being performed through judgment or at most by the aid of 2D sketches [2] because commercial 4D visual planning tools lack effective and holistic workspace management capabilities [4,9].

This study has developed an intelligent BIM-based spatial conflict simulator that integrates physics simulations and Bayesian inference in a serious gaming environment. The tool, aiming to enhance the range of detectable spatial conflicts and filter non-critical ones, has been tested and compared to one of the most popular 4D tools, namely Synchro Pro. The rest of this paper is

structured as follows. In Section 2, the scientific background of construction workspace management and related gaps are discussed. Section 3 reports the methodology adopted for developing the proposed spatial conflict simulator. Section 4 describes the use case and the experiment design, whereas Section 5 discusses the results. Finally, Section 6 is devoted to conclusions.

2 Scientific Background

2.1 Research contributions to workspace management

Nowadays, the need to consider the spatial dimension to ensure schedule feasibility and avoid critical issues, such as safety, productivity, and constructability, is unanimously accepted by field experts. Stemming from this assumption, researchers have spent many efforts on the topic of workspace management.

The authors in [10] proposed a prototype system based on a micro-level discretization (i.e., building component space, labor crew space, equipment space, hazard space, protected space, and, finally, temporary structure space) to detect spatial interferences. The system, implemented using the object-oriented programming language Powermodel, displayed the list of categorized and prioritized time-space conflicts in a 4D CAD simulation environment, namely VRML 2.0 and Excel. Complementarily, the authors in [11], introduced the concepts of macro-level (e.g., storage areas) and paths (e.g., equipment's and crews' paths) discretization. They applied the critical space-time analysis (CSA) approach to model and quantify workspace congestion, developing a computerized tool called PECASO (Patterns Execution and Critical Analysis of site Space Organization) using VBA programming. A macro- and micro-level discretization has been presented in [1], where the labor crew workspaces were differentiated into static (i.e., the entire workspace is required throughout the activity duration) and dynamic (i.e., a specific portion of the workspace is occupied during each time interval). In the same study, the authors have developed a dynamic 4D BIM-based system aiming to detect time-space conflicts and quantify their impacts considering labor crew movements. The proposed system has been implemented using the C# programming language in the .NET framework in a Visual Studio environment. In [7], a micro-level discretization plus the space for material handling path were defined. In the same study, the authors proposed a novel method that collects, formalizes, and reuses historical activity-specific workspace information for congestion identification and safety analysis in BIM. In [3], the workspaces defined by the studies mentioned above have been grouped into two

main categories: entity (i.e., laborers, mechanical equipment, and building components) and working spaces (i.e., spaces required to ensure smooth operation and tasks). This workspace discretization has been applied to develop a workspace conflict detection framework using the Navisworks SDK toolkit. The classification of the workspaces adopted in [4], inherited from the manufacturing industry, considers, in addition to the workspaces occupied by building elements and reserved as safety distance, a workspace discretization depending on the value added by the activities. A “main workspace” is required by activities that add tangible value to the project (e.g., building a wall), whereas a “support workspace” is required by preparatory activities (e.g., transferring materials) supporting the first category. The authors applied the described taxonomy to develop a construction workspace management tool implemented using the XNA game engine.

A gap found from the literature review is that most existing studies [1,7,10,11] consider object-based workspace taxonomies that allocate workspaces for each building element under construction for very specific purposes. This means assuming the strong hypothesis that spatial conflicts are likely to happen only around specific building elements under construction, that is, between their static object-based workspaces. The authors of this study, adopting the workspace taxonomy proposed by [4], have preliminary addressed this issue by developing a spatial conflict simulator in Unity3D™ to reduce threats of COVID-19 transmission related to spatial conflicts among main and support workspaces in construction projects [12]. Another limitation of existing studies [3,4,10,11] is detecting spatial conflicts by simply carrying out geometric intersection tests between defined workspaces. Although this approach has provided early valuable results and enabled process automation, it has also led to overestimating the results and missing incompatibilities that are not purely geometric (e.g., struck-by hazards, electrical hazards, etc.).

2.2 Commercial 4D software

Several 4D modeling software (e.g., Vico Schedule Planner and 4D Player, Innovaya Visual 4D Simulation, Autodesk Navisworks, Synchro Pro, Elecosoft Powerproject BIM, etc.) already provide different capabilities and toolkits [13]. The most popular ones, namely Autodesk Navisworks and Synchro Pro, have been exhaustively compared [13]. Both implement clash detection functionalities with the possibility to define a clearance threshold around objects (i.e., clearance clash test). Whereas Autodesk Navisworks is more indicated for running clash detection tests between single building elements to adjust the building design and avoid spatial issues, Synchro Pro is preferred for checking spatial interferences during the construction process and

adjusting the construction schedule accordingly [14]. As reported in [15], Synchro Pro enables the user to create workspaces from bounding boxes [16] and check for related spatial conflicts. In addition, both software enable the creation of animation for simulating the construction site dynamics, with the not negligible limitation that they must be defined point by point by the user [17]. In other words, these software solutions do not enable to carry out automatic realistic physics simulations. This represents a significant limit, especially if we consider big construction projects where a significant number of agents moving for each time frame must be modeled one by one.

3 Methodology

3.1 Workspace management framework

In order to cover the gaps identified in Section 2, the authors have proposed a workspace management framework that integrates the construction scheduling phase (Figure 1, top lane) with the contribution given by a spatial conflict simulator (Figure 1, bottom lane). The latter, developed by adopting the serious game engine technology, can detect eventual spatial interferences based on both the geometric and semantic information provided by the Building Information Model (Figure 1, green nodes) and the construction process data included in the construction schedule (Figure 1, blue nodes). Based on this data, workspaces can be generated in the gaming environment to carry out geometric computations and physics simulations (Figure 1, red nodes). As a result,

the list of detected spatial conflicts is generated and provided as an input to a Bayesian network (BN) (Figure 2) that, using expert knowledge, can be applied to find out (and filter) non-critical scenarios to avoid conflict overestimations (Figure 1, orange tasks). Afterward, the construction management team, made aware of likely future spatial conflicts, can adjust (Figure 1, blue nodes) or confirm (Figure 1, violet node) the construction schedule.

3.2 Intelligent BIM-based spatial conflict simulator

The spatial conflict simulator proposed by this study has been developed in Unity3D™, a serious gaming environment. As shown in Figure 1 (bottom lane), the BIM model and the construction (work) schedule are provided as input to define the 4D model. BIM models are imported into the gaming environment using the in-house IFC Loader for Unity3D™, developed based on the IFC Engine DLL library [18]. Construction schedules are imported into the serious gaming tool in CSV format using a C# script developed by the authors. Based on the 4D model information, the proposed spatial conflict simulator developed by the authors in another C# script generates main workspaces and detects “direct” and “indirect” spatial conflicts between them. A spatial conflict is detected between two given workspaces assigned to different crews only if their boundaries intersect. “Direct” spatial conflicts are detected by carrying out geometric intersection tests among workspaces in their initial static position, inherited from the corresponding building elements.

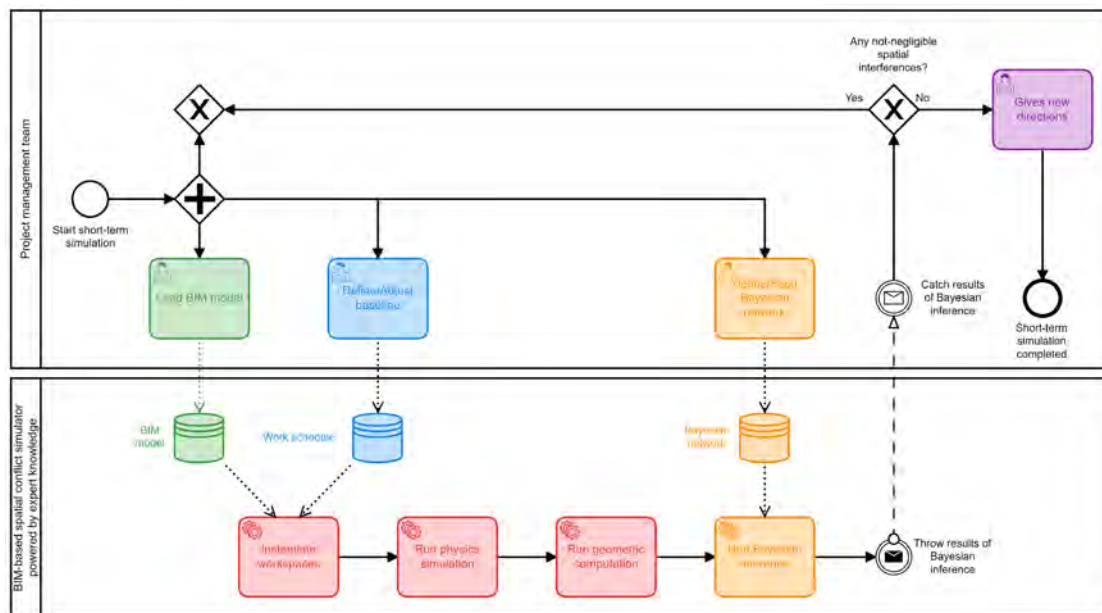


Figure 1. Proposed workspace management framework to consider spatial-temporal conflicts.

Although detecting these kinds of spatial conflicts is included in the state of the art of workspace management, it cannot cover the totality of spatial issues affecting the construction site. To make an example, a main workspace at a higher level (i.e., Cause-of-Risk-Activities or CORA workspace) than another one (i.e., Exposed-to-Risk-Activities or ETRA workspace), even if they do not intersect each other in their initial position, can be affected by spatial conflicts due to falling objects (i.e., struck-by hazards from CORA into ETRA workspaces). This kind of spatial conflict is named, here on for simplicity, as “indirect” or “possible” one. The serious gaming environment, embodying mechanical physics, enables the detection of this kind of spatial conflict, carrying out physics simulations and geometric computation (i.e., dropping down main workspaces according to the gravity law and looking for geometric intersections). The criticality level of “indirect” or “possible” spatial conflicts, being the result of a virtual simulation (e.g., dropping-down physics simulations), needs to be assessed. As indicated in Section 3.1, the criticality level is judged using Bayesian inference. The BN (Figure 2) was developed according to the factors categories (i.e., external, organizational, direct factors) presented in [19] and complemented by expert knowledge elicitation. In this study, the conditional probability tables (CPTs) have been filled according to the authors’ knowledge. BNs can provide valuable support for assessing, based on surroundings conditions, the criticality level of each detected “possible” spatial conflict and filtering the ones that are not going to cause any spatial issue. The struck-by hazard BN has been implemented in Unity3D™ through a C# script based on

the Discrete Bayesian Network library [20]. This enables the variables’ evidence of the BN to be automatically fed by gaming simulation data.

4 Use Case and Experiment Design

The proposed approach provides a valuable contribution to the planning phase of construction works related to medium and large projects. Falling in this category, the construction process of the Eustachio Building, a public building hosting the Faculty of Medicine of the Polytechnic University of Marche, has been selected as an example for validation purposes. The Eustachio Building is located in the extra-urban area of Ancona (Italy), close to the main regional hospital. The mixed-use building is arranged on six floors above ground and has a total area of 16,900 m².

The application of the spatial conflict simulator presented in this study has been defined as the “Enhanced” approach. In order to be validated, it has been compared with the “Benchmark” approach, based on the application of the most popular 4D tool for workspace management, namely Synchro Pro. Three experiments have been set up considering, for simplicity and paper length constraints, only two working days (i.e., 27th and 28th May), highlighted in yellow in Figure 3(c). The four activities shown in Figure 3(c) are the ones that have been scheduled during this time interval. One experiment (i.e., Experiment No. 1, Table 1) out of three is related to the “Benchmark” approach and considers the “Standard” BIM model of the use case (Figure 3(a)). One of the remaining two experiments (i.e., Experiment No. 2,

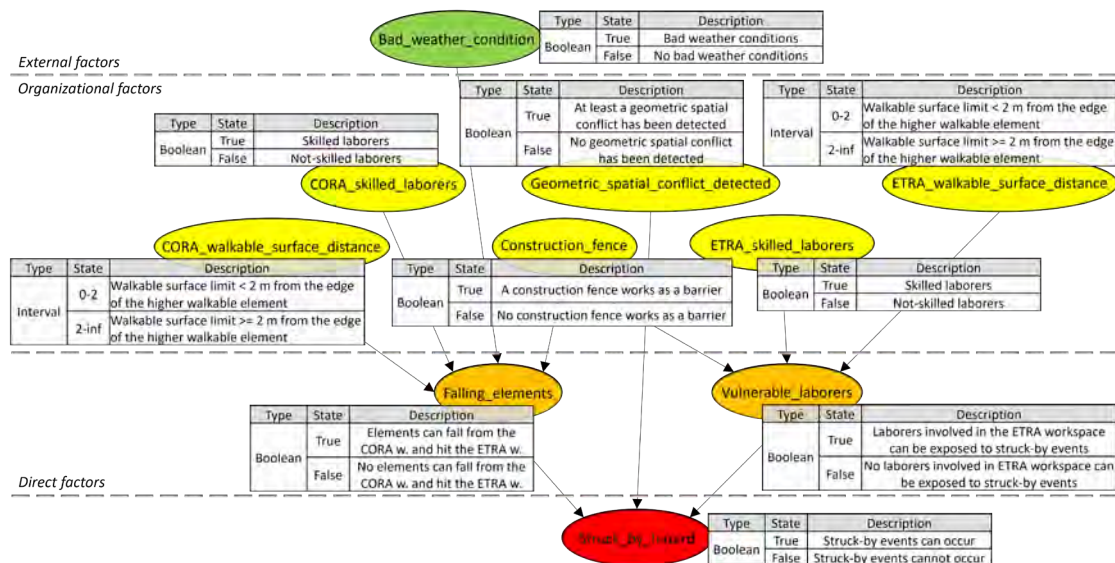


Figure 2. BN to assess the criticality level of “indirect” or “possible” spatial conflicts due to struck-by hazards.

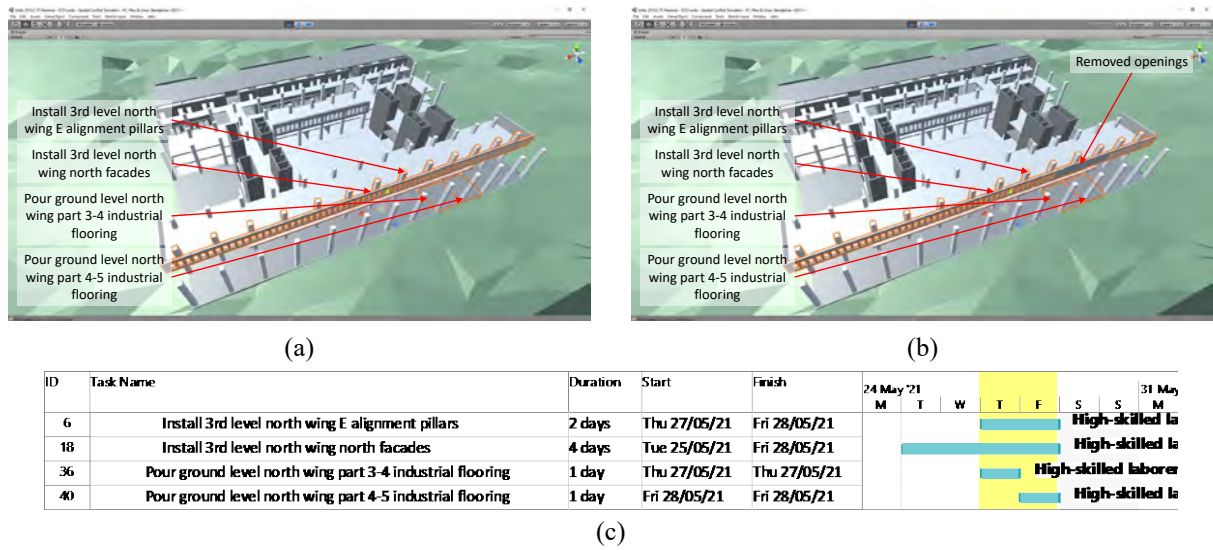


Figure 3. Views of the (a) “Standard” and (b) “Variation A” BIM models. (c) Excerpt of the overall construction schedule reporting the activities scheduled for the selected working days.

Table 1. Overview of the main differences between the three experiments.

Experiment No.	Approach	Construction schedule	BIM model	Tool functionalities				
				Loading BIM model and construction schedule	Generating main workspaces	Carrying out geometric intersection tests	Carrying out physics simulations	Running BN
1	Benchmark (Synchro)	May 27 th and 28 th	Standard model	✓	✓	✓	✗	✗
2	Enhanced (proposed tool)		Standard model	✓	✓	✓	✓	✓
3	Enhanced (proposed tool)		Variation A	✓	✓	✓	✓	✓

Table 1), related to the “Enhanced” approach, considers the “Standard” BIM model, whereas the other one (i.e., Experiment No. 3, Table 1) considers the “Variation A” BIM model (Figure 3(b)).

The “Variation A” BIM model has been obtained from the “Standard” one by removing some of the openings on the 3rd level north façade. This was done to assess the contribution of the Bayesian inference by considering a different spatial conflict criticality level depending on the surrounding conditions. The different

tool functionalities tested in the “Standard” and “Enhanced” approaches have been summarized in Table 1.

5 Results and Discussion

In the Experiment No. 1, the geometric intersection test carried out by Synchro Pro (e.g., “Benchmark” approach) has detected one spatial conflict between two workspaces at the same level (i.e., 3rd level) overlapping

each other in their static position (Figure 4, Table 2). The proposed spatial conflict simulator (i.e., “Enhanced” approach) has been tested in Experiments No. 2 and 3. In both cases, one “direct” and four “indirect” or “possible” spatial conflicts have been detected (Figure 5, Table 2).

The “direct” spatial conflict detected between two workspaces overlapping in their initial static position is the same one reported in Experiment No. 1. The four “indirect” or “possible” spatial conflicts, detected

between workspaces at different levels (i.e., ground level and 3rd level), resulted from the combination of the physics simulation and geometric intersection tests. In Experiment No. 2, that is when the “Standard” BIM model was considered in the Enhanced approach, the Bayesian inference has provided a “high” criticality level. In fact, as shown in Table 2, the “high” state of the “Struck_by_hazard” variable has, for each “indirect” or “possible” spatial conflict, the highest probability value

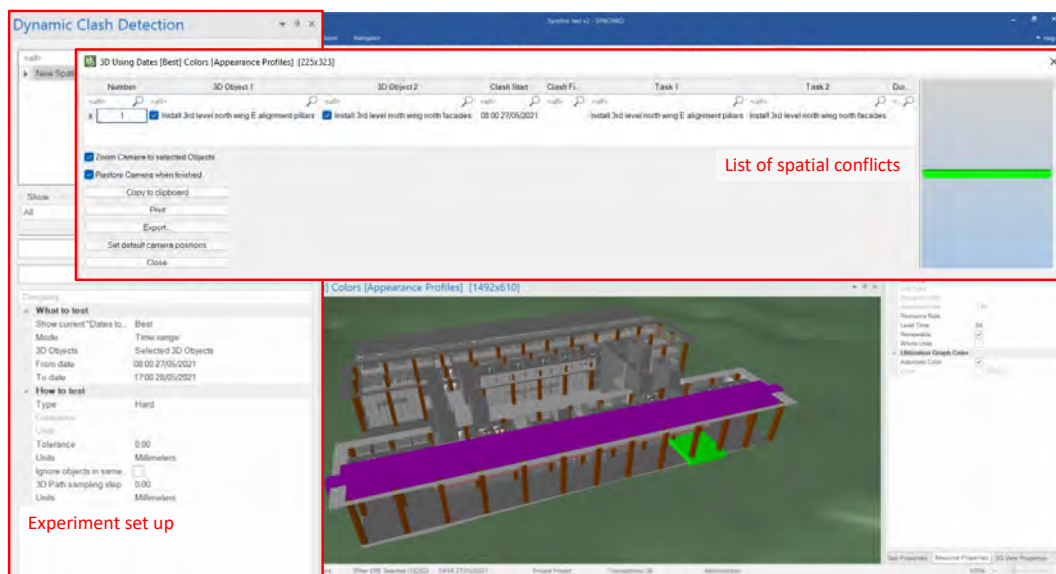


Figure 4. View of Synchro Pro after executing the “Benchmark” approach experiment.

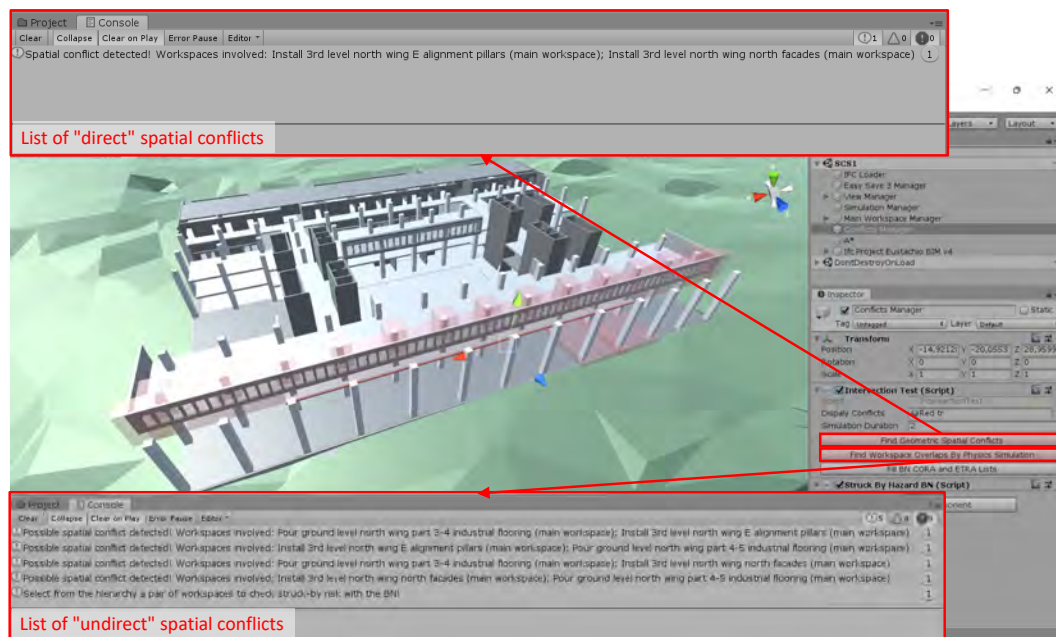


Figure 5. View of the gaming environment after executing the “Enhanced” approach experiments.

(e.g., 78%). The Bayesian inference provided a “low” criticality level when the “Variation A” BIM model was considered. In fact, as reported in Table 2, the “low” state of the “Struck_by_hazard” variable has, for each “possible” spatial conflict, the highest probability value (e.g., 57%). This is because in Experiment No. 3, the 3rd level north façade, in correspondence to the part in which some openings have been removed, works as a barrier protecting workers in the ETRA workspace from eventual objects falling from the CORA workspace. This scenario is modeled by the “Construction_fence” variable of the BN (Figure 2) that, in Experiment No. 3, contrarily to Experiment No. 2, assumes the “True” state.

The three experiments demonstrate the proposed spatial conflict simulator's added value in detecting not-purely-geometric spatial issues and filtering the not critical ones to avoid overestimations.

6 Conclusions and outlook

In construction projects, spatial interferences lead to significant problems such as labor safety hazards, construction delays, and loss of productivity. Although researchers and 4D software developers have spent many

valuable efforts in the field of workspace management, several issues have not been addressed yet. One of the main gaps in past studies is the assumption that spatial conflicts can happen only around specific building elements under construction; that is, between their static object-based workspaces. In addition, spatial conflicts have been detected simply by carrying out geometric intersection tests between workspaces. What has emerged from academic and industry sectors is the need to enhance the range of detected spatial conflicts and avoid overestimations.

The proposed tool combines physics simulations and geometric computations to replicate the construction site dynamics and detect resulting conflicts. Then, an integrated BN fed by expert knowledge and simulation results filters non-critical scenarios. The serious gaming tool has been validated by comparing the results with the ones provided by Synchro Pro for the same scenario. Promising results come from the proposed approach. In fact, combining physics simulations and Bayesian inference enhances the range of detected spatial conflicts (e.g., including the ones between workspaces at different levels affected by struck-by hazards) and filters out unlikely spatial conflicts.

Table 2. Results from the “Benchmark” and “Enhanced” approaches.

Experiment No.	Pairs of conflicting workspaces detected by geometric intersection tests	Pairs of conflicting workspaces detected by physics simulation and geometric intersection tests		Criticality levels by Bayesian inference	
1 (Figure 3(a) and Figure 4)	Install 3rd level north wing E alignment pillars	Install 3rd level north wing north facades	(Functionality not available)	(Functionality not available)	
2 (Figure 3(a) and Figure 5)			Pour ground level north wing part 3-4 industrial flooring	Install 3rd level north wing E alignment pillars	High (78%)
			Pour ground level north wing part 4-5 industrial flooring	Install 3rd level north wing E alignment pillars	High (78%)
			Pour ground level north wing part 3-4 industrial flooring	Install 3rd level north wing north facades	High (78%)
			Pour ground level north wing part 4-5 industrial flooring	Install 3rd level north wing north facades	High (78%)
			Pour ground level north wing part 3-4 industrial flooring	Install 3rd level north wing E alignment pillars	Low (57%)
3 (Figure 3(b) and Figure 5)			Pour ground level north wing part 4-5 industrial flooring	Install 3rd level north wing E alignment pillars	Low (57%)
			Pour ground level north wing part 3-4 industrial flooring	Install 3rd level north wing north facades	Low (57%)
			Pour ground level north wing part 4-5 industrial flooring	Install 3rd level north wing north facades	Low (57%)

In the future, the proposed BN can be extended by adding variables for all the factors presented in [19]. Then, the tool can be applied to other scenarios in which physical simulations provide valuable contributions (e.g., tower cranes lifting heavy loads in windy conditions). Finally, the tool can be improved to address different scenarios. For example, modeling detailed materials properties and implementing electromagnetism rules in the gaming environment would enable addressing electrical hazards.

Acknowledgments

This research has been partially funded by the Graduate and Postdoctoral Programs at New York University Abu Dhabi (NYUAD), the Ph.D. Program at the Polytechnic University of Marche (UNIVPM), and the Italian Ministry of Education, University and Research PRIN 2017 Project entitled “A Distributed Digital Collaboration Framework for Small and Medium-Sized Engineering and Construction Enterprises” (prot. 2017EY3ESB).

References

- [1] A. Mirzaei, F. Nasirzadeh, M. Parchami Jalal, Y. Zamani, 4D-BIM Dynamic Time-Space Conflict Detection and Quantification System for Building Construction Projects, *J. Constr. Eng. Manag.* 144 (2018) 04018056.
- [2] A. Hosny, M. Nik-Bakht, O. Moselhi, Workspace planning in construction: non-deterministic factors, *Autom. Constr.* 116 (2020).
- [3] H. Ma, H. Zhang, P. Chang, 4D-Based Workspace Conflict Detection in Prefabricated Building Constructions, *J. Constr. Eng. Manag.* 146 (2020) 04020112.
- [4] M. Kassem, N. Dawood, R. Chavada, Construction workspace management within an Industry Foundation Class-Compliant 4D tool, *Autom. Constr.* 52 (2015) 42–58.
- [5] A. Hosny, M. Nik-Bakht, O. Moselhi, Workspace management on construction jobsites: An industry survey, *Proceedings, Annu. Conf. - Can. Soc. Civ. Eng.* 2019-June (2019) 1–9.
- [6] S.R. Sanders, *An Analysis of factors Affecting Labor Productivity in Masonry Construction*, (1989).
- [7] S. Zhang, J. Teizer, N. Pradhananga, C.M. Eastman, Workforce location tracking to model, visualize and analyze workspace requirements in building information models for construction safety planning, *Autom. Constr.* 60 (2015) 74–86.
- [8] Z. Mallasi, Dynamic quantification and analysis of the construction workspace congestion utilising 4D visualisation, *Autom. Constr.* 15 (2006) 640–655.
- [9] C. Igwe, F. Nasiri, A. Hammad, Construction workspace management: critical review and roadmap, *Int. J. Constr. Manag.* 0 (2020) 1–14.
- [10] B. Akinci, M. Fisken, R. Levitt, R. Carlson, Formalization and Automation of Time-Space Conflict Analysis, *J. Comput. Civ. Eng.* 16 (2002) 124–134.
- [11] N. Dawood, Z. Mallasi, Construction workspace planning: Assignment and analysis utilizing 4D visualization technologies, *Comput. Civ. Infrastruct. Eng.* 21 (2006) 498–513.
- [12] L. Messi, B. García de Soto, A. Carbonari, B. Naticchia, Addressing COVID-19 Spatial Restrictions on Construction Sites Using a BIM-Based Gaming Environment, *Proc. 38th Int. Symp. Autom. Robot. Constr.* (2021) 521–528.
- [13] Y. Nechyporchuk, R. Bašková, The conformity of the tools of selected software programs for 4D building modeling, *IOP Conf. Ser. Mater. Sci. Eng.* 867 (2020).
- [14] SYNCHRO Construction, 15 Minute Fridays - Using Synchro PRO for Workspace Clash Detection, (2015). <https://www.youtube.com/watch?v=0hDcb9aeUPE> (accessed November 24, 2021).
- [15] Bentley System Inc., Create Workspace from Bounding Box. https://www.bimsdks.com/bentley/Synchro/create_workspace_boundingBox.htm (accessed November 27, 2021).
- [16] SYNCHRO Construction, Creating objects and workspaces within Synchro, (2011). <https://www.youtube.com/watch?v=x0fv3So5Rkc> (accessed November 27, 2021).
- [17] SYNCHRO Construction, Creating 3D Paths for 4D Simulation in Synchro, (2016). <https://www.youtube.com/watch?v=mFIITzqRBWY> (accessed November 27, 2021).
- [18] RDF Ltd., IFCEngine DLL Library, (2006). <http://rdf.bg/product-list/ifc-engine/> (accessed November 27, 2021).
- [19] L.D. Nguyen, D.Q. Tran, M.P. Chandrawinata, Predicting Safety Risk of Working at Heights Using Bayesian Networks, *J. Constr. Eng. Manag.* 142 (2016) 04016041.
- [20] J. Chen, Discrete Bayesian Network, (2017). <https://assetstore.unity.com/packages/tools/ai/discrete-bayesian-network-61312> (accessed November 27, 2021).

BIM-assisted, automated processes for commissioning in building services engineering

R. Becker^a, C. Blut^a, C. Emunds^b, J. Frisch^b, D. Heidermann^c, T. Kinnen^a, A. Lenz^d, M. Möller^d, N. Pauen^b, T. Rettig^f, D. Schlütter^b, M. Wenthe^e, J. Blankenbach^a, G. Bleimann-Gather^c, J. Fütterer^f, J. Jungedeitering^d and C. van Treeck^b

^aGeodetic Institute and Chair for Computing in Civil Engineering & GIS, RWTH Aachen University, Germany

^bInstitute of Energy Efficiency and Sustainable Building E3D, RWTH Aachen University, Germany

^cTEMA Technologie Marketing AG, Germany, ^dDiConneX GmbH, Germany

^eIMS – Internet Marketing Services GmbH, Germany, ^faedifion GmbH, Germany

E-mail: [ralf.becker, blut, kinnen, blankenbach]@gia.rwth-aachen.de,

[emunds, frisch, pauen, schluetter, treeck]@e3d.rwth-aachen.de, [heidermann, bleimann]@tema.de,

[a.lenz, m.moeller, j.jungedeitering]@diconnex.de, wenthe@ims.de, [trettig, jfuetterer]@aedifion.com

Abstract –

State-of-the-art building energy systems exhibit a high technical complexity. In the commissioning phase, technical building elements (TBE) are put into operation trade-by-trade and as linked complete systems. Besides the correct wiring on component level, instantiating the building automation and detecting errors is a cumbersome process in practice. The paper addresses a novel interconnected toolchain to support commissioning energy systems through digital processes in combination with energy system related digital BIM twins– the “energyTWIN”. This energyTWIN digitizes and automates the process chain in the commissioning of TBE and building automation with its highly complex interrelationships in constant exchange between reality and the BIM model (digital twin) as completely as possible.

By increasing energy efficiency with the novel processes of energyTWIN, a contribution to the worldwide goal of reducing energy consumption can be achieved.

Keywords –

Building Information Modeling; Technical building equipment; Data capturing; Indoor positioning; Virtual and Augmented Reality

1 Introduction

Nowadays, TBE are highly complex and interconnected systems. After dimensioning the hydraulic system and planning as Piping and Instrumentation Diagram (P&ID), TBE are today planned as 3D model with semantic (manufacturer) data using Building Information Modeling (BIM). In practice, building automation with its functional descriptions is

handled separately. Both domains are typically not linked and uniform data labels of TBE are defined often separately without using the (naturally existing) classification system of BIM. So, the TBE functions correctly and according to its control strategy during the operational phase, the commissioning must correspond to the planned configuration, or, the TBE control system must be customized to the actual built situation. Therefore, in the energyTWIN project, modern methods for high-resolution as-built data capturing (reality capturing) are being developed and refined with the aid of Artificial Intelligence (AI)-based data filtering and feature extraction for the automated recognition and classification of components and their topological, functional and informational interrelationships.

Crucial for automated workflows are uniform, generic (manufacturer-neutral) Reference Designation Systems (RDS) for the identification of TBE components (section 2). For data capturing of the actual installed TBE, efficient methods based on photogrammetry, laser scanning, infrared measurement technology, etc. will be developed and refined as described in section 3. This also includes methods for indoor positioning (pose tracking) and georeferencing of the captured data. Georeferencing is needed for comparisons to the as-planned BIM model and will be used in Virtual and Augmented Reality (VR/AR) applications of section 4. A comparison between the planned and the actual as-built situation will be realized using AI-based methods for automated recognition and classification of components and their topological relationships (section 5). A cloud-based system will connect all acquired relevant data on field level. Fault detection and diagnosis increases system and supply reliability (section 6). Finally, all developments will be prototyped and evaluated (section 7). Several, partly AI-supported methods will be fused to derive a

BIM model that reflects the actual situation (as-built) for planning and operational processes including aspects of time and costs (5D BIM in the phase of building life cycle following the commissioning of the TBE).

The energyTWIN, for the first time, employs AI-based approaches to combine image- and laser-based geometric and semantic data of building and system components with simultaneously captured data on field level. Furthermore, the modern technologies VR/AR are used for georeferenced and interactive visualizations as well as updating the BIM-based digital twins of the TBE. Finally, the RDS ensures the unique identification of objects during the various processes and data exchanges.

2 Reference Designation Systems

RDS are used for the unique identification of objects at different levels of granularity. Identifiers can be implemented as Global Unique Identifier (GUID) by a combination of alphanumeric characters or by a hierarchical structuring of data according to certain aspects such as location-, function- and signal related structure. GUIDs are used by BIM software applications and in the Industry Foundation Classes (IFC) data model [1]. The IFC schema is a standardized data model that defines the identity, semantics, characteristics, attributes, and relationships of objects, abstract concepts, processes, and people in a logical form. RDS with a hierarchical structuring of data are used by humans to reference objects across different models and documents. The structuring and level of detail of these systems depends on project-specific conditions as well as use cases and can therefore not be defined globally [2]. However, the reference designation should be as short as possible and as detailed as necessary. Aspects such as readability, memorability and uniformity of the reference designation system must be considered.

RDS map a "component-of" structure via hierarchical structuring according to various aspects. Frequently, a distinction is made between location, functional, and product-related aspects [3]. The location-related aspect describes the installation site or the installation location of an object. Entities of the local structure can be, among others, site, building, storey, area, room, or segment as well as outdoor areas. The product-related aspect defines

the composition of the object. It shows the division of an object into individual parts. Entities of the product-related structure are defined, among others, in standardized classification schemes such as Table 02 of DIN EN IEC 81346 - 2 [4] or Table 03 of VDI 3814 - 4.1 [5]. The functional-related aspect describes the respective function or task of a system and subdivides it hierarchically. Entities of the function-related structure are, among others, functional systems, technical systems, and components. These aspects can be used isolated from each other or interrelated to reference objects. Independently of this, the marking in the individual levels of the aspects takes place via a defined sequence of alphanumeric characters [3].

As part of a literature review, 50 RDS were considered, originating from the private (building automation, utilities), public (cities, state offices, federal offices) and scientific (universities, university hospitals) sectors. In addition, common standards were considered. Especially the reference designation systems of DIN EN IEC 81346 [2], VDI 3814 [5] and the Buildings Unified Data Point naming schema for Operation Management (BUDO) [6] were evaluated positively. Based on these findings and aligned with the concepts of DIN EN IEC 81346, a possible structure for a reference designation system is shown in Figure 1.

To implement a RDS into the planning process, the systems must be streamlined with existing digital data models. The digital model, also known as the BIM model, is the central instrument and is considered the "single source of truth". One option for structuring and exchanging the information is the open IFC standard. Depending on the aspect under consideration, structures for mapping these already exist in the IFC data model. There are five classes (*IfcSite*, *IfcBuilding*, *IfcBuildingStorey*, *IfcRoom* and *IfcZone*) describing the location-related aspect, one class with multiple types (*IfcSystem*) to describe the function-related aspect as well as multiple classes and subtypes to describe the product-related aspect. For the description of the signal-related aspects, no classes are currently available in IFC. Therefore, the structures can only be considered as attributes on elements. To map these, the IFC data model would have to be extended to include classes such as *IfcSignal* and the relationship *IfcRelConnectsSignalToElement*.

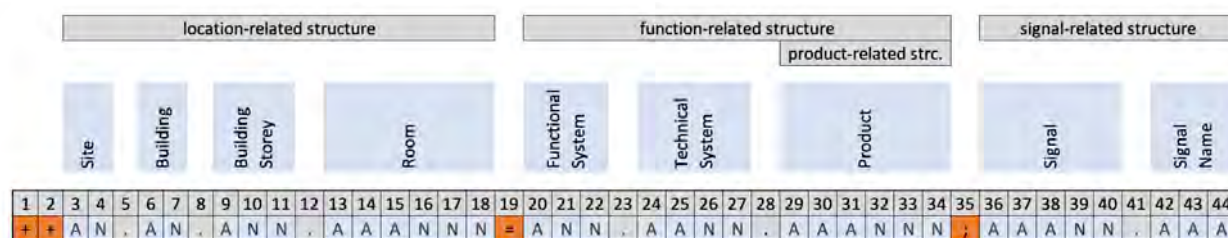


Figure 1: Possible structure of the RDS (A = alphabetic character, N = numeric character)

Corresponding types of *IfcSignal* could be based on existing classifications in the BUDO schema.

3 Efficient methods for reality capturing

For digitizing and automating the process chain during the commissioning of TBE, capturing geometric and semantic data of the realised state is essential. This requires a suitable capturing system, favourably with various sensors and interfaces. State-of-the-art is to use laser scanning and photogrammetry [7]. Often laser scanners are coupled with cameras to create coloured point clouds that represent the environment very realistically. Roughly, laser scanning can be categorized into two types: Terrestrial Laser Scanning (TLS) and Mobile Laser Scanning (MLS). While in TLS the resulting points clouds typically have a higher accuracy in comparison to MLS, MLS is more flexible and time efficient. In the field of photogrammetry, combining images taken from different points of view into a common coordinate system by determining the mutual orientation by means of bundle block adjustment over identical (homologous) points, has long been the standard method [8].

In the energyTWIN project, the goal is to create a flexible and easy to use system that is able to capture geometry, visual information, and additional properties of TBE, such as thermal data. The geometry contains information about size and shape of an object, an image contains information such as colour, and thermal data insight into the functionality of an object. A fusion of multiple data promises an improved and holistic classification of objects, since each part of information provides further indications about an object and reduces the number of object possibilities.

A potentially easier to use and more flexible solution than a stationary TLS or a conventional MLS, is the Microsoft HoloLens 2 (MHL2). The MHL2 is a mobile mixed reality (MR) head-mounted system. Its built-in sensors already include a depth and a colour (RGB) camera to create coloured point clouds. Since the MHL2 is head-worn, it has the same advantages as MLS in terms of flexibility and capturing speeds, compared to TLS. Furthermore, with its fully-fledged MR capabilities, it is a combined data capturing and MR system and enables us to realize the project goals described in section 3 and 4. In a detailed evaluation we found that the MHL2 achieves a sufficient accuracy of 2-5 cm, which is suitable for the project goals. Additionally, we are extending the system with external sensors, such as a thermal camera. We attached a FLIR ADK thermal camera directly to the MHL2 with a 3D-printed mount (Figure 2) and calibrated the system. To incorporate the thermal data into point clouds, we developed a mapping method, which enables generating point clouds coloured

with RGB and thermal camera data.

A prerequisite for fusing data of different sensors is that the data must be transformed into a uniform coordinate system (co-registration). For example, separate point clouds must be registered to each other or thermal images must be projected onto the point clouds.



Figure 2. System setup consisting of the MHL2 in combination with the FLIR ADK thermal camera

The registration of point clouds can be achieved using methods such as RANSAC [9] or Iterative Closest Point (ICP) [10]. These techniques are also used for MLS to register point clouds in real-time. For example, using visual odometry [11] or plane-based methods as shown by Wujanz et al. [12]. Another example of fusing image data and other sensor data is shown in Effkemann et al. [13]. Furthermore, many manufacturers already offer software for the registration process, for example Riegl RiSCAN PRO. Initial methods and results to register thermal images with a model of the outer shell of a building via homogeneous points and project the images onto the facades are shown by [14]. Many methods, however, require lengthy manual steps by the user. Next to constructing a reality capturing system, an important goal of the energyTWIN project therefore is the development of an automated registration process.

Key-information for registering point clouds, georeferencing data or also for visualizing data in VR and AR, are the viewing direction (orientation) and location (position), referred to as pose, of the user or more specifically the device. Any movement of the user must be tracked in real-time, so the pose stays up to date. This is referred pose tracking. There are two basic types of pose tracking: outside-in and inside-out. In outside-in pose tracking, an exterior device, for example attached to a wall, observes the device carried by the user to estimate its pose. In inside-out pose tracking, the tracking device is carried by the user and it observes its surroundings for pose estimation. Outdoors, typically, outside-in methods, e.g. satellite-based localization systems (GNSS), are used. Since GNSS are usually not applicable in covered areas, methods based on the radio technologies Ultra-Wideband (UWB), Bluetooth or WLAN, as well as on infrared or ultrasound, are required for indoor applications [15].

However, these systems need a complex installation. Therefore, the project focuses on inside-out methods, specifically camera-based or laser-based methods, to offer flexibility. In contrast to outside-in, inside-out methods do not require external infrastructure.

A pose requires a reference system, local or global. In VR, typically a local coordinate system is used for the virtual world, oriented during the setup process of the VR equipment. For example, the front facing direction is set, so that the user always starts relative to it. In AR, also often local coordinate systems are used, for instance initialized when the AR system is started. Objects then appear relative to the starting pose of the user. Since AR is much more linked to the physical world than VR, it is beneficial to apply a global coordinate system like World Geodetic System 1984 (WGS 84) or another existing real-world coordinate system, such as a building coordinate system. A common virtual and physical coordinate system, such as a building coordinate system, enables attaching virtual information to physical elements or augmenting these as shown in [16, 17].

Therefore, we will realize a method to transform the MHL2 from its local coordinate system into a building coordinate system and a method to accurately track the pose of the MHL2 in the building coordinate system. For this, we will use the georeferenced BIM-based 3D planning model to register the MHL2 to the building. Building parts of the virtual model can be used as a reference to find the pose of the user in the physical building and then calculate the transformation from the local to the building coordinate system. Afterwards, we will track the user's pose relative to the initialization pose in the building coordinate system. A promising method for this is Simultaneous Localization and Mapping (SLAM) [18]. SLAM uses pose information to generate a local map (3D point cloud) of the environment. The map in turn is employed for pose estimation. Using bundle block adjustment, the relative camera poses are optimized based on the 3D point cloud to obtain a highly accurate local trajectory. For a globally consistent trajectory, a loop closure method is applied. Loop closure refers to returning to a previously visited location and incorporating past pose information into current estimates. The most common types of SLAM are camera-based visual SLAM (V-SLAM) and laser-based light detection and ranging (LiDAR)-SLAM. While V-SLAM uses corresponding salient points in sequences of photographs (feature points) to estimate the motion of the physical camera, LiDAR-SLAM uses sequences of 3D point clouds. While V-SLAM has more information available than LiDAR-SLAM due to the use of cameras, it is more susceptible to different lighting conditions. This must be taken into account, especially for indoor pose tracking in dark rooms.

4 VR/AR for semantic data enrichment

In the energyTWIN project, VR and AR are utilized to visualize a variety of data, to support the user in visually comparing the planned (as-planned/as-designed) and the real-world (as-built) situation, in order to correct the as-planned model with the detected deviations.

VR and AR systems differ by the amount of digital and real-world content the user is presented with. In VR systems, the user can only see digital content and his real-world location becomes less relevant. In room-scale VR, as often possible with modern VR systems, the user is able to move around in a virtual environment and interact with it by moving in his physical environment with hand-held controllers or other devices. In an AR system, a large part of the visual input consists of the real world which the user is surrounded by. The project goal is to develop a system, which allows the user to walk through the building to commission the different parts of the TBE. Therefore, a stationary AR system is not suitable, but a mobile one including georeferencing is needed. For a mobile georeferenced AR system, three subtasks need to be solved: (1) the data needs to be prepared and visualized in real-time, (2) the physical world needs to be observable and (3) methods need to be implemented to combine both worlds.

For energyTWIN, a two-pronged approach has been chosen. In a VR system, the user will be able to interact off-site with purely digital content and use combinations of data from the as-planned model and the captured point cloud data, which represents the as-built situation. In AR, the user will be able to overlay selected data in the real-world on-site and interact with it, for example, the as-planned model, to visually compare it to the actual built situation. As a development framework, we are using Unreal Engine (UE). While UE primarily is a framework for developing computer games, recent additions like interfaces for IFC and point clouds enable the software for business and industry applications. In our VR solution, the user will have two basic datasets available that form the digital world surrounding him: On the one side, the as-planned BIM model in the form of geometry components from IFC and on the other side, point clouds that represent the real-world situation. The user will be able to move around by physically moving and by so called teleporting. For teleportation, the user aims the controller at a point in the virtual space and by pressing a button, he is virtually transported to this location. IFC elements offer much more information than only the geometry. The user will be able to access this information via a User Interface (UI) directly in VR (Figure 3).

The UI consists of a virtual panel attached to the left and a laser pointer to the right virtual hand. The developed RDS (section 1) is used to filter by criteria such as function type of the system.

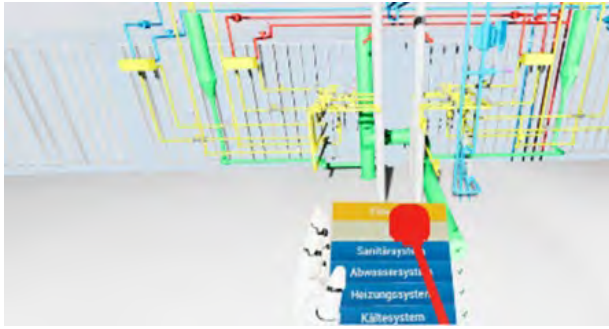


Figure 3. UI in VR

When a deviation between the as-planned and as-built situation is identified, the user can place an issue, using a ticket system. Selected from a predefined list, comments can be attached to individual components of the IFC model. The issues will be exportable using the IFC-based standardized BIM Collaboration Format (BCF), to allow interoperability between the subsystems of energyTWIN.

In the AR, the user will be able to visualize the planned model (IFC model) on-site and interact with it with the MHL2 and its gesture recognition system (Figure 4). For possible performance issues with the MHL2, a solution involving pixel streaming technologies is being tested. With it, heavy processing tasks are outsourced to an external computer or cluster and only the prerendered images are sent to the MHL2 to be displayed. With this AR system, the user will be able to capture data only visible on site and will be guided through processes such as TBE installation or maintenance.

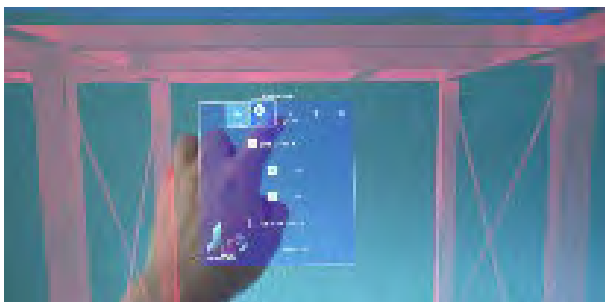


Figure 4. User interaction in AR

5 AI based methods for element recognition and classification

In the first step, we focus on developing methods for the automatic filtering and extraction of features to recognize and classify TBE and its topological, functional and informational relationships from image-, laser- and infrared-based data, to obtain the as-built BIM. In a second step, the extracted TBE should be automatically compared with the as-planned BIM in

order to model the objects geometrically and semantically. In addition, rule-based methods of clash detection are also integrated within this operation to detect differences between the models. This process allows the as-planned BIM to be upgraded to an as-built BIM. A challenge is handling the large data amounts and their complexity, since these are hardly manageable manually, so that automated methods are indispensable. Therefore, we are investigating which approach suits the project's purposes best.

Today, AI enables analysing large and possibly unstructured data sets (big data). Two possible options are classical Machine Learning (ML) methods or more recently Deep Learning (DL) using artificial neural networks (ANN). AI-based methods have already been successfully applied to building element reconstruction, for example of indoor scenes [19]. In this context, the neural networks PointNet [20] and VoxelNet [21] are particularly noteworthy, which differ in the way geometric features are processed and classified. However, modified and refined models such as Voxel-FPN [22] have also proven to be effective in detecting objects in outdoor and indoor environments. Some models such as the MVX-Net [23] also fall back on data from different sensors and merge the separate classification results to achieve a higher level of accuracy. Another possibility for system topology recognition is the classification of building technology time series, as described in their basic functioning in [24]. While DL can handle large amounts of data, its disadvantage is that it also requires large amounts of training data and significant training time. Furthermore, numerous aspects such as the selection of suitable parameters and training features must be considered for preventing problems such as overfitting of the models.

To support the AI, we will integrate the as-planned BIM into the process. This will allow us to apply existing information to the algorithms (knowledge-based). Next to the knowledge-based solution, we will also realize a geometry-based method without prior knowledge, to analyse the data based on geometrical properties. As a third option, we will implement a DL-based solution. Finally, all solutions will be compared, and the most suitable solution for the task will be identified.

The knowledge-based approach is characterized by the fact that the as-planned BIM acts as a reference in the analysis of the captured point cloud data. A prerequisite is a point cloud sampled from the as-planned BIM [25]. To generate such an as-planned point cloud, a ray-casting algorithm will be utilized. Furthermore, the captured point cloud will be segmented object-wise by deriving a bounding box from the corresponding objects from the as-planned BIM and transferring it to the as-built point cloud to cut out an object-specific point cloud. Therefore, a registration of the as-planned BIM and the captured

point cloud must be realized first, so that both are in a common spatial coordinate system. Following this process, certain object data can be transferred directly from the as-planned to the as-built BIM model. To investigate the relative position of the objects, an approach of a two-stage co-registration with coarse and fine positioning will be investigated using known techniques of geometric hashing [26], RANSAC [27] and (modified) iterative closest point (ICP) algorithms [28-30]. On the other hand, the method presented in [31] for the as-built modelling of cylindrical objects using a Hough Transformation will be investigated on its application potential for the described project goal.

For the geometry-based approaches, the detection of corners and edges via the investigation of local point densities and neighbourhood relationships [32] or the skeletonization of object representations in the as-built point cloud [33] will be analysed.

In terms of a DL-based approach, existing methods for point cloud semantic and instance segmentation [34-36] will be evaluated, further customized and advanced to better deal with the particular challenges of TBE.

6 Cloud-based system for data provision and failure diagnostics

The central element of the energyTWIN system architecture is the content server, which stores the BIM model, initially in the planning state, in the form of geometric IFC data of the building and its TBE (Figure 5). Each object of the TBE is uniquely described and locally identifiable via a link to its id in the RDS (see section 2).

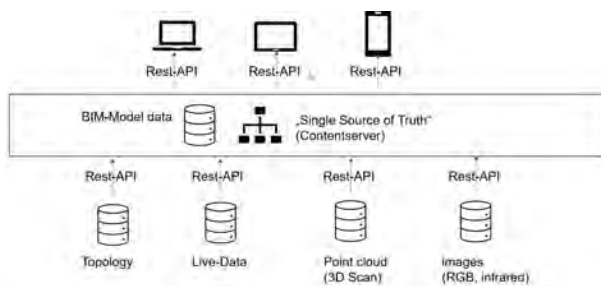


Figure 5. Concept of the cloud infrastructure

The data provided by the project partners, such as topological data, point clouds from laser scanning systems, camera images in the RGB and infrared range, as well as continuous operating data from the field, are not directly stored on the content server, but are held in separate cloud-based systems, each with their own query interfaces. The content server queries these data sources as needed and makes the results available to the user in real time in the context of the BIM model and its RDS.

For this purpose, the content server in turn provides an interface, so that clients, such as web-interfaces or smartphone apps, can be connected and all collected data is visualized, easily comparable and always available.

In this way, the commissioning of the TBE is optimized. Deviations from the planning status, which result from the evaluation of scan and photo data, are documented via a ticket system connected to the content server, with the aim of correcting the as-planned model and, thus, creating a valid modified BIM model that corresponds to the actual conditions.

In the finished model, the sensors located in the field, such as temperature or pressure sensors, are then displayed at the topologically correct location throughout the entire operating time of the building, and the operating data provided by the sensors can be easily queried at any time, which greatly facilitates both fault diagnostics and maintenance. Should maintenance and other tasks become necessary in the future, planning and preparations can be made by combining scans, 3D views and continuous field data with the digital twin, to significantly reduce the number of necessary on-site visits.

7 Evaluation and Demonstration

All the developed methods and the results are evaluated throughout this project. For this purpose, various demonstrators are being set up for testing the relevance of the developed methods in practice. A small-scale demonstrator (Figure 6) was set up on-site, including several technical installations for a drinking water system.



Figure 6. Point cloud (NavVis VLX) of the ViegaCUBE (Viega GmbH & Co. KG)

This is used for all the partners to exploit synergies and to test the already linked work progress. The demonstrator was already put into operation and a first digital twin of it was created. To create the twin, the NavVis mobile mapping system VLX and the terrestrial laser scanner Riegl VZ400 were used as reference

systems, next to the MHL2. First evaluations show that an accuracy of 1-3 cm can be expected from the MHL2, therefore, on a level with dedicated MLS systems. Further, more detailed evaluations will follow later in the project. On the one hand, the digital twin is represented through a point cloud. On the other, the NavVis software offers a web-based viewer, with built-in panorama, point cloud views, and the possibility of creating points of interest. It is used by the project partners to collaborate on the digital twin.

After testing the developed processes on a small-scaled system and evaluating the techniques and methods, they will be evaluated based on their applicability and relevance and subsequently changed or improved. The next step is the preparation of a large-scale demonstrator for further testing and evaluating. Therefore, the associated project partner provides a new constructed office building in Koblenz, Germany. The evaluated methods in the demonstrators provide insights into the accuracy of the applied methodology of the overall project and can, when considered, improve the overall accuracy.

8 Conclusions and Outlook

Within this paper we presented first results of our research. Technologies like indoor positioning, VR/AR, and various sensors (laser scanning, infrared etc.) and mobile devices (MHL2) are combined for data capturing. AI based methods will be used for object element classification and comparison to the as-planned model as well as for detecting and analyzing topological and functional relationships in TBE. A cloud-based system provides and exchanges the data on-site. RDS are used for unique building and TBE element identification. Only the combination of all these technologies enables enough knowledge about deviations between the BIM-planned TBE and its actual commissioning, so that the as-planned BIM model can be updated to the actual, built situation, and subsequently used for optimization of TBE. This promises an improved commissioning for increasing energy efficiency in the building's operation phase.

Future research will focus on increasing the accuracy of the capturing system which also involves increasing the pose tracking accuracy. Further improvements will involve increasing the automation level, for example, the AI-based segmentation and classification of TBE, also for complex elements.

Acknowledgement

The authors gratefully acknowledge the financial support of the German Federal Ministry for Economic Affairs and Climate Action in the project "EnOB: EnergyTWIN - Energiediagnosestecker Digitaler

Zwilling. Neue sensorgestützte und KI-basierte Methoden für die digitale, BIM-basierte Inbetriebnahme von technischen Anlagen in Hochbauwerken und deren energetische Systemoptimierung" (reference number 03EN1026).

References

- [1] buildingSMART International Ltd. Industry Foundation Classes 4.0.2.1. Online: https://standards.buildingsmart.org/IFC/RELEASE/IFC4/ADD2_TC1/HTML/, Accessed: 13/01/2022.
- [2] Industrial systems, installations and equipment and industrial products – Structuring principles and reference designations – Part 1: Basic rules. DIN EN IEC 81346 – 1. International Electrotechnical Commission, London, 2020.
- [3] Essig B. *BIM und TGA – Engineering und Dokumentation der Technischen Gebäudeausrüstung*, v 3. Beuth Verlag GmbH, Berlin, 2021.
- [4] Industrial systems, installations and equipment and industrial products – Structuring principles and reference designations – Part 2: Classification of objects and codes for classes. DIN EN IEC 81346 – 2. IEC, London, 2020.
- [5] Building automation and control systems (BACS) – Methods and tools for planning, building, and acceptance tests – Identification, addressing, and lists. VDI 3814 4.1. VDI. Düsseldorf, 2019.
- [6] Stinner F., Kornas A., Baranski M. and Müller D. Structuring building monitoring and automation system data. *The REHVA European HVAC Journal*, 55(4):10–15, 2018.
- [7] Blankenbach J. Building Surveying for As-built Modeling. *Building Information Modeling*: 393–411, 2018
- [8] Witte, B., Sparla, P., Blankenbach, J.: *Vermessungskunde für das Bauwesen mit Grundlagen des Building Information Modeling (BIM) und der Statistik*. Buch. 9. Auflage. 2020. Wichmann. S. 422-453. ISBN 978-3-87907-657-4
- [9] Fischler M. and Bolles R. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24(6): 381-395, 1981
- [10] Bouaziz S., Tagliasacchi A. and Pauly M. Sparse iterative closest point. *Computer graphics forum* 32(5): 113-123, 2013
- [11] Stumberg, L. von, Usenko V. and Cremers D. Direct Sparse Visual-Inertial Odometry using Dynamic Marginalization. In: *2018 IEEE International Conf. on Robotics and Automation (ICRA)*, pages 2510-2517, Brisbane, 2018
- [12] Wujanz, D., Gielsdorf, F., Romanschek, E., Clemen,

- C.: Ebenenbasiertes Baufortschrittsmonitoring unter Verwendung von terrestrischen Laserscans. Conference: 18. Oldenburger 3D-Tage: Optische 3D-Messtechnik - Photogrammetrie – Laserscanning. Oldenburg, Germany
- [13] Effkemann C., Schwermann R. and Blankenbach J. Kalibrierung und Navigation eines Überwasser-Mapping-Systems für die Erfassung von bildhaften und sensorischen Gewässerparametern. *Ingenieurvermessung* (17): 113-130, 2017
- [14] Zhu J., Xu Y., Hoegner L. and Stilla U. Direct Co-registration of TIR Images and MLS Point Clouds by Corresponding Keypoints. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* (4), 2019
- [15] Blankenbach, J. (2017): Indoor-Positionierung & lokale Positionierungssysteme. In: Brand/Blankenbach/Kolbe (Hrsg.): Leitfaden - Mobile GIS, Von der GNSS-basierten Datenerfassung bis zu Mobile Mapping, vol.6, no.3, Selbstverlag, Runder Tisch GIS e.V., München, Deutschland.
- [16] Blut, C., Blankenbach, J.: Three-dimensional CityGML building models in mobile augmented reality: a smartphone-based pose tracking system, *International Journal of Digital Earth*, 14:1, 32-51, 2021. DOI: 10.1080/17538947.2020.1733680.
- [17] Blut, C., Blut, T., Blankenbach, J.: CityGML goes mobile: application of large 3D CityGML models on smartphones, *International Journal of Digital Earth*, 12:1, 25-42, 2019. DOI: 10.1080/17538947.2017.1404150
- [18] Dai A., Chang A. X., Savva M., Halber M., Funkhouser T. and Nießner M. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In *Proc. of the IEEE conf. on computer vision and pattern recognition*, p. 5929-5839, Honolulu, 2017
- [19] Durrant-Whyte H. and Bailey T. Simultaneous localization and mapping: part I. In *IEEE Robotics & Automation Magazine* 13(2): 99-110, 2006
- [20] Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *arXiv preprint arXiv:1706.02413*, 2017.
- [21] Zhou, Y., Tuzel, O. Voxelnets: End-to-end learning for point cloud based 3d object detection. In *Proc. of the IEEE conf. on computer vision and pattern recognition*, p. 4490-4499, Salt Lake City, 2018
- [22] Wang B., An J. and Cao J. Voxel-FPN: multi-scale voxel feature aggregation in 3D object detection from point clouds, 2019
- [23] Sindagi V. A., Zhou, Y. and Tuzel, O. MVX-Net Multimodal VoxelNet for 3D object Detection. In *2019 International Conf. on Robotics and Automation (ICRA)*, p. 7276-7282, Montreal, 2019
- [24] Fütterer J., Kochanski M. and Müller D. Application of selected supervised learning methods for time series classification in Building Automation and Control Systems. *Energy Procedia* (122): 943-948, 2017
- [25] Koschwitz D., Spinnraker E., Frisch J., van Treek C. Long-term urban heating load predictions based on optimized retrofit orders: A cross-scenario analysis. *Energy and Buildings*, (208): 134-142, 2020
- [26] Wolfson H.J. and Rigoutsos I. Geometric hashing: an overview. In *IEEE Computational Science and Engineering* 4(4), p. 10-21, 1997
- [27] Chen C., Hung Y. and Cheng J. RANSAC-based DARCES: a new approach to fast automatic registration of partially overlapping range images. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21(11), p. 1229-1234, 1999
- [28] Besl P. and McKay N. Method for registration of 3-D shapes. In: *Sensor fusion IV: control paradigms and data structures. International Society for Optics and Photonics*, 1992
- [29] Chen Y. and Medioni G. Object modelling by registration of multiple range images. *Image and Vision Computing* 10(3): 145-155, 1992
- [30] Xie Z., Xu S. and Li X. A high-accuracy method for fine registration of overlapping point clouds. *Image and Vision Computing* 28(4): 563-570, 2010
- [31] Bosché F., Ahmed M., Turkan Y., Haas c., Haas R. The value of integrating Scan-to-BIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: The case of cylindrical MEP components. *Automation in Construction* (49B): 201-213, 2015
- [32] Gumhold S., Wang X., MacLeod R. Feature Extraction From Point Clouds. In *IMR*: 293-305, 2001
- [33] Lee J., Son H., Kim C., Kim, C. Skeleton-based 3D reconstruction of as-built pipelines from laser-scan data. *Automation in construction* (35): 199-207, 2013
- [34] Jiang, L. and Zhao, H. and Shi, S. and Liu, S. and Fu, C. and Jia, J. PointGroup: Dual-Set Point Grouping for 3D Instance Segmentation. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 4866-4875, 2020.
- [35] Han, L., Zheng, T., Xu, L., Fang, L. OccuSeg: Occupancy-aware 3D Instance Segmentation. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2937-2946, 2020.
- Thomas, H., Qi, C., Deschaud, J., Marcotegui, B., Goulette, F., Guibas, L. KPConv: Flexible and Deformable Convolution for Point Clouds. In *Proceedings of the IEEE International Conference on Computer Vision*, 6410-6419, 2019.

Blockchain-supported design tool on building element scale

G. Sibenik^a, M. Sreckovic^a, T. Preindl^b, M. Kjær^b and W. Kastner^b

^aInstitute of Interdisciplinary Building Process Management, TU WIEN, Vienna, Austria

^bInstitute of Computer Engineering, TU WIEN, Vienna, Austria

E-mail: goran.sibenik@tuwien.ac.at, marijana.sreckovic@tuwien.ac.at, thomas.preindl@tuwien.ac.at,
martin.kjaer@tuwien.ac.at, wolfgang.kastner@tuwien.ac.at

Abstract

Sharing building data or building models still represents a problem within design practices in the architecture, engineering and construction industry. Additionally, digitalization, automation or traceability of processes face numerous workflows and changing stakeholder constellations, with multiple software tools escaping the scopes of common data environments or similar digital solutions. Vague standardization regarding data and processes hinders data management technologies from overcoming the design phase digitalization issues. While many central solutions still deal with closed data due to many proprietary tools for domain-specific tasks, each building project requires an inter-domain collaboration. Open solutions to holistically manage projects still lack functionality, even though some existing tools support central data management and process automation. This research investigates data management using popular data exchange formats for coupling with blockchain technology. It establishes a system that can support processes with smart contracts and reference building elements, herewith addressing the question: How to manage data on the building element scale to allow for processes defined with smart contracts and blockchain technology? The resulting system architecture combines Revit as user-local storage and Speckle as an open CDE. Furthermore, it uses the Baseline Protocol for data exchange and as a common point of reference. While data exchange happens off-chain, cryptographic hashes of data are stored on the blockchain to form a single point of reference for process states and all previous versions, creating process chains and allowing data traceability. Data tracing is an essential requirement for building projects, still commonly realized in analog form in practice. This research presents mechanisms for blockchain-based data tracing on a level of granularity required for design processes.

Keywords –

Blockchain; Baseline; CDE; BIM; design workflow

1 Introduction

The building design process is characterized by numerous stakeholders contributing to the design of a single real-world product. However, such a single real-world product is not reflected on a single data repository due to several reasons: models differ between domains, domain-specific models are proprietary, the data exchange process is burdened with difficulties, data ownership is not regulated, existing platforms exclude some stakeholders or software tools, to mention a few. Although ISO standard 19650 [1] suggests using a common data environment (CDE) for the building design process, it does not specify how it should be used and hence differs between projects across the AEC industry. The realization of CDE products is not standardized, and the products provide various functions and solutions [2]. Besides all the problems existing in the data management of a building project, blockchain (BC) is evermore present in all the phases of the building life cycle [3]. The visionary advantages of distributed ledger technologies might be suitable for resolving the AEC industry's communication issues [4]. Therefore, we aim to improve the design process with BC technology. The BC concept for the design phase that we propose allows for better transparency, traceability and data reliability. It can improve the communication between stakeholders with limited trust and individual data management systems (DMS). Building data as the main product of the design phase and additional necessary information must be adequately integrated with the BC concept and further made available in a form suitable for generating added value facilitated by the new technology. However, data required for this purpose has not yet been structured to relate it to BC; even for integrated planning, it has not been organized in a suitable way and on a sufficient level of granularity. Therefore, this paper aims to investigate the organization of information and documents and its appropriateness for the design process with BC. The BC-supported design phase framework needs a critical survey of DMSs.

DMSs considered in this research are:

- a communication platform using documents represented as data files
- a server using IFC standard, which manages data on an object scale
- an existing CDE solution and corresponding data management
- a previously developed solution for data exchange between architectural design and structural analysis using a database MongoDB for storage

These DMSs will be the center of the investigation. We focus on the usefulness and usability of their data organization strategies regarding their integration with a BC solution realized through the Baseline Protocol in the building design process. The following section reviews the existing literature regarding the BIM-based design process, DMSs and versioning, and already recognized relations to BC technology. The methodology used to test the data management solutions is presented in Section 3; findings are presented in Section 4 and the system prototype demonstrated in Section 5.

2 Literature review

The literature review briefly addresses three topics that are relevant for this research:

- Design processes – described as analog, workflows are heterogeneous and in practice rarely automated
- DMSs and versioning – isolated solutions provided for specific practices, where proposed solutions do not form a technological ecosystem for inclusion of all domains
- BC – popular in recent years, has not yet reached a useful solution for broader use, problems and benefits in the construction sector need further exploration.

2.1 BIM-based Design Process

BIM is considered to be a “use of a shared digital representation of a built asset to facilitate design, construction and operation processes to form a reliable basis for decisions” [1]. A built asset is of interest to many stakeholders throughout all stages of its lifecycle. These stakeholders often have different interests and accordingly use domain-specific representations of an asset. They may not even use building models which involve building geometry and may solely use alphanumeric data [2]. Three ways of achieving model-based collaboration are recognized [6]: through separate BIM models, through separate BIM models with an aggregated model, and a single BIM model.

Apart from these technology-related issues, BIM contractual arrangements concerning collaboration among stakeholders need to be taken into consideration

[7]. A high-level design process structuring [8], [9] serves well for establishing contractual relations between the stakeholders, but mostly between hierarchically separated actors. Stakeholders constitute a loosely-coupled system in the most common case [10], [11] where multiple companies cooperate on a single project, based on contractual models arranging their relations. Nevertheless, communication between domain-specific planners is not contractually regulated [12], and results in numerous workflows across the industry. Patterns can be recognized between these activities [13], supporting possibilities for standardization. Attempts to standardize a BIM-based design process originating from the project management domain are based on traditional workflows and do not result in an automatable standard (e.g., [14]). Processes for cost estimation can be investigated through logs [15], which is software-specific and does not show the processes which stand in relation with other stakeholders (communication processes). Such processes are generally underinvestigated, present a research gap [16] and a requirement for workflow automation. Design workflows, and herein communication processes have the potential to be supported through Smart Contracts – i.e. decentralized computer protocols that autonomously, self-execute predefined tasks [17], fostering simplified, interdisciplinary processing of design tasks [18].

The research stemming from the project BIMd.sign [19] describes three scenarios in the design phase for the use realization of smart contracts. Such scenarios will represent the point of departure for this research due to the lack of documented and standardized communication processes between the stakeholders, and the scenarios will be more closely explained in the methodology section.

2.2 DMSs and Versioning

Workflows are heterogeneous, not sufficiently documented, and not automated. However, some technological solutions for collaboration exist on the market which respond to a certain amount to the stakeholders’ requirements. These stakeholders generate, edit and use digital assets in the design phase and materialize them as physical assets in the construction phase. The design phase is focused primarily on digital assets and therefore has potential for digitalization of processes, but its complexity is challenging to keep under control. A CDE [1] recommends four kinds of folders or document containers for digital assets: work-in-progress, shared, published, and archived, describing their state. Using a teleconferencing system to support collaboration is proposed in [6], yet using a single proprietary software tool to realize the communication. The authors suggest several communication patterns, including one-to-one, one-to-many, many-to-one, and many-to-many, depending on how the model is accessed, edited, and

shared across the participants. Authorization regarding model access is a complex topic in the AEC industry. It is most often object-based, depending on the building elements of interest, the domain, and the required task in the domain [20]. Therefore, granting access to an entire model or document might not be desired, especially in the case of a complex building model and design team constellation. Merging fragmented models or data is required for domain-specific processes, as all domains direct the information towards a single resulting product [6]. The design of a building is frequently updated, and documents and models display numerous versions throughout the workflow. Designs change on the object scale, but the technical difficulties in managing the changes on the object scale result in redundantly exchanging models or documents even in the cases of modification of only a few building elements. Four methods can be used to compare building models for design changes [21]: matching-first, comparison-first, hash-code-accelerated and quick hash-code-accelerated. The methods provide different algorithms to compare data, and the comparison accelerated with the hash code shows the most promising results. This approach however does not consider multiple software tools and the small and middle-sized enterprises (SME) involved in projects, which prevail in the AEC industry. Difficulties in addressing classifications and data structures remain, making such approaches much more challenging to implement. The focus of [22] is on the algorithm-based design including systems such as Git, a standard solution for coordinating software tool development. There are two types of version control systems: centralized and distributed. Both are relevant for the BIM-based design process, and as discussed in [22], the centralized systems may be suitable for smaller projects, while the distributed ones are suitable for the larger ones. Their work is a significant contribution to the version control during the design process; however, it only considers designers in the process and not a complete BIM environment which may display additional challenges. The algorithmic design does not correspond to BIM data management due to a core difference between code-based and object-based data management. Although [22] present an algorithmic design oriented towards BIM (and named A-BIM), the relation with BIM lacks clarity.

Automatic versioning of industry foundation classes (IFC) exports, building models defined with the open IFC standard on the object scale, is investigated in [20]. Differences between IFC models are detected in [23], naming them semantic differential transactions (SDT) to record only changed information compared with the base model. These SDT models are referenced with BC to avoid the redundant storing of building models. The authors use IFC models, with all their shortcomings

recognized in the literature [24]. Their work significantly improves how the models are referenced, reducing the size of model versions and giving a basic idea of a BC connection. Still, versioning does not meet the designer requirements described in [21]. Another investigated versioning solution is the ontological representation of building models. Difficulties in assigning unique and stable identities to the numerous anonymous nodes corresponding to a Resource Description Framework (RDF) representation of IFC building models are recognized [25]. Although oriented towards the technical implementation of IFC identifiers, the issue of erroneous exports from heterogeneous software environments towards the open schema [24] is not widely discussed in [25]. Numerous concepts of building data management exist, differing by the ability of technology to record the information of a built asset. Various efforts to analyze and enhance management and versioning of building data can be found, however the ones supporting existing design processes with BC are missing. Therefore, this work considers multiple DMSs to find a suitable answer to the requirements of a BC-supported design workflow.

2.3 BC in the Design Phase

BC represents a distributed ledger technology that may facilitate the exchange of assets without a trusted third party, such as its use for cryptocurrencies. Research on using BC in the AEC industry is mostly not focused on the design but on the construction phase, where a significant amount of monetary exchange occurs [26], [27]. Using BC to support communication between stakeholders is still not sufficiently investigated. One of the papers focusing on communication involving digital information is [12], emphasizing the benefits of BC as a supplementary technology for improving design liability. The work is well appreciated for pursuing BC use in the design phase, but it simplifies data management and workflow heterogeneity. Workflow complexity represents a severe obstacle in communication processes, and improving communication flows is also a motivation for the use of BC. A framework for the design phase which uses BC is developed in [23]. To avoid information redundancy, the authors consider only design changes and store their reference using BC. They recognize problems with object IDs which are also recognized and thoroughly investigated in [25].

Although not explicitly dealing with BC implementations, information such as actor, timestamp, entity name, element ID, type of activity, and name of the modified attribute is recorded in [20] as relevant for the transactions. In their work, the authors use Autodesk Revit, and in that way, problems existing with various IDs and data management problems occurring with multiple tools vanish.

In this research, the Baseline Protocol standard is

used to implement a system on the BC side, defining a framework for data and workflow synchronization called Baseline Protocol Implementation (BPI). Baseline is an emerging industry standard initiated by the Enterprise Ethereum Alliance (EEA), an association of leading organizations from different industries with the aim to drive the use of Ethereum. It defines a way to communicate on the inter-organizational scale by assuring communication correctness and verifiability using BC technologies. While data exchange is facilitated off-chain, meaning without a BC, smart contracts store cryptographic references and verify the correctness of the exchanged data. In this way, data is kept private between the interacting organizations, which is an important aspect not addressed in other solutions. [28].

3 Methodology

This research employs a qualitative analysis of data organization strategies of existing DMSs. The systems to be analyzed result from the investigation of the existing literature in Section 2.2. Several data management approaches exist, and this research aims at investigating different system types, so the following four DMSs are considered: a) file-based platform with the exchange of building-related information through a document exchange, b) IFC-object-based server, c) CDE solution Speckle, d) MongoDB system developed for the exchange between architectural design and structural analysis models.

Data management approaches can be modified to a certain degree if necessary. The novel proposals aim at providing recommendations for the overall system, serving other prototype systems and supporting different scenarios. This work seeks to interrelate and analyze the four existing DMSs - a document base platform, IFC server, Speckle, and MongoDB system - to realize an optimized design process with a BC system delivered with Baseline Protocol. As a result, a traditional planning process will be modified for adoption into the interrelated systems.

Our methodology incorporates two systems, BC-supported communication and shared data management, which will be assessed for their mutual performance in parallel with their suitability to support the design process (Figure 1). The goal here is to investigate the appropriateness of systems to facilitate traceability of shared building representation changes.

The design process might change with new technologies emerging [29], [30], [31]. However, as a starting point, an existing building design process will be considered, herein attributed as traditional, although we investigate a process using BIM authoring tools. Heterogeneity of workflows, numerous stakeholders, and non-standardized processes make it challenging to cover

a whole design process [33], [32]. Therefore, we use planning scenarios from [19] as a base for system testing.

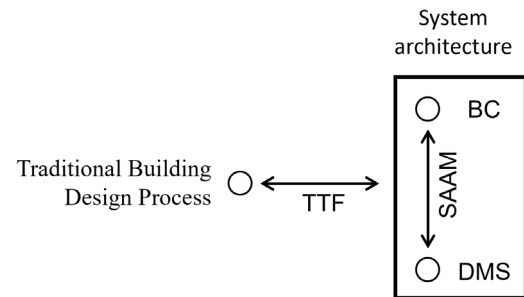


Figure 1. Methodology overview

The scenarios in [19] are described as follows:

1. a typical conceptual design scenario, with an investor, a general planner and a domain-specific planner executing tasks of creating a new design. After the generation of a new design, instructed by the investor, through the general planner to the domain-specific planner, it requires approval by the general planner and finally the investor.
2. scenario involving two domain-specific planners, where a structural building element is changed and sent by an architect for approval to a structural engineer. It involves communication on the building-element level.
3. cost calculation involving a general planner, a BIM manager, as well as a cost department employee with other domain-specific planners. The request originating from the general planner, is further concretized by the BIM manager. The BIM manager indicates required parameters and data for the domain-specific planners so the cost calculation would be optimized. With updated models, the cost department provides the calculations, which are subsequently approved by each domain-specific planner and finally the general planner.

In our methodology, the task-technology fit (TTF) model relates the building design process with system architecture. TTF is defined as the degree to which technology assists an individual in performing their tasks [34]. We identify eleven focus requirements that will subsequently be relevant as BC-supported information from scenarios described in [19]. The scenario analysis identifies a digitalization potential in each step of three scenarios, and identifies which data is required for each particular step. The requirements are investigated within four DMSs in combination with BC technology. This method, called the Software Architecture Analysis Method (SAAM) [35], tests multiple systems for their performance based on the scenarios. After testing the appropriateness of system architecture, we pursued implementing a system involving various technologies,

which is described in Section 5.

4 Results

The evaluated design phase tasks are not entirely digitalized in the existing systems, although BC and various DMSs offer a spectrum of possibilities. Tasks are defined with the help of the literature review [12], [19], [35], [36], [37], [38] and serve as a filter for a detailed listing of requirements from the scenario analysis. The technology combination of BC and DMS could facilitate the following tasks:

- Allow digital assignment of tasks in the design phase
- Partially automate the assignment of tasks
- Authorize actors for activities
- Map performed activities on an independent storage
- Automatically report a performed activity
- Relate activities to a corresponding asset
- Guarantee the communicated asset based on its content
- Address assets on document (file) and building element (object) level
- Validate activity based on predefined rules

The listed tasks would significantly improve the design workflows if realized with any or both technologies. Already the systematic digitalization of processes enables further features like data analysis, which is currently not widely available in the AEC industry. However, new challenges will occur with the realization, and would need to be addressed over time. These are: more extensive energy requirements, expensive transactions, safety and security of information, rigidity of predefined methods, user-friendly interventions, scalability [3]. These challenges open new possibilities for numerous business models in the AEC industry.

With a SAAM method, three scenarios investigated in [19] delivered eleven requirements for the BC-supported design as conceptualized in [39] and requirements derived using TTF. Table 1 lists these requirements, and although a system architecture could be realized with each of the listed solutions after capturing the missing information in an alternative way, the table demonstrates the readiness of the DMSs towards the BC implementation. If a requirement is supported it means that the analyzed system in its existing form captures the needed information or provides a way to perform a specific activity. This does not mean that it is tested with various workflows and can correctly support all BIM models, but that the DMS has that functionality in its current conception, as described by the producers, answered by the users or by analyzing the DMS itself.

Table 1. Requirements for BC-supported communication (y=yes, n=no, p-partly)

Requirement	File-based	BIMserver	Speckle	MongoDB
Sender information	y	y	y	y
Receiver information	y	y	y	y
Relation between an asset and an actor	y	y	y	y
Validation of assets on file level	y	n	n	y
Validation of assets on object level	n	y	y	y
Authorization of an asset as file	y	n	n	y
Authorization of an asset on object level	n	y	p	p
Relation between activities (logs)	p	y	y	y
Report on performed activity (event)	y	y	y	y
Relation of activity and asset on the object level	n	y	y	y
Validation of activity and asset on the object level	n	n	n	n

Table 1 shows that a large part of the required information is or can be easily recorded in a certain form on diverse DMS platforms. However, none of the examined solutions provides a fully suitable scope of information. Additional information could be extracted with each solution with greater effort. An advantage is, however, recognized in open-source solutions, since the additional interventions are easily accessible. Therefore, a prototype system is further developed and demonstrated with Speckle CDE, an open-source development DMS, providing similar concepts for exchanging building data as does Git for software development. The next section demonstrates the prototype on the BC side, and its integration with Speckle.

5 Prototype Demonstration

We chose Ethereum as the preferred BC for our BPI as it is currently the most widely used BC that offers general computation. The ecosystem around Ethereum also provides sufficient additional software components, allowing an integration with various already existing systems and software libraries. Ethereum furthermore supports assigning addresses to actors, references external data with hashes, and is more generally suitable to be used by a Baseline compliant process. However, considering the constantly high transaction fees on Ethereum Mainnet, it has to be mentioned that this approach is currently not sustainable when used in practice. Improvements for using Ethereum on so-called 2nd layer networks, which accomplish a similar level of

security derived from Ethereum itself, are currently in development and should be available soon [40]. These solutions provide much lower transaction fees by batching transactions together and are thus more applicable.

Our proposed BPI uses the Speckle API to implement the functionalities which are necessary for the design process. This system architecture is shown in Figure 2.

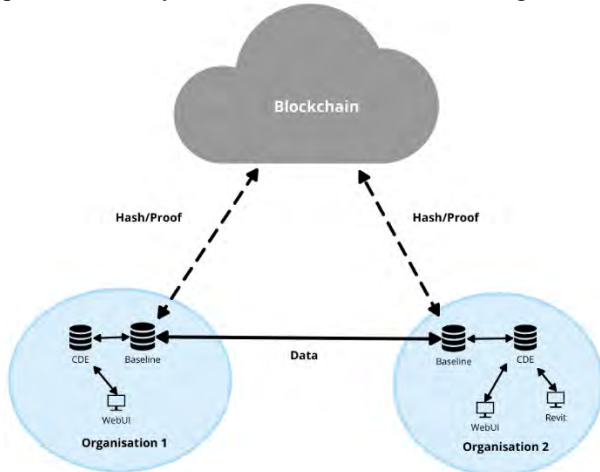


Figure 2. System architecture involving Speckle and BC

The system architecture connects existing systems of record (like a CDE) of each participant involved in the design process with the newly developed Baseline service. The Baseline service manages the data exchange between organizations. Also, it connects to the BC, where privacy-preserving proofs about the current state of both shared data objects and the overall process get stored. Interaction between the CDE, novel Baseline service, and the end-user is organized via a WebUI, which offers an intuitive possibility to follow the design process step-by-step (Figure 3). The WebUI thus abstracts technical details about the BC integration and the Baseline protocol. The novel Baseline service is furthermore able to automatically generate e-mails if

required in specific process steps, e.g., to remind actors that tasks like the rework of a building element are needed. The e-mail service can also be used as an easy way to "trigger" process steps by providing data required within a specific process step. The data enters the system in the form of documents sent by the actor as an attachment. The Speckle Revit plugin provides an additional interface for BIM models and related data. Additional interfaces to the Baseline service can integrate other tools with little effort for different domains. This is especially important as AEC practitioners do not accept well overhead activities.

6 Discussion & Conclusion

The results demonstrate in which amount the tested DMS solutions are ready for implementation with BC. None of the current DMS solutions was entirely suitable for BC employment. This means, a significant effort is required for the adaptation of existing solutions to the novel technology, and the expansion of scope of the captured information for facilitating the expected tasks, ultimately generating novel technology advantages. In this research, the work was pursued with Speckle CDE, hence, giving preference to an open source solution, actively developed at the moment. Testing all four DMS solutions requires advancements of each and eventually integration with the BC system, which was not within the scope of this work.

The newly developed tool integrates BC and Speckle. The prototype is able to support basic design communication, including the model based and e-mail-based communication. The hardest problem for the wider use of the novel tool are numerous activities and interfaces with native tools that need to be correctly defined in advance. Only the predefined SCs and the correct communication with native design and analysis tools can be tested for wider application and brought to the end users.

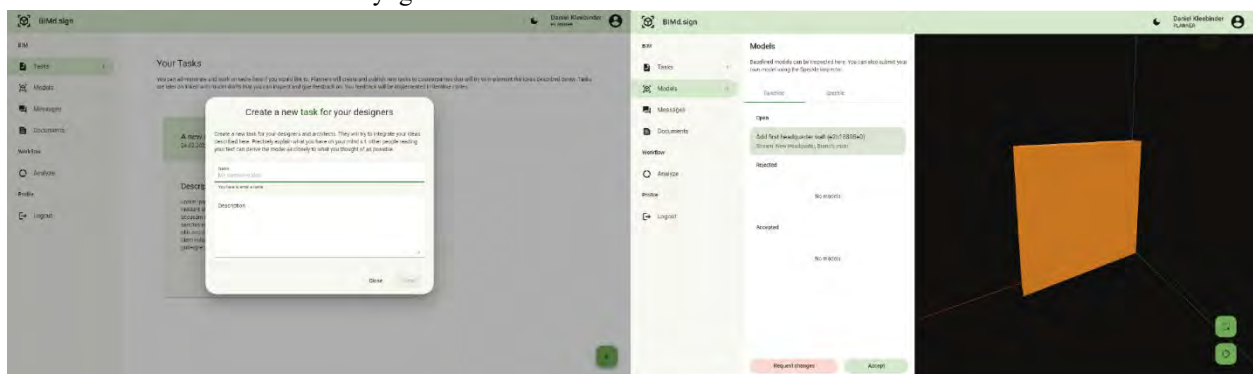


Figure 3 Screenshots of user interface in web browser connecting Baseline with Speckle: assignment of a new task (left) and model viewer (right)

This work demonstrates the use of BC and DMS to support existing design workflows in the AEC industry. Both technologies are used in a more advanced manner than found in practice. Existing research focuses on one or another technology in improving the AEC practices. In this research, both technologies' visionary implementations integrate to complement each-others features. Problems in the AEC industry are multifold, commonly grouped as people, process and technology problems [40]. Our work primarily addresses process and technology problems, and problems related to people are not within the scope. However, the proposed solution aims to minimize the overhead of design stakeholders' involvement and the influence on the existing workflows. Novel technologies are incorporated into current tools and provide additional value by capturing existing information in a suitable form. Merging the potentials of BC and DMS in a single system brings advantages such as verifiability, compliance, traceability, liability and indirectly even standardization. The recorded information further facilitates new processes and data analysis, which is currently lacking on the industry scale. This continuously improves the design process with the help of digital tools.

The main limitation of the research is the still limited applicability of the approach due to the low digitalization of the industry as a whole. DMS solutions employed for communication between stakeholders are used between domains in intra-firm workflows, with high level of trust, meaning that trust between the parties is higher than in inter-firm workflows. Using BC could increase the level of confidence in this area. However, bringing two technologies together might need a paradigm shift that could be difficult to realize, requiring a change of mindset of professionals [3]. Exploring novel business models attached to the proposal is necessary to document benefits for early adopters, thus motivating its implementation.

The following steps in the research involve primarily linking the demonstrated system directly with design tools. More scenarios will be integrated and tested, and the system will be adapted accordingly. The system needs to be verified; it is planned to provide end users with the extensions of existing tools, extract the generated knowledge, and demonstrate its use for increasing planning liability with BC. Aspects of securely managing digital objects could also be covered with non-fungible tokens (NFT) and BC, but it requires further investigation. Object-based use of DMS and process mapping in a digital form opens numerous opportunities for further improvement of the design process to deliver better performing buildings in the future. This research represents an essential initial step in the direction of object-based DMS and BC integrated solutions.

7 Acknowledgements

This research is part of the research project BIMd.sign - *BIM digitally signed with blockchain and smart contracts*, realized in cooperation with the project FMChain - *Automated payment and contract management in construction with Blockchain technology and BIM 5D*. Both projects are funded by the Austrian Research Promotion Agency (FFG) - Program ICT of the Future and the Austrian Federal Ministry for Transport, Innovation and Technology (BMVIT), FFG-Grants 873842 and 873827.

References

- [1] ISO (International Organization for Standardization). *ISO 19650-1:2018: Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) — Information management using building information modelling — Part 1: Concepts and principles*. ISO, Geneva, Switzerland, 2018.
- [2] Stransky M. Functions of common data environment supporting procurement of subcontractors. In Malinovska L. and Osadcuks V. (eds.) *Proceedings of 19th International Scientific Conference Engineering for Rural Development*. Jelgava, Latvia, 2020. <http://www.tf.llu.lv/conference/proceedings2020/Papers/TF186.pdf>
- [3] Figueiredo K., Hammad A.W.A., Haddad A. and Tam V.W.Y. Assessing the usability of blockchain for sustainability: Extending key themes to the construction industry. *Journal of Cleaner Production*, 343, 2022. <https://doi.org/10.1016/j.jclepro.2022.131047>.
- [4] Tao X., Liu Y., Wong P.K.-Y., Chen K., Das M. and Cheng J.C.P. Confidentiality-minded framework for blockchain-based BIM design collaboration. *Automation in Construction*, 136, 2022. <https://doi.org/10.1016/j.autcon.2022.104172>.
- [5] Borrmann A., König M., Koch C. and Beetz, J. Building information modeling: Why? what? how?. In *Building information modeling*, pages 1-24. Springer, Cham, Switzerland, 2018.
- [6] Adamu Z.A., Emmitt S. and Soetanto R. Social BIM: Co-creation with shared situational awareness. *Journal of Information Technology in Construction (ITcon)*, 20:230-252, 2015. <http://www.itcon.org/2015/16>.
- [7] Oraee M., Hosseini M.R., Edwards D. and Papadonikolaki E. Collaboration in BIM-based construction networks: a qualitative model of influential factors. *Engineering, Construction and Architectural Management*, ahead-of-print, 2021. <https://doi.org/10.1108/ECAM-10-2020-0865>.
- [8] RIBA - Royal Institute of British Architects. *Plan of Work 2020 - Overview*. RIBA Publishing, London, United Kingdom, 2020.
- [9] HOAI (Honorarordnung für Architekten und Ingenieure). HOAI 2013 Volltext. Online: https://www.hoai.de/online/HOAI_2013/HOAI_2013.php, Accessed: 20/09/2021.
- [10] Dubois A. and Gadde L.-E. The construction industry as a loosely coupled system: implications for productivity and innovation. *Construction Management and Economics*, 20(7):621-631, 2002. <https://doi.org/10.1080/01446190210163543>.
- [11] Papadonikolaki E. Loosely Coupled Systems of Innovation: Aligning BIM Adoption with Implementation in Dutch Construction, *Journal of Management in Engineering*, 34(6), 2018. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000644](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000644)
- [12] Erri Pradeep A. S., Yiu T. W., Zou Y. and Amor R. Blockchain-aided information exchange records for design liability control

- and improved security, *Automation in Construction*, 126, 2021, <https://doi.org/10.1016/j.autcon.2021.103667>.
- [13] Sibenik G., Sreckovic M. and Radu A. Modular process patterns in the design phase. In *Proceedings of 2021 European Conference on Computing in Construction*, pages 303-311, Rhodes, Greece (Online), 2021. <https://doi.org/10.35490/EC3.2021.157>.
- [14] LM.VM.2014 Lechner H and Heck D (eds) *Leistungsmodelle + Vergütungsmodelle*. Verlag der Technischen Universität Graz, Graz, Austria, 2017.
- [15] Kropp T., Bombeck A. and Lennerts K. An Approach to Data Driven Process Discovery in the Cost Estimation Process of a Construction Company. In Feng C., Linner T. and Brilakis I. (eds.) *Proceedings of 38th International Symposium on Automation and Robotics in Construction (ISARC 2021)*, pages 893-900, Dubai, UAE (Online), 2021.
- [16] Abdirad H. and Pishdad-Bozorgi. P. Trends of Assessing BIM Implementation in Construction Research. In *Proceedings of Computing in Civil and Building Engineering..* Orlando, FL, USA, 2014. <https://doi.org/10.1061/9780784413616.062>.
- [17] Wang S., Ouyang L., Yuan Y., Ni X., Han X. and Wang F. Blockchain-Enabled Smart Contracts: Architecture, Applications, and Future Trends. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(11):2266-2277, 2019. doi:10.1109/TSMC.2019.2895123
- [18] Mason J. and Escott H. Smart contracts in construction: views and perceptions of stakeholders. In *Proceedings of FIG Conference*, Istanbul, Turkey, 2018.
- [19] Breitfuß D. (2021). Smart Contracts für den Planungsprozess [Diploma Thesis, Technische Universität Wien]. reposiTUm. <https://doi.org/10.34726/hss.2021.71465>
- [20] Kouhestani S. and Nik-Bakht M. IFC-based process mining for design authoring, *Automation in Construction*, 112, 2020. <https://doi.org/10.1016/j.autcon.2019.103069>.
- [21] Lin J.-R. and Zhou Y.-C. Semantic classification and hash code accelerated detection of design changes in BIM models. *Automation in Construction*, 115, 2020. <https://doi.org/10.1016/j.autcon.2020.103212>.
- [22] Feistl S., Ferreira B. and Leita A. Collaborative Algorithmic-based Building Information Modelling. In Janssen P., Loh P., Raonic A. and Schnabel M.A. (eds.) *Protocols, Flows and Glitches, Proceedings of the 22nd International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA)*, pages 613-623, Suzhou, China, 2017
- [23] Xue F. and Lu W. A semantic differential transaction approach to minimizing information redundancy for BIM and blockchain integration. *Automation in Construction*, 118, 2020. <https://doi.org/10.1016/j.autcon.2020.103270>.
- [24] Sibenik G. and Kovacic I. Assessment of model-based data exchange between architectural design and structural analysis, *Journal of Building Engineering*, 32, 2020, <https://doi.org/10.1016/j.jobe.2020.101589>.
- [25] Oraskari J. and Törmä S. RDF-based signature algorithms for computing differences of IFC models. *Automation in Construction*, 57:213-221, 2015. <https://doi.org/10.1016/j.autcon.2015.05.008>.
- [26] Di Giuda G.M., Pattini G., Seghezzi E., Schievano M. and Paleari F. The Construction Contract Execution Through the Integration of Blockchain Technology. In: Daniotti B., Gianinetti M., Della Torre S. (eds) *Digital Transformation of the Design, Construction and Management Processes of the Built Environment*, pages 27-36. Springer, Cham, Switzerland, 2020. https://doi.org/10.1007/978-3-030-33570-0_3.
- [27] Sigalov K., Ye X., König M., Hagedorn P., Blum F., Severin B., Hettmer M., Hückinghaus P., Wölkerling J. and Groß D. Automated Payment and Contract Management in the Construction Industry by Integrating Building Information Modeling and Blockchain-Based Smart Contracts. *Applied Sciences*, 11(16):7653, 2021. <https://doi.org/10.3390/app11167653>
- [28] OASIS Open. Baseline protocol. On-line: <https://docs.baseline-protocol.org/>, Accessed: 27/01/2022.
- [29] Du J., Zou Z., Shi Y., and Zhao, D. Zero latency: Real-time synchronization of BIM data in virtual reality for collaborative decision-making. *Automation in Construction*, 85, 51-64, 2018. <https://doi.org/10.1016/j.autcon.2017.10.009>
- [30] Alizadehsalehi S., Hadavi A., and Huang J.C. From BIM to extended reality in AEC industry. *Automation in Construction*, 116, 2020. <https://doi.org/10.1016/j.autcon.2020.103254>
- [31] Pruskova K. and Kaiser J. Implementation of BIM Technology into the Design Process Using the Scheme of BIM Execution Plan. IOP Conference Series: Materials Science and Engineering, 47(2), <https://doi.org/10.1088/1757-899X/471/2/022019>
- [32] Kerosuo H. BIM-based Collaboration Across Organizational and Disciplinary Boundaries Through Knotworking. *Procedia Economics and Finance*, 21: 201-208, 2015. [https://doi.org/10.1016/S2212-5671\(15\)00168-9](https://doi.org/10.1016/S2212-5671(15)00168-9)
- [33] Tribelsky E. and Sacks R. Measuring information flow in the detailed design of construction projects. *Research in Engineering Design*, 21(3):189-206, 2010. <https://doi.org/10.1007/s00163-009-0084-3>
- [34] Goodhue D.L. and Thompson R.L. Task-Technology Fit and Individual Performance, *MIS Quarterly*, 19(2):213-236, 1995. <https://doi.org/10.2307/249689>.
- [35] Kazman R., Abowd G., Bass L. and Clements P. Scenario-based analysis of software architecture. *IEEE software*, 13(6):47-55, 1996.
- [36] Sreckovic M., Sibenik G., Breitfuß D., Preindl T., and Kastner W. Analysis of design phase processes with BIM for blockchain implementation. In Semenov V. and Scherer R.J. (eds.) *ECPPM 2021 – eWork and eBusiness in Architecture, Engineering and Construction: Proceedings of the 13th European Conference on Product & Process Modelling (1st ed.)*. CRC Press, London, UK, 2021. <https://doi.org/10.1201/9781003191476>
- [37] Turk Z. and Klinc R A social-product-process framework for construction, *Building Research and Information*, 48(7):747-762, 2020. <https://doi.org/10.1080/09613218.2019.1691487>.
- [38] Senescu R.R., Haymaker J.R., Meža S. and Fischer M.A. Design Process Communication Methodology: Improving the Effectiveness and Efficiency of Collaboration, Sharing, and Understanding. *Journal of Architectural Engineering*, 20(1), 2014. [https://doi.org/10.1061/\(ASCE\)AE.1943-5568.0000122](https://doi.org/10.1061/(ASCE)AE.1943-5568.0000122)
- [39] Breitfuß D., Sibenik G., and Sreckovic M. Digital Traceability for Planning Processes. In Semenov V. and Scherer R.J. (eds.) *ECPPM 2021 – eWork and eBusiness in Architecture, Engineering and Construction: Proceedings of the 13th European Conference on Product & Process Modelling (1st ed.)*. CRC Press, London, UK, 2021. <https://doi.org/10.1201/9781003191476>
- [40] Barsby O. Planet Crypto: Ethereum 2.0 Release Date: When does Eth2 launch? 2022. Online: <https://www.gfityesports.com/cryptocurrency/ethereum-2-release-date-eth2-roadmap-phases-is-ethereum-2-new-coin-serenity/>, Accessed: 20/09/2021.
- [41] Gu N. and London K. Understanding and facilitating BIM adoption in the AEC industry. *Automation in Construction*, 19(8):988-999, 2010. <https://doi.org/10.1016/j.autcon.2010.09.002>

Digital Transformation in Asset Management – A Case of BIM Adoption in New Zealand Local government

J.N. Jiang, T. F.P. Henning, and Y. Zou

Department of Civil and Environmental Engineering, University of Auckland, New Zealand
E-mail: jjia338@aucklanduni.ac.nz

Abstract –

Digital technologies are taking place in all spheres of the society and affecting every business sector including the construction industry. As a key player of the industry, local government in New Zealand invests in numerous digitalised Asset Management Information Systems (AMIS) in order to efficiently and effectively manage its infrastructure assets that are critical to the local economy and social wellbeing. However, the silo nature of the organisational operations and proprietary constraints of AMIS have created challenges to the system's interoperability and asset information sharing. Recognised as a key driver behind digital transformation in the industry, BIM has the potential to provide a holistic solution to system integration and fundamentally change the current asset management practices. Based on a tree-pillared framework observed in the existing studies, where the three key aspects, i.e., business models, operational processes, and end-user experience are used as a development roadmap, this study creates a BIM-based digital platform, namely the Asset Management Common Data Environment (AMCDE) model that can integrate multiple systems to improve AM process interoperability and data sharing. Using a case study of a BIM implementation project carried out by a local government in New Zealand, this study demonstrates the development and validation of the AMCDE model. The novel BIM use presented in the study contributes empirical evidence in addressing the sector's siloed manner of business operations, as well as sets the foundation for future studies in local government's digital transformation such as developing the true asset Digital Twins.

Keywords – Asset Management Information System; BIM; Digital Transformation; Common Data Environment; Local Government

1 Introduction

The advent of digital technology has dramatically changed routines and practices in most areas of human

activities. Digital technologies bring about “disruption” in the marketplaces where businesses operate [1]. To enable business agility and changing the way people work in order to optimise business performance to stay competitive, companies have undertaken transformation through adopting digital technologies that provide the game-changing opportunities and reduce existential threats to business [2]. Subsequently, digital transformation is generally taking place in all spheres of the society and affecting every business sector [3]. These transformations have had an impact to enable major business improvements by enhancing customer experience, streamlining operations or creating new business models [4, 5].

While technological innovation and disruption is nothing new, it is widely accepted that the increasingly pervasive nature of technology disruptions and the pace of change will create significant opportunities for countries like New Zealand to achieve a productive, sustainable, and inclusive economy [6]. Consequently, as one of the country's largest infrastructure asset owners and the biggest client of the construction industry, New Zealand local government plays a key role in transforming the traditional business models into digitalised forms. Asset managers are increasingly turning to ‘digital’ to help deal with the complexity and multi-disciplinary nature of the infrastructure asset management. In particular, the local government sector invests in numerous computerised Asset Management Information Systems (AMIS) such as Computer Aided Maintenance Management (CAMM), in order to support their infrastructure asset operations and management decision-making. However, the segmental nature of the organisational operations and proprietary constraints of AMIS have resulted in multiple information systems being sourced and operated in different locations across the whole organisation in a silo fashion [7].

In recent years, the emergence of Building Information Modelling (BIM) has made dramatic changes in the design, construction, and operation processes of buildings and other physical assets [8]. BIM has been envisaged uses in each stage of an asset's lifecycle management [9, 10]. The diffusion of BIM

technology has transformed the construction industry towards more digitalised operation environment; therefore, it is not a question of “if”, but rather “when” and “how” to adopt and implement BIM in asset management practices. Despite the urgency that practitioners and researchers present to the industry on taking the digital transformation journey through BIM adoption, however the adoption rate is low where BIM use in asset management practices is absent from both the industry practices and research with few studies showing the “how”. Apart from the various barriers and obstacles such as cultural, organisational, and lack of real cases [11], the missing of a strategic transformation roadmap also contributes to the slow transformation pace.

In past years, various studies have contributed to the theoretical and practical knowledge base in the domain of digital transformation. For example, Fitzgerald, et al. [5], Westerman, et al. [12] argue that digital transformation is more than just a technological shift [4]. Instead, the transformation affects every aspect of an organization, from the business models, operational processes, to the user experience. The transformation building blocks or the “three pillars” as described by [12], are identified as the vital motivational factors for digital transformation [8], thus providing a strategic roadmap for those organisations that plan to take the transformation journey. On the other hand, the existing studies have largely focused on the enterprises or businesses where customer satisfaction, market demand, and business competitiveness are identified as the key drives for technologies adoption, while studies of digital transformation in none profit driven organisations such as the local governments are limited and the research in asset management is even rarer. In addition, although the digital transformation has entailed a wide range form of technologies including, Cloud, Internet-of-Things (IoT), Blockchain (BC), Artificial Intelligence (AI), and Machine Learning (ML), which constitute a bulk of what is being adopted by organisations as part of their innovation effort, yet technologies with specialties such as AMIS and BIM lack for research in the realisation of transformation projects, thus requiring particular attention.

By examining the existing studies and applying the established theories and concepts, this study intends to explore the opportunity for BIM use in asset management and discuss how digitally transform the current management models and process to improve AMIS interoperability and data sharing. Using the three-pillar transformation concept, this study aims to 1) develop a planning model built on a digital platform as the single information source; 2) create a BIM-enabled asset management common data environment (AMCDE) to integrate the existing planning process and AMIS, and 3) to develop an end-user friendly interface within the

AMCDE to support information access, analysis, and decision-making. The development process and research findings are validated through a case study of a BIM implementation project undertaken by a New Zealand local council. In addition to providing a much-needed real project case, which is limited and incomplete in the current BIM/AM studies, this research contributes to the industry’s digital transformation in particular the adoption of BIM by integrating multiple systems that are commonly implemented independently and operated in a silo fashion, thus creating a new business model by digitalising and integrating the existing asset management processes and multiple AMIS deployed across the organisation. The AMCDE model is fully scalable to further include a wider range of digital technologies, such as CAMM, BMS, and IoT applications, that can be used to develop a true asset Digital Twin operation environment, thus helping the organisation achieve the full scope of digital transformation.

2 Background

Local governments and other public entities worldwide play an important role in taking stewardship and custodianship responsibilities of public infrastructure assets that are fundamental to the local economy [7]. Their responsibilities have also been reinforced further to promote the local communities’ well-being beyond the economic aspect by satisfying social, environmental and cultural requirements. Collectively, local governments in New Zealand owned \$123.4 billion worth of the built assets and had an annual operating expenditure of \$10.3 billion in 2018 alone [13]. Among those assets, the public building and community facility assets accounted for 9% of the total capital expenditure. The public buildings and community facilities, defined as the “social infrastructure”, form part of the ecosystem of the overall construction industry, which constituted 13% of the global GDP [14], is ranked the largest industry in the world. Consequently, adopting a systematic and structured approach helps asset owners manage these public assets more efficiently and effectively.

In the past 20 years, the concept and best practices of asset management have been widely adopted in the local government sector in New Zealand and other developed countries. In particular, long-term lifecycle planning has been recognised as one the key asset management principles to support a resilient and sustainable built environment. Furthermore, with the long-term uncertainties, asset lifecycle planning becomes increasingly complex and sophisticated. Thus it requires the right asset information and tools to enable asset managers to make the right decisions, at the right times, and for the right reasons [15]. Subsequently, local

governments are obligated to adopt and implement digital technologies such as a AMIS to support decision making [7].

2.1 Asset Management Information Systems (AMIS)

An AMIS is a computer-based system which is designed to assist the user for the asset management function [16]. It is a combination of process, data, software, and hardware applied to provide the essential outputs for effective asset management [15]. Most asset intensive organisations such as the local government that own or operate buildings or facilities in long term have a significant existing portfolio, thus it requires some forms of AMIS to manage the FM/AM information. The purpose of implementing an AMIS is to enable asset managers to “plan, monitor, and control” all maintenance that impact on staff and financial resources and the lifecycle performance of the fixed assets. These AMIS serve asset owners with different functions such as asset registers, Computer Aided Maintenance Management (CMM), Geographical Information Systems (GIS) and asset lifecycle planning [17]. The AMIS are also specialised in different types of assets, such as roading, water supplies, and public buildings. Given the speciality of AMIS, asset owner is relying heavily on different and incompatible systems to various management activities, ranging from operations, maintenance and repairs, space management, to asset valuations and long-term strategic planning etc, [18]. In addition, the fast-growing asset data in both the quantity and variety terms have resulted in multiple information systems being sourced and operated in different locations across the whole organisation in a silo fashion [7].

The diversity in software tools, which are proprietary in nature, has not only created the interoperability and collaboration issues, but also presented challenges to the end-users that have to face multiple systems with various level of technical knowledge and skills required. Although using a consolidated enterprise management system based on the existing business model to integrate the financial, customer service, and asset management functions into a single interface have been adopted by some, the integrated system is not yet widely adopted simply because of the cost and complexity to operate and maintain the system [19]. Therefore, there is a great desire to create a unified business model based on a centralised information platform that can bring disparate information and systems together to enable asset managers and end-users to coherently manage their assets.

2.2 Building Information Modelling (BIM)

In recent years, BIM has become more and more used

in the construction industry. BIM is described as a coordinated set of processes [20], supported by technology that add value through the sharing of structured information for buildings and infrastructure assets. In a more thorough description, Sacks, et al. [21] claim that: “*BIM supports support design through its phases, allowing better analysis and control than manual processes. When completed, these computer-generated models contain precise geometry and data needed to support the construction, fabrication, and procurement activities through which the building is realized.*” As a shared knowledge resource for information, BIM forms a basis for decisions in each stage of the facility’s lifecycle [9], ranging from the phases of design, construction, to operation, and maintenance [10]. Benefited from the enhanced system interoperability and asset lifecycle synchronous collaboration, the adoption of BIM has resulted in various management improvements, from cost and time savings, reduced errors and omissions, reduced rework, to maintained repeat business and enhanced construction productivity [22]. Unlike most of the specialised AMIS, BIM’s capabilities of object 3D modelling, open access to information, data visualization, and multidisciplinary integration [23], enhance the interoperability of different systems and collaboration among the stakeholders, thus streamlining operational processes and management activities at the strategic, tactical and operational levels [24].

Nevertheless, despite BIM’s immense technical advantages and potential benefits, the use of BIM worldwide still falls considerably short of its capabilities, and the rate of BIM adoption is much lower than expected [25], in particular in the operation stage where BIM is still perceived by many as merely a design assistance tool, rather than a lifecycle management system [26]. Overall, the current practices in asset management have not fully embraced BIM technology. Further observations from the authors indicate that an asset planning process integrated with BIM has not yet been developed, nor the related literature been published. While various barriers such as cultural and organisational and lack of real cases reasons for BIM adoption have been extensively studied, however, developing a road map to guide how BIM can be adopted is missing from the practices and BIM studies.

2.3 Digital Transformation and the Three Pillars

Digital technology has become central to how the business operates, thus leading to organizations to effectively re-think and possibly re-invent their business models in order to remain competitive [27]. Defined as a process that aims to improve an entity by triggering significant changes to its properties through the integration of information, computing, communication,

and connectivity technologies [3], the term of Digital Transformation refers to a broad concept affecting a wide knowledge domain concerning politics, business, and social issues [28]. Various studies expose that digital transformation affects every aspect of an organization. For example, [5, 12] summarise that digital technology can expect to improve business performance in at least three areas, namely, better customer experiences and engagement, streamlined operations, and new lines of business or business models. In addition to these three key areas, [4] also emphasise digital capability as the fourth element, of which the capability of Information Technology (IT) enables managers to adapt their business strategy to the digital reality, by integrating new technologies in their business models and operation processes. While studying digital transformation in the public sector, [29] acknowledge a set of five factors, namely, strategy, leadership, workforce skills, digital culture, and user focus that shape the path to digital transformation from an organisational perspective. Digital transformation's success thus depends on operational process and management changes, not only by doing something completely new, but also by taking advantage of digitising the existing operations and processes, i.e. to turn existing products or services into digital variants, and thus offer advantages over tangible product and services [30].

By examining the existing studies, it can conclude that, although the current trend of research is largely focusing on the enterprises or businesses where customer satisfaction, market demand, and business competitiveness are identified as the key drives for technologies adoption, it does however set a roadmap to guide digital transformation not only for businesses that are market and profit-driven, but also for non-profit organisations such as local governments, by emphasising session, followed by the case study that demonstrates and validates the findings of the AMCDE framework. the three key elements described as building blocks or “pillars” of transformation [12], they are: business models, operational processes, and experiences of users that can be either external customers or the employees who make the business operate and have first-hand insights on where processes need to improve [31]. The Three Pillars provide a theoretical framework to support organisations to shape their transformation strategies.

3 Research Methodology

In this paper, we investigate local governments as the public client responsible for managing large public property portfolio and community facilities and examine how they use the “three pillars” as the organisation's digital transformation roadmap for their BIM adoption. This research applies a qualitative method with a mixed

approach to include literature review, workshop, and case study methods.

The research data was collected through author's observations and a project stakeholder workshop during a BIM implementation project carried out in the case study as described in Section 4. The first author of this research had observed the overall implementation process, as well as reviewed the relevant documentation including the case owner's asset management policies, procedures, database, and the organisation's BIM implementation strategies and plans that help authors foster the conceptual framework of AMCDE. The first author also participated in the workshop designed for understanding the owner's current asset management practices, establishing project information requirements, and obtaining feedbacks of the proposed AMCDE during the workshop, where discussions and outcomes were noted. The participants of the workshop included the owner appointed project manager, a BIM manager from the consultancy firm, and the owner's facility management staff as the potential user group.

Using the “three pillars” as the transformation development reference, this research describes Pillar One to represent the AMCDE framework, a new business model in asset management decision making. Then the process of developing the AMCDE shows the integrated BIM and AMIS as Pillar Two. And last, Pillar Three is a machine/user interface created using a novel BIM-GIS interoperability solution that enhances user's experience with 3D visualisation and easy data access portals. Each of the three pillars is described in Figure 1 to 3, followed by the case study that demonstrates and validates the findings of the AMCDE.

3.1 Pillar One – Business Model

Pillar One of the transformations is to develop the business model built on a digital platform as the single information source. Figure 1 illustrates a conceptual model of AMCDE, which enables various building asset management functions to be performed through an integrated platform. The AMCDE determines and describes asset data flows and the relationship between building hierarchy, BIM, AMIS, and asset management decision-making framework throughout the in-use stage of an asset lifecycle planning system. The conceptual framework will enable the geometric models and asset database to be mapped to the CDE, of which the data then are fed to AMIS such as Computer Added Facility (Maintenance) Management (CAFM & CAMM) systems, Building Management Systems (BMS), or other Application Programming Interface (APIs) etc., for analysing, processing, and optimizing to support an

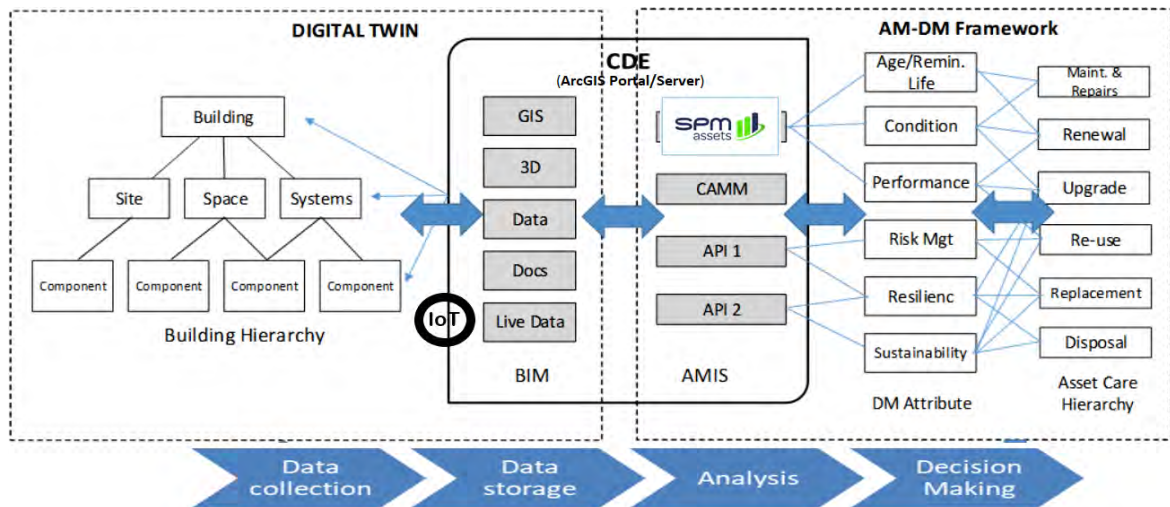


Figure 1. Pillar One. AMCDE and Asset Management Decision-Making Framework

evidence-based decision-making for asset care interventions.

The outputs defined in the proposed decision-making process (AM-DM) provide the various options for management's consideration based on the asset lifecycle interventions with a selection of various construction activities, such as maintenance and repairs, like-for-like renewals, upgrades, replacement, and even the disposals as described in the list of asset care hierarchies.

3.2 Pillar Two – Operational Process

Pillar Two of the transformations involves the creation process of a BIM-enabled AMCDE to integrate the existing planning process and multiple AMIS. The core process is to develop the CDE that consists of five steps, namely, planning, data collection, modelling, CDE and interface development, and project peer review. The

breakdown steps including data flows are illustrated in Figure 2.

The BIM implementation project provides the organisation with the opportunity to create a formalised process that supported the changes induced by the introduction of BIM. In particular a BIM-based asset planning process is created through the following steps: 1) as-built data capture can be conducted by utilising a combination of point cloud scanning, photogrammetry mapping, and site surveying; 2) 3D geometry models is produced using BIM authoring tool such as Autodesk Revit from the data extracted in the capturing process; 3) while the non-geometry data, including building element metadata such as age, cost, condition, capacity, and criticality, etc., are captured in a separate database aligned to the classification defined by standard formats such as those defined in NAMS property manual [17]; 4) a single sourced data system acting as the central data

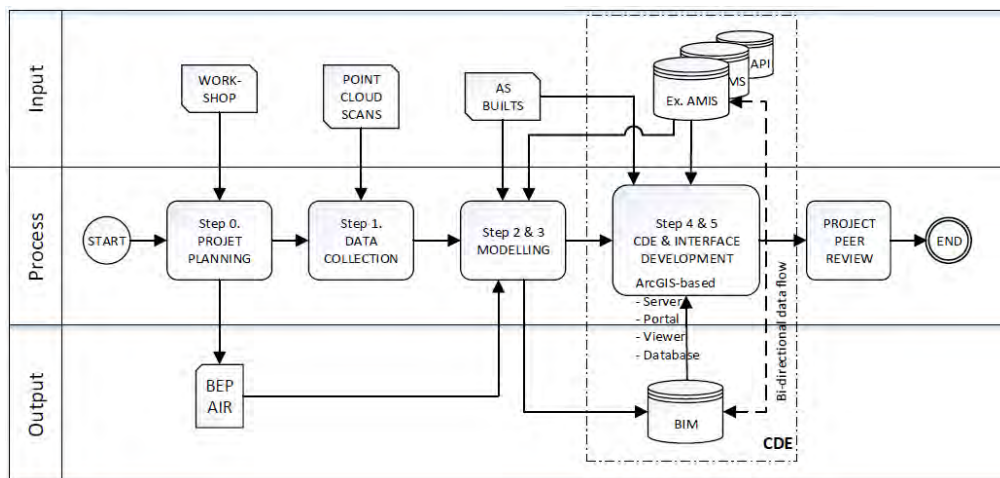


Figure 2. Pillar Two - Integrated AMCDE development process

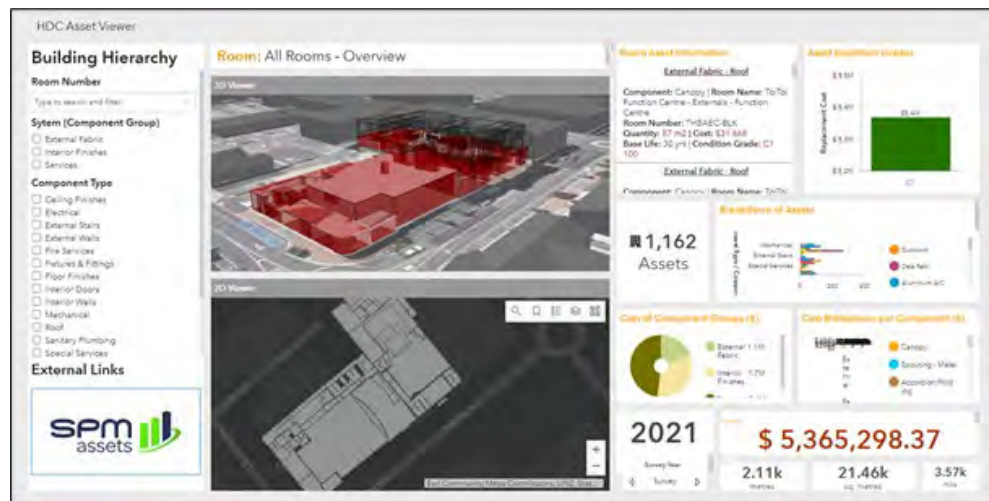


Figure 3. Pillar Three - ArcGIS User Webviewer/Interface

repository is created and hosted in a model and data integrated platform such as GIS-based interface; and 5) asset information dashboards were established to provide management insight into the building data, such as physical condition, space usage and surface areas etc. It is worth noting that while step 4) and 5) can be one-off and unique to this project, step 1) to 3) is repetitive and applicable to the future assets for digitalisation.

3.3 Pillar Three – End-user Experience

Pillar Three of the transformations represent the ultimate goal of this project, which is to develop a user-friendly interface within the AMCDE to support information access, analysis and reporting, as shown in Figure 3. The BIM model created from the previous steps is integrated with the existing AMIS, along with asset data, photographs, GIS data, O&M manuals, and other as-built documentations, using a selected database and portal. In this study, ArcGIS Pro of Esri is adopted as the CDE architecture to store, retrieve, analyse, and visually represent asset data and information. The ArcGIS-BIM

integration enables a standards-based Extract-Load-Translate (ELT) to extract the geometry and data from the BIM models and translate into 3D and location enabled GIS data, via data common denominators. For example, using MSSQL as a back-end repository for data storage and management to enable different asset data formats to be integrated with Esri GIS data. ArcGIS' Rest API allows the integration of various data to remain in-sync across multiple platforms and the specific AMIS.

4 Case Study

The building chosen for the case study is a historical opera house complex (aka Toitoti Arts & Events Centre) built in Hastings, New Zealand. Figure 4 illustrates a sample floor plan and the interior image of the theatre. The Toitoti building was constructed in 1915 and has had numerous renovations and modifications over its life. With a combination of new and old building blocks, the complex has a total floor area of 5,000m² that comprising of theatre, exhibition, office, and amenity facility spaces. The project was driven by the owner's decision to adopt

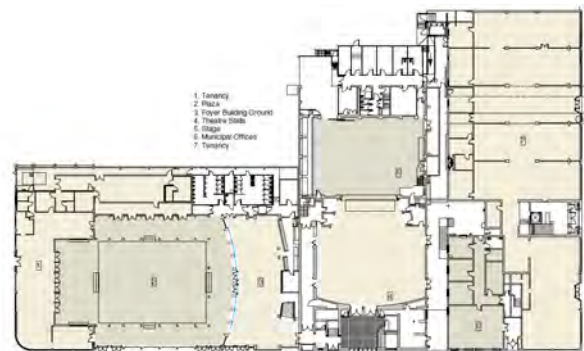


Figure 4. Toitoti Arts & Events Centre (Source: Hastings District Council)

a BIM-based planning platform to support the integration of the existing AMIS. The scope of BIM implementation in this project includes the process of data collection, modelling and the creation of an easy-to-use interface that can bring asset information, including static data such as drawings, documents, asset data, as well as geospatial information on a single digital platform AMCDE, as shown in the above Figure 1 to Figure 3.

5 Discussion and Conclusions

Digital transformation can effectively enhance the performance of organisational management and business operations in various aspects such as better customer experiences and engagement, streamlined operations, and new lines of business or business models. Digital technologies such as BIM as illustrated in this study help asset owners improve their management practices in numerous perspectives, ranging from automating data collection, enhancing analysis, to optimising asset performance. The study demonstrates an efficient method of using BIM to improve system interoperability and information sharing by populating multiple AMIS, where asset data can be stored and exchanged on a centralised digital platform, i.e., the AMCDE model, to provide the insight of asset management from a single source of information normally stored in multiple systems and locations. In addition, the development of a user-friendly interface supported by the AMCDE, of which the enhanced end-user experience helps remove obstacles in accessing asset data across the whole organisation and shared by different stakeholders in a remote manner, is particularly useful in the situation such as Covid-19 pandemic, where working remotely has become the new norm. The case study demonstrates that digital technologies not only enhance the efficiency of the operational process for data collection, but also improve data accuracy and reduces human errors that are commonly occurring in a manual process. Overall, local governments in New Zealand will benefit from adopting digital technologies including BIM to improve their asset management efficiency through time and cost savings, a more collaborative operation environment, and visualised information analysis and reporting.

It is noteworthy to mention that, the existing studies have largely focused on the enterprises or businesses where customer satisfaction, market demand, and business competitiveness are identified as the key drives for technologies adoption. Using these drives to guide the organisations such as the local government, which are non-profit driven business and whose operation objective is more towards improving management efficiency, is up to debate. Furthermore, unlike the private sector, the local government has its own unique operation processes, management models, and digital technologies (e.g.,

AMIS), thus using the existing transformation theories as the roadmap without adaption presents challenges due to the lack of studies, theories, and empirical evidence.

While BIM research and practice is currently intensifying in project-oriented design and construction activities, there is still an untapped potential to view BIM as a catalyst to develop new business models from a service innovation perspective. Hence, this study focuses on the BIM use in asset management planning in a built environment. In addition to contributing a much-needed real project case, this study takes a holistic transformation approach to the challenge of integrating multiple systems that are commonly implemented independently and operated in a silo fashion. The novel business model based on AMCDE platform that is developed for integrating the existing asset management process and the multiple AMIS deployed across the organisation, has the potential to further include a wider range digital technology, such as CAMM, BMS, and IoT applications, in order to form true asset Digital Twins and further advance organisations towards a full digital transformation.

Nonetheless, this research is limited by several factors. Firstly, BIM is such a disruptive technology as previously mentioned, its adoption and implementation requires changes in workflow, practices, and most of all, human habits and ways of working that form the key pieces of an organisation's culture, which requires further research. Secondly, asset management is a complex process. To digitally transform the management practice requires a balance of people, processes, and technology, where the three-pillar theory only address the latter two elements. The people factor deserves a deep investigation and certainly warrants a fresh new chapter and requires further research. And finally, the research uses a single case study to provide a way of concept validation, which has its limitation. Future studies using multiple real cases, reviews, and other validation methods need to be explored and examined.

References

- [1] Mithas S., Tafti A., and Mitchell W., How a firm's competitive environment and digital strategic posture influence digital business strategy, *MIS quarterly*, pp. 511-536, 2013.
- [2] Sebastian I.M., Ross J.W., Beath C., Mocker M., Moloney K.G., and Fonstad N.O., How big old companies navigate digital transformation, in *Strategic information management*: Routledge, 2020, pp. 133-150.
- [3] Vial G., Understanding digital transformation: A review and a research agenda, *Managing Digital Transformation*, pp. 13-66, 2021.
- [4] Henriette E., Feki M., and Boughzala I., The shape

- of digital transformation: a systematic literature review, *MCIS 2015 proceedings*, vol. 10, pp. 431-443, 2015.
- [5] Fitzgerald M., Kruschwitz N., Bonnet D., and Welch M., Embracing digital technology: A new strategic imperative, *MIT sloan management review*, vol. 55, no. 2, p. 1, 2014.
- [6] New Zealand Productivity Commission, Local government funding and financing: Issues Paper., ISBN: 978-1-98-851926-5, 2018.
- [7] Office of the Auditor-General, Reflections from our audits: Investment and asset management, 978-0-478-44271-7, 2017.
- [8] Solas A.R.a.M.Z., Shaping the Future of Construction A Breakthrough in Mindset and Technology, Cologne/Geneva Switzerland, 2016.
- [9] Olawumi T.O., Chan D.W.M., and Wong J.K.W., Evolution in the intellectual structure of BIM research: a bibliometric analysis, *Journal of Civil Engineering and Management*, vol. 23, no. 8, pp. 1060-1081, 2017/11/17 2017.
- [10] Li X., Xu J., and Zhang Q., Research on Construction Schedule Management Based on BIM Technology, *Procedia Engineering*, vol. 174, pp. 657-667, 2017/01/01/ 2017.
- [11] Becerik-Gerber B., Jazizadeh F., Li N., and Calis G., Application Areas and Data Requirements for BIM-Enabled Facilities Management, *Journal of Construction Engineering and Management*, vol. 138, no. 3, pp. 431-442, 2012.
- [12] Westerman G., Bonnet D., and McAfee A., The nine elements of digital transformation, *MIT Sloan Management Review*, vol. 55, no. 3, pp. 1-6, 2014.
- [13] Department of Internal Affairs, Local authorities' economic contribution to New Zealand, 2021.
- [14] Manyika J. *et al.*, A future that works: automation, employment, and productivity, McKinsey Global Institute, McKinsey Global Institute, 2017.
- [15] Institute of Public Works Engineering Australasia, *International infrastructure management manual*. IPWEA Sydney, NSW & Wellington, N.Z. NAMS Group, 2015.
- [16] Hastings N.A.J., *Asset Management Information Systems - Physical Asset Management*, Second edition ed. Berlin: Springer, 2015.
- [17] National Asset Management Steering Group (NAMS), NAMS Property Manual for the Strategic to Tactical Long-term Planning for Property Assets. V2.0, 2014.
- [18] Cheng J.C., Chen W., Tan Y., and Wang M., A BIM-based decision support system framework for predictive maintenance management of building facilities, in *The 16th International Conference on Computing in Civil and Building Engineering (ICCCBE2016)*, Osaka, Japan, 2016.
- [19] Rashid M.A., Hossain L., and Patrick J.D., The evolution of ERP systems: A historical perspective, in *Enterprise resource planning: Solutions and management*: IGI global, 2002, pp. 35-50.
- [20] BIM Acceleration Committee, *The New Zealand BIM Handbook*, 3rd ed., 2019.
- [21] Sacks R., Eastman C., Lee G., and Teicholz P., *BIM handbook: A guide to building information modeling for owners, designers, engineers, contractors, and facility managers*. John Wiley & Sons, 2018.
- [22] Zhao X., A scientometric review of global BIM research: Analysis and visualization, *Automation in Construction*, vol. 80, pp. 37-47, 2017/08/01/ 2017.
- [23] Whitlock K. and Abanda H., Making a Business Case for BIM Adoption, 2020, pp. 231-246.
- [24] Talamo C. and Atta N., FM Services Procurement and Management: Scenarios of Innovation. Cham: Cham: Springer International Publishing, 2018, pp. 201-242.
- [25] Oraee M., Hosseini M.R., Edwards D.J., Li H., Papadonikolaki E., and Cao D., Collaboration barriers in BIM-based construction networks: A conceptual model, *International Journal of Project Management*, vol. 37, no. 6, pp. 839-854, 2019/08/01/ 2019.
- [26] Flores M. *et al.*, The Construction Value Chain in a BIM Environment, in *IFIP International Conference on Advances in Production Management Systems (APMS)*, Seoul, South Korea, 2018-08-26 2018, vol. AICT-536, no. Part II: Springer International Publishing, pp. 255-262.
- [27] Carcary M., Doherty E., and Conway G., A dynamic capability approach to digital transformation: a focus on key foundational themes, in *The European Conference on Information Systems Management*, 2016: Academic Conferences International Limited, p. 20.
- [28] Reis J., Amorim M., Melão N., and Matos P., Digital transformation: a literature review and guidelines for future research, in *World conference on information systems and technologies*, 2018: Springer, pp. 411-421.
- [29] Kokkinakos P., Markaki O., Koussouris S., and Psarras J., Digital Transformation: Is Public Sector Following the Enterprise 2.0 Paradigm?, in *Digital Transformation and Global Society*, Cham, A. V. Chugunov, R. Bolgov, Y. Kabanov, G. Kampis, and M. Wimmer, Eds., 2016// 2016: Springer International Publishing, pp. 96-105.
- [30] Gassmann O., Frankenberger K., and Csik M., *The St. Gallen business model navigator*, 2013.
- [31] Bonnet D. and Westerman G., The New Elements of Digital Transformation, *MIT Sloan Management Review*, vol. 62, no. 2, pp. 82-89, 2021 2021.

Project-based curriculum for teaching construction robotics

C. Brosque and M. Fischer

Civil and Environmental Engineering, Stanford University, USA

cbrosque@stanford.edu, fischer@stanford.edu

Abstract -

As innovations in construction robotics are being tested and deployed on site thanks to technological advancements in computing and sensing, Civil Engineering researchers must consider how to connect real-world innovations with research and teaching at a much faster pace. Observing the practice helps identify problems and test engineering solutions and models through research. As the research develops, engineering courses can foster innovation adoption in the industry. This cycle leads to a new practice and the recognition of new problems that feed the research and teaching. This paper focuses on teaching construction robotics through a project-based curriculum as an essential mechanism to enhance this ecosystem from research to practice. The project involves the collaboration of construction companies, robot companies, and students to analyze the potential Safety, Quality, Schedule, and Cost impacts of at least ten construction robots. The anticipated benefits for the students are engaging in real engineering problem solving and synthesizing academic and industry experience. At the same time, the collaboration between the students and the industry helps validate the research generality and contributions. This paper does not claim that this represents the only or best way to teach this topic but aims to open the subject for discussion.

Keywords -

Construction Robotics; Education; Automation; Construction Management

1 Introduction

Civil Engineering and construction practice still involve many aspects of a craft. Hence, real-world practice can help inform and test research. The Center for Integrated Facility Engineering (CIFE) at Stanford University focuses its research efforts on the built environment with a careful observation of practice to identify problems. These problems generate a solution intuition backed by a theoretical point of departure. Following research questions and methods inform the research tasks to solve the industry problem. The validation of the research results leads to contributions and practical impacts in the industry, like the adoption of cutting-edge technologies (Fig. 1). Examples are found in many CIFE Ph.D. projects and thesis [1, 2, 3, 4] and the integration of courses into the CEE curriculum at the graduate level. Examples are Building Information Modeling, Parametric Design and Optimization, Virtual Design and Construction (VDC), Industrialized Construction, Managing Fabrication and Construction, Project Assessment and Budgeting, and Computer-Integrated AEC.

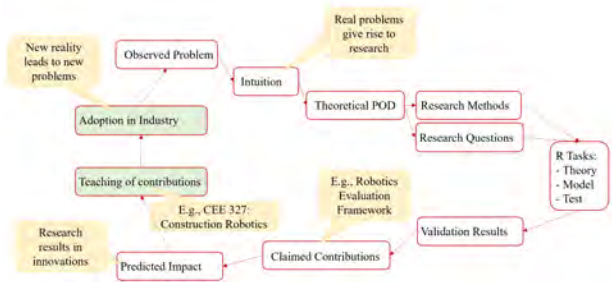


Figure 1: Connection of real-world innovations with research and teaching cycle based on CIFE's academic research method. Focus on the first project-based curriculum to teach Construction Robotics for Civil Engineering graduate students: CEE 327: Construction Robotics.

These courses foment innovation adoption and diffusion in the industry. According to Navon [5], such courses could reduce resistance to new technologies by enriching the knowledge and understating of future industry leaders. Custovic et al. [6] also highlight the role of construction management curricula in researching new methods and technologies and demonstrating their applicability and benefits for the industry. Once the adoption matures, it can lead to a new reality and new problems or observations that feed into the research cycle and the following teaching of the new concepts, methods, or techniques.

1.1 Construction Robotics

The Construction Automation and Robotics field has developed significantly in the last decade thanks to advances in the internet of things, artificial intelligence, sensors, and the use of Building Information Models (BIM). These advances materialize in new construction robots being developed and tested on construction sites. The International Federation of Robotics forecasts 4,200 construction robotics units to be sold from 2019 to 2021 [7], and Bock and Linner [8] outlined 24 categories of on-site task-specific construction robots. Given this new practice, researchers in Civil Engineering must consider how to connect real-world applications with research and teaching.

As robots mature and become suitable construction methods, innovation managers must consistently evaluate

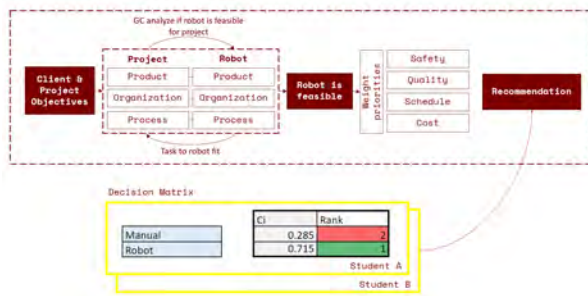


Figure 2: Project-based approach to analyze the Safety, Quality, Schedule, and Cost impacts of ten robots independently by two students.

the impact of deploying robots compared to traditional construction methods. However, construction robotics courses are not traditionally included in Civil Engineering, with only a few courses and programs available worldwide [9, 10, 11, 12, 13].

Based on the importance of teaching the academic foundations about innovation approaches [6] to support the cycle between research and practice, this paper focuses on a project-based construction robotics class for Civil Engineering. The course pairs real construction projects with on-site robots for about ten tasks. Two students are assigned per case to assess each robot's potential Safety, Quality, Schedule, and Cost impacts for the selected project. The 10-week course allows the students to work closely with industry partners from the General Contractor and the robot start-up or manufacturer. The collaboration in the class between the students and the industry helps validate the research generality [14] with real-world test cases. For the students, it directly connects the theory and practice. And finally, for the industry partners, the class presents an objective and repeatable method to approach robot evaluations, together with the direct collaboration with a student to resolve an industry problem.

In addition to the course project, we include five modules that cover 1) an introduction to construction robotics and a Robotics Evaluation Framework (Fig. 2), 2) robot examples, 3) the sustainability perspective, 4) human-robot collaboration, and 5) robots in the context of Virtual Design and Construction (VDC).

This paper aims to share the learning objectives and course structure of the proposed curriculum to open the subject for discussion with researchers and lecturers in the field. We present the course feedback from the first implementation in Winter 2021.

2 Related Work

A few universities have identified the need for including construction robotics education in Civil Engineering and Architecture courses. These courses introduce automation strategies for construction processes and their implications in the design outcome. A common thread of the existing studies is the interdisciplinary Architecture, Engineering, and Construction (AEC) approach that combines expertise from various sources.

The University of Maryland, A.J. Clark School of Engineering, offers a *Construction Automation and Robotics course*. The course aims to redesign traditional construction processes to utilize state-of-the-art automation and robotics technology [15]. Prof. Skibniewski's 12-week course covers 1) the history of construction automation and robots; 2) construction projects including ergonomics and the physical and cognitive requirements of construction labor; 3) an introduction to industrial robotics; 4) robot components; 5) feasibility of robot applications; 6) calculation of costs and benefits in robot assessment; 7) integration issues; 8) additive manufacturing; 9) safety and ethical issues.

The University of Sydney's School of Architecture, Design, and Planning provides a *Robotics in Architecture and Construction* masterclass introducing robotics for the fields of architecture and construction. The goal of the class is to understand what a robot is and how to consider robot methods in the design workflow [13].

ETH Zurich has also developed several courses focusing on digital and robotic fabrication in architecture, including a one-year *Master of Advanced Studies in Architecture and Digital Fabrication* to teach the fundamentals of technologies and methods of digital design and fabrication for architecture and construction [9].

The Institute for Advanced Architecture of Catalonia developed a *Master's in Robotics and Advanced Construction (MRAC)*. The program seeks to train professionals on the emerging design and market opportunities of deploying new robotic and advanced manufacturing systems. The curriculum involves seminars and studio projects investigating how robots and automation will change the existing building methods. A goal of the program is to develop processes and design tools that address these new methods from the engineer, designer, architect, workforce, and academic perspectives [12].

The Technical University of Munich, Department of Architecture, includes a newly established *Master's study in Advanced Construction and Building Technology* within the Chair of Building Realization and Robotics. The mission of this program is to design and build a future robotic society. A cross-disciplinary approach focuses on finding and creating technologies for robotic construction by target-value design [11].

Lastly, the Faculty of Architecture at RWTH Aachen University designed a Master's program in *Construction Robotics* to shape students in automated construction machinery and robotics. The program combines prototyping of machinery and processes with virtual design and simulation [16].

These courses and graduate programs evidence the importance and growing demand to teach a new generation of AEC students about automation and construction robotics. Most of the existing effort is led by Architecture schools, focusing on new design and fabrication methodologies enabled by robots. Civil Engineering programs should similarly address the potential on-site robot uses, especially from a construction managerial perspective, as outlined by Skibniewski [15].

3 Proposed Curriculum

3.1 Context

Construction Robotics (CEE 327) is part of Stanford's Civil and Environmental Engineering (CEE) graduate studies. The target students for the class include the Sustainable Design and Construction (SDC) students within CEE and graduate researchers at the Center for Integrated Facility Engineering (CIFE). CIFE aims to improve the built environment's planning, design, construction, and operation across all sectors and scales, including buildings, industrial plants, urban districts, manufacturing facilities, infrastructure, and cities. The CEE-SDC curriculum includes courses on engineering and management methods that improve the built environment's reliability, productivity, innovation, and sustainability.

The new Construction Robotics course adds to the existing engineering and construction management curriculum, focusing on evaluating on-site construction robots. An in-depth analysis of off-site automation examples is outside the class's scope, as an Industrialized Construction class and annual industry forum covers this topic.

Any CEE graduate student can enroll in the class. It is assumed that students are familiar with BIM, project specifications, and basic scheduling principles. However, students are not expected to have any prior programming or robot design experience, as the course focuses on the management of robot technologies.

3.2 Lecture Plan and Objectives

The first goal of this course is to introduce on-site construction robotic applications through academic and industry practice.

The second goal is to assess the applicability and potential impact of promising construction robots available in the market against traditional construction methods using a consistent, repeatable evaluation method.

Third, to gain a good understanding of basic robot principles and become familiar with state-of-the-art research in the field, including human-robot collaboration approaches.

Finally, to connect robot applications to the broader project and client objectives deploying the Equity, Environment, Economy (EEE) framework [17] and the Virtual Design and Construction (VDC) methodology [18] to manage projects.

The developed curriculum includes two 90-minute lectures per week over ten weeks. Table 1 provides an overview of the lectures' content divided into five modules: 1) Introduction to evaluating construction robots, 2) Robot examples, 3) The sustainable perspective, 4) Human-robot collaboration, and 5) Robotics in the context of VDC.

Table 1: CEE 327 Construction Robotics Curriculum

#	Module 1: Intro to Construction Robots	Assignment
1	Intro to Construction Robots	
2	Robots 101	A1. Product
	Introduce class robots	
3	POP analysis	
4	SQSC and Decision Matrix	A2. Org and Process
	Lessons learned from drilling robot	
5	Meet industry partners	A3. Safety, Quality, Schedule
	Off-site vs. on-site robots	
6	Industrialized Construction	A4. Cost and Decision Matrix
	Closing framework reflections	
	Layout robot guest lecture	A5. Off-site vs. on-site robots
	Module 2: Robot Examples	Assignment
7	Single-task robots intro	A6. REF Conclusions
	ETH Zurich's novel robotic processes	
	Hilti's Jaibot	
8	Canvas and Swinerton case	A7. Robots as a Service
9	Obayashi's robot examples	
10	Robot examples: TyBot, Kewazo	A8. Product Draft (Project)
11	Robot examples: Boston Dynamics, SafeAI	A9. Prep for guest lectures
12	Robot example: Civ Robotics	A10. Org and Process draft (Project)
13	TUM Guest Lecture: Thomas Bock	A11. Prep for guest lecture
14	Robot examples: Shimizu, Exoskeletons	A12. SQSC draft (project)
	Module 3: The sustainable perspective	Assignment
15	Overview of EEE	A13. Prep for guest lecture
	Equity and Economic perspectives	
	Silicon Valley Robotics	
16	Ecology perspective	A14. Individual REF template
	Demolition and urban mining	
	Economics, productivity, and employment	
	Module 4: Human-Robot Collaboration	Assignment
17	HRC Stanford Robotics Lab (haptics)	A15. Elevator pitch
18	Autodesk's perspective	A16. HRC Ocean One
	Module 5: Robotics in the context of VDC	Assignment
19	Robotics in the context of VDC	
	Expert panel discussion	A17. Opportunities and challenges
20	Project presentations	A18. Final report and slides

Module 1: Introduction to Evaluating Construction Robots

This module introduces the course objectives and key concepts to analyze robots in construction. First, we cover an overview of construction robotics history from a literature review of the International Association of Automation and Robotics in Construction (IAARC), Automation in Construction, the Cambridge Handbooks of Construction Robotics, and the American Society of Civil Engineers (ASCE), specifically the Journals of Computing in Civil Engineering and the Journal of Construction Engineering and Management. Second, we introduce basic robot definitions and principles, bearing in mind that the class audience is not likely to have any prior robotics background.

A Robotics Evaluation Framework (REF) based on case studies and the literature review is presented as a basis for

the class project [14]. We introduce Assignments 1 to 4 as a mock-up of the class project with data provided from a previous layout case study [19]. Each of the four first assignments focuses on a different aspect of the robot evaluation. These assignments aim to deploy a consistent robotic evaluation method, develop process modeling skills, explore schedule representations, and study cost and benefit scenarios to purchase or use a robot as a service. This example allows students to exchange answers and assess their evaluation against the teaching team solution.

Finally, this module introduces the main differences between applying robots on site and off site.

Module 2: Robot Examples

The second module covers robot application examples. Guest lectures with industry leaders and start-up founders showcase the applications included in the projects and other relevant examples. We also engaged robot examples developed in academia, such as the construction robotics research at ETH Zurich and TUM.

Module 3: The Sustainable Perspective

This module introduces the Economic, Ecologic, and Equity perspectives through the triple "E" Sustainability framework. Topics addressed in this module include the role of robots in the circular economy, robotic deconstruction, statistical analysis of jobs in industry 4.0, and building new skills to grant access to construction robots. The guest lectures include different aspects of robots' sustainability to optimize revenue growth, public good, and health [20].

Module 4: Human-Robot Collaboration

This module addresses state-of-the-art approaches in human-robot collaboration from the academia and industry perspective. Topics include wearable technologies, haptics, virtual reality, and augmented reality. Robot examples include underwater humanoid bi-manual exploration robot: Ocean One [21] and haptic bolting, welding, and joint sealing [22].

Module 5: Robotics in the Context of VDC

The final module discusses the opportunities and challenges of robots within the broader context of Virtual Design and Construction (VDC) (Fig. 3). VDC is the management of integrated multi-disciplinary performance models of design-construction projects, including the product, work processes, and organization of the design, construction, and operation team to support project and client objectives [18]. Course materials present construction robotics as a controllable factor for construction managers to achieve desired project and client outcomes.

This module includes a final panel discussion with VDC experts that previously deployed construction robots in the field. The module aims to understand synergies between deploying the VDC methodology and construction robots and share lessons learned by GCs using robots in the field.

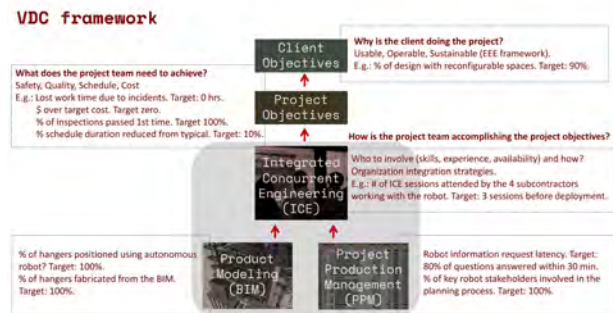


Figure 3: Managing the deployment of construction robots in the context of the VDC methodology.

3.3 Class Project

Finally, the industry project represents a significant component of the course. The teaching team matches pairs of students with a construction industry professional and a robot developer. In our experience, the student-professional collaboration enriches both parties as the students provide an objective lens to a robot evaluation problem. On the other side, the industry professionals on the robotics and the construction site helps connect the class concepts with reality.

Each student receives a blank Robotics Evaluation Framework template accessible online and four case study examples. The pair of students evaluating the same robot must do so independently without looking at their partner's results and tracking the hours taken in each analysis step. In the final two weeks of the class, the students share their finished templates and document differences in results and recommendations. Further, they prepare a joint final presentation, attended by the industry partners.

The robot evaluation method is based on prior work from the researchers [19, 14]. The main evaluation parameters are the analysis of the robot and construction site Product, Organization, and Process. Second, the breakdown of the robot's potential Safety, Quality, Schedule, and Cost impacts to the traditional method for the selected project.

4 Implementation

We offered the class for the first time in the Winter quarter of 2021, including ten Civil Engineering graduate students from Stanford University and nine from the University of Lima working in 12 case studies (Table 2). Due to Covid-19 restrictions, the course was held entirely

online in Discord.

Table 2: Industry Partners Winter quarter 2021

#	GC	Robot
1	DPR	Hilti Jaibot
2	Obayashi	Material handling robot
3	Bechtel	Kewazo scaffolding
4	Megacentro Lima	Exyn autonomous drones
5	Produktiva	SafeAi autonomous machinery
6	NCC	Boston Dynamics Spot
7	Swinerton	Canvas drywall finishes
8	DPR	Canvas drywall finishes
9	HDlab	SuitX exoskeleton
10	MT Højgaard	Civ Robotics layout
11	Implenia	TyBot rebar tying
12	Traylor Brothers	TyBot rebar tying

The robot industry partners included five founders and CEOs, a CTO, one Vice President, a Chief Operating Officer, a Construction Technology Manager, and a Business Unit Manager. The GC partners included one Owner, five Innovation Managers/Directors, one Chief Financial Officer, three heads of VDC, and two project directors.

Even though construction robots on the field are at infancy in their deployment worldwide, the consistent student analysis of these examples showed how promising the technology already is for a range of robot types, mobility, autonomy, scale, business models, and locations.

4.1 Case Study Example: Material Handling Robot

One of the case studies analyzed Obayashi's interior material handling robot developed with Stocklin (Fig. 4). This subsection addresses the main conclusions of the comparative analysis to illustrate the type of insights achieved by the students in the course project. Future work will cover in detail the Safety, Quality, Schedule, and Cost impacts of the ten robots under study.

Material handling is a time-consuming critical task that is also hazardous for labor. The Ministry of Health, Labor, and Welfare in Japan identified 1,256 cases of material-related injuries in 2020, 15% of all injuries that occurred in the year. Labor shortages also motivated the Obayashi Logistics System to reduce the workload and burden of interior material handling to the desired floor, typically part of Obayashi's work scope.

The robot system consists of an Automated Guided Vehicle (AGV) that can carry palletized materials, a custom elevator, and a logistics management system that guides the robot's work. The robot requires a flat and clear floorplate with a maximum deviation of 40 mm. Obayashi developed a small lift for steps up to 350 mm. Following AGV ISO regulations, the robot detected humans with an onboard 2D Lidar scanner. Whenever a human is closer than 500 mm, the AGV reduces the speed to 0.3 m/s and stops if closer than 300 mm.



Figure 4: Interior robotic material handling on site.

The site elevator autonomously interacts with the robot without a human operator and can fit two AGVs. Each AGV can handle pallets up to 1,200 x 1,800 mm.

The logistics management system replaced 68% of manually transmitted material orders with digitally transmitted data, reducing rework from 3% to 1%. The robot failure rate reported was zero, but the 1% rework includes mistakes in communication from the operator regarding the placement of materials.

Obayashi's robot was designed to have the same productivity as the laborers. However, the site measurements showed it took 50% more time to carry the same amount of material with the robot than manually because it took more time to find the materials. Hence, the development team looked at ways to deploy the robot with human labor. A squad of two robots and two crewmembers achieved the same productivity as five workers. The robots worked during the day and night shifts with one operator, while the two laborers worked only during the day. This decision increased 68% of the traditional total daily carrying capacity of 125 tons/day by five workers.

Finally, according to the student analysis, the material handling robot achieved a 41% cost reduction when using two robots simultaneously. The conventional material handling team of five crewmembers costs \$1,325/day. In comparison, the cost for a hybrid team of two crewmembers working during the dayshift and two robots with an operator working both the day and nightshift was \$1,328/day. The robot's autonomous elevator added \$12,000 per project, but the robot reduced coordination needs by the superintendents with \$188/day savings. Overall, the conventional material handling cost was \$10.6/ton, and the hybrid robot method cost was \$6.32/ton. The robot service cost included maintenance and transportation to the project.

In this case, the hybrid use of two robots and two crewmembers outperformed the traditional team of five workers in the four key variables (Safety, Quality, Schedule, and Cost).

4.2 Course Feedback

During the first implementation, the teaching team collected student feedback to improve future offerings of the course.

The course evaluation inquired about: (1) the initial goals to join the class, (2) desire to participate in future robotics research opportunities, and (3) curriculum improvements to advance their goals.

4.2.1 Students' Goals

The 19 students stated achieving their course goals. The students' goals included:

- "To learn about the exciting robots being used in construction, all of which I was not familiar with coming into the class."
- "Learn more about how robotics is applied to construction tasks and the unique challenges that construction poses."
- "Acquire knowledge about many robots and the mindset of evaluating something through a framework."
- "I wanted to learn more about how the construction industry is trending with automation and technology and hear first-hand from the people involved in decision-making and bringing about this goal."
- "My main objective was to learn about the use of robots in the construction industry, and I came away with the necessary knowledge for feasibility analysis, so it was very satisfying to participate in the course."
- "Information on robotics available and their design. Impacts on traditional organizations, contracting methods and work break down structure."

4.2.2 Future Research Engagement

The general experience of the students in the course was positive. Thirteen of the 19 students stated as "highly likely" their participation in future research opportunities on this topic (Fig. 5).

4.2.3 Curriculum Improvements

The curriculum improvements suggested by the students included the desired for additional room "to exchange experiences and to know how other colleagues are progressing." Similarly, others argued that the online offering constrained the in-class discussions among students working on different robots.

Finally, students inquired about the possibility of observing the actual robots in use. Although visiting sites

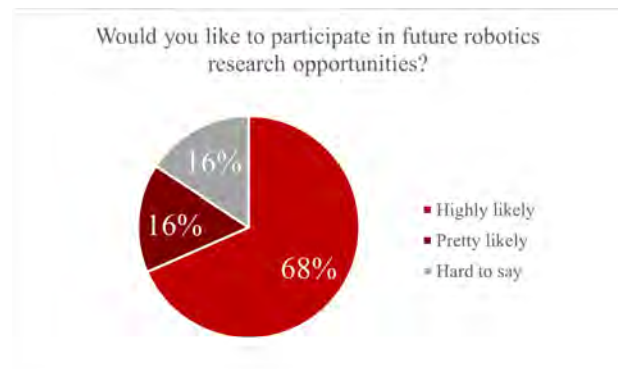


Figure 5: Future research participation.

or labs was not possible during the first offering due to Covid-19 restrictions, the second iteration of the class already included four in-person robot demonstrations.

5 Conclusions and Future Work

The researchers describe Stanford's first Construction Robotics curriculum to encourage discussion with other researchers and colleagues in Civil Engineering and Construction Management fields. We proposed a project-based course to connect real-world applications with the construction robotics research carried out at CIFE. Most of the existing construction robotic programs address on-site robots from a design and fabrication perspective. However, this curriculum emphasises the construction managerial perspective by matching the perspectives of construction managers and robot companies.

Our project-based curriculum included four main course learning goals (Table 3).

Table 3: CEE 327: Construction Robotics learning goals

#	Course Objectives
1	Introduce on-site construction robot applications
2	Assess the applicability and potential impacts of promising construction robots available in the market against traditional construction methods using a consistent evaluation method
3	Gain an understanding of basic robot principles and state-of-the-art research in the field, including human-robot collaboration approaches
4	Connect robot applications to broader project and client objectives deploying the EEE and VDC frameworks

Students without prior knowledge of construction robotics achieved comprehensive analyses and recommendations on using construction robots for the given projects. By collaborating directly with the industry professionals leading the development of construction robots, students learned a great deal about the strengths and limitations of

the practical application of the technology. The students' effort helped the professionals identify opportunities related to construction robot adoption with a consistent, repeatable, and exportable process.

Reflecting on the first implementation of the proposed curriculum, the researchers noted one key limitation. The curriculum places high demands on the teaching team to engage many industry professionals and develop the initial pairing between robot technologies and construction projects. If this collaboration with the industry is not possible, the curriculum cannot be successfully implemented.

Our future research will expand on the Robotics Evaluation Framework deployed in the class and the quantitative results achieved by consistently evaluating ten robot cases. Moreover, the second offering of the course involved students in Computer Science and Mechanical Engineering majors to complement the analysis from a multidisciplinary perspective.

Acknowledgments

The authors wish to thank the industry partners and colleagues who took part in the first edition of the class. Especially: Bechtel, Boston Dynamics, Canvas, Civ Robotics, DPR Construction, Exyn, HDlab, Hilti, Implenia, Kewazo, Megacentro, MT Højgaard, NCC, Obayashi (especially the contributions of Marcel Nagatoishi), Produktiva, SafeAI, Swinerton, Traylor Brothers, TyBot, and the University of Lima (with Prof. Alexandre Almeida) which provided us with the case studies. Special thanks are also due to the guest speakers: Andra Keay, Alex Walzer, Daniel Hall, Elena Galbally, Gunnar Skeie, Janis Pieterwas, Kiyoaki Okasita, Marco Annunziata, Masahiro Indo, Michal Wojtak, Nick Cote, Shuntaro Kano, Takehiko Nakaya, Tessa Lau, Thomas Bock, and Oussama Khatib.

References

- [1] Hesam Hamledari and Martin Fischer. The application of blockchain-based crypto assets for integrating the physical and financial supply chains in the construction & engineering industry. *Automation in Construction*, 127:103711, 7 2021. ISSN 0926-5805. doi:10.1016/J.AUTCON.2021.103711.
- [2] Timo Hartmann, William E. Goodrich, Martin Fischer, and Doug Eberhard. TR170: Fulton Street Transit Center Project: 3D/4D Model Application Report. 2007. URL <https://cife.stanford.edu/fulton-street-transit-center-project-3d4d-model-application-report>.
- [3] A Khanzode, M Fischer, D Reed, and G Ballard. WP093: A Guide to Applying the Principles of Virtual Design and Construction (VDC) to the Lean Project Delivery Process. 2006. URL <https://cife.stanford.edu/guide-applying-principles-virtual-design-and-construction-vdc-lean-project-delivery-process>.
- [4] Atul Khanzode. TR187: An Integrated, Virtual Design and Construction and Lean (IVL) Method for Coordination of MEP. 2010. URL <https://purl.stanford.edu/db106qv2525>.
- [5] Ronie Navon. Teaching automation as part of the construction curriculum. *Automation in Construction*, 5(4):301–311, 1996. ISSN 09265805. doi:10.1016/S0926-5805(96)00155-0.
- [6] I Custovic, R Sriram, and D M Hall. A Proposed Curriculum for Teaching the Tri-Constraint Method of Generative Construction Scheduling.
- [7] IFR. Executive Summary World Robotics 2018 Service Robots. Technical report, 2018. URL https://ifr.org/downloads/press2018/Executive_Summary_WR_Service_Robots_2018.pdf.
- [8] Thomas Bock and Thomas Linner. *Construction robots: Elementary Technologies and Single-Task Construction Robots*. Cambridge University Press, Cambridge, 2016. ISBN 9781107075993. doi:<https://doi.org/10.1017/CBO9781139872041>.
- [9] Architecture and Digital Fabrication – ITA Institute of Technology in Architecture | ETH Zurich. URL <https://ita.arch.ethz.ch/chairs/architecture-and-digital-fabrication.html>.
- [10] Laura Demsetz and Ronie Navon. Curriculum Development in Construction Automation. In *Proceedings of the 7th International Symposium on Automation and Robotics in Construction (ISARC)*. International Association for Automation and Robotics in Construction (IAARC), 11 2017. doi:10.22260/ISARC1990/0004.
- [11] Lehrstuhl für Baurealisierung und Baurobotik der TUM. URL <http://www.br2.ar.tum.de/>.
- [12] Master in Robotics and Advanced Construction - MRAC. URL <https://iaac.net/educational-programmes/masters-programmes/master-in-robotics-and-advanced-construction-mrac/>.

- [13] Robotics in Architecture and Construction - School of Architecture, Design and Planning. URL <https://lms-catalog.sydney.edu.au/browse/architecture-design-planning/courses/robotics-in-architecture-and-construction>.
- [14] C. Brosque, G. Skeie, and M. Fischer. Comparative analysis of manual and robotic concrete drilling for installation hangers. *Journal of Construction Engineering and Management*, 147(3):05021001, 3 2021. ISSN 0733-9364. doi:10.1061/(ASCE)CO.1943-7862.0002002. URL <http://ascelibrary.org/doi/10.1061/%28ASCE%29CO.1943-7862.0002002>.
- [15] Construction Automation & Robotics | University of Maryland Project Management, . URL <https://pm.umd.edu/course/construction-automation-robotics/>.
- [16] Construction Robotics M.Sc., . URL <https://cr.rwth-aachen.de/>.
- [17] Ben Purvis, Yong Mao, and Darren Robinson. Three pillars of sustainability: in search of conceptual origins. *Sustainability Science*, 14(3):681–695, 5 2019. ISSN 18624057. doi:10.1007/S11625-018-0627-5/FIGURES/1. URL <https://link-springer-com.stanford.idm.oclc.org/article/10.1007/s11625-018-0627-5>.
- [18] Martin Fischer, Atul Khanzode, Dean Reed, and Howard W. Ashcraft. *Integrating Project Delivery*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2017. ISBN 9780470587355. doi:10.1002/9781119179009.
- [19] C. Brosque, G. Skeie, J. Orn, J. Jacobson, T. Lau, and M. Fischer. Comparison of construction robots and traditional methods for drilling, drywall, and layout tasks. In *HORA 2020 - 2nd International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proceedings*, 2020. ISBN 9781728193526. doi:10.1109/HORA49412.2020.9152871.
- [20] Mi Pan, Thomas Linner, Wei Pan, Huimin Cheng, and Thomas Bock. A framework of indicators for assessing construction automation and robotics in the sustainability context, 2018. ISSN 09596526.
- [21] Oussama Khatib, Xiyang Yeh, Gerald Brantner, Brian Soe, Boyeon Kim, Shameek Ganguly, Hannah Stuart, Shiquan Wang, Mark Cutkosky, Aaron Edsinger, Phillip Mullins, Mitchell Barham, Christian R. Voolstra, Khaled Nabil Salama, Michel L'Hour, and Vincent Creuze. Ocean one: A robotic avatar for oceanic discovery. *IEEE Robotics and Automation Magazine*, 23(4):20–29, 2016. ISSN 10709932. doi:10.1109/MRA.2016.2613281.
- [22] Cynthia Brosque, Elena Galbally, Yuxiao Chen, Ruta Joshi, Oussama Khatib, and Martin Fischer. Collaborative Welding and Joint Sealing Robots With Haptic Feedback. *Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC)*, (Isarc):1–8, 2021. doi:10.22260/isarc2021/0003.

A case-based reasoning technique for evaluating performance improvement in automated construction projects

S. Krishnamoorthi^a and B. Raphael^a

^aBuilding Technology and Construction Management Division,
Department of Civil Engineering,
Indian Institute of Technology Madras, India
E-mail: ce15d012@smail.iitm.ac.in, benny@iitm.ac.in

Abstract –

Automation of construction processes facilitates increased productivity and overall higher project performance. This paper presents a methodology for comparative assessment of different construction processes and selection of an optimal solution based on appropriate automation implementation. Construction processes are quantitatively evaluated using a methodology combining case-based reasoning and compositional modeling.

Through the generation of many combinations of process fragments that are compiled from case libraries, potential solutions are explored and evaluated. An example involving solutions such as RCC frame construction, precast construction, and modular steel frame construction is described in this paper. The study demonstrates the possibility of selection of suitable construction processes based on the quantitative assessment of a large number of potential solutions. Processes are modelled by decomposing them up to the elementary tasks and appropriate level of automation is identified in all the tasks.

Keywords-

Case Based-Reasoning; Compositional Modeling; Discrete Event Simulation; Construction automation; Level of Automation; Therbligs; Project Performance Improvement.

1 Introduction

Construction industry, due to its complexity and diversity, has high potential for the application of automation. Over the years, various automation systems have been developed [1-2] that demonstrate their capabilities to perform many of the basic building activities [3]. However, implementing automation systems in construction is quite challenging since the

nature of construction is highly heterogeneous with high variability in the working environment [4]. Construction automation has the capacity to eliminate manual-labour oriented non-value adding processes, and improve those manual-labour processes that contribute to value addition. Thus, construction automation can deliver improvement in project performance. This can be achieved with relevant process planning methods and tools [5]. Though there is broad consensus that construction automation achieves reduction in time and cost, the research based on quantitative methodologies that evaluate the improvement of productivity by implementation of suitable automation is not sufficiently explored [6].

Recently, the authors have developed a systematic methodology for productivity analysis of automated construction processes [7]. The study specifically explores the use of simulation tools to predict productivity improvement through creating multiple processes by combining process fragments from a case base. The aim of this paper is to demonstrate the methodology of generating simulation models of construction processes for identifying optimal levels of automation.

2 Literature Review

Study of various literature by authors [8] reveal that analysing the performance of construction projects is very powerful and advantageous using discrete event simulations [9-14], especially those involving time-cost based evaluations [15-18]. This gives rise to a research question:

How to select an optimal typology of construction process that maximises the overall project performance from a given set of possibilities?

Based on the literature review, the following specific knowledge gaps are identified.

1. There exists limited research that evaluates the productivity of automated construction processes.

2. The prevalent discrete event simulation models are based on static process models in which the structure of the process is fixed a-priori. They are not capable of evaluating multiple options for the activities in the process.

The challenge of optimizing processes using appropriate automated methods is addressed by the authors through a new methodology [7]. This research work is inspired from the concept of compositional modelling from the domain of computer science [19]. Previously, it has been applied for design problem solving [20-21] and system identification [22].

The authors followed design science research methodology [23] approach to develop this framework. This involved problem identification and objectives development. This was followed by search for solutions from the literature and exploratory implementation. Demonstration and evaluation of simulation methodologies. The issues and challenges were identified.

Through iterations, modifications to the solution were done.

This paper demonstrates the evaluation of different construction process typologies showcasing its application for complex problems in construction. For demonstration, the authors have considered three types of construction process typologies, with different levels of decompositions with different automation possibilities (For example, rebar cutting process has possibilities of execution through mechanical saw, electrical circular saw, and automated rebar cutter machine among systems). The methodology is able to evaluate the potential for different types of automation through search and exploration approach.

The remaining portion of this paper explains the methodology, its application through demonstration, results of study and evaluation, and lastly summary and conclusion.

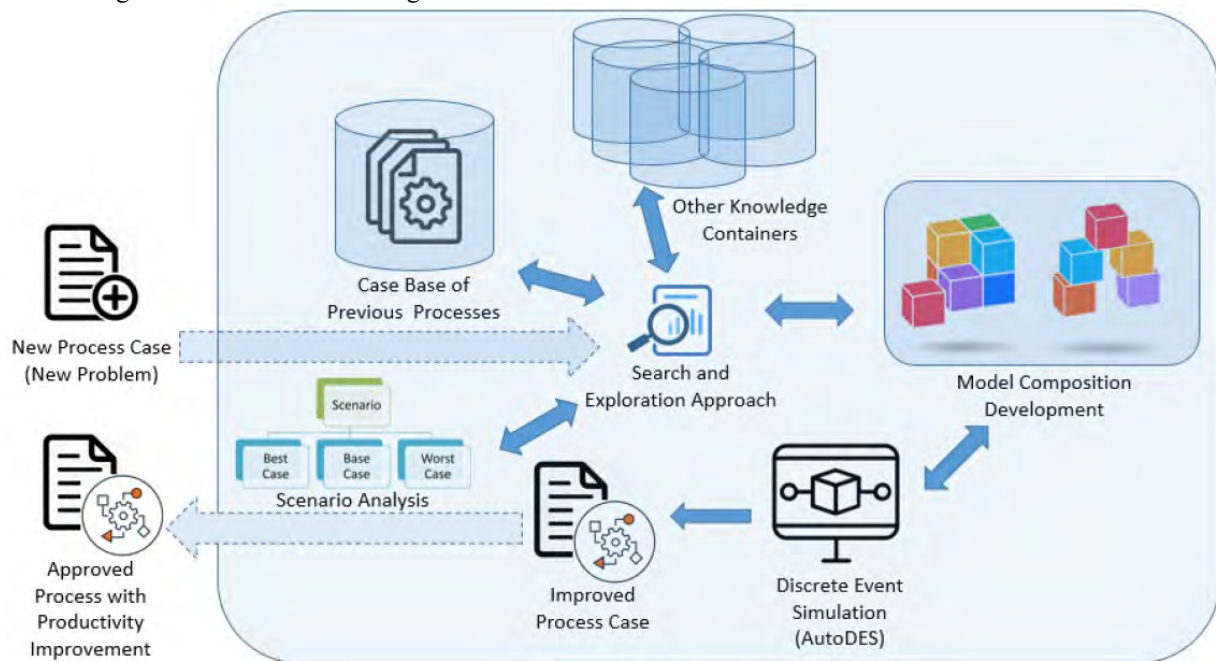




















Figure 1. A case-based reasoning methodology using compositional modelling for construction process evaluation [7]

3 Combining compositional modeling and case-based reasoning

The present methodology (Figure 1) involves a combination of case based reasoning and compositional modelling (This is explained in detail by authors in [7], is briefed in the following portion). Multiple discrete event simulation models are generated by combining fragments of models using the methodology of compositional modelling. Model

fragments are compiled from cases. A search and exploration approach using time and cost as parameters is used to identify optimal level of automation for construction process. The construction process is decomposed into fragments up to the basic task level or therbligs (Table 1). This is useful for the assessment of manual labour and suitability of introducing appropriate automation as replacement.

Table 1. Therbligs and their symbols

Sl. No.	Therblig	Symbol	Sl. No.	Therblig	Symbol
1	Search		10	Use	
2	Find		11	Disassemble	
3	Select		12	Inspect	
4	Grasp		13	Preposition	
5	Hold		14	Release Load	
6	Transport Loaded		15	Unavoidable Delay	
7	Transport Empty		16	Avoidable Delay	
8	Position		17	Plan	
9	Assemble		18	Rest	

Past construction processes that are decomposed in this form are stored in case libraries (Figure 2). In the present work, the libraries are created from cases with data collected from various sources such as construction sites, laboratory experiments, and videos from the world-wide-web. The sources for the web-based videos are primarily from 81 videos taken the YouTube involving the RCC construction process execution through different execution modes. This exercise was performed to increase the multitudes of variations of process fragment cases to generate combinatorial solutions during simulation.

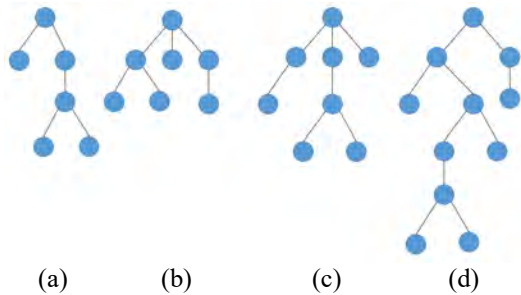


Figure 2. A schematic representation of multiple activity cases in case library

The objective is to capture various modes of performing activities with different levels of automation.

For each construction process fragment, there are multiple possible modes of execution; a top-level process fragment may be performed through several sub-processes. An object oriented representation is used for process fragments. Fragments are grouped into classes based on standard object oriented representation concepts such as inheritance.

Fragments that are at the top of the inheritance hierarchy represent generic activities and those at the lower levels represent specialized modes of execution of these activities. The inheritance hierarchy represents “type-of” relationships. Children classes are special types of parent classes.

The work breakdown structure [24] of the overall construction process contains fragments of process that are in “part-of” relationship (Table.2). This decomposition is down to the level of basic tasks or therbligs (also refer section 4.1).

Table 2. Construction process typology-1: RCC frame assembly: portion of work breakdown structure

ACTIVITY CASES AND DESCRIPTION
<!-- 1.1.1 RCC COLUMN CONSTRUCTION -->
<!-- 1.1.1.1 STEEL REINFORCEMENT ASSEMBLY -->
<!-- 1.1.1.1.1 (A) FABRICATION -->
<!-- 1.1.1.1.1.1 (A.1) MAIN REINFORCEMENT -->
<!-- 1.1.1.1.1.1.1 (A.1.1) TRANSPORTING -->
.
.
<!-- 1.1.2 RCC BEAM CONSTRUCTION -->
<!-- 1.1.2.1 STEEL REINFORCEMENT ASSEMBLY -->
.
.
<!-- 1.1.2.2 CONCRETING -->
<!-- 1.1.2.2.1 (B.1) CONCRETE MIX PREPARATION -->
<!-- 1.1.2.2.2 (B.2) TRANSPORTING TO .. LOCATION -->
<!-- 1.1.2.2.3 (B.3) CONCRETE POURING -->
<!-- 1.1.2.2.4 (B.4) COMPACTING -->
.
.

During case adaptation, similar fragments that belong to the same parent class are interchanged provided the constraints are satisfied. Therefore, it becomes feasible to generate multiple processes for any top-level task by using different combinations of fragments for sub-activities at lower levels.

For example, a combination of different case options involving, say, transportation of resources from stockyard to site location, assembly of column components, stirrup bending, and more cases, give rise to millions of combinations with varying degrees of automation containing manual, mechanical, electromechanical, electronic, and automation systems.

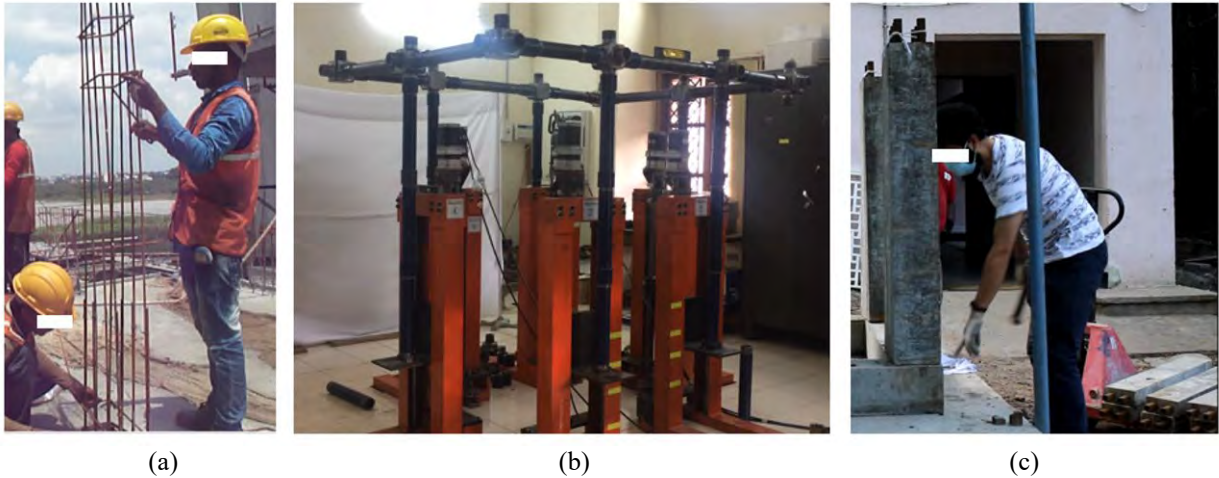


Figure 3. Construction process typology cases: (a) RCC frame assembly, (b) steel frame assembly, and (c) precast frame assembly

3.1 Description of Construction process typology cases

An example involving the construction of a structural frame is considered here. Figure 3 shows three cases having different construction typologies.

The first one is a conventional RCC frame construction involving tasks such as preparation and assembly of the rebars and stirrups, formwork preparation, and preparation of concrete mix, transporting, and casting. The second one involves a steel structural frame in which mild steel pipes are used to construct columns. Cylindrical couplers with internal threading are used for connections. Beam modules are also made of steel pipes and connected using couplers mentioned before. The third case involves a precast frame assembly. Eight RCC columns containing hollow square steel tubes instead of conventional reinforcement are used. The steel tubes project 75 mm outwards to create a moment resisting connection with the beam.

Certain assumptions are made in the simulations: the foundation work is completed, and the frame above is assembled. All the required resources are available at project location.

3.2 Data collection

Data was collected from multiple sources such as onsite project construction activities, laboratory based construction assembly experiments, and construction videos from multitudes of videos from world-wide web resource, such as YouTube. For each activity, the time is compiled from process fragment cases. Moreover, for each activity, the resources required were identified: such as the labour, helper, tools and equipments utilised, operating staff, and materials handled. The costs for activities were based on CPWD rates [25] and cost details

collected from other sources such as brochures and online sources.

4 Evaluation of the performance of the process typologies

A new project scenario is taken to demonstrate case-based model composition. Cases are adapted through substitution of similar and suitable cases based on the consideration of inheritance relationship. An in-house software (called AutoDES) developed for this research performed the discrete event simulation for estimating the time taken for completing the construction. The detail of implementing the model and performing the simulation is described below.

4.1 Generation of model cases

Cases are represented in XML format. Each case is not only a decomposition of the process but it also contains the various resources involved. The decomposition is performed up to the basic level tasks or Therbligs. For example (Figure 4), in the rebar bending process fragment, the task is broken down into basic fragments or therbligs comprising search, find, select, grasp, hold, position, use, and so on. Each therblig has time data in seconds.

These activities may be run in parallel or in sequence depending on the case being considered. Their durations are stored in the cases. The specification of the number of cycles of therbligs is also included in case.



Figure 5. A portion of RCC frame construction process with search and substitution of process fragments

```

<bending_stirrup_using_MechanicalBender1>
  <!-- bending stirrup using Mechanical Bender1,

  <Resource>
    <!-- Resource required for this activity -->
    <labor> 1 </labor>
    <MechanicalBender1> 1 </MechanicalBender1>
  </Resource>

  <Activities>

    <therblig_Search>
      <Duration> 0,2 </Duration>
    </therblig_Search>

    <therblig_Find>
      <Duration> 0,2 </Duration>
    </therblig_Find>

    <therblig_Select>
      <Duration> 0,2 </Duration>
    </therblig_Select>

    <therblig_Grasp>
      <Duration> 0,3 </Duration>
    </therblig_Grasp>

    <therblig_Hold>
      <Duration> 0,3 </Duration>
    </therblig_Hold>
  
```

Figure 4. A portion of schematic of therbligs based XML coding for rebar bending activity

The classification of all the activities is done using inheritance relationships. For example, bending stirrups using mechanical bender is specified as a special case of bending stirrups. Similarly, manually bending bars to make stirrups is also a special case of the same parent class. Both the children classes might be replaced with each other to create new processes having different sets of requirements for resources, resulting in different times of execution.

4.2 Generating solutions for a new problem

To the case based reasoning module, sufficient data with respect to a new project is input. This includes the resources (such as labour, and equipment), and inheritance relationships (part-of relationships of the group cases).

Through case adaptation, multitudes of processes are generated through a random search algorithm. For generating a new solution, every adapted case undergoes a combination of portion of activities from the previous cases. This combinatorial model generation approach end in generating new processes based on the combinations of process fragments in the case library.

For example, in the RCC frame construction process (Figure 5), each fragment has multiple case possibilities of substitution containing manual, mechanical, electromechanical, and automated systems.

AutoDES, a simulation tool developed in-house for this research, is applied for the performance of discrete

event simulation. The AutoDES executes simulations involving scanning of activities. It contains a simulation clock which, at a given time step, checks the availability of any activities that requires to be scheduled. Once, all the activities are completed, the simulation stops.

The output of the simulations provide with the results in the form of time durations for every process fragment. For each process solution, the capital and resource costs are derived separated.

5 Results

After simulations of processes generated through compositional modeling, potential solutions are selected based on their time and cost. Millions of combinations of process fragments are generated during this procedure.

In the illustrative example considered in this study, for RCC frame construction, the stages up to concreting is included. The curing stage is excluded. The manufacturing of the steel modules of the steel assembly process typology is excluded from the scope. The three construction process typologies contain a combination of activities that involve manual, mechanical, and electromechanical, tools and devices. The time duration for each stirrup rebar bending (process fragment) using mechanical plier type-1, type-2, and automatic bending machine are: 84 sec, 24 sec, and 04sec, respectively. The time is drastically reduced by using an automation machine, however, it comes with high investment cost. Similarly, in another instance, during concrete mix preparation (process fragment), the time varied with different mixer types: 84, 59, and 38 sec. The first two durations were from drum mixers, while the third duration was from handheld power mixer. The substitution of process fragments from the above mentioned examples and others process fragments in to the overall RCC construction process gave significant performance improvement in the results.

The durations for RCC frame assembly, steel frame assembly, and precast assembly (in minutes) are 1214.07, 68.23, and 984.90 respectively. The optimal solution obtained for the selected example, involves process typology-2 (steel frame assembly) based on time alone. The most appropriate process based on combined time and cost performance involves pre-cast frame assembly.

6 Summary and Conclusion

This paper demonstrated the application of a methodology combining case-based reasoning and compositional modeling for identifying the optimal level of automation in construction. Cases of construction processes consisting of three process typologies were used: RCC frame construction, precast frame construction, modular steel construction.

For exploring the different approaches of construction with various levels of automation, case libraries were generated based on quantitative data inputs from site, laboratory experiments, and web-based resources. Every case is unique with the construction methods and tools adopting manual, mechanical, electro-mechanical or other approaches. Multiple cases are represented through breaking down the entire construction process into fragments of activities to the level of elementary tasks or Therbligs.

It is concluded that a comparative analysis of different typologies of construction processes leads to better decision making for appropriate level of automation in construction implementation and achieving project performance improvement. Results show that implementation of automation facilitates significant saving of construction time through selecting an appropriate process typology. The savings in time and cost can be quantified and could be used in decision making depending on the preferences of the user.

Research Contributions:

This research facilitates early decision making of appropriate automation to achieve overall construction process enhancement. This study demonstrates the application of compositional modeling, a concept in computer-science domain, to the field of construction automation.

Limitations and future work:

Present study focuses on low-rise residential construction projects and has limited number of cases at present. Future work involves better algorithms for time-cost trade-off.

The stages that are currently not included in the present study of overall construction processes, such as curing of RCC, are considered for accommodating in the future research work for a more comprehensive study.

Acknowledgements

The authors wish to thank (1) Mr. Rakshith (M.Tech, BTCM, IIT Madras), and team members of Project-A (name not disclosed due to confidentiality) for the field data of construction operations; (2) Project staff team of Building Automation Laboratory, BTCM, IIT Madras for the lab-based construction operations.

This project is funded by the Department of Science and Technology (DST) and Science and Engineering Research Board (SERB), India, through the grants DST/TSG/AMT/2015/234 and IMP/2018/000224. The Institute of Eminence Research Initiative Project on Technologies for Low-Carbon Lean-Construction (TLC2) has also provided financial support for this research.

References

- [1] Bock T. Construction automation and robotics. *Robotics and Automation in Construction*, 21-42, 2008.
- [2] Neelamkavil J. Automation in the prefab and modular construction industry. In *Proceedings of the 26th International Symposium on Automation and Robotics in Construction*, pages 299-306, Austin, USA, 2009.
- [3] Morales G., Herbzman Z. and Najafi FT. Robots and construction automation. *Automation and Robotics in Construction XVI UC3M*, 283-288. 1999.
- [4] Krishnamoorthi, S. and Raphael, B. A methodology for analysing productivity in automated modular construction. In *Proceedings of the 35th International Symposium on Automation and Robotics in Construction*, pages 1030-1037, Berlin, Germany, 2018.
- [5] Martinez J.C. EZStrobe - general-purpose simulation system based on activity cycle diagrams. In *Proceedings of Winter Simulation Conference*. Pages 341 – 348. Washington DC, USA, 1998.
- [6] Krishnamoorthi, S. and Raphael, B. A Theoretical Framework to assess Project Performance through Construction Automation. In *Proceedings of the 4th International Conference of The Robotic Society*, pages 1-8, Chennai, India, 2019.
- [7] Krishnamoorthi, S. and Raphael, B. A Case Based Reasoning Approach for Selecting Appropriate Construction Automation Method. In *Proceedings of the 38th International Symposium on Automation and Robotics in Construction*, pages 880-885, Dubai, UAE, 2021.
- [8] Krishnamoorthi, S. and Raphael, B. A review of methodologies for performance evaluation of automated construction processes. *Built Environment Project and Asset Management*, Vol. ahead-of-print No. ahead-of-print. Article publication date: 31 December 2021 <https://doi.org/10.1108/BEPAM-03-2021-0059>.
- [9] Poshdar M., González V. A., Sullivan M., Shahbazpour M., Walker C. G. and Golzarpoor H. The role of conceptual modeling in lean construction simulation. In *Proceedings of 24th Annual Conference of the International Group for Lean Construction*, pages 63-72, Boston, Massachusetts, USA, 2016.
- [10] Dang T. Modeling Microtunnelling Construction Operations with WebCYCLONE. *Journal of Geological Resource and Engineering*, 4: 88-196, 2017.
- [11] Charinee L. and Nathee A. An Application of Discrete-Event Simulation in Estimating Emissions from Equipment Operations in Flexible Pavement Construction Projects, *Engineering Journal*, 21 (7): 197-211, 2017.
- [12] Muhammet E.A., Sadik A., Yiğit K. and Erkan D. Simulation Optimization for Transportation System: A Real Case Application, *TEM Journal*, 1: 97-102, 2017.
- [13] SangHyeok H., Shafiul H., Ahmed B., Mohamed A. and Joe K. An integrated decision support model for selecting the most feasible crane at heavy construction sites. *Automation in Construction*, 87 (3): 188-200, 2018.
- [14] Chijoo L. and Sungil H. Automated system for form layout to increase the proportion of standard forms and improve work efficiency. *Automation in Construction*, 87 (3): 273-286, 2018.
- [15] Shide S., Mohammed M. and Amin H. Performance analysis of simulation-based optimization of construction projects using High Performance Computing. *Automation in Construction*, 87 (3): 158-172, 2018.
- [16] Hamed G., Vicente A. G., Michael S., Mehdi S., Cameron G.W. and Mani P. A non-queue-based paradigm in Discrete-Event-Simulation modelling for construction operations. *Simulation Modelling Practice and Theory*, 77 (9): 49-67, 2017.
- [17] Kaira S. and Tezeswi T. P. A Study on Simulation methodology in Construction Industry. *Imperial Journal of Interdisciplinary Research*, 2 (8): 962-965, 2016.
- [18] Fahimeh Z. and James O. B. R., An ACD diagram developed for simulating a Bridge construction operation. *International Journal of Construction Supply Chain Management*, 4 (2): 34-50, 2014.
- [19] Falkenhainer B. and Forbus K.D. Compositional modeling: Finding the right model for the job. *Artificial intelligence*, 51 (1-3): 95-143, 1991.
- [20] Raphael B. Reconstructive Memory in Design Problem Solving. *PhD thesis*, University of Strathclyde, Glasgow, UK. 1995.
- [21] Kumar, B. and Raphael, B., CADREM: A case-based system for conceptual structural design. *Engineering with computers*, 13(3): 153-164, 1997.
- [22] Robert-Nicoud, Y., Raphael, B. and Smith, I.F. System identification through model composition and stochastic search. *Journal of computing in civil engineering*, 19(3): 239-247, 2005.
- [23] Wieringa, R.J., Design science methodology for information systems and software engineering. *Springer*. 2014.
- [24] Sanchez, J.D.F., A guide to the project management body of knowledge (PMBOK Guide) –Fifth Edition. *Project Management Inc*. 2013.
- [25] Central Public Works Department, Delhi Schedule of Rates, *DG-CPWD*. (1). 2021.

The Optimized Process for Effective Decision Makings to Minimize Fall From Height (FFH) Accidents on Construction Site

Q. Zeng and J.H. Yoon

School of Building Construction, Georgia Institute of Technology, United States of America

E-mail: qzeng41@gatech.edu, jyoon337@gatech.edu

Abstract –

As one of the most hazardous industries, the construction sector is specifically dangerous due to its dynamic and unexpected working conditions. Among all the occupational risks, Fall From Height (FFH) accidents remain the riskiest ones for a long time, resulting in a huge loss for projects. Although numerous technologies have been adopted to reduce the occurring rates of FFH accidents, the results are not as satisfying as expected. This research aims to figure out both the casual factors or indicators of the FFH accidents and the limitations of the current sensing system for preventing FFH accidents by conducting a systematic literature review. Based on the findings, this study provides potential solutions to each limitation. In addition, this study proposes the optimized process of FFH accident prevention that enables the relevant stakeholders to make effective decisions on the safety issues associated with FFH accidents. This proposed process will set the foundation for system development to minimize FFH accidents, serving as a contribution to the body of knowledge.

Keywords –

Fall from height; Construction safety issues; Technological solutions; Improved system process flow

1 Introduction

Considered as one of the most dangerous and risky industries, construction sector is extremely hazardous due to its dynamic and complex inheritance [1] caused by poor working conditions, direct exposure to outdoor environments, and unpredictable change orders [2]. Interactions between workplaces and workforces are complicated but unorganized under some circumstances. As a result, various occupational injuries are identified and addressed [3], causing by trapping, slipping,

overexertion, electrical struck, and especially falling from height (FFH) of both objects and human. Based on the reflections from the NIOSH FACE reports [4], FFH is considered as the most significant occupational risk due to its frequently fatal consequences. Additionally, according to the statistics from Occupational Safety and Health Act (OSHA) 2021, the number of violation cases of fall protection is 5295, much more than the number of respiratory protection violations which is 2527. Consequently, there is an urgent requirement to mitigate the problem and decrease the fatality of FFH.

After introducing the significance of FFH to the safety management of construction site, the general types and leading causes for FFH accident will be discussed. As summarized by Liy, Ibrahim [5], six major types of FFH are presented as: falling from roof, falling from scaffolding, falling through holes, falling from ladder, falling from aerial lift platforms, and falling from building girders or other structures. Among all the types, “falling from roof” is considered as the most hazardous type on the basis of a questionnaire for construction professionals and several secondary data such as literature review and case studies. Moreover, four main kinds of causes for FFH are concluded, including “dangerous actions”, “dangerous working conditions”, “communication barriers”, and “failed management commitment”. Among all the causes, the highest contributors are “failed to wear Personal Protection Equipment (PPE)” for dangerous actions, “poor site housekeeping” for dangerous working conditions, “information unclearness” for communication barriers. Therefore, these are the key topics under discussion about solving problems on construction sites using technologies currently.

Based both on experiences and regulations, traditionally, three key approaches for FFH protection are setting safety nets, installing guardrails, and wearing personal protection equipment according to OSHA rules. However, these measures are not accurate which including locally specific and worker specific, leading to

low efficiency and effectiveness of worker protection. Nowadays, numerous technologies such as radio-frequency identification (RFID) [6], ultra-wide band (UWB) [7], global navigation satellite system (GNSS) [8], and Bluetooth low energy (BLE) beacons are applied to increase the accuracy and effectiveness for solving safety issues [9]. Here are some current functions and purposes of those technologies concluded from previous research (see Table 1).

Table 1 Purposes of some technological implementations on site for falling prevention

Technology	Purpose	Source
RFID	Detect and control PPE on site.	[6]
UWB	Identify, track and locate equipment and materials.	[7]
GNSS	Detect, categorize and organize risk area for workplace.	[8]
BLE	Monitor and harness on site.	[9]

However, although numerous technologies are presented to construction site managers, due to the uncertainty of device effectiveness as well as the limitation and prerequisites for some solutions, many site managers remain suspicious and conservative. In order to explore more, a systematic literature review is conducted for understanding the specific reasons. After figuring out why most site managers refused to adopt technologies for mitigating negative impacts of FFH accidents, an improved process flow aiming at assisting stakeholders with decision making for safety issues on construction site will be proposed, based on the benefits of high-tech devices, prerequisites and capital costs, as well as the conditional effectiveness. Additionally, this proposed process flow will set the foundation for system development of minimization of construction risks, serving as the contribution to the body of knowledge.

2 Research Methodology

The research methodology selected for this research is a systematic literature review, including necessary procedures that defines the key words for searching as well as selection and exclusion strategy for extracting valuable literature information are required. Therefore, this research follows the guidelines created by Thom  , Scavarda [10] which contains four steps as: determine the purpose, select the database, select the keywords, review the relevant articles.

2.1 Determine the Purpose

In order to explore the reasons why site managers are

reluctant to implement technological devices in a large scale, the process and mechanism of FFH accident prevention must be understood in advance. Also, the limitation, advantages, and shortcomings of current solutions are under investigation. Based on the results from literature review, elements and indicators that influence the decision-making process of site managers would be acknowledged, served as the foundation for framework establishment.

2.2 Select the Database

Based on the analysis of journal coverage, topic relevance and academic influence from the research done by Mongeon and Paul-Hus [11], Web of Science is chosen due to its significant influences on natural science and engineering.

2.3 Select the Key Words

Since the purpose of literature review is investigating the process and mechanism of FFH accident prevention as well as finding out the pros and cons of current solutions, the topic key words for searching are selected as “fall from height”, “technology”, and “accident”. After first round selection, 81 articles from all time are shown (see Figure 1).

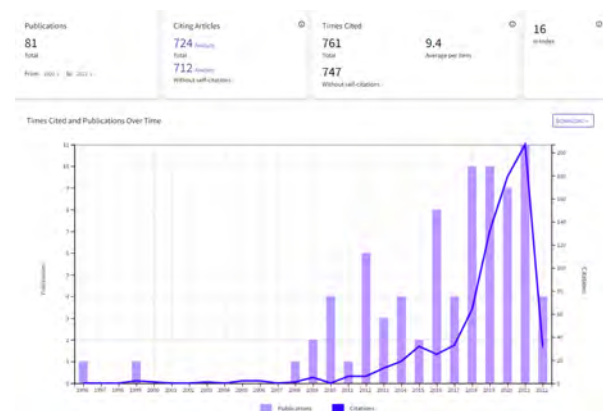


Figure 1 Results from first round selection of literatures for reviewing

2.4 Review the Relevant Articles

Based on the purpose, number of citations and topic relevance, numerous articles are selected for reviewing. Among them six publications are considered as the main materials for literature review due to their relevance and contribution (see Table 2). Three categories are set for the articles as: (a) FFH casual factors, (b) FFH prevention process, and (c) introduction of technical solutions about FFH. The categories are listed along with the titles of literatures.

Table 2 Final literature selection for reviewing

Title	Reason of Selection
<i>(a) Falls from Height in the Construction Industry: A Critical Review of the Scientific Literature [12]</i>	A total of 297 articles concerning fall incidents were reviewed, which are enough to support a critical review.
<i>(a) Analysis of the Contributing Factors for Fatal Accidents due to Falls from Heights in Malaysia and the USA [13]</i>	Official data from the department of Occupational Safety and Health (DOSH) in Malaysia and Occupational Safety and Health and Administration (OSHA) in the USA are the research objects.
<i>(c) Improving the prevention of fall from height on construction sites through the combination of technologies [14]</i>	This research provide several alternative solutions found in the literature about the integrations between different technologies for FFH prevention.
<i>(b) Intelligent Platform Based on Smart PPE for Safety in Workplaces [15]</i>	This research proposes a comprehensive architecture about data transmission process with PPE.
<i>(b) Leading Indicators-A Conceptual IoT-Based Framework to Produce Active Leading Indicators for Construction Safety [16]</i>	This research presents a methodology about utilizing IoT system to collect quantifiable data for actionable response in real time, serving as another approach for optimizing FFH prevention process.
<i>(b) (c) A Fall from Height prevention proposal for construction sites based on Fuzzy Markup Language, JFML and IoT solutions [17]</i>	This research illustrates the benefits that IoT and Fuzzy Logic Systems (FLS) could provide to support decision-making process for FFH prevention.

Three main purposes are set for the systematic literature review: determining casual factors and indicators of FFH, understanding current cyber infrastructure and system architecture, and discovering

the limitations and prerequisites of current system and technologies for falling protection. The next step would be the utilization of these information. As the research goal is to assist stakeholders with decision making for safety issues on construction site, the methodology for this research is designing a system architecture for them, helping them reduce the loss caused by FFH accidents while controlling the budget required for technological implementations (see Figure 2).

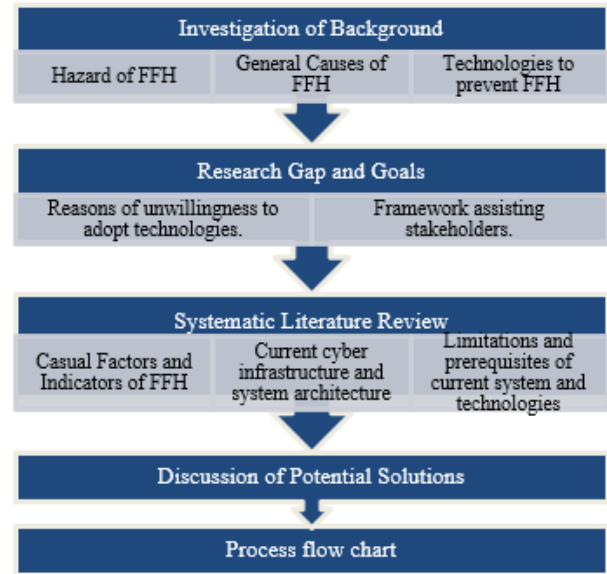


Figure 2 Research Methodology and Process

3 Literature Review Findings

3.1 Casual Factors and Indicators of FFH Accidents

In order to provide the research background about causal factors and indicators leading to FFH accidents, a proven classification obtained from Nadhim, Hon [12] has been applied for this part, explained as below (see Table 3).

Table 3 Classification of risk factors and indicators for FFH accidents from the work of Nadhim, Hon [12]

Risk factors	Indicators	Quantitative variables (unit)
Risk Activities	Working at height	Workplace level or height (ft)
	Complex, difficult, and prolonged tasks	Estimation of working time (hr/person)

Individual Characteristics	Demographic information	Age, gender, weight, ethnicity
	Knowledge level	Education level, working and training time (hr)
	Human behaviours	Times of misjudgement and unsafe behaviours
	Worker's well-being	Absence time due to fatigue, illness, average sleeping time (hr)
Site Conditions	Insufficient illumination	Number of lights, illumination level (lux or lumen/sqft)
	Unprotected platform and surface	Coverage area under protection (sqft)
Organization, Management	Lack training programs	Compulsory training time (hr/person)
	Improper self-protection	PPE availability (counts) and ratio of wearing
	Improper shifting	Shifting times and break periods (min)
	Pressure from schedule	Time until the end date (day)
Weather, Environmental conditions	Bad weather	Temperature (F), humidity (%), noise (dB)
	Untidy workplace	Empty place in a certain area (sqft)

Indicators from Table 2 demonstrate the possible reasons which are responsible for the increasing number of FFH accidents based on the descriptions mainly from OSHA. Also, it would be considered as the criteria for evaluating FFH risk level, a significant part for the establishment of the conceptual framework.

3.2 Cyber Infrastructure and System Architecture for Preventing “Fall from Height” Accidents

By investigating the prototype system from [16], [17], and [15], a normal FFH accidents sensing system contains three major layers as: perception and physical detection layer, transmission and computing layer, and service and cyber-secure applications layer (see Figure 3). This whole system is relying on the Bluetooth Low Energy (BLE) beacon monitoring system.

- Perception and physical detection layer: As the lowest level of the whole architecture, all the physical devices such as sensors, actuators on the helmet, belt, and even bracelet are covered in this layer. Not only aiming at collecting data and gathering status information of both workers and workplace conditions, but also this elementary layer is also responsible for sending the data to the upper layers.
- Transmission and computing layer: The purpose of this layer is to transfer and process the primitive data coming from the first layer and send the transformed data which is considered as the input data to the application layer. This layer contains three key components as local aggregator, service gateway and transmission protocol. The local aggregator is responsible for initially organizing and filtering raw data from sensors, making it transmittable. Service gateway is considered as the “gate” for service and application layer due to its function of receiving and securing data transmission. Normally, there's a modification and repetition loop between the local aggregator and the service gateway, managing to adapt and optimize the system based on user preferences and environmental conditions. As for the last component, one of the widely adopted protocol is the Message Queue Telemetry Transport (MQTT), which is established on the basis of TCP/IP [18].
- Service and cyber-secure applications layer: Considered as the upper part of the structure, this layer is importing information and store them in the cloud environment for both visualization and data analysis such as deep learning models and machine learning algorithms.

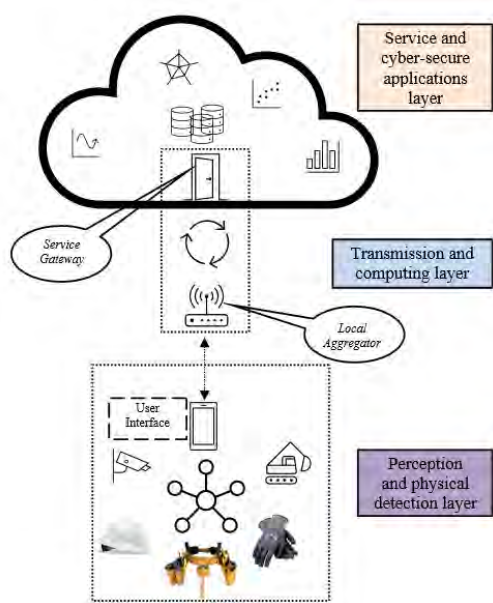


Figure 3 Cyber infrastructure for FFH accidents sensing system

The procedures of FFH accident prevention of this sensing system could be summarized as the following steps: equip workers with self-protection BLE lifeline system, detect and categorize safety area in workplace, send warnings to both workers and management system, and record the parameters from the hazardous situation and save them at the database (see Figure 4).

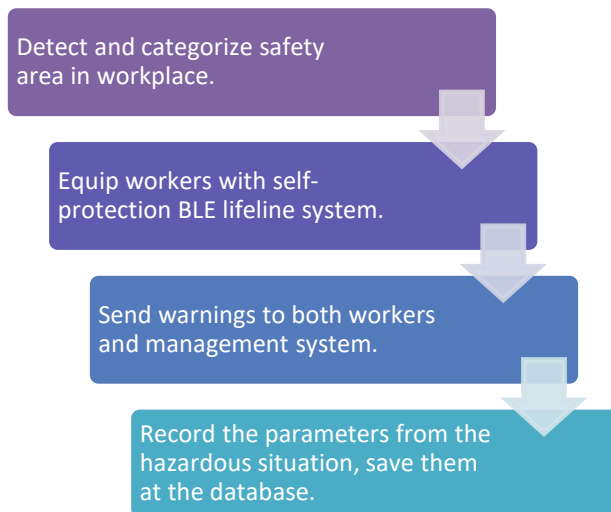


Figure 4 Procedures of FFH accidents prevention

3.3 Limitations and Prerequisites of Current System and Technologies for Falling Protection

Although BLE monitoring system demonstrated strengths in the robustness in risk detection, convenience in installment, calibration and relocation, as well as availability of external support of map updates [9], several important limitations which can't be ignored are presented to construction site managers, too (see Table 4).

Table 4 Limitations and prerequisites of BLE monitoring system

Items	Prerequisites	Limitations	Ref
Risk zone identification, categorization	Analysis of certain areas, locational information	Inaccuracy of area identification due to insufficient number of BLE beacons and opaque or dim conditions.	[9]
Wearable BLE devices	Preinstalled devices and routinely maintenance	Cheating detection required, lag of warning due to poor interference of signal, lack of real-time feedback of risk exposure due to remote monitoring	[19]
Data transfer through Wi-Fi system	Reinstalled Wi-Fi routers	Unavailable in isolated zones, interfered by metal materials	[20]
Computational analysis of worker and workplace data	Sufficient locational and worker data, cloud-based database for storing	Estimation of worker placement with respect to beacons has to be simplified, low accuracy for insufficient data.	[14]

4 Discussion of Potential Solutions

After figuring out the limitations and prerequisites of current sensing system, potential solutions should be identified and analyzed to mitigate the negative impacts caused by limitations. Based on the previous Table 4, keys to solving the limitations are proposed (see Table 5).

Table 5 Potential Solutions for limitations/problems of current sensing system

Items	Problems	Solutions
Risk zone identification, categorization (1)	- Inaccurate area detection & categorization. (1.1)	- GPS system + BIM model + K-means clustering [21]
	- Determination of required beacon number. (1.2)	- Simulation on BIM model
	- Potential FFH risks are not on building plans. (1.3)	- Automatic-BIM based FFH identification and planning tool [22]
Wearable BLE devices (2)	- Cheating detection. (2.1)	- Computer vision-based movement detection
	- Warning lag, lack of real-time feedback (2.2)	- Real-time location system to detect the proper use of PPE [23]
Data transfer through Wi-Fi system (3)	- Unavailable in isolated zones. (3.1)	- Global Navigation Satellite System (GNSS)
	- Metal material interference. (3.2)	- Second RFID signals
Computational analysis of worker and workplace data (4)	- Localization of workers in the certain workplace. (4.1)	- BIM model + Simultaneous location and mapping (SLAM) algorithm [24]

Explanations of the solutions:

- (1.1): In order to categorize the area and set the boundaries accurately, Global Positioning System (GPS) is applied for locating the coordinates of essential points on site. By combining the data with 3D BIM model rather than traditional 2D building drawings, a more accurate risk distribution map

could be made for construction managers, while the K-means algorithm could categorize the whole project area into different clusters based on their historical data about FFH accidents, risk factors, demographical information of workers, etc. Results from the analysis will be directly shown to site managers for further operations.

- (1.2): One of the applications of BIM modeling is the 4D simulation for clash detection and making scheduling alternations. Hence, if 4D BIM is used for determining the exact location and number of sensors that are required on site, the cost for real-time experiment could be diminished since the managers could experience the real-time situation simply during the simulation.
- (1.3): According to the paper by Zhang, Teizer [25], they created an automatic BIM-based FFH identification and planning tool, which could help with the beacon harness system to confine the inspection areas for risk detection, better than general inspections that are time and money costing.
- (2.1): In some cases, all the BLE devices are connected properly, but some workers undress the harness and leave them to the lifeline since it's not comfortable to wear them all the time, cheating the system eventually. With the help of computer-vision-based movement detection, even if the devices are connected, warnings will be sent if cameras detect suspicious cheating movements.
- (2.2): In order to have real-time feedbacks between workers and the system, Dong, Li [23] designed a real-time location system for detecting proper use of PPE and giving feedbacks to workers under urgent circumstances.
- (3.1): For some specific projects, including

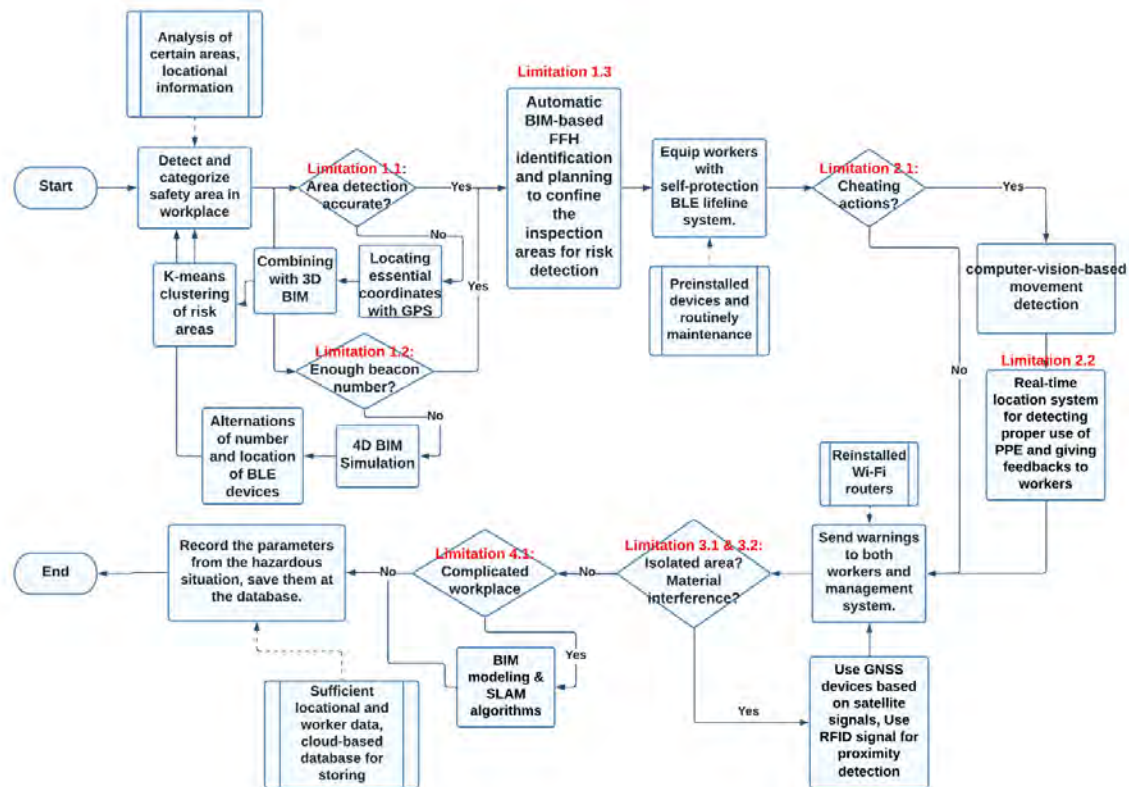


Figure 5 Optimized procedures of FFH accidents prevention

highways, railways, and tunnels, Wi-Fi signals are not always consistent, resulting in lag of communications or ineffectiveness of Wi-Fi relied on devices. Thus, GNSS devices based on satellite signals are especially useful under this circumstance.

- (3.2): If the BLE or Wi-Fi signals are blocked by metal materials such as steel beams and steel nets, the second RFID signal is needed for proximity detection, controlling the use of PPE in some hazardous areas.
- (4.1): If the working environment is complicated, by using BIM modeling and SLAM algorithms, the map and environment could be worker-based, meaning the localization is based on worker's position. The locational relationship between workers and workplace is adapting automatically.

After applying those potential solutions to the current system processing procedure, several alternations could be made (see Figure 5).

5 Conclusion and Future Research

This study identified and analyzed the casual factors and indicators of FFH accidents as well as the limitations of the current sensing system for preventing FFH

accidents by conducting a systematic literature review. Based on the analysis, this study also proposed a novel process flow that can support site managers' decision making for minimizing FFH accidents. The proposed process flow will serve as the foundation for future system development of FFH accident minimization software and cloud-based management platform. The findings will contribute to the body of knowledge by 1) enabling the construction industry to deeply understand FFH accident and how to prevent it and 2) providing a practical framework for optimized procedures of FFH accidents prevention.

Future studies will focus on the justification of the effectiveness for each solution, validating that the improved system can actually reduce the FFH accident rates on site. What's more, the realizability of different integration and implementation of technologies should be under further investigation, along with the guarantee of cyber security during the data transmission process. Last but not least, comparison studies for the cost of the system implementation between previous versions and the improved versions will be made to demonstrate that the proposed improved process flow not only satisfies the goal of accident rates reduction but also controls the total budget of the project.

References

- [1] Gibb, A., et al., *Construction accident causality: learning from different countries and differing consequences*. Construction Management and Economics, **32**(5):446-459, 2014.
- [2] Pinto, A., I.L. Nunes, and R.A. Ribeiro, *Occupational risk assessment in construction industry—Overview and reflection*. Safety science, **49**(5):616-624, 2011.
- [3] Im, H.-J., et al., *The characteristics of fatal occupational injuries in Korea's construction industry, 1997–2004*. Safety Science, **47**(8):1159-1162, 2009.
- [4] Dong, X.S., et al., *Fatal falls and PFAS use in the construction industry: Findings from the NIOSH FACE reports*. Accident Analysis & Prevention, **102**:136-143, 2017.
- [5] Liy, C.H., et al., *Causes of fall hazards in construction site management*. International Review of Management and Marketing, **6**(8):257-263, 2016.
- [6] Kelm, A., et al., *Mobile passive Radio Frequency Identification (RFID) portal for automated and rapid control of Personal Protective Equipment (PPE) on construction sites*. Automation in construction, **36**:38-52, 2013.
- [7] Teizer, J., D. Lao, and M. Sofer. *Rapid automated monitoring of construction site activities using ultra-wideband*. in *Proceedings of the 24th International Symposium on Automation and Robotics in Construction, Kochi, Kerala, India*. pages 19-21. 2007.
- [8] D'Arco, M., et al., *Enhancing workers safety in worksites through augmented GNSS sensors*. Measurement, **117**:144-152, 2018.
- [9] Gomez-de-Gabriel, J.M., et al., *Monitoring harness use in construction with BLE beacons*. Measurement, **131**:329-340, 2019.
- [10] Thomé, A.M.T., L.F. Scavarda, and A.J. Scavarda, *Conducting systematic literature review in operations management*. Production Planning & Control, **27**(5):408-420, 2016.
- [11] Mongeon, P. and A. Paul-Hus, *The journal coverage of Web of Science and Scopus: a comparative analysis*. Scientometrics, **106**(1):213-228, 2016.
- [12] Nadhim, E.A., et al., *Falls from height in the construction industry: A critical review of the scientific literature*. International journal of environmental research and public health, **13**(7):638, 2016.
- [13] Zermane, A., et al., *Analysis of the Contributing Factors for Fatal Accidents due to Falls from Heights in Malaysia and the USA*. Safety Engineering Interest Group, Department of Chemical and Environmental Engineering, 2020.
- [14] Rey-Merchán, M.d.C., et al., *Improving the prevention of fall from height on construction sites through the combination of technologies*. International journal of occupational safety and ergonomics:1-10, 2020.
- [15] Márquez-Sánchez, S., et al., *Intelligent Platform Based on Smart PPE for Safety in Workplaces*. Sensors, **21**(14):4652, 2021.
- [16] Costin, A., A. Wehle, and A. Adibfar, *Leading indicators—A conceptual IoT-based framework to produce active leading indicators for construction safety*. Safety, **5**(4):86, 2019.
- [17] del Carmen Rey-Merchán, M., A.L. Arquillos, and J.M. Soto-Hidalgo. *A Fall from Height prevention proposal for construction sites based on Fuzzy Markup Language, JFML and IoT solutions*. in *2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. pages 1-6, IEEE. 2021.
- [18] Standard, O., *MQTT version 3.1.1*. URL <http://docs.oasis-open.org/mqtt/mqtt/v3,1,2014>.
- [19] Abbasianjahromi, H. and E. Sohrab Ghazvini, *Developing a Wearable Device Based on IoT to Monitor the Use of Personal Protective Equipment in Construction Projects*. Iranian Journal of Science and Technology, Transactions of Civil Engineering:1-13, 2021.
- [20] Park, M., et al. *IoT-based safety recognition service for construction site*. in *2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN)*. pages 738-741, IEEE. 2019.
- [21] Dhanachandra, N., K. Manglem, and Y.J. Chanu, *Image segmentation using K-means clustering algorithm and subtractive clustering algorithm*. Procedia Computer Science, **54**:764-771, 2015.
- [22] Zhang, S., et al., *BIM-based fall hazard identification and prevention in construction safety planning*. Safety science, **72**:31-45, 2015.
- [23] Dong, S., H. Li, and Q. Yin, *Building information modeling in combination with real time location systems and sensors for safety performance enhancement*. Safety science, **102**:226-237, 2018.
- [24] Bailey, T. and H. Durrant-Whyte, *Simultaneous localization and mapping (SLAM): Part II*. IEEE robotics & automation magazine, **13**(3):108-117, 2006.
- [25] Zhang, S., et al., *Building information modeling (BIM) and safety: Automatic safety checking of construction models and schedules*. Automation in construction, **29**:183-195, 2013.

Construction Robotics Excellence Model: A framework to overcome existing barriers for the implementation of robotics in the construction industry

J.-I. Jäkel^a, S. Rahnama^a and K. Klemt-Albert^a

^a Institute for Construction Management, Digital Engineering and Robotics in Construction,
RWTH Aachen University, Germany

E-Mail: jaekel@icom.rwth-aachen.de, rahnama@icom.rwth-aachen.de, klemt-albert@icom.rwth-aachen.de

Abstract –

The construction industry is characterized by low productivity and inefficient processes. In addition, the level of digitalization and automation in this industry is low compared to other industries. Although the added values of using robots in the industry have been known for a long time and the technology has been available for several decades. Nevertheless, the use and implementation of robotic systems is still slow.

The aim of this work is to identify the barriers to the introduction of robots in the existing literature and to supplement them with an empirical study focused on the German construction industry, and subsequently to derive meta-barriers. In addition, a holistic model for overcoming barriers is developed

Based on a literature review, the already identified barriers for implementation in the construction industry are identified, clustered and five meta-barriers are derived. The literature review is complemented by an empirical study in the German construction industry. Based on the results, the Construction Robotics Excellence Model is presented. The model serves as a generic framework for overcoming existing barriers and promoting the implementation of robotics systems in the construction industry.

The results of the article show the versatility of the existing barriers in the construction industry and the need for a framework to support the implementation and use of robotics in the construction industry.

Keywords –

Robotics, Barriers, Construction Robotics Excellence Model

1 Introduction

The construction industry is a significant industrial sector in the world due to its large economic output and high

social relevance [1–3]. Nevertheless, the industry has been characterized by inefficiency and low productivity for many decades [4,5]. In addition, the industry still has a low level of digitalization and automation in direct comparison with other business sectors [6]. This is attributed to the limits reached by the construction industry [7].

In other industries, for example the automotive, manufacturing and aerospace industry, the added value of robotic systems has been known for several decades and has been successfully implemented in existing process structures [8]. In the construction industry, robotic systems have been developed since the 1960s [7,9,10]. Several fields of application [11] and added values have been identified [8,12], but the implementation of the technology is still progressing very slowly [7]. One reason for this are the specifics of the construction industry, such as the heterogeneous production environments and many unique processes in the construction projects, etc.. [6].

This article examines and identifies these barriers to the introduction of robots into the construction industry in a global and national context. As a result, a general framework, the Construction Robotics Excellence Model (ConRoX), for companies to overcome these existing barriers and introduce robots is derived.

2 Related Work

In addition to the possibilities and potential of robots in the construction industry, the barriers to the introduction and use of robot technology in this sector need to be considered. In a literature analysis, a total of nine scientific papers were identified that address existing barriers to the introduction of robots in the construction industry. The articles were analyzed, the identified barriers extracted and grouped into five higher-level clusters. These five clusters represent generic meta-barriers of the construction industry (s. Table 1). The

identified generic meta-barriers include social, economic and technical aspects. The most concise meta-barrier is the I. Adoption & Implementation with a total of 15 subordinate individual barriers. This meta-barrier refers to the complexity of the construction industry, its heterogeneous environment and the low level of digitalization and productivity. The second cluster (II. Skeptical Attitude) considered from the social perspective refers to the skeptical attitude of the construction industry and its fundamental resistance to innovation and new technologies and includes 12 individual barriers. Another meta-barrier (IV. Lack of knowledge) from the social perspective considers the lack of knowledge and expertise of the personnel about the subject area of robotics and automation as well as the lack of competences to deal with the technologies. A concise economic meta-barrier (III. High Costs) is derived from the still high acquisition, usage and maintenance costs for robots. From a technical perspective, meta-barrier V. No Standardization results from the lack of uniform standards at the organizational, process and information technology levels.

Table 1 List of identified meta-barriers

Barriers	Source	List of quant y
<i>I. Adoption & Implementation</i>		15
Difficulties of implementation in complex structures	[18]	
Resources limitation of the companies	[17]	
Lack of interoperability between organizational units and the general fragmentation of the construction industry	[17]	
To integrate the automation flexibly into the overall process from the start	[15]	
Traditional procurement methods that need to be adapted	[14]	
Changes are associated with risks and uncertainties	[14]	
Conflicts of interest when the contractor is added during the design	[14]	
Incompatibility with existing construction processes	[13]	
Inconsistency in the structure of the construction industry	[13]	
The complexity of the supply chain with different players in the implementation	[8]	
Unstructured nature of	[8]	

construction site	
Different requirements	[8]
characterize market diversity	[8]
Variability of building types	[8]
Limitations of adopting new technologies in the construction industry	[3]
Adoption inefficiencies and low productivity of robots	[3]
<i>II. Skeptical attitude</i>	
12	
Concerns of the workers	[19]
Basic rejection of new technologies of workers	[18]
The pronounced concern for safety	[17]
Resistance to change	[17]
The attitude of the management and team level	[16]
The emotional stress of replacing humans with robots	[15]
The disinterest of designers in disruptive technologies,	[14]
Resistance to new construction designs	[14]
Lack of worker acceptance	[13]
Skeptical attitudes of stakeholders towards innovative technologies	[8]
The expectation of a new technology	[3]
Concerns about the adoption of a new technology	[3]
<i>III. High costs</i>	
10	
High acquisition costs	[19]
High costs for acquisition, maintenance, and operation of robots	[18]
High investment and maintenance costs in the company	[15]
Automation is associated with high costs	[14]
Ownership and operation are associated with high costs	[14]
Cost intensity	[13]
High acquisition, operation, and maintenance costs	[13]
High cost of adoption	[3]
Declining public funding	[3]
Insufficient economic efficiency and associated profitability of introducing a robot to achieve	[3]

real cost reductions	
<i>IV. Lack of knowledge</i>	9
Adapting new process structures	[19]
Low level of competence in the use of technology	[17]
Access to technologies knowledge	[16]
Limited experience in automation	[14]
Limited practical experience of designers on site	[14]
Lack of knowledge of designers about construction methods in automation	[14]
Difficulty in handling	[13]
Low technological competence of project stakeholders	[13]
Unproven effectiveness and immaturity	[3]
<i>V. No standardization</i>	6
Lack of standardized construction elements	[14]
No standardized processes	[14]
Lack of references for the design	[14]
Rapid replacement and changes due to the high technological progress	[14]
The new roles and responsibilities that are emerging among planners	[14]
Difficulty in procuring robots	[13]

The results of the previously published articles on the barriers to the use of robot systems in the construction industry include many barriers from different perspectives. These perspectives were investigated, specified and clustered into generic meta-barriers. In the next chapter, the existing conclusions from the literature analysis are extended by an empirical study focusing on the German construction industry.

3 Research Design

3.1 General Methodology

In the previous chapter the general state of the art on the already identified barriers for the use of robotics in the construction is elaborated based on a literature review.

Based on the results, the general approach is set up with two consecutive tasks, (i) the empirical quantitative study and (ii) the derivation and presentation of the framework. In a first step, an empirical study will be

carried out as a quantitative survey with an inductive approach. The topic of the study focuses on the identification of challenges for the use of robotics in the German construction industry. Additionally, the organizational, process and information levels are considered from a strategic and operational perspective. The evaluation of the survey provides barriers for the implementation of robot systems in the German construction industry. The barriers are specified into the three categories - social, technical and economic. Based on the results, a general framework for the qualitative implementation and effective and efficient use of robotic systems is designed. This framework is termed the Construction Robotics Excellence (ConRoX) model. For a better understanding of the ConRoX model, the individual components are explained in more detail. This is followed by a discussion of the added value of the ConRoX model, its limitations, and its future potential for the construction industry.

At the end of the article, a critical review of the results is provided and an outlook on the resulting research is presented.

3.2 Structure of the empirical quantitative study

The structure of the empirical part is developed according to Diekmann and divided into the (a) data collection and (b) data evaluation. [13]

The primary question for the empirical study focuses on the barriers for the implementation of robotics in the German construction industry. The survey expands on the knowledge already gained from the preceding literature analysis.

The data collection method of the empirical study is conducted as an internet survey [13,14]. The survey is explicitly aimed at professionals in the construction industry. The conscious selection of the interviewed group is chosen with care and a total of three criteria that have to be fulfilled are defined as markers. The first criterion is several years of professional experience and/or an academic title of the respondent in the field of civil engineering as well as initial experience in robotics. The second criterion is the existence of the company's main business field in the German construction industry. The third and last criterion is that the respondent's field of activity must be in the construction execution phase.

For the experimental design of the Internet survey, the ex-post-facto design is chosen and conducted as a cross-sectional survey in a period of six weeks [15–17]. To obtain measurable and representative results, no limit is set on the number of participants. [13,18]. For the analysis of the collected data, an inductive approach using the statistical method of inferential statistics is provided. The method of inferential statistics for analyzing the survey data is used to derive general

statements about the barriers of the robot implementation in the German construction industry from the results [17,18].

The survey is structured according to Raithel and is implemented as a written, structured and standardized online survey. In the process, individual persons are questioned in a structured manner [15]. In total, the survey contains 26 questions with a mixture of closed-ended (73%) and hybrid (23%) and one open-ended questions (4%). The questions are divided into five main categories - General information of the respondent (category 1), the status quo of robotics in the respondent's company on the topic area (category 2), barriers (category 3) and drivers (category 4) of the implementation of robotics and the approaches to overcome the barriers (category 5) [19]. The main focus in this paper is on categories 2 and 3 due to only capturing the existing experience of German enterprises with robotics and elaborate the significant barriers. The other categories are not considered in this research paper.

In category 3, respondents were given 32 different hypotheses to choose from. These hypotheses were divided into three categories - social, technical, and economic - for further specification [13,14,19]. The basis of the 32 hypotheses findings from the knowledge is derived from the identified barriers from the previous literature analysis. To assess the importance of all predefined hypotheses, a unipolar ordinal scale with a total of six response options ranging from “disagree” (1) to “strongly agree” (6) is provided for the evaluation. The ranking of the most important barriers is made based on the cumulative sum (presented as a cumulative percentage) of the level of agreement S_z (S_z : sum of the positive responses of option 5 & 6). The barriers with the highest level of agreement are presented in Table 2 as survey category results.

4 Empirical Quantitative Study

4.1 Data Collection

The survey data was collected using a web survey over a period of six weeks, from August 16th, 2021 to September 27th, 2021. The respondents were limited to specialists and managers in the construction industry. The focus of the study was on the German construction sector. A total of 130 companies and associations were contacted. However, merely 65 questionnaires were completely answered. The overall response rate of the surveys is 50.0%. According to Diekmann, a response rate of over 20% are rarely achieved, tending to be around 5% [13]. For this reason, the existing response rate of 50.0% can be classified as representative.

4.2 Data Evaluation

Due to the focus in this scientific article, the following evaluation of the study will focus on category 3 (barriers to implementing robotics). In addition, Category 2 (status quo of robotics in German enterprises) is evaluated at the beginning of the evaluation. The other categories will not be considered within the scope of the research article.

4.2.1 Results: Experience with Robotics in German construction companies

This section evaluates the questions in Category 2 – the status quo of robotics in German enterprises. The result was that the majority of the respondents in total 73.7% have not yet gained knowledge about the use of robotics and automation technologies. In contrast, 26.3% of the respondents already gained experience using robots, either directly in the traditional in-situ execution (11.6%) or in industrial prefabrication (14.7%) (s. Figure 1). Next, the expected time horizon for the implementation of robotics and automation technology was asked. Only 5.30% are already planning implementation. For 23.2 %, the implementation is planned in 5 years, for 12.5 % in 10 years, for 14.3 % in 15 years and for 17.9 % not at all. construction companies

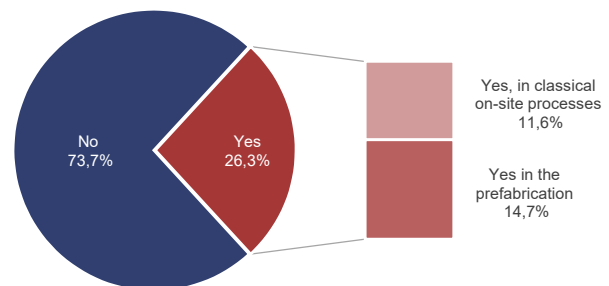


Figure 1 Use of robotic systems in German

4.2.2 Results: Barriers of robotic implementation

The ranking of the most important barriers (s. Table 2) is based on the cumulative sum (represented as cumulative percentage) of the level of agreement S_z (S_z : sum of positive responses of option 5 & 6). The barriers with a degree of agreement $S_z > 55$ % were characterized as main barriers after the evaluation. In total, 12 main barriers on a social, technical and economic level were identified from the 32 hypotheses with the corresponding level of agreement.

Seven barriers were identified in the economic area. Hence, this area presents the most barriers. The barriers relate on the one hand to monetary aspects, such as a lack of resources in the company, high acquisition costs and a

lack of incentives from the state. On the other hand, they address strategic issues, such as a missing consistent implementation strategy or a lack of best practice. The high number of economic barriers illustrates the importance of the field for future construction robotics. Above all, clear overarching strategies and financial support are needed in this context. On a social level, three barriers were identified. These are all in the context of a lack of expertise and training opportunities, as well as a lack of knowledge about the potential uses of robotics. Without increased knowledge of the technology, companies do not yet see the point of implementing it, as the long-term added value is not yet obvious. On the technical level, two barriers were identified. The respondents consider the dynamic and heterogeneous construction site environment and the lack of

standardization of processes to be the main obstacles.

In the comparison of the results of the literature analysis and the empirical study, it can be deduced that each identified individual barrier of the studies can also be classified in the meta-barriers. This strengthens the generality of the meta-barriers and its broad applicability.

The other 20 hypotheses from the survey consider, for example, aspects relating to the use of data in the construction industry, human-machine collaboration on the construction site, or a lack of acceptance and resistance among employees. These aspects were rated as manageable by the respondents.

Table 2 Identified barrier to construction robotics implementation.

Barriers	Results of the Evaluation [%]						S _z (5+6) [%]	Meta-Barrier
	1	2	3	4	5	6		
A. Social Level								
A.1 Lack of advanced training	1.0	8.0	6.0	10.0	40.0	36.0	76.0	IV.
A.2 Low level of employee expertise	1.0	9.0	7.0	11.0	27.0	45.0	72.0	IV
A.3 Lack of knowledge about possible applications slows down implementation	0.0	1.0	11.0	16.0	39.0	33.0	72.0	IV
B. Technical Level								
B.1 Lack of standardized processes	0.0	8.0	12.0	20.0	37.0	23.0	60.0	V.
B.2 Problems due to dynamic construction site environment	7.0	5.0	14.0	19.0	38.0	17.0	55.0	I.
C. Commercial Level								
C.1 Limited resources for SMEs	1.0	3.0	4.0	18.0	41.0	33.0	74.0	I.
C.2 High acquisition costs	0.0	1.0	13.0	20.0	27.0	39.0	66.0	III.
C.3 No consistent implementation strategy	0.0	0.0	12.0	27.0	35.0	26.0	61.0	V.
C.4 Lack of government support for the use of robotic systems	0.0	11.0	4.0	24.0	33.0	28.0	61.0	III.
C.5 Lack of skilled workers in the construction industry for the implementation	0.0	4.0	11.0	24.0	34.0	27.0	61.0	II.
C.6 Tight project timeline leaves little time to implement new technologies	7.0	4.0	4.0	24.0	29.0	32.0	61.0	I.
C.7 No existing best practice	0.0	1.0	11.0	31.0	34.0	23.0	57.0	V.

5 Construction Robotics Excellence Model

In order to compete in the increasingly dynamic and complex economy, continuous improvement of the company at the strategic and operational levels is essential [20–23]. This continuous optimization goes hand in hand with the pursuit of excellence of the company [22,24]. As a support and framework for companies to cope with complexities, adapt to constant changes as well as increase performance, a variety of

excellence models have been developed in the last decades for a wide range of sectors [25,26]. There are generic, cross-industry models of excellence, such as the Malcolm Baldrige National Quality Award (Baldrige), European Foundation for Quality Management (EFQM), Deming Prize Japan or the National Quality Award and Swedish Institute for Quality (SIQ) [22,27,28]. In addition, there also exist discipline-specific excellence models, such as the Service Excellence Model [29], the Marketing Excellence Model [30], the Sustainable Excellence Model [31], the Data Excellence Model [32] or the Construction Excellence Model [33] and many

more.

Derived from the twelve identified barriers in the literature review in conjunction with the defined meta-barriers from the literature review, the Construction Robotics Excellence Model is presented as a holistic solution approach. The Construction Robotics Excellence Model (ConRoX Model) is a generic framework to overcome the existing barriers and for the qualitative implementation of robot systems in the construction industry (see Figure 2). The ConRoX model considers all necessary levels - organization, processes and IT - of a construction company. Furthermore, the model shows which criteria are necessary as enablers for the use of the existing interface potentials between the robot systems and the business capabilities of the company.

The ConRoX model is divided into three interdependent areas - potentials, enablers and results. The first area of potentials serves to create knowledge about the robotic capabilities, to identify synergies with the company's own corporate capabilities and to identify fields of application. The step of identifying potentials results from the defined economic and social barriers of the empirical investigation. Thereby the social barrier A.3 is overcome, since basic knowledge is identified before an acquisition and/or employment of the robotics in the enterprise. At the same time, potentials are identified to support the minimum resources in the company (C.1) and the basis for an implementation strategy (C.3) and new required organizational and process structures (B1) are defined.

Table 3 Derivation of the enablers

Barriers from survey	Meta-Barrier	Resulting enabler
A2, B1, C1, C5	I.; V.	Organization
	II; V.	Roles & Responsibilities
B1, B2, C6	V.; III.	Processes & Methods
A1, A2, A3, C5	II; IV.	Training & Education
C4, B1, B2	I.; V.	Legal & Regulations
Manageable barriers according to survey results, yet an important technological component for the companies for robotic implementation	I.; V.	Data
	I.; V.	Application Systems

The second stage, “Enablers and Implementation”, lists seven relevant parameters for a successful implementation of robotic systems. These seven parameters include an implementation from different perspectives and cover all relevant company levels - organization, processes, data and application systems. Table 3 shows from which identified barriers the individual enablers were derived. The enablers support the overcoming of the barriers to the multi-beneficial use of robotics when applying the ConRoX-model.

All defined enablers result from the indexed barriers of the survey and the defined meta-barriers. Therefore, they are directly related to the barriers and consequently contribute to their direct elimination (see Table 2). In addition, the important technical component of data and application systems is included in the model. Thus, all essential aspects on the social, economic and technical level are considered in the model.

Furthermore, the ConRoX model refers to the definition of a company-wide robotics strategy as well as to the definition of concrete goals and use cases for the future use of robotics systems in the company. The parameters located in the inner circle of the model (Robotic strategy, Robotic Objectives and Robotic Use Cases) result from the barriers C.3-No consistent implementation strategy and C.7-No best practice. They provide the company with a guideline for the development of a company-wide robotics strategy and the definition of concrete objectives as well as use cases for the use of robotics in the company. Since the enabler level is a key component, the level is further integrated with a continuous improvement (CIP) approach. The third level of the model are the results of the successful use of the identified potentials (level 1), considering the enablers for implementation as well as the precise definition of strategy, goals and application fields. The results of the third level and the simultaneous successful implementation and use of robots in the company counteract the barriers C.2-High acquisition costs and C.7 n-o best practice. Through the positive results, new best practices are defined in the company, new knowledge is generated and the added value of the use of robots compared to the high acquisition costs becomes apparent. The intended result through the use of the model is the increase of the business value of the company and the achievement of excellence in the company in the economic (Business Excellence) and technical (Robotic Excellence) aspects. The foundation of the model is the company's general and permanent pursuit of excellence and continuous self-optimization.

6 Discussion

The aim of the article was to develop the Construction Robotics Excellence Model based on the previously

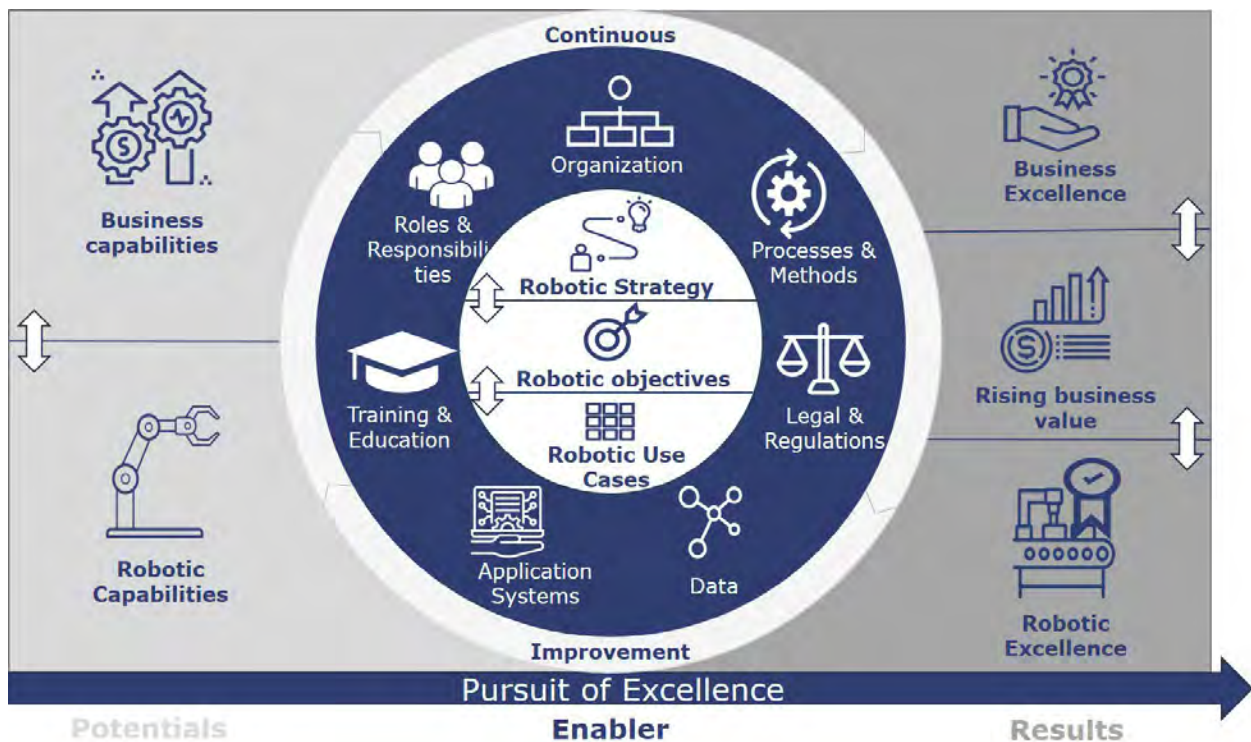


Figure 2 Construction Robotics Excellence Model (Own representation based on [32])

identified barriers and defined meta-barriers for the implementation of robotics in the German construction industry. As a result of the literature analysis, five meta-barriers were defined and enriched by the results of the survey with a total of twelve significant barriers on a social, technical and economic level. The barriers identified in the literature and the empirical study show the need for a generic framework for the implementation of robotics in the construction industry. This gap is addressed by the model. It serves as a guide for construction companies that want to implement robotics systems in their enterprise structures in the future. In addition, the model identifies seven essential enablers at the organizational, process and information levels in the company. These enablers serve a company before and during the implementation phase for the successful implementation of robotic systems and the achievement of resulting added values. The model closes a significant gap and minimizes existing obstacles and corporate reluctance to adopt robotics. In addition, the use of the ConRoX model will simplify the implementation of robotic systems and make it more efficient. In this regard, the model will also further promote interest in robotic systems in the construction industry. Likewise, the model will help to increase the level of automation in the industry as a whole, thereby increasing its productivity in the long term.

The ConRoX model is comprehensively described as a strategic framework for implementing robotics systems

in the construction industry based on identified barriers in this article, but it still requires consideration of limitations. For example, the survey only considered the German construction industry and placed it in context with existing findings from other literature. In addition, 65 people participated in the survey. For further specification of the model, the next step should be to obtain further expert opinions, for example through interviews. The model is a first generic approach to support the implementation of robotic systems in the construction industry. The ConRoX was developed based on the identified barriers in conjunction with knowledge of existing models of excellence from other industries. As a result, further scientific review and further specification of the model's content is needed. In addition, the model should be tested for practicality in defined, realistic use cases in collaboration with companies.

In future research, drivers for the implementation of robotics in the construction industry in a national as well as international context should be investigated in addition to the already identified challenges. Moreover, the holistic model should be checked for validation. Furthermore, an additional specification of the contents of the generic model is required. To ensure that the theoretical model can also serve for companies in practice in the future, its practical usability should be investigated in specific use cases with partners from the industry. In addition, research should be conducted to identify key performance indicators and application areas

for robotic systems in the construction industry.

7 Conclusion

This paper presents a holistic framework for the adoption of robotics in the construction industry derived from the identified barriers. At the beginning, a literature review is conducted to analyze the existing barriers to the adoption of robotics in the construction industry and the derivation of meta-barriers. This is followed by an empirical study to identify barriers that are specific to the German construction industry. As a result, a total of twelve barriers are identified. Based on the identified barriers the Construction Robotics Excellence Model is developed. The model serves as a generic framework for companies to implement robotic systems. The ConRoX model considers three different areas of robotics implementation - the potentials, the enablers and the results for qualitative implementation. Furthermore, the model presents the expected results of successful implementation at the end.

Despite the ConRoX model, further research activities in the field of construction robotics are needed in the future to explore the existing technical, economic and social problems and to further minimize the existing passive mindset of the industry. In this way, the added value of robotics technology can also find its way into the construction industry in the medium and long term, increase productivity and counteract the shortage of skilled workers

8 Reference

- [1] R. Bogue, What are the prospects for robots in the construction industry?, *IR* 45 (2018) 1–6. <https://doi.org/10.1108/IR-11-2017-0194>.
- [2] K. Hampson, J.A. Kraatz, A.X. Sanchez (Eds.), *R&D Investment and Impact in the Global Construction Industry*, Taylor and Francis, Hoboken, 2014.
- [3] J.M. Davila Delgado, L. Oyedele, A. Ajayi, L. Akanbi, O. Akinade, M. Bilal, H. Owolabi, Robotics and automated systems in construction: Understanding industry-specific challenges for adoption, *Journal of Building Engineering* 26 (2019) 100868. <https://doi.org/10.1016/j.jobbe.2019.100868>.
- [4] *Construction Industry Handbook 2012* [Research report], 2012.
- [5] T. Linner, *Automated and Robotic Construction: Integrated Automated Construction Sites*, Technical University Munich, Munich, 2013.
- [6] McKinsey Global Institute, *Reinventing Construction through a Productivity Revolution*, 2017. <https://www.mckinsey.com/business-functions/operations/our-insights/reinventing-construction-through-a-productivity-revolution>.
- [7] T. Bock, The future of construction automation: Technological disruption and the upcoming ubiquity of robotics, *Automation in Construction* 59 (2015) 113–121. <https://doi.org/10.1016/j.autcon.2015.07.022>.
- [8] G. Carra, A. Argiolas, A. Bellissima, M. Niccolini, M. Ragaglia, Robotics in the Construction Industry: State of the Art and Future Opportunities, in: *Proceedings of the 35th International Symposium on Automation and Robotics in Construction (ISARC)*, Taipei, Taiwan, International Association for Automation and Robotics in Construction (IAARC), 2018.
- [9] B. Chu, D. Kim, D. Hong, Robotic automation technologies in construction: A review (2008) 85–91.
- [10] K.S. Saidi, T. Bock, C. Georgoulas, Robotics in Construction, in: *Springer Handbook of Robotics*, 2016, pp. 1493–1520.
- [11] S.M.S. Elattar, Automation and robotics in construction: opportunities and challenges (2008) 21–26.
- [12] S. Martinez, C. Balaguer, A. Jardon, J. Navarro, A. Gimenez, C. Barcena (Eds.), *Robotized lean assembly in the building industry*, 2008.
- [13] A. Diekmann, *Empirische Sozialforschung - Grundlagen, Methoden, Anwendungen*, Rowohlt Taschenbuch Verlag, Renbek bei Hamburg, 2008.
- [14] R. Schnell, P.B. Hill, E. Esser, *Methoden der empirischen Sozialforschung*, tenth. Aufl., Oldenbourg, München, 2014.
- [15] J. Raithel, *Quantitative Forschung: Ein Praxiskurs*, VS Verlag für Sozialwissenschaften, Wiesbaden, 2006.
- [16] J. Bortz, N. Döring, *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler ; mit 87 Tabellen*, fourth., überarb. Aufl., [Nachdr.], Springer-Medizin-Verl., Heidelberg, 2006.
- [17] E. Raab-Steiner, M. Benesch, *Der Fragebogen: Von der Forschungsidee zur SPSS-Auswertung*, fourth., aktualisierte und überarb. Aufl., facultas.wuv; Facultas-Verl., Wien, 2015.

- [18] M. Eisend, Grundlagen empirischer Forschung: Zur Methodologie in der Betriebswirtschaftslehre, Springer Gabler, Wiesbaden, 2017.
- [19] H.-B. Brosius, Methoden der empirischen Kommunikationsforschung: Eine Einführung, VS Verlag für Sozialwissenschaften, Wiesbaden, 2016.
- [20] B. Enquist, M. Johnson, Å. Rönnbäck, The paradigm shift to Business Excellence 2.0, *International Journal of Quality and Service Sciences* 7 (2015) 321–333.
<https://doi.org/10.1108/IJQSS-03-2015-0032>.
- [21] J. Nenadál, D. Vykydal, D. Waloszek, Organizational Excellence: Approaches, Models and Their Use at Czech Organizations, *QIP Journal* 22 (2018) 47.
<https://doi.org/10.12776/QIP.V22I2.1129>.
- [22] A. Muhammad Din, M. Asif, M.U. Awan, G. Thomas, What makes excellence models excellent: a comparison of the American, European and Japanese models, *The TQM Journal* 33 (2021) 1143–1162.
<https://doi.org/10.1108/TQM-06-2020-0124>.
- [23] O.K. Al Zawati, O.Y. Hammad, I. Alsyouf (Eds.), Assessing organizations against excellence models: will subject matter experts add value to the assessment outcomes? An exploratory study, 2018.
- [24] P. Sampaio, P. Saraiva, A. Monteiro, A comparison and usage overview of business excellence models, *The TQM Journal* 24 (2012) 181–200.
<https://doi.org/10.1108/17542731211215125>.
- [25] R. Williams, B. Bertsch, A. van der Wiele, J. van Iwaarden, B. Dale, Self-Assessment Against Business Excellence Models: A Critique and Perspective, *Total Quality Management & Business Excellence* 17 (2006) 1287–1300.
<https://doi.org/10.1080/14783360600753737>.
- [26] S.-G. Toma, P. Marinescu, Business excellence models: a comparison, *Proceedings of the International Conference on Business Excellence* 12 (2018) 966–974. <https://doi.org/10.2478/picbe-2018-0086>.
- [27] B. Talwar, Comparative study of framework, criteria and criterion weighting of excellence models, *Measuring Business Excellence* 15 (2011) 49–65.
<https://doi.org/10.1108/13683041111113240>.
- [28] J. Gómez-Gómez, M. Martínez-Costa, Á.R. Martínez-Lorente, Weighting the dimensions in models of excellence – a critical review from a business perspective, *Measuring Business Excellence* 20 (2016) 79–90.
<https://doi.org/10.1108/MBE-01-2016-0007>.
- [29] M. Gouthier, A. Giese, C. Bartl, Service excellence models: a critical discussion and comparison, *Managing Service Quality* 22 (2012) 447–464.
<https://doi.org/10.1108/09604521211281378>.
- [30] C. Homburg, M. Theel, S. Hohenberg, Marketing Excellence: Nature, Measurement, and Investor Valuations, *Journal of Marketing* 84 (2020) 1–22.
<https://doi.org/10.1177/0022242920925517>.
- [31] S. Amir Bolboli, M. Reiche, A model for sustainable business excellence: implementation and the roadmap, *The TQM Journal* 25 (2013) 331–346.
<https://doi.org/10.1108/17542731311314845>.
- [32] T. Pentek, A capability reference model for strategic data management, University St.Gallen, St. Gallen, 2020.
- [33] H.A. Bassioni, T.M. Hassan, A. Price, Evaluation and analysis of criteria and sub - criteria of a construction excellence model, *Eng, Const and Arch Man* 15 (2008) 21 – 41.
<https://doi.org/10.1108/09699980810842043>.

State of Advances in Reality Capture for Construction Progress Monitoring and Documentation

J. Grogan^a, A. Sattineni^a and J. Kim^a

^aMcWhorter School of Building Science at Auburn University, Auburn, AL, USA

E-mail: jng0006@auburn.edu; sattian@auburn.edu; jsk0022@auburn.edu

Abstract

Construction documentation is necessary on every project. Information captured must be accurate, timely and actionable. With several technologies and techniques available, the aim of this study is to determine the current state of the industry on this topic. The objectives of this study are to investigate the status of construction documentation, understand different technologies efficiency and practicality for their deployment to monitor construction progress. A mixed methods approach used a survey and interview instrument to distribute to professionals within the construction industry. Data from surveys was analyzed and later validated with thematic analysis of interview data. Main findings include that a single documentation technique or technology is not available to fit all scenarios and solve all documentation problems. Most of the documentation is done on the job site by project team members that are there, including superintendents and project managers. Still imagery and video are being captured daily on almost every job site. It takes more technically trained professionals to deploy more advanced technologies like laser scanning and drones. The primary decision for when a technology is deployed is determined by the type of data to be collected or the problem to solved. Within each of these groups, the decision is made based on four key issues including cost, speed, data value, and training required.

Keywords –

Progress Monitoring; Reality Capture; As-Built Documentation; 360 photography; Drones; Laser Scanning

1 Introduction

Progress monitoring is a critical and challenging task required of a project manager. Inspections of construction progress are often manual and time intensive. Delivering a project on time and within budget requires active management of unique trades, all

operating on different timelines toward a common goal. Therefore, it is imperative to have constant progress data available to ensure a project is on schedule and within budget [1]. Progress monitoring data must be accessible, actionable, and easily comparable so that corrective steps can be taken when problems arise. Progress monitoring is ineffective if its review cannot result in accurate remedial action.

Capturing the data required comes in many formats. Several technological advances have moved the industry away from the original method of progress tracking of doing it manually and recording it on paper. These include Light Detection and Ranging (LiDAR), photogrammetry, static imagery, and radio-frequency identification (RFID). Research by Omar and Nedi [2] classify technologies used for progress monitoring data collection into four categories: Enhanced IT, Geo-spatial, Imaging, and Augmented Reality [2]. These disciplines have matured over the past several years through increases in accuracy, reduction in time, and decreased cost. However, there is not a consensus on what technologies solutions solve which specific construction problems [3].

As-built documentation is an important part of the construction ecosystem. Being able to document the location of key systems within a building with supplemental data is a huge asset for owners and facilities managers. Being able to communicate that information with all the stakeholders in a project is equally valuable. How this is best accomplished has not been agreed upon within the AEC industry. There is a lack of standardization in how this data is collected, stored, and distributed to various stakeholders. This is a result of the numerous variables used to determine how progress monitoring data is captured. The sheer quantity of data is also a hindrance to valuable data collection and dissemination. It is estimated that the average construction project will generate more than 400,000 images [4].

The ever-evolving nature of technology necessitates periodic review of best practices and recommendations from previous research. Sensors become more accurate, computers calculate faster, and storage has moved to the

cloud. The recent attention on technological solutions to the construction industry's problems has stimulated the growth of progress and improvement. Companies that specialize in reality capture are competing for better accuracy, range, and speed which is driving down costs. Leica, a prominent manufacturing of LiDAR technology, now sells a laser scanner for \$18,000, compared to an entry level model sold three years ago for £30,000 [1]. Technologies have the potential to be combined as they continue to mature. For example, the combination of small unmanned aerial systems (sUAS) and LiDAR. sUAS are currently being used for photogrammetry, because the photography industry is mature enough to combine high resolution in a small form factor. Research is being conducted on improving the navigation and autonomy of sUAS [5]. How all these technologies come together to solve problems in progress monitoring is not standard. The progression of technology and its integration into construction progress monitoring needs to be evaluated and understood.

2 Literature Review

Construction progress monitoring is a core component to managing any construction project. Knowing the progress of a project is integral to its success, much like managing quality standards and a budget. Construction progress monitoring gives project managers the ability to control schedule, budget, and unforeseen problems by comparing current conditions with a benchmark [1]. This task is often manual, involving multiple personnel from various disciplines. Not only does it require the attention of the general contractor, but also the design team and owner. All parties involved in a construction project require accurate and timely information on the status of the project to make informed decisions. How this data is collected and shared is constantly evolving due to the strides in advancing technology. Kopsida et al. [1] present a review of literature on the status of progress monitoring technologies by evaluating the following metrics: utility, time efficiency, accuracy, level of automation, required preparation, user's training requirements, cost, and mobility. They conclude that due to all metrics being considered and the complexity of construction projects, no general approach to construction progress monitoring can be recommended at this time [1].

Untimely detection of discrepancies between as-built conditions and as-designed plans are far too common on construction projects. Managing reliable progress data and integrating it into other project management systems is critical to remaining on schedule and avoiding delays. Unfortunately, traditional manual approaches do not integrate well [2]. Traditional methods are being replaced by various technologies, which are benefiting from the

move toward automation. Laser scanning and photogrammetry are promising for indoor applications, but due to the lack of automation in object recognition they do not meet the researcher's threshold for an ideal use case [2]. "The future of the construction industry is of a highly automated project management environment integrated across all phases of the project lifecycle" [2].

Categorizing the various progress tracking technologies allows for greater comparison and application for their potential problem-solving ability. These categories each have their strengths and weaknesses, but all lack in automation. Manual input is required at some level in each of these areas [2]. Further advances in holistic automation and integration with current construction monitoring systems is required.

Collecting data for progress monitoring creates the secondary problem of data storage and collaboration. Multiple independent systems are utilized on construction projects creating enormous amounts of digital data. For example, on a 20,000m² building project in Champaign, IL, the construction management team collects an average of 250 photos per day [6]. Storing and sharing data when the file count is in the hundreds of thousands become a problem in and of itself, separate from progress monitoring [4]. With several file formats and interested parties, sharing data that is easily organized and distributed is a problem. "The AEC industry does not have a comprehensive visual dataset" [4]. Finding a universally accepted format and organizational structure is paramount to continue down the path of automation in progress monitoring. Taking photos or video on a jobsite is essential, as well as the use of BIM. However, there is not a systematic way for integrating the two. Several factors were found causing this problem: image to BIM alignment, unordered images, and distortion. Han [4] also notes several issues with point cloud to BIM technologies trying to integrate big visual data, object detection and agent localization. Specifically, solving the perspective-n-point problem would allow for greater automation in processing visual data from multiple sources [4].

2.1 Automation

Part of automating progress monitoring is recognizing work activities. Difficulties in distinguishing various stages of work activities create problems when automating the process of assessment. "*It is difficult to differentiate between forming, placing, and back-filling a concrete foundation wall and inferring the current state of progress*" [7]. Recognizing what is happening in collected data and comparing to what was designed is still being developed. Yang et al. [7] discuss two methods, occupancy-based assessment and appearance-based method. The first method uses point clouds of as-built conditions and overlays the data on the design model to

detect discrepancies as well as utilize 4D BIM to track scheduling progress. The second method uses photogrammetry to create point clouds from images to identify materials, their quantities, and compare to the schedule [7].

Structure-from-Motion (SfM) is another technique used to automate progress monitoring. SfM uses as-built point cloud data and automatically registers photographs over the mesh model [6]. “SfM aims to reconstruct the unknown 3D scene structure and estimate camera positions and orientations from a set of image feature correspondences” [8]. This technique is key to the registration of photos to visualize progress compared to as-planned models. All of the methods presented by Golparvr-Fard et al. use this technique [8].

2.2 Challenges

One of the main obstacles to as-built documentation and progress monitoring is the accuracy of the data collectors. Laser scanning is becoming more precise and photogrammetry algorithms are processing point clouds with greater detail [4]. However, different scenarios require various levels of accuracy. Photogrammetry is not ideal for interior work [1]. Fine details that distinguish various levels of completeness of interior work are not easily determined from photogrammetry data. Laser scanning is better suited for high accuracy and range [4]. *“If the accuracy and quality level desired for a particular application is not high (i.e., error < 10 cm, and completeness rate > 80%), image-based methods constitute a good alternative for time-of-flight-based methods”* [3].

The second obstacle in effective progress monitoring is the lack of consistency in data collection. Not only does the accuracy of the data collection vary, so does the type and frequency of data collection. Individuals collecting the data must interpret in real time what needs to be captured, how it needs to be captured, and the way it should be presented [6]. Progress monitoring data can be collected through numerous types of equipment such as cellphone photography, LiDAR, 360 photography, sUAS, and written daily reports. All of these methods require extensive as-planned and as-built data extraction from construction drawings, schedules, and daily construction reports produced by superintendents, subcontractors, and trade foremen [6]. All of these variables combine to create a unique situation on every construction site based on the people involved, type of project, and equipment available. Not only is the data itself inconsistent, but so is the interpretation of the information. Data may be collected with the intent of highlighting one thing while it is interpreted in another way as it is distributed to various stakeholders.

The third obstacle in effective progress monitoring is the level of automation of a system. Scan to BIM level of

automation is currently being researched in the industry [9]. It is cumbersome and resource intensive to interpret visual data manually [10]. Bosché et al. [9] introduced a system to further automate the Scan-vs-BIM technique, specifically for MEP components. The original approach of Scan-vs-BIM can only recognize objects within 5 cm of their as-planned locations. The new approach proposed by Bosché et al. [9] is more accurate, and lends itself to higher degrees of confidence at greater distances. The second finding of the study is in regard to *“pipe completeness recognition”*. Recognition is defined as, *“the type of object can be discerned. More specifically here, this means that the analysis of the features enables discerning objects of a specific type (e.g. pipes with circular cross-sections)”* [9]. The new method proposed by the researchers can match cross-sections with greater degrees of confidence by matching at regular intervals of 10 cm. Other methods do not measure at regular intervals, which can lead to errors in detection and recognition. The new method performed better and validated elements at higher levels of confidence compared to the original method in identifying as-built MEP pipes to as-planned. Scan-vs-BIM or any proposed approach to compare as-built conditions vs a BIM model assumes a model that is continually updated with every architect’s supplemental instructions (ASIs), requests for information (RFIs), requests for proposal (RFPs), or change orders [6]. An incomplete model will sabotage any attempt at progress monitoring that relies on BIM to produce timely and informative decisions. Additional work is needed in this area to reliably interpret visual data in complex environments such as a construction jobsite [11].

2.3 Progress Monitoring

Progress monitoring of construction activities takes on many forms. Karsch et al., [10] conducted interviews of experts and describe a method of progress monitoring consisting of notes and photos taken on smart phones. Their participants describe this method of documentation “subjective” or “unreliable” [10]. With the prevalence of smartphones, it can be reasonably concluded that pictures and video in combination with email is the most common form of documentation and progress monitoring because of its familiarity and deployment among the general population. However, this does not mean this system is the most effective. Several interview participants commented the need for a much higher level of automation and analysis [10].

Image based systems have become the primary technology for progress monitoring of construction sites. Several LiDAR based methods have been proposed by previous researchers, but the most recent research reveals a turn toward image-based methods. The low cost of digital cameras and the implementation of high-resolution cameras on smartphones has enabled the

capture and sharing of construction photography to become more relied upon [6]. Prices on laser scanners will, drop yet they are unlikely to catch up with those on cameras soon since their manufactures do not respond to the competitive mass market of digital cameras [6]. In general, many recent methods for monitoring construction projects have moved to image-based techniques. Just like laser scanning techniques, image-based methods capture occlusions during the documentation stage. Static occlusions, which are a product of progress itself, and dynamic occlusions, which are the capture of moving objects such as people, vehicles, and equipment [6].

3 Methodology

The research methodology used for this study is mixed methods. A sequential explanatory design was chosen to further explain and interpret the quantitative data from survey via semi-structures interview based qualitative data. This approach is in two phases, first a survey instrument to collect quantitative data and then semi-structured interviews to collect qualitative data. Interview participants were chosen from those who participated in the survey in order to further investigate the motivations behind the survey results. Preliminary data for this research was collected through a literature review. Literature review main findings were used to develop a survey instrument. The survey was first distributed to a small group of five industry professionals as a pilot to receive feedback. The survey was then distributed by email with an anonymous link to industry Virtual Design and Construction (VDC) personnel to determine the deployment of specific technologies. This survey was web-based and included questions that were open-ended and scaled. A snowball method was used to distribute the survey to construction professionals. The survey did not include personal demographic questions to maintain confidentiality and limit any potential of matching responses with participants. Two definitions were given to participants for consistency in terminology, construction documentation - data captured in various formats that describes the current status of a construction activity and progress monitoring - the process of comparing current construction documentation data with past data to compare and identify the progression of construction activities. The survey consisted of thirty-two questions covering the themes of departmental demographics, technology specific questions, progress monitoring, value, and general remarks. Once survey data was analyzed, main findings were used to develop questions for a semi-structured interview format.

4 Results and Discussion

4.1 Quantitative Results

The survey portion of the mixed method study was answered by 56 respondents, 48 of whom answered all the questions. 53% of respondents were employed at companies larger than \$1.5 Billion in revenue per year. This indicates that a majority of the respondents worked in large companies. Survey participants showed a high level of experience with 32% indicating more than 10 years' experience and another 25% between 5 and 10 years of experience in the construction industry. Due to space restrictions, limited amount of data is presented here.

The importance of progress monitoring is underscored by responses to a survey question asking respondents about on how often they document it, as shown in Figure 1. The overwhelming majority of the participants, 28 out of 48 indicated that they updated the progress schedule on a daily basis.

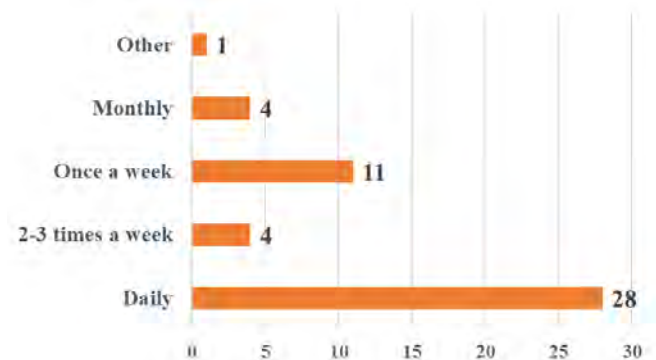


Figure 1. Frequency of Project Progress Monitoring

Survey participants were given the choice to select all of the technologies that they use to document construction project progress. This was mapped against how often they use them on a project. Not surprisingly, still imagery and video were chosen as the technology that is most commonly used on a daily basis, as shown in Figure 2. Those participants who noted they document daily, also reported that they use drone footage about half of the time. One respondent noted that they use laser scanning on every project, and document about 2-3 times per week.

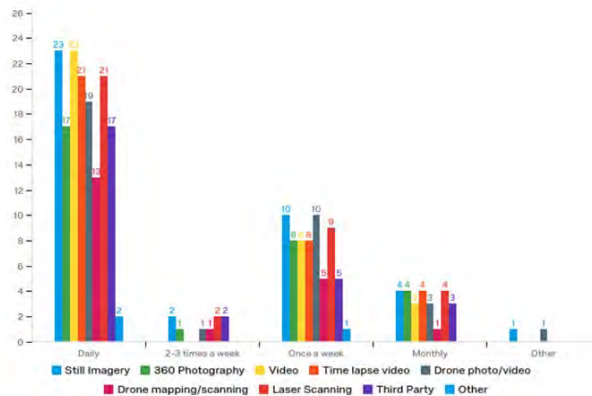


Figure 2. Frequency of project progress monitoring and technology tools used

Respondents were asked to rate how satisfied they were with the data collected by various methods to capture construction progress. The choices presented to the respondents were on a 5-point Likert scale ranging from 'Extremely Satisfied' to 'Extremely Dissatisfied'. The numbers of respondents who indicated 'Extremely Satisfied' about a particular technology/tool is shown in Figure 3. Results indicate that respondents felt that laser scanning and drone mapping provided the most accurate data to document construction progress monitoring. The results are in keeping with other researchers' findings that indicate project managers prefer a variety of ways to document visual project data [12].

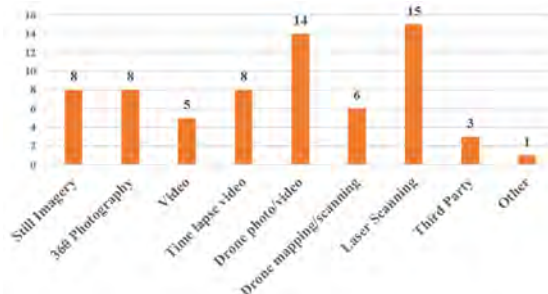


Figure 3. Respondent satisfaction with progress monitoring technology tools used

Respondents were further probed about what technologies they used to document construction progress, based on the scope of work being recorded. The technologies chosen by scope of work varied and no single technology available currently was identified by the respondents as the best solution for documenting all areas of a project. Concerning the documentation of site work, drone mapping/scanning was selected as the best technology for this scope of work.

The technology selected by participants for the documentation of structure was laser scanning, followed by 360-degree photography. Documenting the skin of a structure produced the choice of still imagery followed

by drone photo/video and then 360-degree photography. For interior framing, participants selected 360-degree photography as the best technology with 11 total responses and still imagery with 10 was a close second. For finishes, participants selected still imagery as the best technology with 17 responses followed by 360-degree photography with 8 responses.

Participants noted several different software platforms they use for monitoring progress of construction projects. The two mentioned most frequently were Procore and StructionSite. These platforms allow for seamless dissemination of information, especially media to various stakeholders. This is useful as the majority (91%) of participants commented that they do share documented activities with other project stakeholders. Participants did note that they have not received feedback from project stakeholders on a preferred format for documentation data, 58% vs. 42%. However, more participants chose that they do not use any form of documentation in lieu of in-person visits to the jobsite, 61% vs. 39%. Conversely, 73% of participants said they do use digitally documented data for inspections or verification of work performed, while 23% said they did not.

Survey participants ranked their preference on the technology that provided the best cost/benefit ratio. Only 2 responses separated the first three choices, still imagery, 360-degree photography, and laser scanning. 56% indicated that choosing a third party is the worst choice based on a cost/benefit ratio. Not surprisingly, 59% of survey participants chose still imagery as the technology that requires the minimum amount of effort to produce a finished product for documentation. Almost 50% of respondents said laser scanning provides the most valuable information for documentation see Figure 4.

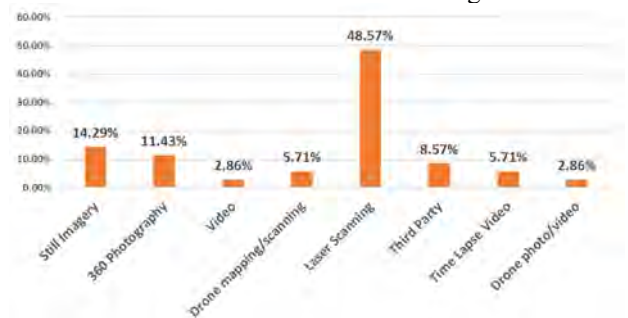


Figure 4. Most valuable information for documenting project progress

Participants were asked about which tool seemed to be the most promising one for future development to document project progress monitoring. The results indicate drone mapping, laser scanning and 360-degree photography are expected to get better and have the greatest potential to document construction progress in

future projects.

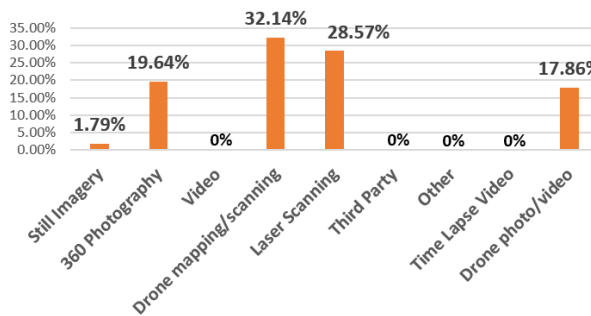


Figure 5. Most promising technology to document construction progress in future

4.2 Qualitative Results

Eight participants were selected from the researchers' established professional network, who had completed the survey to be interviewed. Research has shown that 8-interviews are appropriate to reach 80% saturation rate in qualitative interview based studies [13]. Interview questions were designed to explore in further detail the reasons behind certain choices being made by VDC personnel in the construction industry. No attempt was made to connect survey responses with interview participants to maintain confidentiality of survey results.

Interviewees responded to which currently is best for documenting construction by saying laser scanning and 360-degree photos. Laser scanning is the most comprehensive technology while 360-degree photos and videos are good for easily digestible data. One interviewee said, *"Laser scanning is best overall. But a lot of people aren't willing to pay, you know, \$30,000 for a laser scanner. The 360-degree photos are a quick and simple tool that is a lot cheaper. And you can have normal people look at the data that you get from it and they understand what it is."*

When interview participants answered what is lacking from current documentation technology, two themes that arose are training and education. Both of these are in regard to construction industry personnel and other project stakeholders. Interviewees commented that while the technology could be better, there is slow adoption and training on how to implement the technology within the construction industry. They also commented that project stakeholders are unaware of the capabilities of current technology and how to leverage the data collected to improve their projects.

When asked about waste, three themes were discovered, time, money, and labor. Time is being wasted in processing captured data, especially laser scan data. Post processing and point cloud registration is consuming value time and requires skilled and trained professionals

to handle and interpret the data. Trained personnel cost money to imbed on site with a project team, or to travel to and from the project to collect and process data

Several problems are able to be solved through technology available for documentation. Primarily, referenced by interviewees, were work verification and as-built point clouds. Work verification is primarily being done with 360-degree cameras as a fast way to capture the status of a scope of work for quality assurance and quality control (QA/QC) purposes. One interviewee even mentioned using the information capture for determining liability of a problem. Several interviewees mentioned using laser scanning for verifying slab penetrations before concrete is poured. They compare scan data with models previously built to verify the work is correct. Another use of laser scanners mentioned by participants is in scanning and verify overhead mechanical, electrical, and plumbing (MEP) rough-ins. One interviewee said they specifically use laser scanning to verify overhead MEP on their large healthcare projects. Another example of problem-solving laser scanners are being used for is documenting as-built conditions for renovations. Several interviewees mentioned using laser scanners to document buildings they were doing renovations on. However, they were able to successfully leverage that information, but other project stakeholders they shared the scan data with were less capable.

Interview participants gave several reasons why they think project stakeholders do not utilize documentation data. The primary reasons given were a poor understanding of the technology and poor communication between project stakeholders about the technology and data. While tech savvy project stakeholders are able to leverage the data, the typical stakeholder cannot. This leads to a lack of understanding about what the technology is and does, but importantly the value it brings to justify its cost. Poor communication between stakeholders also leads to them not leveraging the data collected. One interviewee mentioned he has worked with very few architects who can take laser scan data he provides from a renovation project and model off of the data. Another told the story of an owner who hired a third-party laser scan the entire project several times, but through conversations with the owner, he realized the owner did not fully understand why he was having the building laser scanned.

The two major improvements described by interview participants about how to make laser scanning more accessible are cost reduction and time reduction. Every interviewee mentioned the high cost of laser scanning. The equipment is expensive as is the training and time required to use it properly. This was described as the main hurdle of implementing laser scanning on smaller projects because it simply costs too much.

When describing what could improve the progress

monitoring process, interviewees spoke about data integration and training. Data is currently fragmented throughout different formats and platforms that it is difficult to digest the information, especially when it is not in one place. One interview participant described a scenario where subcontractors and trades were better involved in the documentation process to better capture data. He said having trades document their own work would enable a less adversarial relationship between the subcontractor and general contractor while also enabling better information sharing. Another interviewee spoke about lowering the burden on technically trained people to perform the data collection by training others to perform the work.

Interview participants had several ideas on where they wanted the documentation process to be in the next five years. The primary responses included better automation, greater speed of data capture, better data integration, and greater jobsite connectivity to enable all of these advances.

All technologies are currently being utilized on construction projects, often multiple technologies on the same project. There is not a one size fits all solution for construction documentation technologies. Projects are being documented on a daily or multiple time per week basis with technologies that are the most accessible. Superintendents and project managers or combination thereof are the most common project team member to document construction. Therefore, less technically trained team members are using the technology they are most familiar with and comfortable using, their phones to capture still imagery and video. This leads to a disconnect between the cost of more advanced capture methods and the value of the data they provide. Technically trained team members are documenting projects that require technical solutions provided by more advanced equipment. Laser scanning and drone mapping/scanning/photo/video are being performed less regularly by highly trained professionals because high effort is required to plan, perform, and post process data collection with these methods.

Low acceptance and deployment of highly technical equipment is more common with smaller companies. The hurdles of cost and low understanding of value from project stakeholders is limiting deployment with all company sizes. Simpler technologies are more efficiency because they require less training, the equipment is less sensitive to the job site, and they are cheaper. Practicality of highly technical equipment has not reached a sustainable level due to the lack of understanding of the value of those technologies and the lack of data integration.

A graphic is proposed to better understand the situation of numerous documentation technologies available, as shown in Figure 6. Technologies are initially

split into two groups, status/progress and technical solutions. These categories inform the reader based on the type of data they are attempting to collect. This leads the user down the path to technologies that can satisfy their data needs while also informing them of the effort required and quality of data provided by each technology. The graphic also shows the technology best suited for their scope of work, as indicated by the participants of the study. By working from the outside in, a scope of work is selected which leads to the appropriate technology determined by industry survey and interview data.



Figure 6: Summary of thematic analysis of interview data

5 Conclusions and Recommendations

Progress monitoring and construction documentation are vital to every project. Monitoring progress to update budgets and schedules help project stay on track. Successful documentation must be accurate, timely, and actionable. Unfortunately, no single documentation technique or technology fits all scenarios and solves all problems. Each technology has its strengths and weaknesses. The majority of documentation is done on the job site by project team members that are there, superintendents and project managers. They often reach for the most accessible technology that is comfortable for them, their phones. Still imagery and video are being captured on a daily basis on almost every job site. It takes more technically trained professionals to deploy more advanced technologies like laser scanning and drones. The primary decision for when a technology is deployed is determined by the type of data to be collected, technical solutions or status/progress. Within each of these groups, the decision is made based on four key issues including cost, speed, data value, and training required. Smaller companies trend to not use the most advanced and expensive technologies compared to large.

Further investigation is recommended for the integration of construction documentation data. Data is currently fragmented based on data type and capture

method. While individual platforms have reduced the burden of access for the non-technically trained professions, there is not consensus or common access. A type of cloud-based dashboard to host different data formats with one common access would be the goal.

Several platforms and devices are presently being developed and deployed in the construction industry for 360-degree photography. The potential value of data collected in 360-degree photography might not fully be realized until the technology has had more time to mature, but data captured now could be mined in the future. How this data could be applied to construction documentation needs to be further investigated.

Further investigation into why there is a lack of buy in and understanding of documentation by the architect and engineering community is needed to identify the problem and develop a strategy for better include this group with the valuable information being collected. The data being collected is valuable, but if it is not easily used by project stakeholders its costs cannot be justified. Overcoming this hurdle would be a step forward in standardizing the technology and techniques used to document construction projects. Finally, further investigation into computer vision and machine learning is needed to address the capability of autonomously identifying construction material, equipment, and activities. This area of research is being advanced in other industries and needs to be applied to construction. With greater processing power and artificial intelligence, identifying problem areas and scopes of work that are behind schedule autonomously will become the future.

References

- [1] Kopsida, M., Brilakis, I., & Vela, P. A. (2015). A Review of Automated Construction Progress Monitoring and Inspection Methods. *Proceedings of the 32nd CIB W78 Conference 2015*, 421–431. Retrieved from <http://itc.scix.net/cgi-bin/works/Show?w78-2015-paper-044>
- [2] Omar, T., & Nehdi, M. L. (2016). Data acquisition technologies for construction progress tracking. *Automation in Construction*, 70, 143–155. <https://doi.org/10.1016/j.autcon.2016.06.016>
- [3] Rebolj, D., Pučko, Z., Babič, N. Č., Bizjak, M., & Mongus, D. (2017). Point cloud quality requirements for Scan-vs-BIM based automated construction progress monitoring. *Automation in Construction*, 84, 323–334. <https://doi.org/10.1016/j.autcon.2017.09.021>
- [4] Han, K. K., & Golparvar-Fard, M. (2017). Potential of big visual data and building information modeling for construction performance analytics: An exploratory study. *Automation in Construction*, 73, 184–198. <https://doi.org/10.1016/j.autcon.2016.11.004>
- [5] Asadi, K., Ramshankar, H., Pullagurla, H., Bhandare, A., Shanbhag, S., Mehta, P., ... Wu, T. (2018). Vision-based integrated mobile robotic system for real-time applications in construction. *Automation in Construction*, 96, 470–482. <https://doi.org/10.1016/j.autcon.2018.10.009>
- [6] Golparvar-Fard Mani, Peña-Mora Feniosky, & Savarese Silvio. (2015). Automated Progress Monitoring Using Unordered Daily Construction Photographs and IFC-Based Building Information Models. *Journal of Computing in Civil Engineering*, 29(1), 04014025. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000205](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000205)
- [7] Yang, J., Park, M.-W., Vela, P. A., & Golparvar-Fard, M. (2015). Construction performance monitoring via still images, time-lapse photos, and video streams: Now, tomorrow, and the future. *Advanced Engineering Informatics*, 29(2), 211–224. <https://doi.org/10.1016/j.aei.2015.01.011>
- [8] Golparvar-Fard Mani, Savarese Silvio, & Peña-Mora Feniosky. (2009). Interactive Visual Construction Progress Monitoring with D4 AR ? 4D Augmented Reality ? Models. *Building a Sustainable Future*. [https://doi.org/10.1061/41020\(339\)5](https://doi.org/10.1061/41020(339)5)
- [9] Bosché, F., Ahmed, M., Turkan, Y., Haas, C. T., & Haas, R. (2015). The value of integrating Scan-to-BIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: The case of cylindrical MEP components. *Automation in Construction*, 49, 201–213. <https://doi.org/10.1016/j.autcon.2014.05.014>
- [10] Karsch, K., Golparvar-Fard, M., Forsyth, D., 2014. ConstructAide: analyzing and visualizing construction sites through photographs and building models. *ACM Trans. Graph.* 33, 176:1-176:11. <https://doi.org/10.1145/2661229.2661256>
- [11] Nath, N.D., Behzadan, A.H., 2020. Deep Convolutional Networks for Construction Object Detection Under Different Visual Conditions. *Frontiers in Built Environment* 6.
- [12] Pal, A., Hsieh, S.-H., 2021. Deep-learning-based visual data analytics for smart construction management. *Automation in Construction* 131, 103892. <https://doi.org/10.1016/j.autcon.2021.103892>
- Namey, E., Guest, G., McKenna, K., Chen, M., 2016. Evaluating Bang for the Buck: A Cost-Effectiveness Comparison Between Individual Interviews and Focus Groups Based on Thematic Saturation Levels. *American Journal of Evaluation* 37, 425–440. <https://doi.org/10.1177/1098214016630406>

An Approach for Estimation of Swing Angle and Digging Depth During Excavation Operation

A. Molaei^{1,2}, M. Geimer¹ and A. Kolu²

¹Institute of Mobile Machines, Karlsruhe Institute of Technology, Karlsruhe, Germany

²Radical Innovation Research Group, Novatron Ltd., Pirkkala, Finland

amirmasoud.molaei@partner.kit.edu, marcus.geimer@kit.edu, antti.kolu@novatron.fi

Abstract -

This paper discusses the estimation of the swing angle and digging depth during the excavation operation. The ability to calculate the excavator's productivity is an essential step toward autonomous excavators. The swing angle and digging depth have significant effects on the excavator's productivity and must be taken into account for the productivity estimation. Two approaches are proposed to estimate these variables. The first method estimates the swing angle using cabin encoder measurements. The local minimum and maximum points are found, and then Otsu's method is exploited to detect the points that are representative of scooping and dumping positions. The second method utilizes the bucket position to estimate the digging depth. The bucket position is calculated using Inertial Measurement Units (IMUs) measurements and the forward kinematics of the excavator. Otsu's method is used to distinguish the local minimum points that are representative of the digging depth during the operation. Moreover, the algorithms are computationally efficient. Finally, the performance of the proposed methods is studied using real measurements. The results show that the methods can effectively estimate the swing angle and digging depth under different working conditions such as various materials, swing angles, and digging depths.

Keywords -

Excavator; Swing Angle; Digging Depth; Productivity; Otsu's method

1 Introduction

Heavy-duty mobile machines (HDMMs) play significant roles in the world and are exploited in many fields such as construction, forestry, and mining industries. These industries are one of the highly increasing industries and have significant challenges such as lack of skilled human operators, high productivity, harsh environment, and safety [1]. It has been analyzed that the 5% to 10% of direct costs in building projects and up to 40% of direct costs in highway construction projects are related to equipment costs [2]. These industries are highly competitive, and companies must try to improve their products to remain

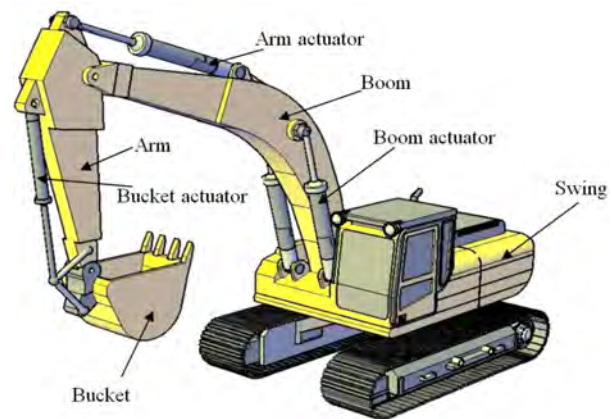


Figure 1. A typical hydraulic excavator and its different parts [4].

in business against other competitors. In order to reduce costs, the effects of the challenges, and also to increase the productivity of mobile machines, autonomous HDMMs are one main solution. Proposing one autonomous method for all HDMMs is highly complicated since there are a lot of mobile machines in different sizes, shapes, and functions [3].

1.1 Excavator's Productivity

There are different types of HDMMs, and the hydraulic excavator is one of the most used machines in this field. The excavator is a human-operated machine that is mostly driven by using a hydraulic system. Fig. 1 shows a typical hydraulic excavator. The excavator is one of the primary earth-moving machines in various construction projects such as the construction of highways, airports, industrial and residential buildings. Almost all construction projects require various types of excavation work [5]. An excavator is a multi-functional machine that can easily do different tasks such as dig & dump, trenching and leveling cycles, and utilize different tools. The traveling body, swing body, and front digging manipulator are three main parts of the hydraulic excavator. The manipulator consists of three

links: boom, arm, and bucket. The links are manipulated by hydraulic cylinders. Also, the excavator has revolute joints between the swing body, boom, arm, and bucket [6].

The ability to calculate the productivity of hydraulic excavators during different operations can be an essential step toward autonomous excavators. Monitoring the productivity of excavators can reduce the operation time, fuel consumption, and optimize the planning and working parameters. Also, the estimation of the excavator's productivity has major effects on the management and economic aspects. The performance of excavators highly depends on the skills of the human operator, therefore a method for productivity monitoring is significantly required. In addition, human operators can improve their skills by using feedback from the excavator's productivity. Furthermore, since the excavator has repetitive duty cycles, a slight improvement in the operation cycle time or fuel efficiency can bring about huge improvements in the overall performance [3].

Generally, the quantity of material and the operation cycle time are the main factors for the productivity of most cyclical types of machinery. The excavator's productivity means the quantity of transferred material per unit of time. The quantity can be the weight or volume of material. This is the simplest definition of the excavator's productivity. There are different parameters and working conditions such as the swing angle, digging depth, size of the excavator, bucket capacity, dumping conditions, type of materials, weather conditions, and operator's skill that can significantly increase or decrease the productivity of the excavators [6]. Dig & dump duty cycle is one of the most important tasks in different construction projects. This duty cycle consists of four main sub-tasks: 1) digging, 2) swinging loaded, 3) dumping, and 4) swinging empty. Digging depth and type of material are two of the essential factors in the digging sub-task. When the soil becomes harder or the location of material gets deeper, it takes longer to fill the bucket. Moreover, the swing angle is another variable that can increase or decrease the time of swinging loaded/empty and subsequently the overall cycle time [7, 8]. Also, the cycle time is highly dependent on the machine's size because small machines can cycle faster than large machines. Another challenge is continuous variations in environmental and load conditions. These variations can substantially change the productivity of excavators. Furthermore, the cycle time can be influenced by dumping conditions. There are different dumping conditions such as trucks in various sizes, and large or small dump targets [9].

1.2 Literature Review

Publications show the prediction of productivity is done by analyzing the effects of different parameters. Thereby

the use of special datasets and methods of companies' handbooks is common. In [10], the authors proposed a deterministic multiple regression model to predict the excavator's cycle time as a measure of productivity. The machine's weight, swing angle, and digging depth are inputs or predictor variables in the regression model. The dataset was obtained from companies' performance handbooks. In [11], the authors propose an artificial neural network by using the same dataset of [10]. The proposed ANN has a higher performance than the MR model in [10]. In [12], an artificial neural network combined with queuing theory is designed to predict the productivity of earthwork machinery including several excavators and haulers. Because of the lack of real data, a computer simulation is utilized to generate data. In [13], the operator competence is presented as a modifying factor in the productivity estimation. The authors model the operator competence and then analyze the effects of this variable on the productivity estimation. In [14], a deep neural network (DNN) model is presented to predict the productivity of excavators by using telematics data. Deep neural networks require a huge dataset and have a high computational load. These challenges can limit the efficiency and practicability of the model. In [15], different deterministic productivity estimation methodologies are introduced and compared with each other. It has been studied that the productivity of the excavator is highly influenced by the swing angle and digging depth. The methods shown only try to approximately predict the operation cycle time, and subsequently, the excavator's productivity and cannot calculate the productivity in real time. These methods are very dependent on datasets and cannot be used for different machines, job conditions, and operators. The most significant challenge is that these methods cannot estimate the swing angle and digging depth, and these parameters should be anticipated by managers or operators at the beginning of operations. During the operation, the swing angle and digging depth change and the assumption of constant values for these variables cannot be a correct solution. Furthermore, these methods cannot be easily utilized for different duty cycles such as trenching.

1.3 Objectives

The focus of this paper is to propose novel frameworks to estimate the swing angle and digging depth during the dig & dump duty cycle. There are several methods to calculate the excavator's productivity, and all of them highly depend on the swing angle and digging depth. These variables can significantly affect the operation cycle time and subsequently the productivity of the excavator. In conventional methods, managers consider only constant values as the swing angle and digging depth at the beginning of operations, but in the proposed algorithm, these variables

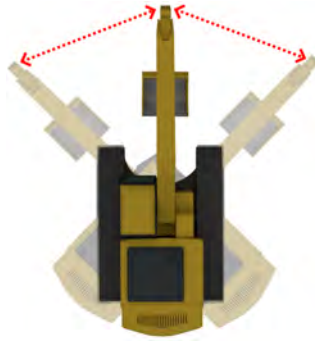


Figure 2. A visualization of swinging movements.

are not considered constant values. In the paper, the swing angle and digging depth are estimated based on the measurements from the excavator and are updated during the operation. The presented algorithms are computationally efficient and use common sensors such as the incremental encoder and Inertial Measurement Units (IMUs) that are affordable and can be easily installed on different excavators. Also, the method can be easily utilized for different duty cycles and also extended for other heavy-duty machines. Currently, there is no automated algorithm for the swing angle and digging depth estimation in commercial automated machine guidance systems. The proposed method can be an interesting feature for the new generation of excavators' automated machine guidance systems. The information about the swing angle, digging depth, and productivity estimation can be utilized as feedback to analyze and improve the skill of human operators in machine guidance systems. Furthermore, the excavator's productivity can be used for the optimization of worksites.

This paper is organized as follows: the methods to estimate the swing angle and digging depth are introduced in Section 2. Section 3 briefly describes the collected datasets and measurements in the experiments. The results of the methods are explained in Section 4. Finally, Section 5 concludes the paper.

2 Methodology

The estimation of swing angle and digging depth is certainly required to calculate and analyze the excavator's productivity during different operations and conditions. In Sections 2.1 and 2.2, two methods are proposed to estimate the swing angle and digging depth, respectively. The proposed methods use measurements from common sensors in the excavator and also are computationally efficient.

2.1 Swing Angle

The swing angle can significantly affect the operation cycle time during the dig & dump duty cycle. In this paper,

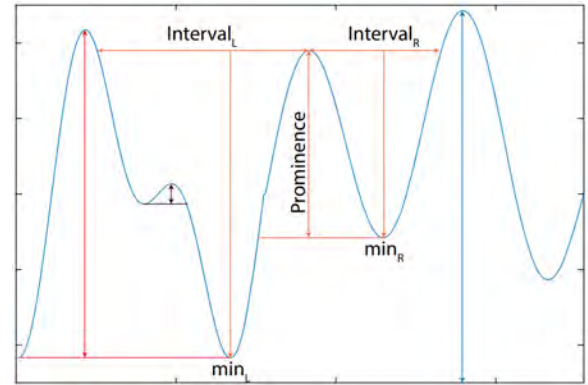


Figure 3. The prominence value [16]

a novel framework is proposed to estimate and update the swing angle by using the measurement of the cabin encoder. The swinging movements are shown in Fig. 2. Measurements of the cabin encoder during the previous T seconds are considered as input data in this algorithm. The length of the input vector is equal to $T \times f_s$, where f_s is the sampling frequency of measurements.

Firstly, a moving average filter is used to reduce the effects of noises and sudden movements and variations. In this filter, each element of the output is computed by using an equal number of input data on either side of the central value. Actually, the number of samples in one sliding time window is equal to $T_{filter} \times f_s$, where T_{filter} is the length of the sliding time window.

Secondly, local minimum and maximum points are detected to specify the scooping and dumping positions. A prominence value is defined for each local minimum or maximum point. In fact, the prominence of a local minimum point (or a valley) determines its depth compared to other local minimum points. To calculate the prominence of a local minimum point, a horizontal line from the local minimum point is extended to the left and right of the point. Where the horizontal line intersects the data can be another local minimum point or the end of the data. The intersections are outer end-points of the left and right intervals. In the next step, the highest peaks in both the left and right intervals are found, and only the smaller peak is considered. The vertical distance between the local minimum point and the peak is called the prominence value. Also, there is a similar definition for the prominence of local maximum points. The prominence of a local maximum point (or a peak) specifies the height of the point with respect to the other local maximum points. To calculate the prominence of a local maximum point, firstly a horizontal line from the local maximum point is extended to the left and right of the point. The intersections of the line with data can be another peak or the end of data. The intersections specify outer end-points of the left and right

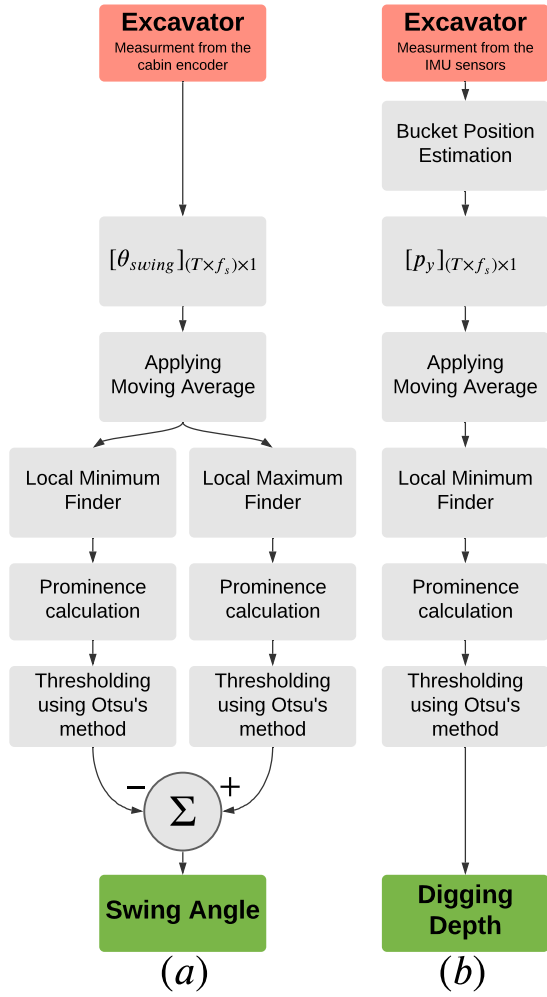


Figure 4. The flowcharts of the proposed methods: (a) swing angle estimation, (b) digging depth estimation.

intervals. After that, the lowest valleys in both the left and right intervals are detected, and only the larger valley is taken into account. The prominence is defined as the vertical distance between the valley and the local maximum point [17]. Fig. 3 shows an example for the prominence calculation of a local maximum point. Firstly, a horizontal line from the local maximum point is extended to the left and right of the peak. The left interval lies between the peak and crossing due to another peak, and the right interval lies between the peak and crossing due to another peak. The lowest points on the left and right intervals are shown by min_L and min_R , respectively. The reference level (highest minimum) is min_R . The prominence is the vertical distance between the reference level and the local maximum point.

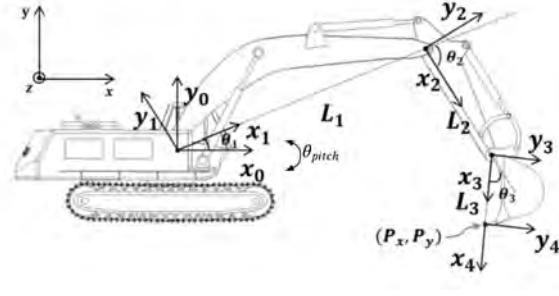


Figure 5. The forward kinematics of the excavator [19].

Probably, all local minimum or maximum points are not acceptable and cannot be considered as the scooping or dumping positions. Otsu's method is exploited to automatically distinguish the valid local extremum points. Otsu's method is an optimum thresholding method by maximizing the variance between classes. This method is mainly used for image segmentation [18]. Finally, Otsu's method diagnoses the valid local extremum points that are representative of the scooping and dumping positions based on their prominence values. The swing angle is defined as the difference between the minimum and maximum angles. The flowchart of the proposed method is presented in Fig. 4.

2.2 Digging Depth

The digging depth is another essential parameter that must be taken into account for the productivity analysis. In this paper, the digging depth is estimated based on the bucket position. The position of the bucket is calculated by using the forward kinematics of the excavator and measurements from four Inertial Measurement Units (IMUs) that were installed on the different parts of the excavator such as the swing body, boom, arm, and bucket. In this part, the swing of the cabin is not considered since it does not have any effect on the digging depth. The axis and frame of forward kinematics of the excavator based on the two-dimensional space is provided in Fig. 5. The end-point position of the excavator is calculated by using the following equations

$$P_x = L_1 \cos(\theta_{pitch} + \theta_1) + L_2 \cos(\theta_{pitch} + \theta_1 + \theta_2) + L_3 \cos(\theta_{pitch} + \theta_1 + \theta_2 + \theta_3) \quad (1)$$

$$P_y = L_1 \sin(\theta_{pitch} + \theta_1) + L_2 \sin(\theta_{pitch} + \theta_1 + \theta_2) + L_3 \sin(\theta_{pitch} + \theta_1 + \theta_2 + \theta_3) \quad (2)$$

where P_x and P_y are x and y -components of the bucket position, respectively.



Figure 6. The excavator used in the data collection phase. In the picture the cabin (1.), boom (2.), arm (3.) and bucket (4.) are highlighted with red boxes.

The estimation of digging depth is performed by using the y -component of the bucket position. This algorithm is similar to the swing angle estimation approach introduced in Section 2.1. In each iteration, the length of input data is equal to $T \times f_s$. Firstly, a moving average filter is utilized to reduce the effects of noises and meaningless movements, and variations of the bucket. Secondly, to find the depth of digging, local minimum points are detected. There are a lot of local minimum points that are not representative of actual digging depth. Otsu's method is applied to the prominence values of local minimum points to find the valid local minimum points. The flowchart of the presented method is shown in Fig. 4.

3 Data Collection

In this paper, the dataset was collected from a Komatsu PC138US excavator. The crawler excavator used in the experiments is shown in Fig. 6. Although this excavator is old, it has been well-maintained, and it is in good condition. The inspection and maintenance are performed every 500 working hours. The operating weight of this medium-rated excavator is 14000Kg and has a standard mono boom, arm, and bucket. The bucket is attached to the arm by using a quick coupler, and also, the excavator has a tiltrotator. The heaped capacity of the bucket is $0.37m^3$ based on the standard of the Society of Automotive Engineers (SAE). The MATHWORKS SIMULINK model was used to collect data from the excavator. Measurements of different sensors are transmitted over the controller area network (CAN) bus. The model is connected to the CAN bus utilizing a Kvaser leaf light CAN to USB interface. The sampling frequency f_s is equal to 200Hz. The Inertial Measurement Units (IMUs) and an incremental encoder are utilized to measure the orientation and rotation of moving parts of the excavator. The configuration of the sensors on the excavator is shown in Fig. 7. The data collection was done in a private worksite where there was no active construction work in the worksite. In fact, there is no unexpected factor that suddenly stops the operation.



Figure 7. The configuration of the sensors on the excavator [20].

In the experiments, the dig & dump duty cycle is done by an inexperienced operator, and also two types of materials such as sand and rough gravel are used to show the robustness of the methods. The dig & dump duty cycle is one of the main tasks in all worksites, and it comprises four sub-tasks such as filling the bucket, swinging loaded, dumping, and swinging empty. The operator has practiced less than 30 hours to drive the excavator which this factor can bring about more vibrations and meaningless movements of the excavator and subsequently can increase the challenges of swing angle and digging depth estimation. There are two different scenarios in the dataset. The duration of each scenario is approximately 6 minutes. The start and end positions in both scenarios are near the digging position. In the first scenario, the type of material is rough gravel, and the swing angle of operation is around 60° . In the second scenario, the type of material is sand, the swing angle is approximately 120° , and the digging depth is higher than in the first scenario. To increase the digging depth in the second scenario, the excavator goes on top of a small pile to reach a higher position.

4 Results

The performance of the proposed methods is illustrated by using real measurements. The algorithm was implemented using MATHWORKS MATLAB R2021a on a laptop with a 1.8 GHz Intel Core i7 CPU and 16 GB of RAM. The proposed algorithms are computationally efficient. The average required time for the computation at each time step in the swing angle and digging depth estimation algorithms are 0.0034 and 0.0024 seconds, respectively. The length of input data T is equal to 60 seconds. The output variables such as the swing angle and digging depth are updated in each iteration. The updating rate can easily change based on the application and final goal. Firstly, the performance of the swing angle estimation is analyzed

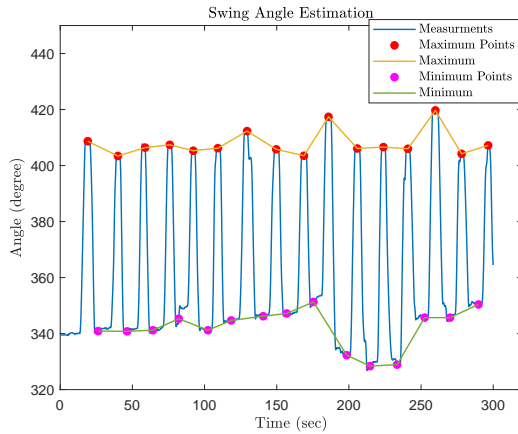


Figure 8. The estimation of maximum and minimum boundaries in the first scenario.

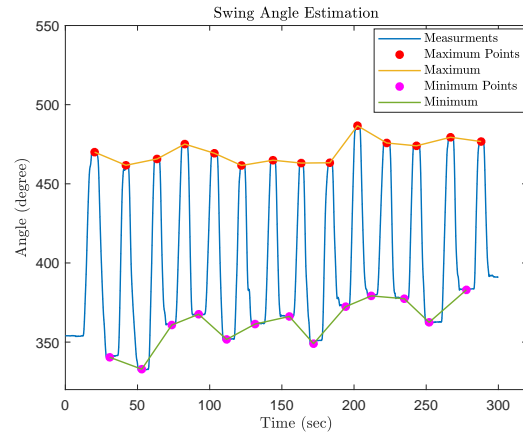


Figure 10. The estimation of maximum and minimum boundaries in the second scenario.

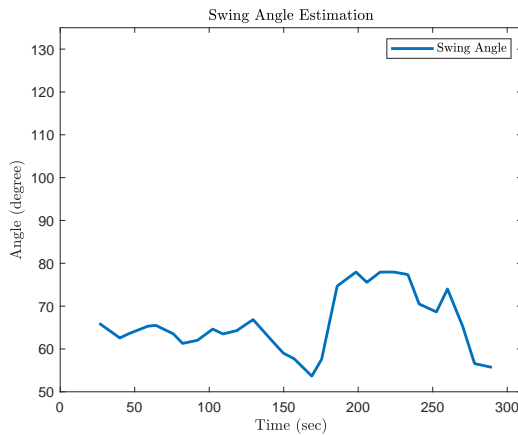


Figure 9. The estimation of swing angle in the first scenario.

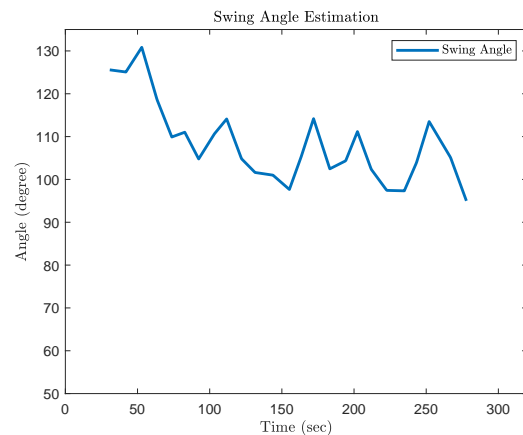


Figure 11. The estimation of swing angle in the second scenario.

in two scenarios. In the algorithm, the length of the time window in the moving average filter T_{filter} is equal to 5 seconds. In the first experiment, the dig & dump duty cycle with a swing angle of 60° is done, and the type of material is rough gravel. The results of the method are shown in Fig. 8 and Fig. 9. The results show the method efficiently estimates the swing angle and can track the changes during the operation. Moreover, the method is evaluated by using another experiment. In the second experiment, the swing angle is approximately 120° , and the type of material is sand. The results are presented in Fig. 10 and Fig. 11. The method can effectively estimate the swing angle.

In the next phase, the performance of the digging depth estimation is investigated in two scenarios. In this algorithm, the length of the time window in the moving average T_{filter} is equal to 2.5 seconds. In the first experiment, the digging depth is lower than in the second

experiment. The type of used material in this experiment is rough gravel, and the swing angle is approximately 60° . The estimation of digging depth is shown in Fig. 12. The method can estimate the digging depth, and it is robust against sudden movements of the bucket. In the second experiment, the type of material is sand, and the swing angle is approximately 120° . The result is shown in Fig. 13. The algorithm accurately estimates the current digging depth based on the bucket position. The results prove the presented methods can easily be utilized for real-time productivity estimation of excavators in different working conditions.

5 Conclusion

In this paper, two novel frameworks are presented to estimate the swing angle and digging depth of the dig &

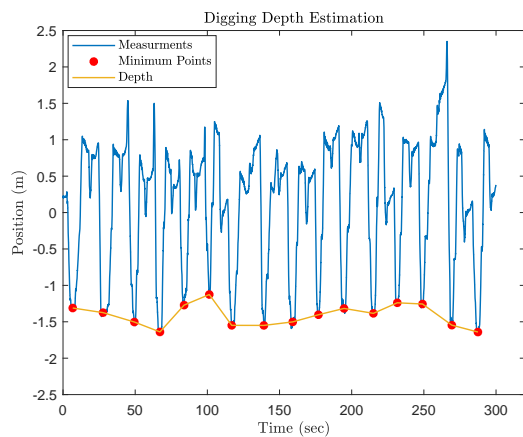


Figure 12. The estimation of digging depth in the first scenario.

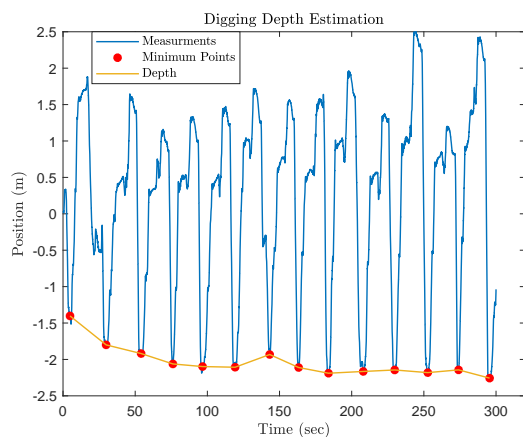


Figure 13. The estimation of digging depth in the second scenario.

dump duty cycle in real time. These variables have significant effects on the excavator's productivity and must be taken into account. Firstly, an algorithm is proposed to estimate the swing angle using the cabin encoder measurements. A moving average filter is used to reduce the effects of noises and sudden movements of the cabin, and then local minimum and maximum finders are utilized to find the extremum points. Afterward, Otsu's method is exploited to find the extremum points that are representative of scooping and dumping positions. Secondly, a similar approach is proposed to estimate the digging depth during operations based on the bucket position. The bucket position is calculated by using the measurements of IMU sensors and the forward kinematics of the excavator. After using the moving average filter, the local minimum points of the bucket position are found, and then Otsu's method is used to recognize the local minimum points that are

representative of the digging depth. Finally, the methods are tested by using the collected dataset in a private work-site. The dataset includes two scenarios including different materials such as sand and rough gravel, different swing angles, and digging depths. The results prove the methods can be used in the productivity estimation of excavators.

Acknowledgment

This project is part of the MORE-project [21], which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 858101.

The MORE project is an innovative European Industrial Doctorate (EID) research and training programme that tries to answer the challenges of heavy-duty mobile machines (HDMMs). The MORE project contributes to the development of state-of-the-art HDMMs by proposing effective solutions driven by digitalization and Artificial Intelligence (AI). There are eight researchers in the MORE project that work on three main work packages: 1) Process, 2) Machine, and 3) Control. In the process work package, we investigate a set of tasks that are performed by a machine or machines to achieve the purpose of work [21].

We would like to express our sincere gratitude for the continued support of Janne Koivumäki, Kalle Lahtinen, and Johanna Ylisaari during the data collection phase.

References

- [1] Marcus Geimer. *Mobile Working Machines*. SAE International, Warrendale, PA, 2020. doi:10.4271/9780768094329.
- [2] Dushyant A Deshmukh and Parag S Mahatme. Factors affecting performance of excavating equipment: An overview. *International Journal of Science and Research (IJSR) ISSN (Online)*, pages 1250–1253, 2016.
- [3] Tyrone Machado, David Fassbender, Abdolreza Taheri, Daniel Eriksson, Himanshu Gupta, Amir-masoud Molaei, Paolo Forte, Prashant Kumar Rai, Reza Ghabcheloo, Saku Mäkinen, et al. Autonomous heavy-duty mobile machinery: A multidisciplinary collaborative challenge. In *2021 IEEE International Conference on Technology and Entrepreneurship (ICTE)*, pages 1–8, Kaunas, Lithuania, August 2021. IEEE.
- [4] NG Hareesha and KN Umesh. Kinematic and isotropic properties of excavator mechanism. *Int. J. Eng. Res. Technol.-IJERT*, 3(17):1–6, 2015.

- [5] R Mundane Sagar and R Khare Pranay. Comparative study of factors affecting productivity and cycle time of different excavators and their bucket size. *International Journal on Recent and Innovation Trends in Computing and Communication*, 3:6518–6520, 2015.
- [6] Mario Klanfar, Vjekoslav Herceg, Dalibor Kuhinek, and Kristijan Sekulić. Construction and testing of the measurement system for excavator productivity. *Rudarsko-geološko-naftni zbornik (The Mining-Geological-Petroleum Bulletin)*, 34(2), 2019.
- [7] Khandakar M Rashid and Joseph Louis. Times-series data augmentation and deep learning for construction equipment activity recognition. *Advanced Engineering Informatics*, 42:100944, 2019.
- [8] Armin Kassemi Langroodi, Faridaddin Vahdatikhaki, and Andre Doree. Activity recognition of construction equipment using fractional random forest. *Automation in construction*, 122:103465, 2021.
- [9] *Caterpillar Performance Handbook*. Cat® publication by Caterpillar Inc., Peoria, Illinois, USA, 43 edition, 2013.
- [10] David J Edwards and Gary D Holt. Estivate: a model for calculating excavator productivity and output costs. *Engineering, Construction and Architectural Management*, 2000.
- [11] CM Tam, Thomas KL Tong, and Sharon L Tse. Artificial neural networks model for predicting excavator productivity. *Engineering, construction and architectural management*, 2002.
- [12] Krzysztof Schabowicz and Bozena Hola. Mathematical-neural model for assessing productivity of earthmoving machinery. *Journal of Civil Engineering and Management*, 13(1):47–54, 2007.
- [13] Gary D Holt and David Edwards. Analysis of interrelationships among excavator productivity modifying factors. *International Journal of Productivity and Performance Management*, 2015.
- [14] Mohamd Kassem, Elham Mahamedi, Kay Rogage, Kieren Duffy, and James Huntingdon. Measuring and benchmarking the productivity of excavators in infrastructure projects: A deep neural network approach. *Automation in Construction*, 124:103532, 2021.
- [15] A Panas and JP Pantouvakis. Comparative analysis of operational coefficients’ impact on excavation operations. *Engineering, Construction and Architectural Management*, 2010.
- [16] Benjamin Hughes and Tilo Burghardt. Automated visual fin identification of individual great white sharks. *International Journal of Computer Vision*, 122(3):542–557, 2017.
- [17] *MATLAB version 9.10.0.1613233 (R2021a)*. The Mathworks, Inc., Natick, Massachusetts, 2021.
- [18] Deng-Yuan Huang and Chia-Hung Wang. Optimal multi-level thresholding using a two-stage otsu optimization approach. *Pattern Recognition Letters*, 30(3):275–284, 2009.
- [19] Dongik Sun, Changuk Ji, Sunghoon Jang, Sangkeun Lee, Joonkyu No, Changsoo Han, Jeakweon Han, and Minsung Kang. Analysis of the position recognition of the bucket tip according to the motion measurement method of excavator boom, stick and bucket. *Sensors*, 20(10):2881, 2020.
- [20] Novatron ltd. <https://novatron.fi/en/>. Accessed: 2022-02-01.
- [21] MORE-ITN project. <https://www.more-itn.eu/>. Accessed: 2022-02-01.

Automated Valve Detection in Piping and Instrumentation (P&ID) Diagrams

M. Gupta^a, C. Wei^a and T. Czerniawski^a

^aSchool of Sustainable Engineering and Built Environment, Arizona State University, USA

E-mail: mgupta70@asu.edu, cwei32@asu.edu, Thomas.Czerniawski@asu.edu

Abstract – For successfully training neural networks, developers often require large and carefully labelled datasets. However, gathering such high-quality data is often time-consuming and prohibitively expensive. Thus, synthetic data are used for developing AI (Artificial Intelligence) /ML (Machine Learning) models because their generation is comparatively faster and inexpensive. The paper presents a proof-of-concept for generating a synthetic labelled dataset for P&ID diagrams. This is accomplished by employing a data-augmentation approach of random cropping. The framework also facilitates the creation of a complete and automatically labelled dataset which can be used directly as an input to the deep learning models. We also investigate the importance of context in an image that is, the impact of relative resolution of a symbol and the background image. We have tested our algorithm for the symbol of a valve as a proof-of-concept and obtained encouraging results.

Keywords –

Piping and Instrumentation Drawings; Yolo; Symbol Detection; Convolution Neural Network; Engineering Drawings; Symbol Classification; Deep Learning

1 Introduction

A P&ID diagram depicts the logical flow of information about physical processes and plant components with help of lines and symbols. Lines of varying thicknesses are used to represent different pipelines and each symbol represents a unique item like pressure sensor, temperature sensor, gate valve, floor drain, etc. To exemplify, a sample P&ID diagram is shown in Figure 1. A typical P&ID diagram can have more than 30 different symbols and thus, are information-rich. These drawings are analyzed manually for the purpose of estimating the quantities of various equipment while placing a purchase order and even when project teams are planning their work schedule. This analysis is highly dependent on the subjective

knowledge of the person who is reviewing these drawings and thus, can be time-consuming and prone to human errors. This task can become even more challenging and complex when there are symbols on P&ID diagrams which are functionally different but visually similar, as shown in Figure 2. Thus, differentiating one symbol from another can become extremely important and challenging. Additionally, misinterpreting or overlooking any information can prove detrimental to a project's progress and can result in serious internal conflicts.

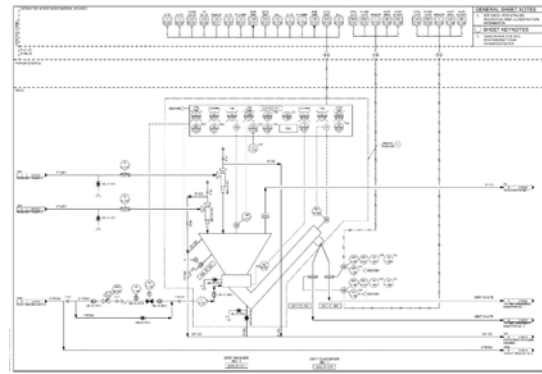


Figure 1. Sample P&ID diagram.



Figure 2. An example of similar symbols

It is safe to assume that companies have these drawings in legible electronic format for their ongoing projects which can be manipulated using state-of-the-art software. But there are still many companies who have these drawings in form of hard copies that is, paper-format or in scanned-format, especially for their older projects. So, digitization of these drawings in a format

which enables easy and user-friendly extraction of information can prove beneficial [1]. This can enable easy rectification of old drawings where the plant components have been replaced due to maintenance over time. Thus, with digitized and updated P&ID diagrams, it will be easier for the project teams to track their instrumentation inventory during the construction phase and develop an up-to-date drawings repository for the maintenance in the post-construction phase.

Currently, the construction industry does not possess a large dataset which is labelled and is publicly available. Creating such a large and real-world labeled dataset requires dedicated efforts from the experts to carefully annotate them. Thus, this process can be time-consuming and can become prohibitively expensive. Therefore, synthetic data can be used for training deep learning models.

In our research, we have created a synthetic dataset which was used for training our object detection algorithm. The approach has significance as it obviates the requirement of human annotation. We try to identify a particular symbol on a P&ID diagram which we define as the 'Target Symbol'. While doing so, we investigate the importance of contextual information while developing an object detection deep learning model. Contextual information in this paper refers to the part of an image which does not include the object of interest (OOI). With our work, we also try to bring the applications of modern technology like machine learning and deep learning into the industries like construction, oil and gas which rely on engineering drawings for their operation and functioning. Deep Learning algorithms and frameworks like Convolutional Neural Network (CNN) [2], Histogram of Oriented Gradients (HOG) [3], You only Look Once (Yolo) [4], [5], etc. enable the processing of image data and hence, the engineering drawings. For our research project, we have applied Yolo version 2 which is an object detection algorithm.

We believe the method described in our paper can be applied to the cases which have class-imbalance [6], [7] problem in an image dataset. In the context of our paper, the class-imbalance problem would mean that there could be a symbol on P&ID sheets whose total number of occurrences is far less than the occurrence of another symbol. In such cases, data augmentation strategies can often help in improving a neural network's performance. As our method allows user to control the number of occurrences of less frequent symbols in an artificially generated dataset hence the distribution of the minority class can be balanced. This can result in increased performance while training a neural network.

2 Literature Review

In this section, we will discuss some of the previously published papers related to processing of the engineering drawings and their proposed frameworks. Papers such as [8] present work in symbol detection and [9] in symbol classification.

[10] highlights the challenges in the successful classification of symbols. The biggest challenge is the unavailability of a labelled public dataset. They discussed intra-class and inter-class similarity amongst symbols as a major obstruction for detection algorithms. Their results found that class-decomposition helped in increasing the classification accuracy. Class-decomposition is the process of breaking down labelled datasets to a larger number of subclasses by means of clustering the instances that belong to one class at a time. [10], [11] developed heuristics-based rules for extraction and localizing symbols. Generally, these methods are highly data-dependent and could not be used if there are slight variations in symbols or P&ID diagrams as a whole.

[9], [10] present works in symbol classification and [8], [12] in symbol detection. [12] performed symbol detection on a dataset which had class-imbalance problem. The dataset had 29 different symbols in P&ID diagrams and the distribution of these symbols was non-uniform. They first applied Yolo model for symbol recognition and found 3 least occurring symbols were missed by their network. Then, they performed data augmentation for 8 minority classes (that is, 8 least occurring symbols) using MFC-GAN (Multi Fake Class Generative Adversarial Network). Results showed that MFC-GAN improved the accuracy of their model in detecting even the least occurring symbols. They stated Yolo is a simple framework consisting of a single convolutional neural network which can be used for detecting multiple bounding boxes for objects belonging to different classes. That is why we have also implemented Yolo for our research.

For an extensive analysis of P&ID diagrams, [11], [13], and [14] developed frameworks to recognize symbols, texts, and lines. In an inspiring work stated in [13], they presented a proof-of-concept for identifying symbols, interpreting component connections, and representing those connections graphically for a P&ID drawing. Their model required data in a vectorized format and had an assumption that P&ID symbols are all 'blocked' in DXF files. Similarly, [11] developed a methodology for a complete analysis of P&ID diagrams encompassing symbol recognition, pipeline identification, and text localization. They used a pre-trained CTPN (Connectionist Text Proposal Networks) network for text detection. Shape properties like, the number of sides in a polygon, length to width ratio, etc. were used as features to detect a pipe's inlet and outlet.

Identified pipelines, tags, symbols, and texts were associated with each other based on the Euclidean distance. However, they annotated symbols by completely masking their pixel values. So, essentially annotations were based on the shape of a symbol's boundary or outline. We believe this approach would not be able to give accurate results when 2 or more symbols have the same shape or the same 'outer boundary'. For instance, the 3 symbols shown in Figure 3 are distinct while their outline/ boundary is the same. Nevertheless, the work in [11] is significant for the complete analysis of P&ID diagrams.

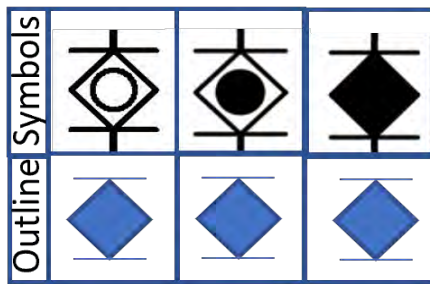


Figure 3. Showing that 'Target Symbol' has same boundary as its close neighbors

3 Experiment and Problem Setup

An industry partner provided us with P&ID diagrams, an example of which has been shown in Figure 1. The sheets have various symbols and some of them are shown in Figure 4. The main goal is to generate an algorithm that can correctly identify all the instances of these symbols. However, the current work has been performed for one symbol as a proof-of-concept and will be extended in the future for all the different symbols.

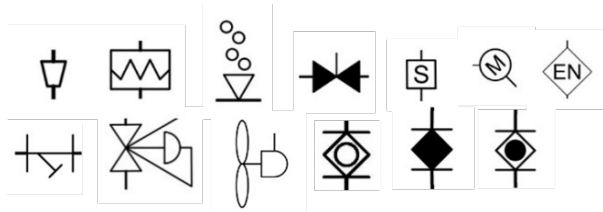


Figure 4. Few symbols in the P&ID diagrams

Thus, in this project, we are detecting one symbol representing a valve which we call as the 'Target symbol', and is shown below in Figure 5. The reason for choosing this symbol as our 'Target symbol' is that it has look-alike symbols which are the last 2 symbols

shown in Figure 4. The existence of look-alike symbols makes our symbol recognition a relatively challenging case-study, and so representative of one of the more difficult cases. So, we want our algorithm to identify the Target symbol successfully on a P&ID diagram which can also have other similar looking (but, functionally different) on it.



Figure 5. The 'Target Symbol'

Visual assessment of given P&ID diagrams reveals that the Target Symbol occurs in different orientations and has a slight variation in its appearance across the P&ID drawings. These changes in appearance are attributed to the pipeline connections and rotation of the symbol. All the possible variations of the Target Symbol have been shown in Figure 6. These configurations of the Target Symbol are stored as 8 different images in a folder whose name is 'Target_Symbol_folder' (suppose). This forms the first step of our experimental setup. Thereafter, we develop a labeled synthetic data through random cropping.

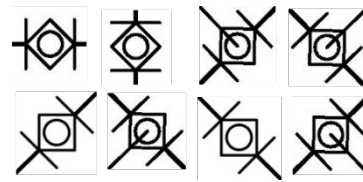


Figure 6. 'Target Symbol' and its different configurations in the P&ID diagrams.

3.1 Data Augmentation

We received 3 P&ID sheets from our industry partner. We performed data augmentation using 2 of these 3 sheets and kept 1 sheet for testing/inferencing. Original sheets have a resolution of 10100-by-6600 pixels. Thus, we decided to split it into sub-images of a smaller size of 256x256 pixels. The sub-image/ 'crop' is extracted randomly as shown in Figure 7, and the number of sub-images can be defined by the user. Script for processing the P&ID sheets to generate random crops was written in Matlab where the user also has an option to control the total number of crops to be generated from the 2 original sheets.

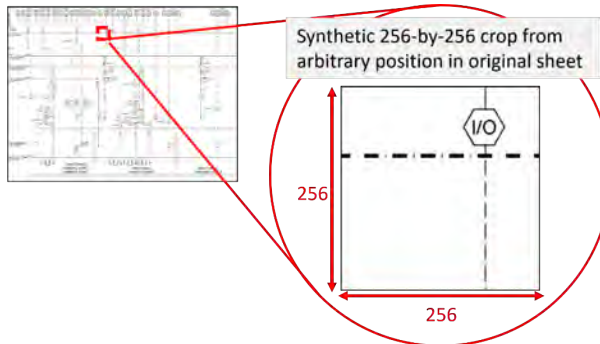


Figure 7. Sample of a 256-by-256 random crop

The idea is to generate several thousand sub-images from 2 P&ID sheets which can then be used for training the Yolo v2 neural network.

3.2 Labelled Data Generation

To create a labelled dataset, we need to have information about the Object of Interest (OOI). Because we are using Yolo v2 for object detection, therefore, we would require the location information that is, the bounding box coordinates of the Target symbol in each sub-image. To assess the impact of the context on our symbol detection we have partitioned our project into 2 phases namely, Phase-1 and Phase-2. In Phase-1, a neural network was trained on images which only had our Target symbol while in Phase-2 the input images also had one of the look-alike symbols in addition to the 'Target symbol'. We wanted to assess the impact of this change and compare the performance of these two networks. We believe this can enable us to draw insights about strategies to create an effective synthetic dataset. Detailed methodology for Phase-1 and Phase-2 are discussed below.

3.2.1 Phase-1

In Phase-1 of the project, we generated a total of 1000 random crops of size 256-by-256 pixels through the same approach as defined in Section 3.1. Then, our custom-built Matlab function will access the 'Target_Symbol_folder', and then, randomly select any one of the 8 Target symbol images and, place it randomly on 256-by-256 crop as shown in Figure 8. Our Matlab function allows us to save the position of the Target Symbol in the $[x,y,w,h]$ format. 'x' and 'y' represent the coordinates for the upper left corner, and 'w' represents the width and 'h' represents the height of the Target Symbol crop. Essentially, this is the bounding box information for the Target Symbol on each sub-image. Thus, we are able to produce a labelled dataset which can be used for training. The format of

$[x,y,w,h]$ is selected because this is the default input format for Matlab. Hence, the dataset generated can be used directly as an input for training the neural network in Matlab.

However, minor manual updates were required in a special case where the crop already had the 'Target Symbol' on it. Thus, when our algorithm pasted one more instance of it, the sub-image would have 2 Target symbols on it. An example is shown in Figure 9. Hence, bounding box information for the pre-existing symbol was added manually to the list where our Matlab program was storing the position coordinates $[x,y,w,h]$ of all the Phase-1 crops.

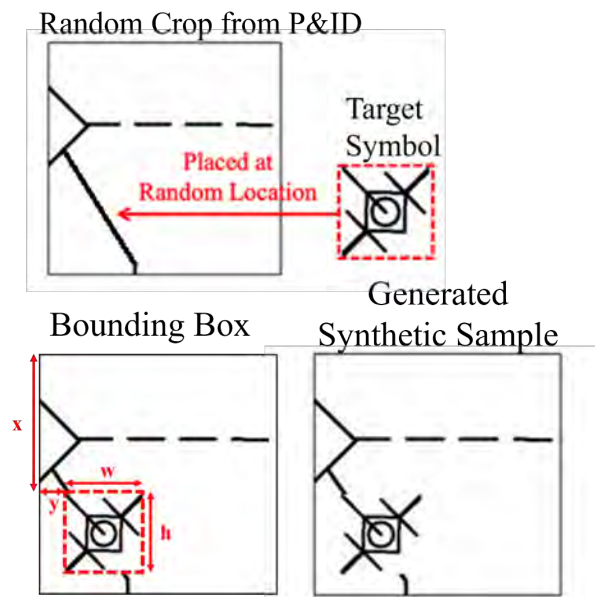


Figure 8. Sample of a 256-by-256 random crop with 'Target symbol'

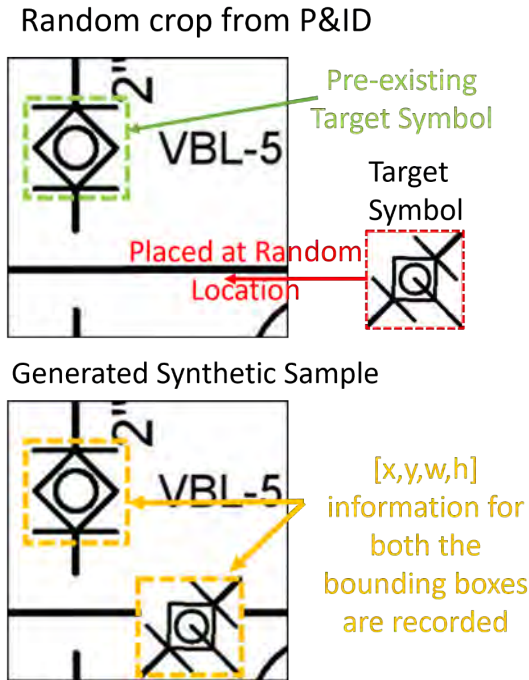


Figure 9. Crop having both the pre-existing and newly pasted 'Target Symbol'

Now this labeled dataset can be used as input to our object detection model. Similar to [12], we also used Yolo algorithm as the choice of our object detection model. Our Yolo v2 network has Resnet-50 [15] as a classifier for feature extraction and a CNN as our detection network.

It is to be noted that a detector trained as described above can only be used for testing on 256-by-256 images. Hence, we could not use this detector for testing directly on the third original sheet whose size is 10100-by-6600 pixels. So, after training the Yolo detector, we developed a custom pipeline that would run our trained detector in a sliding window manner of size 256-by-256 pixel with a stride of 256 to test any P&ID diagram whose resolution is more than 256-by-256. Results of Phase-1 are summarised in Section 4.1.

3.2.2 Phase-2

In Phase-2, we seek to improve the performance of the system presented in Phase-1 by (1) increasing the amount of contextual information in each sample and (2) adding adversarial symbols (i.e., look-alike symbols) into the training symbols. We think one way to achieve (1) is by decreasing the resolution of the original P&ID sheet which in turn increases the amount of 'background' for the Target symbol. For (2) we modified the crops so that our sub-image comprised of one Target Symbol and one of its look-alike symbols on the same 256-by-256 image crop. We wanted to study how these variations from Phase-1 affects the model performance. As already

mentioned, our Target Symbol has a close resemblance with 2 other symbols. The various configurations in which these 2 symbols occur in the drawings set are shown in Figure 10. These symbols are saved as 14 different images in a folder which is named as 'Similar_to_Target' (suppose). In other words, the images in Similar_to_Target folder form a close neighbour group for our Target Symbol.

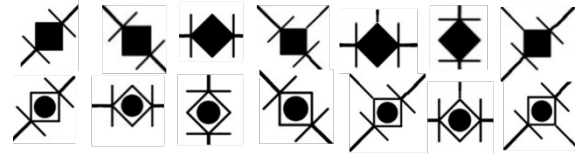
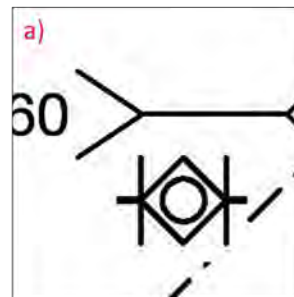


Figure 10. Collection of symbols configurations that resemble the 'Target Symbol' (referred to as close neighbors)

Hence, for Phase-2 we formulated a few changes in Phase-1 which are listed below:

1. Resize the original P&ID sheets from 10100-by-6600 pixels to 3800-by-2450 pixels. The reason for doing this is that a crop of the 256-by-256 crop out of a 10100-by-6600 pixels sheet didn't capture the background (contextual) information well. Thus, all the 1000 crops of Phase-1 didn't have a well-representative background. Whereas crops generated from 3800-by-2450 pixels sheets fairly capture the background information in them. Thus, these new crops are more representative of the original sheets. This is demonstrated in Figure 11. It can be observed from the figure that very less background is captured in crop (a) while crop (b) has significantly more context/ information about the background.

Crop from 10100-by-6600 pixels sheet



Crop from 3800-by-2450 pixels sheet

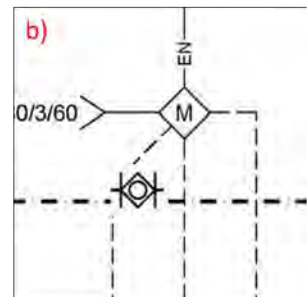


Figure 11. Comparison of background information in crops generated from a) 10,100-by -600 pixels sheet and b) 3800-by-2450 pixels sheet

- In addition to pasting only the Target Symbol, we also randomly pasted one of the 14 images of the “look-alike” symbols from Similar_to_Target folder. An example is shown in Figure 12 This is done to increase the ability of the model to effectively differentiate amongst similar-looking symbols.

The process of manually updating the bounding box information for the crops having pre-existing Target symbol is the same as Phase-1. However, on average only 1 in 70 crops had a pre-existing Target symbol and thus, our dataset requires minor manual updates. Hence, Phase-2 has background information-rich 1500 sub-images of size 256-by-256. The results are discussed in Section 4.2

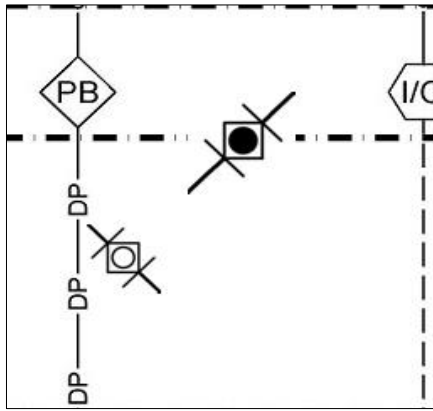


Figure 12. 256-by-256 crop in Phase-2 having both the ‘Target Symbol’ and one of its close neighbours

4 Results

As mentioned, 2 out of 3 original sheets are used for data augmentation and training while the remaining one sheet is kept reserved for testing our model and assessing its performance. The trained model is run in the sliding window manner on the third original sheet for inferencing and the results are discussed below.

4.1 Phase-1 results

The detection results for Phase-1 are shown in Figure 13 and Figure 14. It is observed that out of 5 occurrences of our Target Symbol on this sheet, our model is able to correctly identify 3 occurrences while wrongly classifying a different symbol 2 times and completely missing it 2 times as well.

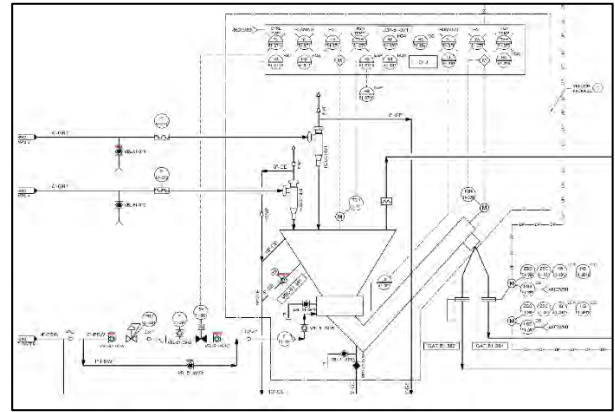


Figure 13. Phase-1 detection results on 10,100-by-6600 pixels sheet

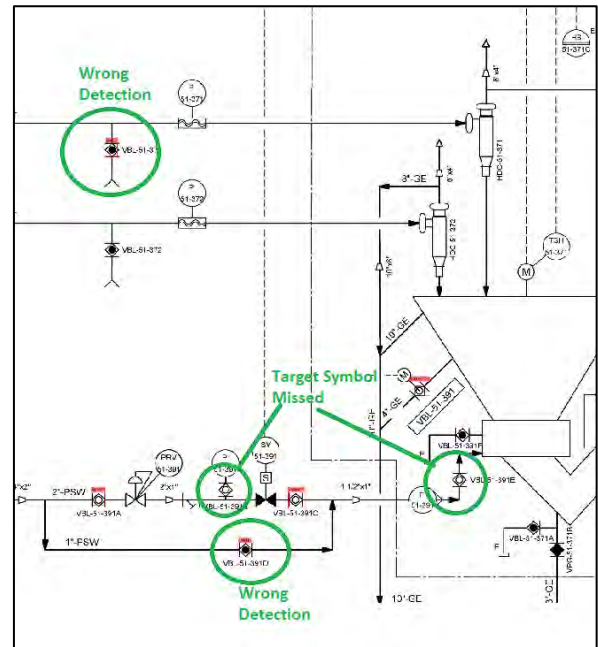


Figure 14. Zoomed-in view of Phase-1 results

Thus, our results of Phase-I have both the Type-I error and Type-II error [16]. The type-I error refers to the False positives which in our case is the “Wrong Detection” as highlighted in Figure 14. And, Type-II error is the False-negative which in our case is “Target Symbols missed” as highlighted in Figure 14. It can be noted that the wrongly classified symbol for both the occasions is one of the close resembling symbols as identified in Figure 10. This is closely related to the argument made in [10] regarding the presence of look-alike symbols.

4.2 Phase-2 results

Results for Phase-2 are shown in Figure 15 and Figure 16. It can be verified that all the 5 instances of the Target Symbol are now correctly identified without any error. It means that increasing the contextual information and introducing the look-alike symbols in 256-by-256 sub-images enhanced the discriminative ability of the neural network.

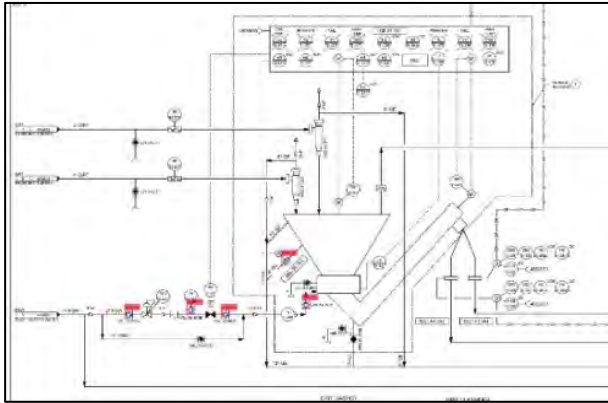


Figure 15. Phase-2 detection results

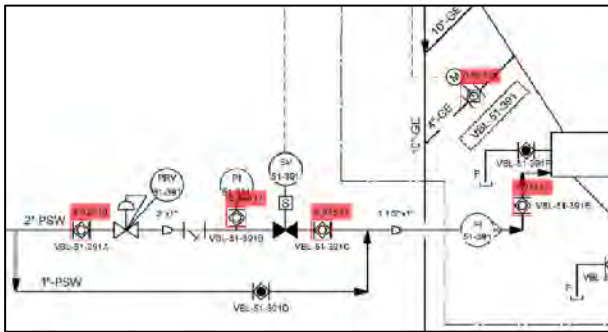


Figure 16. Zoomed-in view of Phase-2 results

5 Conclusion

In this paper, we investigated the effect of contextual information in boosting the discriminative ability of a neural network among similar-looking objects. Our developed model can differentiate among close resembling symbols and is able to find out all the instances of the desired symbol. The method for labeled data generation is simple and time efficient. It gave promising results on our dataset along with bypassing the requirement of an expert annotating the data. Therefore, it is cost-efficient too. We opine that with our approach users can generate labeled synthetic data with minimal effort. This project can also find applications in electronic circuit designing.

However, there are a few limitations to our work that we will be overcoming in our future endeavors. First, the model is trained for detecting only one symbol. In the future, we will be scaling up the project to detect all the symbols. Second, current data is from a single contractor hence, highly subjective to their 'style' of P&ID drawings. To make the program more robust and ready-to-use by contractors we will be collecting data from multiple companies to account for variations in designing layout and styles. Third, we will be increasing the scope to detect the pipelines and associated tags to graphically represent these connections. We will be developing an application to enable exporting the schedules of various pipes/symbols in a .csv format. This would be helpful in preparing BOQs (Bill-of-Quantities) and efficiently keeping track of the inventories by the project team.

Acknowledgment

We would like to thank Sundt Construction Company & General Contractor for providing us with the P&ID sheets for analysis.

References

- [1] S. v. Ablameyko and S. Uchida, "RECOGNITION OF ENGINEERING DRAWING ENTITIES: REVIEW OF APPROACHES," 2007. [Online]. Available: www.worldscientific.com
- [2] K. Fukushima, "Biological Cybernetics Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position," 1980.
- [3] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, 2005, vol. I, pp. 886–893. doi: 10.1109/CVPR.2005.177.
- [4] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," Dec. 2016, [Online]. Available: <http://arxiv.org/abs/1612.08242>
- [5] X. Huang *et al.*, "PP-YOLOv2: A Practical Object Detector," Apr. 2021, [Online]. Available: <http://arxiv.org/abs/2104.10419>
- [6] N. Japkowicz, "The Class Imbalance Problem: Significance and Strategies."
- [7] M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," Oct. 2017, doi: 10.1016/j.neunet.2018.07.011.

- [8] L. P. Cordella and M. Vento, “Symbol recognition in documents: a collection of techniques?”
- [9] J. Lladós, E. Valveny, G. Sánchez, and E. Martí, “Symbol recognition: Current advances and perspectives,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2002, vol. 2390, pp. 104–128. doi: 10.1007/3-540-45868-9_9.
- [10] E. Elyan, C. M. Garcia, and C. Jayne, “Symbols Classification in Engineering Drawings,” in *Proceedings of the International Joint Conference on Neural Networks*, Oct. 2018, vol. 2018-July. doi: 10.1109/IJCNN.2018.8489087.
- [11] R. Rahul, S. Paliwal, M. Sharma, and L. Vig, “Automatic Information Extraction from Piping and Instrumentation Diagrams,” Jan. 2019, [Online]. Available: <http://arxiv.org/abs/1901.11383>
- [12] E. Elyan, L. Jamieson, and A. Ali-Gombe, “Deep learning for symbols detection and classification in engineering drawings,” *Neural Networks*, vol. 129, pp. 91–102, Sep. 2020, doi: 10.1016/j.neunet.2020.05.025.
- [13] C. Howie, J. Kunz, T. Binford, T. Chen, and K. H. Law, “Computer interpretation of process and instrumentation drawings.”
- [14] L. Boatto *et al.*, “Detection and Separation of Symbols Connected to Graphics in Line Drawings.”
- [15] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Dec. 2016, vol. 2016-December, pp. 770–778. doi: 10.1109/CVPR.2016.90.
- [16] Dekking, Michel. A modern introduction to probability and statistics: understanding why and how. Springer, London. 2005.

Suitability and Effectiveness of Visualization Platform-based Construction Safety Training modules

K. Bhagwat^a and V.S.K. Delhi^a

^a Department of Civil Engineering, Indian Institute of Technology Bombay, India.
E-mail: kishorsbhagwat@iitb.ac.in, venkatad@iitb.ac.in

Abstract –

The construction industry is one of the most hazardous industries in the world. In addition, past literature highlighted that about 50% of the hazard remains unrecognized in the construction work environment, resulting in catastrophic consequences. Construction Safety Training (CST) is one of the best safety interventions to deal with this. The traditional classroom-based CST method is fraught with several non-interactive, non-engaging, and ineffective limitations. With the advent of visualization platform-based technologies, researchers have introduced various CST delivery modules using digital environments. However, in-depth investigation of visualization platform-based safety training modules from a suitability and efficacy perspective for the construction industry is understudied. Therefore, this study aimed to identify a suitable and effective safety training delivery module for the construction industry. To this end, three visualization platform-based CST modules were developed, namely, Image, Virtual Tour (VT), and Mobile Virtual Reality (MVR), and introduced to the construction professionals. The feedback analysis highlighted that VT CST is suitable for the industry. Next, the effectiveness of the VT CST modules was determined using Hazard Recognition Score (HRS). The results highlighted that the overall pre-training HRS of the construction professionals was 56.83%, and post-training HRS was 88.56% which shows significant ($p < 0.00$) enhancement in the HRS. In conclusion, this study noticed the VT CST module as a suitable and effective training module for the industry. Further, theoretical and practical implications of the study and future research directions were discussed.

Keywords –

Construction; Hazard; Safety; Training; Visualization

1 Introduction

The construction industry continues to be one of the hazardous industries in the world due to associated risky work practices and complex and dynamic nature [1]. On one end, it is one of the significant pillars in the economic development of several countries [2]. On the other end, globally, it is responsible for more than 60,000 mortalities every year [3]. Accident numbers vary from country to country [4]. For example, the construction industry in the developed economy such as the United States of America is responsible for over 900 mortalities and 200,000 non-fatal injuries in 2016 [5]. At the same time, the construction industry in the developing economy such as India is responsible for mortalities ranging from 11,614 to 22080 [6]. These statistics provide a snapshot of the hazardous nature of the construction industry worldwide. Consequently, construction safety management research has attracted the attention of academicians and practitioners from all over the world [7].

The root cause analysis of the industrial accident was performed by H. W. Heinrich and identified unsafe behavior (88%), unsafe conditions (10%), and natural disaster (2%) as root causes of the industrial accidents [5], [8]. 98% of accidents can be eliminated by practicing safe behavior and maintaining safe site conditions [9], [10]. For instance, by focusing on major victims of construction accidents, i.e., workers, they can secure their safety by safe behavior and either accepting or rejecting the risk involved in the work assigned [11]. In the current scenario, even though the employer is safety conscious, workers can decide to behave either safely or unsafely [11]. Also, workers' ability to recognize the hazard and understand the magnitude of risk involved determines their behavior and safety [11]. However, studies highlighted that construction workers do not possess the essential skill sets to recognize hazards correctly [7], [12], [13]. Consequently, poor hazard recognition leads to construction accidents [14]. More specifically, about 50% of the hazards remain unrecognized in the construction work environment [5], and such poor hazard recognition

and assessment are responsible for more than 42% of the construction accidents [15].

Past literature has highlighted that safety training is one of the best methods of improving hazard recognition skills of the construction workforce [1], [11]. It can be ascertained that several employers invest millions of dollars in training their employees for hazard management and preventing accidents [1]. Notwithstanding such initiatives from the employers, the desirable level of hazard recognition has not been reached [13]. Past studies have highlighted that only 10% to 15% of training investments result in expected outcomes in the jobsite [1], [16] and have not noticed a positive correlation between implementing traditional safety training methods and safety performance [17]. Whereas on-site demonstrations of hazards for training purposes are subjected to injury risk, often time-consuming and costly [18], [19].

In recent years, visualization platform-based safety training modules have been introduced in the safety training domain to overcome the drawbacks of the traditional safety training methods. The visualization platform-based safety training modules depict actual construction sites in the digital environment. Such a digital environment replicates hazardous construction site conditions that are impossible to observe, unsafe behavior that is dangerous to perform, and costly to develop in an actual construction work environment [20]. Consequently, visualization platform-based safety training modules such as Image, VT, and MVR CST methods have captured the attention of the researchers. Past research has explored the usability of visualization platform-based safety training modules [19]. However, the detailed investigation of the effectiveness of a suitable visualization platform-based CST method is understudied. Therefore, this study aimed to identify the effectiveness of the suitable visualization platform-based CST module. The objectives of the study were two-fold,

- To identify a suitable visualization platform-based CST module,
- To determine the effectiveness of a suitable visualization platform-based CST module.

2 Background of the study

Hazard is the potential of something to cause harm [21]. Hazards may be predictable as part of planned tasks or emergent due to the industry's dynamic nature [11]. At present, hazard recognition is solely relied on safety officers [22]. At the same time, a past study highlighted that about 50% of the hazards remain unrecognized in the construction work environment [5] due to the industry's dynamic nature, varied safety perception levels, and limited numbers of safety officers [22]. On one end, studies reported that hazard recognition is the safety

officer's responsibility [22], and on the other end, safety is the responsibility of all stakeholders. In addition, small-budget construction projects and safety-insentient organizations often do not engage safety officers. Therefore, developing and enhancing the hazard recognition abilities of the major stakeholders is imperative [23]. As a first step, this study considered major stakeholders such as construction workers and site supervisors as target subjects. Past literature has continuously highlighted that safety training is one of the best interventions to develop and enhance hazard recognition skill sets among construction stakeholders [1].

Safety training enhances employee safety knowledge, perception, behavior, compliance, and safety culture and performance [24], [25]. As a result, safety training interventions have been a point of attraction for the last four decades. Some of the studies focused on safety training delivery methods, whereas some on safety training knowledge transfer. For instance, Cromwell and Kolb [16] highlighted that trainees with a high level of organization, supervisor, and peer support reported higher safety knowledge transfer. Past studies [25], [26] also highlighted that traditional safety training delivery method such as classroom-based safety training needs to be improved for effective knowledge transfer. Consequently, researchers have introduced visualization platform-based safety training modules [5], [7], [11], [17], [19], [20], [27]–[30]. However, a systematic investigation on the suitability of the visualization platform-based safety training modules for the construction industry and the effectiveness of such a suitable safety training method is understudied. This investigation can strengthen safety training delivery and boost knowledge transfer by adopting a suitable and effective safety training module for the industry.

3 Research methodology

The aim of the study was achieved in two phases. The first phase focused on developing visualization platform-based CST modules and identifying a suitable training module for the construction industry. The second phase focused on investigating the effectiveness of the suitable visualization platform-based CST module.

3.1 Visualization platform-based construction safety training modules and their suitability

This study developed three CST modules: Image, VT, and MVR CST. Here, the Image-based CST module includes virtual photographs along with associated safety information. The VT CST module contains a virtual tour around the simulated construction scenario. Here, users

were provided with safety information while exploring the digital construction work site. The third CST module was MVR, which was based on the gaming platform. A simulated construction scenario was introduced to users through a virtual reality gearbox using a mobile device. In this module, the user had the freedom to comprehend safety practices while exploring simulated construction scenarios using a gaming joystick.

The content for CST modules was based on general safety practices related to personal protective equipment (PPE), scaffolding, construction machinery, and material storage. A total of 21 safety practices were considered from Bhagwat et al. [19], adopted from various Bureau of Indian Standards and a safety handbook. One of the examples out of the considered 21 safety practices was 'Worker on site with proper PPE'. Simulated construction scenarios were developed considering 21 safety practices and using Trimble SketchUp, Autodesk Revit, and Unity game engine. Based on developed simulated construction scenarios, Image, VT, and MVR CST modules were developed as shown in Figures 1, 2, and 3, respectively.

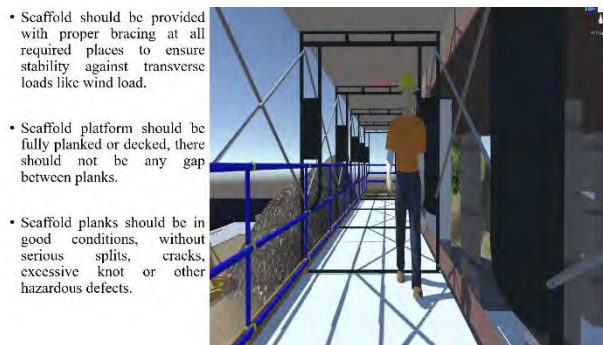


Figure 1. Screenshot of Image CST module



Figure 2. Screenshot of VT CST module

Developed CST modules were introduced to construction professionals, and their views regarding suitable safety training module for the construction industry was identified through a questionnaire survey.

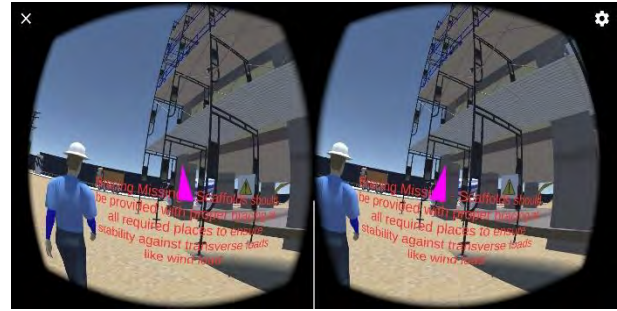


Figure 3. Screenshot of MVR CST module
Note: Bracing missing- scaffolds should be provided with proper bracing at all required places to ensure stability against transverse loads like wind load.

3.2 Effectiveness of the suitable CST module

The effectiveness of the suitable CST modules was evaluated using the HRS of the construction workers and supervisors. To do so, adopted 21 safety practices were negatively coded and considered for the HRS assessment. The relative average hazard weight associated with 21 unsafe practices was adopted from Bhagwat et al. [19] to evaluate the HRS, as shown in table 1. For example, the relative average hazard weight of 'Worker on site without proper PPE' was 5.12%. The sum of all relative average hazard weights for all unsafe practices was 100%. If any respondent identified all hazards during the HRS assessment, then the respondent was awarded 100% HRS. If any respondent did not identify a single hazard was awarded 0% HRS. Here, to investigate the effectiveness of the suitable CST module, pre-training and post-training HRS of construction workers and supervisors were determined through an experimental study.

Table 1. Relative average hazard weight (%) of the unsafe practices (adapted from Bhagwat et al., [19] with permission from ASCE)

Unsafe practices	Weight (%)
Workers without safety helmet	5.58
Crane member close to overhead power lines	5.58
Person being lifted by crane on its hook or boom	5.53
Unstable or uneven scaffold footing	5.48
Scaffold bracing is missing	5.38
working on the scaffold without safety belt	5.33
Working platform with cracks	5.17
Worker on site without proper PPE	5.12
Nails/Bars being projected out	4.97
Working platform is missing	4.87
Guardrail not provided	4.61

Obstruction on working platform	4.61
Scaffolding ladder without railing	4.56
Use of ladder on scaffold	4.56
Worker sitting on gas cylinder.	4.51
Toe Boards is missing	4.46
Stack of cement bags higher than 15	4.10
Guardrail top railing missing	4.00
Scaffolds platform not fully planked	4.00
Cement bags stack higher than 8 without crosswise pattern	3.79
Worker handling cement without goggles and dust mask	3.79

4 Data collection, analysis and results

4.1 Suitable CST module

As mentioned earlier, this study developed three visualization platform-based CST modules. Further, a questionnaire survey was conducted through construction professionals to identify a suitable CST module for the industry. A total of 47 construction professionals were approached using the snowball sampling technique. Out of which, 45 responses were recorded, with a response rate of 95.74%. All responses were collected through face-to-face interactions, which resulted in a higher response rate and precise inputs. The respondents' designation (as shown in Figure 4) and work experience (as shown in Figure 5) details highlighted that the respondents had varied roles in the construction industry with varied work experiences. The total experience of the respondents was 352 years, and the average experience was more than 7.5 years. Such diverse roles and work experiences helped to maintain quality and unbiased responses.

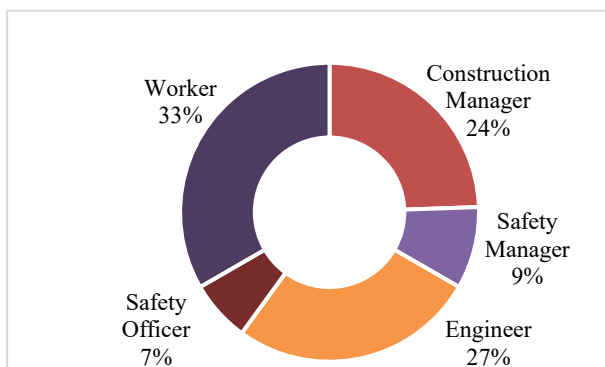


Figure 4. Construction professionals' role in the construction industry

Further, the data analysis was performed, and a suitable safety training module was identified for the

construction industry. Out of 45 respondents, eight (18%) respondents selected Image, eight (18%) respondents selected MVR, and 29 (64%) respondents selected VT as a suitable safety training module for the construction industry, as shown in Figure 6. Broadly, according to construction professionals, the VT CST module is a suitable module for the industry. However, more in-depth investigations in this domain are warranted to confirm the generalizability of the results.

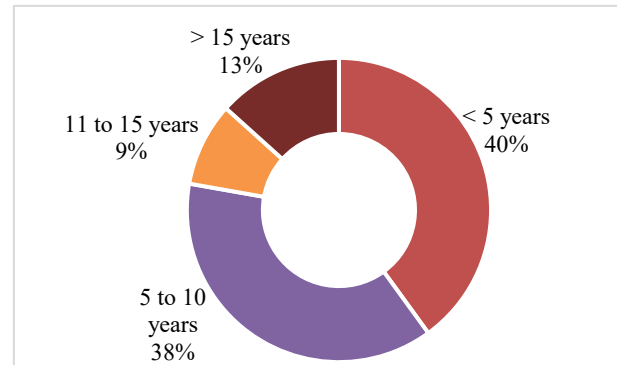


Figure 5. Construction professionals' work experience in the construction industry

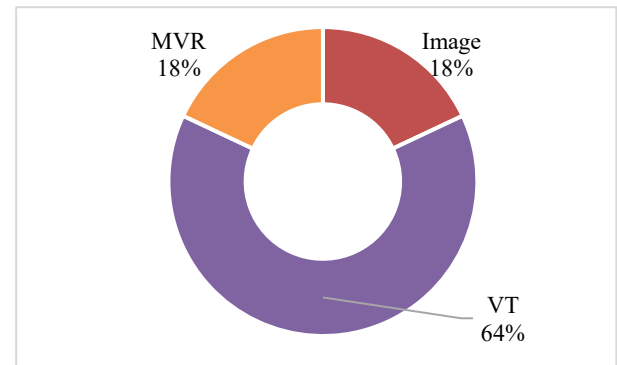


Figure 6. Construction professionals' percentage preferences to the CST modules

4.2 Effectiveness of VT CST module

The study's second objective was achieved using pre-training and post-training HRS of construction workers and supervisors. In this experimental study, fifteen respondents participated from three different building construction projects. Out of fifteen, twelve respondents were workers, and three were site supervisors. The total experience of the respondents was 104 years, and the average experience was more than 6.5 years. Out of three projects, two were residential, and one was a commercial building construction project. All three projects were located in Mumbai city, Maharashtra, India. Before initiating the experimental study, all the respondents

were introduced to the study's objective, and their consent was obtained. Further, respondents' pre-training HRS assessment was conducted. A sample photograph of the pre-training HRS assessment is shown in Figure 7. The pre-training HRS were calculated using Equation (1).

$$HRS = \sum_{i=1}^{21} (p_i) (q_i) \quad (1)$$

Where, p_i = relative average hazard percentage weight, i = 21 unsafe practices considered in the assessment, and $q_i = (0, 1)$ if unsafe practice i was identified, then $q_i = 1$, else 0. All pre-training HRS are given in Table 2. For example, the pre-training HRS of respondent 1 was 70.80%. After the pre-training HRS assessment, safety training was provided to all respondents using the VT module. The training was provided with the help of a tablet. A sample photograph of the VT safety training is shown in Figure 8. The training time for each respondent was seven minutes. After safety training, a post-training HRS assessment was conducted, and HRS were calculated using Equation (1). All post-training HRS are given in Table 2.



Figure 7. Sample photograph of pre-training HRS assessment

4.3 Statistical differences between pre-training and post-training HRS

Based on the initial observation in Table 2, there were differences in construction professionals' pre-training and post-training HRS. Therefore, these HRS differences were further statistically evaluated with the help of hypothesis testing. The null hypothesis (H_0) was set as construction professionals' post-training HRS were lesser than pre-training. The alternative hypothesis (H_1) was set as construction professionals' post-training HRS were greater than pre-training. A One-tail t-test for paired two samples for means was performed at the confidence interval of 95%. The hypothesis test result rejected H_0 ($p < 0.00$), i.e., the post-training HRS were greater than the pre-training HRS. This highlighted the effectiveness of

the VT CST module.



Figure 8. Sample photograph of VT safety training to construction professional

Table 2. Construction professionals' HRS in pre-training and post-training safety assessment

Respondent code	Pre-training HRS (%)	Post-training HRS (%)
Worker 1	70.80	90.29
Worker 2	32.37	82.35
Worker 3	47.74	81.64
Worker 4	56.37	95.83
Worker 5	88.40	90.70
Worker 6	59.73	95.83
Worker 7	45.31	69.19
Worker 8	36.91	73.69
Worker 9	50.48	85.93
Worker 10	72.20	95.72
Worker 11	49.41	81.99
Worker 12	53.97	92.81
Supervisor 1	60.93	100.00
Supervisor 2	67.59	98.11
Supervisor 3	60.29	94.32

5 Discussion

This study performed a systematic investigation of visualization platform-based CST delivery methods based on suitability and efficacy for the construction industry. A total of 21 practices were considered for the investigation purpose. As a next step, this study developed three visualization platform-based CST modules such as Image, VT, and MVR. According to the data analysis, 64% of the construction professionals favored the VT module as a suitable CST module for the industry. Construction professionals have preferred the VT based on time (quick safety training compared to Image and MVR), cost (cost-effective compared to MVR), ease to use compared to MVR, and little trainer intensive [19].

Further, the levels of hazard (percentage weights) were evaluated with the help of experienced construction professionals' inputs. The effectiveness of the VT CST module was investigated through workers and supervisors. The lowest HRS for workers pre-training was 32.37%, and the highest HRS was 88.40%. The post-training lowest HRS for workers was 69.19%, and the highest HRS was 95.83%. On average, the pre-training HRS for workers was 55.31%, and the post-training HRS was 86.33%, as shown in figure 9. The pre-training lowest HRS for supervisors was 60.29%, and the highest HRS was 67.59%. The post-training lowest HRS for supervisors was 94.32%, and the highest score was 100.00%. On average, the pre-training HRS for supervisors was 62.93%, and the post-training HRS was 97.47%, as shown in figure 9. The overall pre-training HRS was 56.83%, and post-training HRS was 88.56%, as shown in figure 10.

Broadly, the results revealed that with a single training session of seven minutes, the HRS for the workers was enhanced by 31.02%, for supervisors by 34.54%, and overall by 31.73%. The statistical analysis also supported noticed HRS enhancement and concluded that pre-training and post-training scores are significantly different at the confidence interval of 95%.

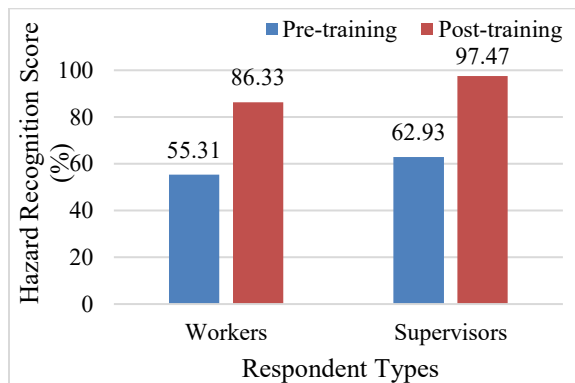


Figure 9. Pre-training and post-training HRS of workers and supervisors

This study also noticed interesting findings that construction professionals identified 56.83% of hazards, and about 43.17% of hazards were unrecognized. These findings are somewhat in line with the past study highlighting that about 50% of hazards remain unrecognized in the construction environment [5]. In fact, there is an improvement in the HRS and a percentage reduction in unrecognized hazards, which is a good sign for the construction industry. To sum it up, this study noticed that VT CST is a suitable safety training delivery method for the industry and an effective safety training method for knowledge transfer.



Figure 10. Overall HRS for pre-training and post-training HRS assessment

6 Conclusion

This study aimed to identify a suitable and effective visualization platform-based safety training module for the construction industry. As a first step, three visualization platform-based CST modules were developed, and VT was identified as a suitable safety training module for the construction industry. Here, the suitability of the modules presented an initial Next, the effectiveness of the VT CST module was evaluated. The efficacy of the VT CST was confirmed based on significantly ($p < 0.00$) enhanced HRS from 56.83% to 88.56% using single-time safety interventions.

As mentioned earlier, findings on the suitability of the CST modules need to be sharpened in the future with more in-depth investigations. Next, recently, 360-degree panorama, Mixed reality, and other visualization platform-based safety training methods have been introduced to the construction safety domain. Future research studies can focus on mentioned advanced safety training delivery methods for further in-depth investigation. The theoretical contribution of the study is to perform a comparative analysis of advanced visualization platform-based CST modules and identify suitable and effective safety training delivery and knowledge transfer module for the construction industry. The practical contribution of the study is construction professionals can develop and adopt personalized VT CST modules for training construction stakeholders and to enhance their hazard recognition capabilities and safety performance of the project.

7 References

- [1] Jeelani I., Albert A., Azevedo R., and Jaselskis E. J. Development and Testing of a Personalized Hazard-Recognition Training Intervention. *Journal of Construction Engineering and*

- Management*, 143 (5): 04016120, 2017.
- [2] Sekar G., Sambasivan M., and Viswanathan K. Does size of construction firms matter? Impact of project-factors and organization-factors on project performance. *Built Environment Project and Asset Management*, 11(2):174–194, 2021.
 - [3] Lingard H. Occupational health and safety in the construction industry. *Construction Management and Economics*. 31(6):505–514, 2013.
 - [4] Okorie V. N. and Musonda I. An investigation on supervisor's ability and competency to conduct construction site health and safety induction training in Nigeria. *International Journal of Construction Management*. 20(5):357–366, 2020.
 - [5] Jeelani I., Han K., and Albert A. Development of virtual reality and stereo-panoramic environments for construction safety training. *Engineering, Construction and Architectural Management*, 27(8):1853–1876, 2020.
 - [6] Patel D. A. and Jha K. N. An Estimate of Fatal Accidents in Indian Construction. In *Proceedings of the 32nd Annual Association of Researchers in Construction Management (ARCOM) Conference*, pages 577–587, Manchester, United Kingdom, 2016.
 - [7] Bhandari S. and Hallowell M. R. Emotional Engagement in Safety Training: Impact of Naturalistic Injury Simulations on the Emotional State of Construction Workers. *Journal of Construction Engineering and Management*. 143(12):04017090, 2017.
 - [8] Heinrich H. *Industrial accident prevention approach*, Second Edi. New York: McGraw-Hill, 1980.
 - [9] Chen D. and Tian H. Behavior based safety for accidents prevention and positive study in China construction project. *Procedia Engineering*, 43: 528–534, 2012.
 - [10] Kalteh H. O., Mortazavi S. B., Mohammadi E., and Salesi M. The relationship between safety culture and safety climate and safety performance: a systematic review. *International Journal of Occupational Safety Ergonomics*, 2019. doi: 10.1080/10803548.2018.1556976.
 - [11] Sacks R., Perlman A., and Barak R. Construction safety training using immersive virtual reality. *Construction Management and Economics*. 31(9):1005–1017, 2013.
 - [12] Albert A., Hallowell M. R., and Kleiner B. M. Enhancing Construction Hazard Recognition and Communication with Energy-Based Cognitive Mnemonics and Safety Meeting Maturity Model: Multiple Baseline Study. *Journal of Construction Engineering and Management*, 140(2):04013042, 2014.
 - [13] Carter G. and Smith S. D. Safety Hazard Identification on Construction Projects. *Journal of Construction Engineering and Management*, 132(2):197–205, 2006.
 - [14] Albert A., Hallowell M. R., and Kleiner B. M. Experimental field testing of a real-time construction hazard identification and transmission technique. *Construction Management and Economics*, 32(10):1000–1016, 2014.
 - [15] Haslam R. A. *et al.* Contributing factors in construction accidents. *Applied Ergonomics*, 36, 401–415, 2005.
 - [16] Cromwell S. E. and Kolb J. A. An examination of work-environment support factors affecting transfer of supervisory skills training to the workplace. *Human Resource Development Quarterly*, 15(4):449–471, 2004.
 - [17] Li H., Lu M., Chan G., and Skitmore M. Proactive training system for safe and efficient precast installation. *Automation in Construction*, 49:163–174, 2015.
 - [18] Li X., Yi W., Chi H. L., Wang X., and Chan A. P. C. A critical review of virtual and augmented reality (VR/AR) applications in construction safety. *Automation in Construction*, 86:150–162, 2018.
 - [19] Bhagwat K., Kumar P., and Delhi V. S. K. Usability of Visualization Platform – Based Safety Training and Assessment Modules for Engineering Students and Construction Professionals. *Journal of Civil Engineering Education*, 147(2):04020016, 2021.
 - [20] Eiris R., Jain A., Gheisari M., and Wehle A. Safety immersive storytelling using narrated 360-degree panoramas: A fall hazard training within the electrical trade context. *Safety Science*, 127:104703, 2020.
 - [21] Luo X., Li H., Dai F., Cao D., Yang X., and Guo H. Hierarchical Bayesian Model of Worker Response to Proximity Warnings of Construction Safety Hazards: Toward Constant Review of Safety Risk Control Measures. *Journal of Construction Engineering and Management*, 143 (6): 1–11, 2017.

- [22] Yang K. and Ahn C. R. Inferring workplace safety hazards from the spatial patterns of workers' wearable data. *Advanced Engineering Informatics*, 41:100924, 2019.
- [23] Namian M., Albert A., Zuluaga C. M., and Jaselskis E. J. Improving Hazard-Recognition Performance and Safety Training Outcomes: Integrating Strategies for Training Transfer. *Journal of Construction Engineering and Management*, 142(10):1–11, 2016.
- [24] Loosemore M. and Malouf N. Safety training and positive safety attitude formation in the Australian construction industry. *Safety Science*, 113:233–243, 2019.
- [25] Wilkins J. R. Construction workers' perceptions of health and safety training programmes. *Construction Management Economics*, 29: 1017–1026, 2011.
- [26] Sinnott W. R. The application of safety teaching to practical work in further education establishments: With reference to basic engineering training courses. *Journal of Occupational Accidents*, 1(1):69–84, 1976.
- [27] Nykänen M. *et al.*. Implementing and evaluating novel safety training methods for construction sector workers: Results of a randomized controlled trial. *Journal of Safety Research*, 75, 205–221, 2020.
- [28] Teizer J., Cheng T., and Fang Y. Location tracking and data visualization technology to advance construction ironworkers' education and training in safety and productivity. *Automation in Construction*, 35, 53–68, 2013.
- [29] Clevenger C., López del Puerto C., and Glick S. Interactive BIM-enabled Safety Training Piloted Construction Education. *Advances in Engineering Education*, 1–14, 2015.
- [30] Eiris R., Gheisari M., and Esmaeili B. Desktop-based safety training using 360-degree panorama and static virtual reality techniques: A comparative experimental study. *Automation in Construction*, 109, 102969, 2020.

Semiarid Terrain Alteration for Converting Dryland into Arable Land – Construction and Earthmoving Perspectives

M. Alamaro¹, J. Louis² and J. Teizer³

¹ Department Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, USA

² School of Civil and Construction Engineering, Oregon State University, USA

³ Department of Civil and Mechanical Engineering, Technical University of Denmark, Denmark

E-Mails: alamaro@alum.mit.edu, joseph.louis@oregonstate.edu, teizerj@byg.dtu.dk

Abstract –

Close to 15% of the world's surface area is semiarid where 1.1 billion of the most vulnerable people on earth live. In contrast to arid, semiarid regions have substantial rainfall albeit it quickly evaporates and therefore vegetation and biomass growth are limited. A new concept calls for terrain alteration by earth moving to construct north-south slopes where on the northern slope in the northern hemisphere solar irradiation is reduced, causing in turn a reduction in evapotranspiration (ET), allowing for cultivation. The dimension of the slopes (or ridges) is on the order of 5-10 meters and its slope angle is in the range of 15-20 degrees. This paper provides the motivation for such development and an overview of the necessary research and development for implementation. It also focuses specifically on the topic of earthmoving operations that are involved, the equipment it requires, and discusses the potential for employing robotics and automated construction methods to improve the economic and environmental feasibility of such an endeavor.

Keywords –

Cultivation; Earth moving equipment; Food and agriculture; Automatic and robotic construction

1 Introduction

The world arable land in 2020 has been reduced to 44% of the arable land that was once available in 1960 (Figure 1). In the meanwhile, the world population has increased during this period by a factor of 2.63. This means that in 2020 each unit of arable land must produce $2.63/0.44 = 5.97$ times more in comparison to a unit of land in 1960 and such an increasing productivity is hard to achieve. Subsequently, products like highly efficient fertilizers or weed killers have emerged, yielding gains in cultivation but impacting the environment. Therefore, to assure world food security more arable land is required. This

could be achieved by creating a new “Green Revolution 2.0” for the semiarid world by the alteration of its terrain to enable a new extensive agriculture mode by retaining rainfall through reduction of ET. The artistic conception of earthmoving and construction of the slopes can be seen in Figure 2. As the proposed concept will be explained in the subsequent chapters, it imitates vegetation patterns that exist in nature. We hereby contribute an overview and a preliminary assessment from the construction and earthmoving point of view.

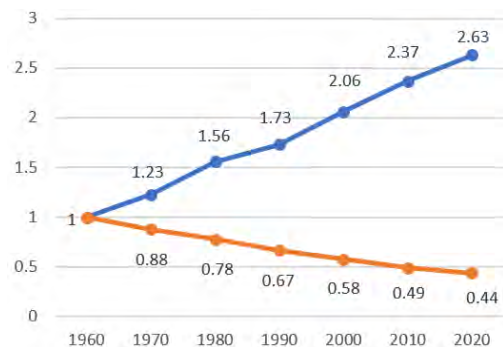


Figure 1. Increase in world population by 2.63 (blue curve) and decrease of arable land to 44% (orange curve), 1960-2020 [1].



Figure 2. Artistic conception of earthmoving and construction of the North-South slopes.

This paper presents first the related work, then introduces the topics of the design and the construction of the slopes, the suitability of sites for application, and the factors and degrees how construction can be automated and robotized. The paper concludes with a brief discussion on potential impact and an outlook of future work.

2 Related work

Research on semiarid agriculture is limited in comparison to that of agriculture in temperate zones [2, 3]. Only a few industrialized countries have significant semiarid cultivation expertise such as in Australia, US Southwestern United States and Israel. There, semiarid land is used intensively in agriculture because large amounts of resources such as water, labor and fertilizer cannot be used per land unit. In contrast, however, most of the world's food is produced by extensive agriculture using large, accessible land areas where the output provides human food and animal feed.

As shown in Figure 3, the Northern slope in the northern hemisphere is lush, moist and green; the Southern slope is dry and unusable. Only the Northern slope is potentially cultivable. The proposed concept imitates vegetation patterns on natural north-south slopes.

Moreover, consider the following example (Figure 4): The annual precipitation in London, UK is 24 inches, while the annual precipitation in Dallas, TX is 35 inches. And yet London (51 degrees latitude) is green and lush while Dallas at 31.5 degrees is semi-arid. A small difference in latitude and solar irradiation flux can cause a dramatic difference in vegetation and biomass productivity.

The concept outlined hereby is new and requires research and development on agronomic, plant and soil sciences, hydrology, climate and rainfall patterns, and atmospheric land interaction. Furthermore, the main variables for a return on investment analysis (ROI) are the cost to construct the slopes and the water “cost avoidance” by the water gain by rainfall retention. This paper explores the technologies that could be used to create and construct the slopes to provide cost-effectiveness for new semiarid arable land that can be cultivated and yield results as shown in Figure 5.

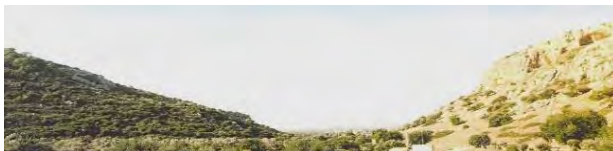


Figure 3: Typical natural North-South slopes in semiarid areas.



Figure 4: Public Work, San Angelo, Texas. Grass appeared spontaneously on the Northern slope two weeks after creation.



Figure 5: Examples of cultivation on the Northern slope.

3 Man-made slopes

3.1 Slope design

The Northern and Southern slopes' widths are divided in a ratio of 2:1 or 3:1 in order to maximize the cultivable northern slope area (Figure 6). The Southern slope is steeper and therefore, to prevent soil erosion it will be covered by a plastic sheet. The sloping angle of the Northern slope will be determined by the "necessary" reduction of evapotranspiration required for different climates and crops intended for cultivation.



Figure 6: Cross-sectional view: Steep southern slope is covered with a plastic sheet to prevent soil erosion; the topsoil below the plastic sheet is scrubbed and placed on the northern slope.

The amount of earth moved by the construction of the slopes is proportional to the cross-section area of the slope triangles, which is proportional in turn to the square of the dimensions of the slope. To minimize the amount of earth moved and its associated cost, narrow slopes are preferable. But from the agronomic point of view wider slopes are more practical. In this effort we will determine the optimal dimensions to address these considerations.

The reduction of solar energy on the slope is expected to reduce evapotranspiration (ET). A shallow sloping angle will not produce enough reduction in solar input and reduction of ET. On the other hand, an overly steep angle will excessively reduce solar radiation and photosynthesis and also increase soil erosion and water runoff. An optimal altered ET will define the sloping angle for the "necessary" reduction of ET in comparison to that on flat terrain. The necessary ET will depend on local climate, growing seasons, and intended crops.

Any reduction of ET will decrease the difference between local precipitation and ET and the irrigation water requirement. The ideal case is when the new altered ET is less or equal to local precipitation. The altered new ET will determine the required sloping angle.

3.2 Relevant other factors

One concern is that the sloping terrain's soil will erode, requiring frequent correction and earthmoving. The steeper Southern slope, which is more susceptible to erosion, could be covered by a plastic sheet. As for the northern slope, we can learn from how hillsides are stabilized for crops such as grapes and winter wheat (Figure 7).

The topsoil on the Southern slope below the plastic sheet could be scrubbed and placed on the Northern slope.

Also, if soil is eroded on the Northern slopes, it is not necessarily lost; it will accumulate in the areas between the slopes and if necessary, it could be spread back on the slope every few years.

Mathematical models of soil erosion and water runoff are difficult to create and run with acceptable accuracy, so these topics could be studied based upon the vegetation grown on natural slopes.

Danger to wildlife, birds and insects will be assessed and passage for desert animals will be created every few slopes to allow easy passage. An assessment of the impact on biotics will also be done by the agronomists and soil scientists taking part in this program.



Figure 7: Practiced hillside cultivation demonstrates how soil erosion and water runoff are prevented in the cultivation of natural slopes and can be implemented in the construction of man-made slopes.

4 Survey of potential sites

4.1 Geoinformation

Detecting differences between vegetation growth on natural sloping terrains with different inclinations could be done using a space-borne Synthetic Aperture Radar (SAR). Differences in vegetation between adjacent areas can be detected using SAR remote sensing techniques. This may be done for testing and demonstrating that indeed natural sloped terrains have different vegetation covers [4].

Spaceborne SAR has been used extensively to map and monitor vegetation since the 1990s, when the first SAR missions were launched. SAR sensors are active systems that transmit a coherent radar signal to the Earth's surface and measure the characteristics of the response backscattered from the target area. The interaction of the radar signal with vegetation is a function of the vegetation type and the radar wavelength. Studies of forest biomass are typically carried out using longer wavelengths such as L-band as the signal is able to penetrate the canopy and interact with the branches and trunks of the forest [5]. Shorter wavelengths (e.g., X- and C-band) are less able to penetrate vegetation as they interact with leaves and smaller branches.

4.2 Relationship between topography and vegetation growth

Vegetation growth is affected by many factors such as soil moisture, air temperature, light, nutrients, soils, competition, predation, disturbance, species composition, and more. Variations in many of these variables are associated with attributes of local topography such as aspect, elevation, slope and inclination, as well as climate drivers. Topography strongly affects the distribution of insolation. Patterns of incoming solar radiation affect energy and water balances within a landscape, resulting in changes in air temperature, humidity and soil moisture, which in turn impact vegetation attributes.

In the Northern hemisphere, it is common sense that a southwestern slope is sunnier, hotter and drier than a northeastern slope because the apex of the sun is perpendicular to south-facing slopes. From the air, it is often easy to see the denser vegetation on north-facing slopes. However, in some parts of the world with plenty of moisture, vegetation may prefer south-facing slopes where they can thrive in the relative warmth.

Other factors such as climate, elevation, and species composition may further complicate these relationships. Jin et al. [6], for example, showed for the Qilian Mountain area of China that such aspects have a large impact on the vegetation at certain slopes and elevations due to changes in evapotranspiration. Figure 8 shows such a complex relationship between elevation, aspect, and Normalized Difference Vegetation Index (NDVI) based on remote sensing and elevation data. The NDVI is a dimensionless index that describes the difference between visible and near-infrared reflectance of vegetation cover and can be used to estimate the density of green on an area of land.

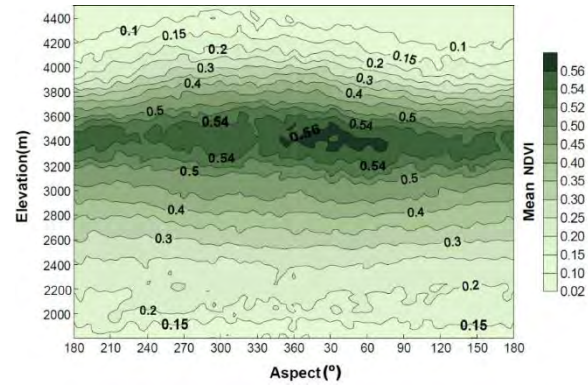


Figure 8: The change of the mean NDVI values with elevation and aspect in the northern part of Qilian Mountain. A Gaussian smooth filter was used and a low pass convolution was performed on the grid data to present a more consistent and smoother map. Note: a refiner scale (0.02) was used when the NDVI value is larger than 0.5 [6].

5 Construction of the slopes

The cost of earth moving cannot be compared to the cost of earthmoving done on relatively small areas such as for the foundations of buildings and road construction. In these operations earth is dug by bulldozers and excavators and is transported to certain distances by trucks.

In contrast, the proposed terrain alteration does not require actual moving of earth (Figure 9). The average level of the terrain remains the same as before the operation. It requires relatively shallow digging and landscape alteration without transporting earth. Despite this relatively localized work scope, the areal spread of the project makes earth moving a very important consideration. Furthermore, given the sustainability goals that are being targeted by the conversion of arid to semi-arid land, the earthmoving operations need appropriate planning to ensure that their cost and impact on environment are minimized in order to ensure the long-term economic and environmental feasibility of this project. Towards that end the following aspects are proposed for planning the earthmoving in this paper: (1) earthwork planning and analysis, (2) equipment selection, (3) construction and maintenance operations, (4) automation of operations, and (5) public perception. These topics are discussed in the following sections.



Figure 9: Earthmoving for north-south slopes construction.

5.1 Earthwork planning and analysis

The primary metrics of interest towards obtaining the cost estimate and schedule for the re-slope operations is determining the following metrics about earthwork: volume of earth to move, haul distances, and haul grades. These will inform the selection of equipment fleets and the total time and cost for project completion. A mass haul diagram will be employed for earthwork planning to obtain these measures.

- *Earthwork planning with mass haul calculations:* In order to convert a terrain of arbitrary topography to a sloped terrain, the contractor determines the amount of earth to be moved and the distance over which such movement occurs. This requires capturing the Digital Terrain Model (DTM) of the area, preferably using an aerial LIDAR (light detection and ranging) survey; and overlaying it with the proposed sloped terrain to identify where the cut and fill sections of the project would be. For purposes of planning, the entire area would be subdivided into a suitable grid (e.g., 5x5 meters). The difference between existing and planned terrain will provide an identification of cut and fill grid sections, cut and fill elevation differences, and volume to be moved between grid sections. These factors will inform the choice of equipment to be employed.
- *Contract administration:* The contract for the earthwork will expect to be paid on a unit-cost basis per unit volume of soil. The exact amount of the unit cost will be determined based on existing project site conditions.
- *Operation time:* Because the work is expected to be conducted in a semi-arid location, the most effective time for performing the primary earthwork tasks would be during or (preferably) just after the wet season. This is because the rainwater that has

permeated the soil can help improve the plasticity of soil, and thereby ease the workload on the earthwork equipment.

5.2 Equipment selection and development

The primary earthwork tasks that are expected to be performed are earthmoving, ploughing, grading to the desired slopes. Given that these projects will mostly occur in relatively non-undulating terrain, it is expected that there will be little need for moving earth across grid sections. Thus the primary equipment needs will be for ploughing the soil for mixture of nutrients and grading.

- *Soil Grading:* Thus, the primary equipment for shaping the terrain into slopes is expected to be motor-graders after the soil has been sufficiently ploughed and infused with nutrients. Typical motor grader blade widths range within a few meters.
- *Ploughing the soil:* The task of deep-ploughing (up to 2 m depth) could be done with rippers attached to dozers, while shallow plowing (up to 0.3 m depth) and mixing could be done with the scarifier implement attached to the motor-grader. Given that it is expected to require ploughing to greater depths, custom-built ploughs and implements that attach to the grader or a dozer could be considered.
- *Soil Stabilization:* It may be required to perform methods to stabilize the soil to maintain the slopes. This will be done by mechanical means, rather than chemical due to the need to grow crops on the slopes.

Given the unique nature of the project, there is quite a considerable potential for the development of novel equipment: for example size and scalability are expected that this project can necessitate the development of novel implements for deep ploughing, nutrient mixing, and large-width grading.

5.3 Construction and maintenance of slopes

The primary construction of slopes will be performed using motor graders that are equipped with APG-enabled automated machine guidance and control (AMG and AMC). This technology uses real-time data from GPS sensors of receivers attached to the graders moldboard to automatically adjust the angle and pitch of the moldboard to obtain the desired shape of slope in the given terrain. This technology augmented the operator's capabilities and ensures adheres to plans that are uploaded to the motor graders AMG system.

Apart from the initial construction of the slopes, annual maintenance may be required to offset any soil erosion caused by wind or other forces. It is expected that for a slight grading a single grader will suffice for an entire area (e.g., 10 square kilometers). Agricultural

operations related to sowing and harvesting can be performed using conventional agricultural equipment.

5.4 Automation of operations

Since the majority of earth moving contractors already have difficulty finding skilled equipment operators or turn down work, it is posited that a project of such large scope and relatively simplistic operations in remote locations provides the ideal testing and proving grounds for the automated construction equipment and their operations. Apart from the previously mentioned automation methods such as AMG and AMC for grader control that will be heavily used in this project, these operations will also enable the testing of autonomous graders and soil stabilizers.

Such automated operations could use robotic graders (such as are used for road leveling or snow plowing). Robots can solve the problem because automatic steering of vehicles in this use case does not require to frequently turn, stop or drive in reverse; they can cruise relatively fast in straight lines making rapid inexpensive terrain alteration and do this automatically in repetitive cycles of typical operations. This eliminates waste like it is common in general construction projects (e.g., schedule and cost overruns, over-allocation of resources, not optimal maintenance tasks). Likewise, high demand is set on a variety of technology that needs to safely steer the vehicles in rough(er than expected) terrain. Safety in equipment operation as it relates to detecting pedestrians, animals, or other obstructions close or nearby to it, is always a concern and specifically when it comes from the exploration or use of driverless machines [8]. There are numerous other challenges in the automation of earth-moving machines [9, 10], including but not limited to emissions that are generated by fossil fuel-powered engines [11].

5.5 Public perception of operations

Given the large-scale nature of operations involved in natural terrain, it will be necessary to clearly communicate the cost and benefits of this project to the public. Metrics such as construction duration and costs, emissions produced, along with the fruits of the project such as crops grown must be clearly understandable to the public. Towards this end, novel tangible user interfaces of Augmented Reality (AR) sandboxes to plan the project and visualize the metrics of interest on the project in response to user input [12].

Previous experiences and works in developing earthwork calculation applications for the AR sandbox for highway construction (Figure 10) and traffic analysis purposes can inform non-literate bystanders of the development. For example, novel visualizations and tangible interactive planning tools will result for the

project. These can serve as a public interaction and dissemination tool to engage with the public and improve their perception about this project while sensors on unmanned aerial vehicles (UAVs) provide run-time project state data [13] that can be used in digital terrain modeling or earthwork planning [14-15] and earthwork progress [16] or productivity analysis [17-19].

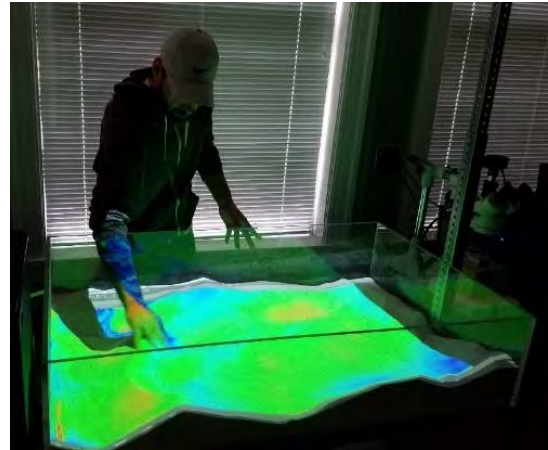


Figure 10: Interactive terrain visualization using a AR sandbox.

6 Conclusion

15% of the world surface area is semi-arid where 1.1 billion people live. In the Northern hemisphere the North-South slopes have the potential to produce jobs, to promote economic development and to provide food security for developing countries. In some cases, agribusiness corporations are likely to invest, buy land, upgrade and cultivate it for export.

The presented work is a preliminary study on revitalizing the concept of North-South slopes for the application of converting dryland into arable land. Explanations to the design and construction methods were presented as they relate to the topic of automatizing and robotizing earth movement equipment in construction. A variety of topics like the impact on converting naturally-built states in the environment and the potential impact on climate and food supply were touched upon but deserve much more attention in future studies.

References

- [1] Higgins, G.M. and Kassam, A.H. (1981). "The FAO agro-ecological zone approach to determination of land potential." *Pedologie*, 31, 147–168.
- [2] Scholes, R.J. (2020). "The Future of Semi-Arid Regions: A Weak Fabric Unravels." *Climate*, Vol.

- 8(3), 43, <https://doi.org/10.3390/cli8030043>.
- [3] Ricart, S., Villar-Navascués, R.A., Hernández-Hernández, M., Rico-Amorós, A.M., Olcina-Cantos, J., Moltó-Mantero, E. (2021). "Extending Natural Limits to Address Water Scarcity? The Role of Non-Conventional Water Fluxes in Climate Change Adaptation Capacity: A Review." *Sustainability*, 13(5), 2473, <https://doi.org/10.3390/su13052473>.
- [4] Desta, F., Colbert, J.J., Rentch, J.S. Gottschalk, K.W. (2004). "Aspect induced differences in vegetation, soil, and microclimatic characteristics of an Appalchian watershed." *Castanea*, 69(2): 92-108, https://www.fs.fed.us/nrs/pubs/jrnl/2004/ne_2004_desta_001.pdf.
- [5] Beaudoin, A., Le Toan, T., Goze, S., Nezry, E., Lopes, A., Mougin, E., Hsu, C.C., Han, H.C., Kong, A., Shin, T. (1994). "Retrieval of forest biomass from SAR data." *International Journal of Remote Sensing*, 15(14), 1994, <https://doi.org/10.1080/01431169408954284>.
- [6] Jin, X.M., Zhang, Y.K., Schaepman, M.E., Clevers, J.G.P.W., Su, Z. (2008). "Impact of elevation and aspect on the spatial distribution of vegetation in the Qilian mountain area with remote sensing data." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS)*, 1385-1390, <http://dx.doi.org/10.5167/uzh-77426>.
- [7] Rabus, B., Eineder, M., Roth, A., Bamler, R. (2003). "The shuttle radar topography mission - a new class of digital elevation models acquired by spaceborne radar." *ISPRS Journal of Photogrammetry and Remote Sensing*, 57(4), 241-262, [https://doi.org/10.1016/S0924-2716\(02\)00124-7](https://doi.org/10.1016/S0924-2716(02)00124-7).
- [8] Golovina, O., Teizer, J., Johansen, K.W., König, M. (2021). "Towards Autonomous Cloud-based Close Call Data Management for Construction Equipment Safety." *Automation in Construction*, 132, 103962, <https://doi.org/10.1016/j.autcon.2021.103962>.
- [9] S. Dadhich, S., Bodin, U., Sandin, F., Andersson, U. (2018). "From tele-remote operation to semi-automated wheel-loader." *Int. J. Electr. Electron. Eng. Telecommun.*, 7(4), 178-182, <https://doi.org/10.18178/IJEETC.7.4.178-182>.
- [10] Dadhich, S., Bodin, U., Andersson, U. (2016). "Key challenges in automation of Earth-moving machines." *Automation in Construction*, 68, 212-222, <https://doi.org/10.1016/j.autcon.2016.05.009>.
- [11] Teizer, J., Wandahl, S. (2022). "Simplified Emissions Measurement System for Construction Equipment." *Construction Research Congress*, Arlington, Virginia, USA, 474-482, <https://doi.org/10.1061/9780784483961.050>.
- [12] Joseph Louis, J., and Lather, J. (2020). "Augmented Reality Sandboxes for Civil and Construction Engineering Education." *Proceedings of the 37th International Symposium on Automation and Robotics in Construction*, 750-756, <https://doi.org/10.22260/ISARC2020/0104>.
- [13] Siebert, S., Teizer, J. (2013). "Mobile 3D mapping for surveying earthwork using an unmanned aerial vehicle (UAV)." *Proceedings of the 30th International Symposium on Automation and Robotics in Construction*, Montreal, Canada, August 11-15, 2013, <https://doi.org/10.22260/ISARC2013/0154>.
- [14] Teizer, J., Green, A., Hilfert, T., Perschewski, M., König, M. (2017). "Mobile Point Cloud Assessment for Trench Safety Audits." *34th International Symposium on Automation and Robotics in Construction*, Taipei, Taiwan, <https://doi.org/10.22260/ISARC2017/0021>.
- [15] Teizer, J. (2007). "Rapid Surveillance of Trenches for Safety." *Proceedings of the Construction Research Congress*, Freeport, Bahamas, May 6-8, 2007.
- [16] Bügler, M., Borrmann, A., Vela, P.A., Teizer, J. (2017). "Fusion of Photogrammetry and Video Analysis for Productivity Assessment of Earthwork Processes", *Computer-Aided Civil and Infrastructure Engineering*, Wiley, 32(2), 107-123, <http://doi.org/10.1111/mice.12235>.
- [17] Pradhananga, N., Teizer, J. (2013). "Construction Site Population Density Analysis using GPS Technology for Automated Safety and Productivity Assessment in Earth-Moving Operations." *13th International Conference on Construction Applications of Virtual Reality*, London, October 30-31, 2013.
- [18] Pradhananga, N., Teizer, J. (2014). "Development of a Cell-based Simulation Model for Earthmoving Operation using Real-time Location Data." *2014 Construction Research Congress*, Atlanta, Georgia, May 19-21, 2014, <https://doi.org/10.1061/9780784413517.020>.
- [19] Pradhananga, N. and Teizer, J. (2015). "Cell-based Construction Site Simulation Model for Earthmoving Operations using Real-time Equipment Location Data", *Visualization in Engineering*, Springer Verlag, 3:12, <http://dx.doi.org/10.1186/s40327-015-0025-3>.

Action Sequencing in Construction Accident Reports using Probabilistic Language Model

Q. Do^a, T. Le^a and C. Le^b

^aThe Glenn Civil Engineering Department, Clemson University, USA

^bDepartment of Civil, Construction and Environmental Engineering, North Dakota State University, USA

E-mail: gdo@clemson.edu, tuyenl@clemson.edu, chau.le@ndsu.edu

Abstract –

Construction remains among the most hazardous workplaces, thus a significant amount of time and effort in reporting and investigating the accident occurrences has been done in the past decades by government agencies. In light of construction safety, analyzing textual information in construction accident records may assist in our comprehension of past data and be used to minimize future risks. Many attempts have been made in previous studies to identify causes and related entities but yet consider worker activities and behaviors. This study presents a framework that adopts a Probabilistic Language Model to sequence actions taken by workers that depict construction scenarios from unstructured accident narrative reports. The proposed approach achieved outstanding performances with the highest sequence accuracy and pairwise sequence accuracy of 84.81 % and 89.12%, respectively. Moreover, an action sequence database that can explain the relationship between workers' actions was created. This research is anticipated to contribute to enhancing understanding and establishing safety management systems to actively forecast and prevent accidents.

Keywords –

Construction safety; Accident reports; Probabilistic Language Model; Natural Language Processing

1 Introduction

Construction sites remain among the most hazardous work environments for laborers. Despite several attempts driven by this relatively low-performance level, accident statistics have never really improved appreciably over the last decade [1]. These accidents raised major health and safety issues and substantial financial loss [2]. Hence, it is critical to gain a deeper insight into construction accidents to enhance safety performance. Over time, a vast amount of detailed information in the form of data

would be gathered. Accident reports are essential but underutilized due to the difficulty of extracting data from unstructured text. Due to the critical significance of accident reports, more focus has lately been put on ensuring the reliability of the data collected and report organization. Construction accident reports are valuable sources of information, and the process of assessing them may provide critical insight into previous events to avoid future recurrences.

Numerous scholars have worked to build automated models accompanied with Natural Language Processing (NLP) that might be used to analyze textual data contained in construction accident reports with the minimum human intervention. Several studies [3]–[8] adopted a supervised approach to classifying the causes, types of injury, and injured body parts. Besides, another study [9] examined the efficacy of an unsupervised approach for clustering accident reports. Afterward, the authors monitored the results and retrieved pertinent information about the objects and factors contributing to the incidents. The studies, as mentioned above, have achieved outstanding success, and their findings have made significant contributions to advancing safety knowledge and improving safety plans. However, a crucial factor that has not been taken into consideration is the worker's sequence of actions that reflects the construction scene at the time of an accident.

The key motivation for this paper is to address that research gap. This study adopted the Probabilistic Language Model to develop an Action Sequencing Model that can sequence the actions taken by workers from unstructured accident reports obtained from the Occupational Safety and Health Administration (OSHA). Accident reports were split into separate sentences, and each sentence was annotated to a particular action label before feeding as input for the model. The major contribution of this work is an action sequence database that can explain the relationship between actions showing the scene of the construction accident. Exploring this database can help widen the horizons and develop a safety management system to predict and prevent catastrophes actively.

2 Background

2.1 Probabilistic Language Model

Models that can be utilized to assign a probability to a sentence or a sequence of words are called Probabilistic Language Models [10]. Probabilistic Language Models have been employed in a variety of research fields to date in numerous NLP applications, such as Handwriting Recognition [11], Machine Translation [12], Speech Recognition [13], Spelling Correction [14], and Information Retrieval [15]–[17]. N-gram models, commonly referred to as Markov models, are detailed in the following section.

2.1.1 N-Gram Language Models

Given a sequence of words $W(w_1, w_2, \dots, w_n)$, a model that calculates the probability of either $P(W)$ or $P(w_n | w_1, w_2, \dots, w_{n-1})$ is called a Probabilistic Language Model.

To decompose these probabilities, the chain rule of probability is applied. The chain rule of probability is a theory that allows calculating any member of a joint distribution of random variables using conditional probabilities. Given n event (i.e., x_1, x_2, \dots, x_n), the probability $P(x_1, x_2, \dots, x_n)$ is

$$P(x_1, x_2, \dots, x_n) = P(x_1)P(x_2 | x_1) \dots P(x_n | x_1, x_2, \dots, x_{n-1}) \quad (1)$$

The sequence event x_1, x_2, \dots, x_n can be represented as $x_{1:n}$. The equation (1) is rewritten:

$$\begin{aligned} P(x_{1:n}) &= P(x_1)P(x_2 | x_1) \dots P(x_n | x_{1:n-1}) \\ &= \prod_{k=1}^n P(x_k | x_{1:k-1}) \end{aligned} \quad (2)$$

Applying the chain rule to the sequence of words W :

$$\begin{aligned} P(w_{1:n}) &= P(w_1)P(w_2 | w_1)P(w_3 | w_{1:2}) \dots P(w_n | w_{1:n-1}) \\ &= \prod_{k=1}^n P(w_k | w_{1:k-1}) \end{aligned} \quad (3)$$

The chain rule emphasizes the link between calculating the joint probability of a sequence and computing the conditional probability of a word given previous words. Equation (3) suggests estimating the joint probability of an entire sequence of words by multiplying the number of conditional probabilities together. However, it is challenging to calculate the exact probability of a word given a long sequence of preceding words $P(w_n | w_{1:n-1})$.

The assumption that the probability of a word depends solely on the preceding word(s) is known as the Markov assumption. Markov models, also known as N-gram models, are the class of probabilistic models that presume that we can estimate the probability of some future items without referring too far into the past [18]. Then we approximate the probability of a word given its entire context as follows:

$$P(w_n | w_{1:n-1}) \approx P(w_n | w_{n-N+1:n-1}) \quad (4)$$

What method do we use to calculate N-gram probabilities? An intuitive method to estimate probabilities is called Maximum Likelihood Estimation (MLE). We obtain the MLE estimation for the parameters of an N-gram model by getting counts from a corpus and normalizing the counts so that they lie between 0 and 1 [10].

$$P(w_n | w_{n-N+1:n-1}) = \frac{C(w_{n-N+1:n-1} w_n)}{C(w_{n-N+1:n-1})} \quad (5)$$

Equation (5) estimates the N-gram probability by dividing the observed frequency of a particular sequence by the observed frequency of a prefix. This is known as a relative frequency ratio.

Because probabilities are less than or equal to one, multiplying probabilities together results in a smaller product. In practice, using log probabilities rather than raw probabilities can assist obtain figures that are not as small.

2.1.2 Smoothing Techniques

Smoothing is a technique for creating an approximation function that tries to capture essential patterns in data while eliminating noise and other fine-scale structures/rapid events [19]. In Probabilistic Language Model, the MLE of probabilities generally results in overfitting training data and poor performance on unseen data. It is preferable to utilize smoothed estimates of these values instead [20]. In some cases, an N-gram is never observed in the training data, resulting in the zero probability of a sequence of words. To avoid the model from assigning 0 probability to these unseen items, we must take a bit of probability from some more frequent items and give it to the items that have never been observed. This modification is called smoothing. A large number of other smoothing techniques for N-gram models have been proposed, such as Laplace Smoothing [21], Add-k smoothing [22], Stupid backoff [23], and Kneser-Ney smoothing [24].

2.2 Action Sequencing

Sequencing actions from natural language text intended for human consumption is difficult since it does not contain a time series attribute and needs agents to comprehend complicated contexts of actions. Husari et al. [25] proposed a framework called *ActionMiner* that combined Entropy and Mutual Information with some basic NLP techniques to extract threat actions from Cyber Threat Intelligence reports and achieved good performance. However, this study only extracted all actions to the list and cannot analyze their relationship or sequence. Manshadi et al. [26] developed a probabilistic language model and used the predicate-argument pair

(verb-object; E.g., got-tire) to represent an action. This model can capture the expected sequences in simple narrative texts which have very few verbs in the corpus of Weblog Stories. In other studies, Feng et al. [27] proposed a novel approach EASDRL to automatically extract action sequences from texts based on deep reinforcement learning, and Mei et al. [28] adopted an encoder-decoder model with long short-term memory recurrent neural networks (LSTM-RNN) translates natural language instructions to action sequences. These works can extract meaningful action sequences from complicated sentences in free natural language; however, input data require that the order of sentences corresponds to the sequence of actions. Due to this limitation, it is hard to apply unstructured textual data such as accident reports. To deal with these restrictions, various significant efforts [29]–[33] employing the state-of-the-art machine learning algorithms for the task Sentence Ordering and Coherence can be taken into account before extracting action sequences. Nevertheless, these models only performed well for judging the order of sentence pairs and achieved relatively poor performance on the whole paragraph; hence, the application of models [27] and [28] would not really be feasible.

2.3 Related Studies

In the construction domain, numerous studies were conducted by researchers to explore the accident reports. Tixier [3] developed an automated model based on keyword dictionaries and R functions. This model is capable of scanning textual injury reports and extracting precursors, injury types, energy sources, and body parts with an accuracy of 95%. Goh et al. [4] adopted six machine learning algorithms, including support vector machine (SVM), linear regression (LR), random forest (RF), k-nearest neighbor (KNN), decision tree (DT), and Naive Bayes (NB), to classifying accident reports into 11 predefined labels of causes. This research indicated the good performance of SVM compared to others; however, the performance metrics were not good. In other research, Cheng [5] proposed a hybrid supervised machine learning named Symbiotic Gated Recurrent Unit (SGRU) for the task of categorizing 1000 construction reports into 11 unique label causes; the result exhibited significant improvements to the previous study. Recently, Zhong et al. [6] employed Convolutional Neural Network to classify accident narratives automatically. The authors later used The Latent Dirichlet Allocation (LDA) model to analyze and visualize the relationship of causes and related objects. The results provide valuable insights from text data. Chokor et al. [9] conducted a K-means clustering unsupervised approach to classify construction injury reports. Four types of accident causes were identified, including fall, struck by objects, electrocutions, and trench collapse. The aforementioned

research is solely concerned with determining the causes and frequent objects causing accidents, not extracting the sequence of actions taken by workers associated with accidents which might be crucial to enhance safety management.

This study addresses the research gap in previous studies; we developed an Action Sequencing Model that can sequence actions from accident reports. Our model can deal with the problem of complicated sentences and unstructured text without any effort of reordering actions and sentences. The result is able to identify potential relationships concerning the occurrences and describe the associated behaviors of workers that reflect the construction scene at the time of an accident.

3 Methodology

This study adopted the Probabilistic Language Model for developing an Action Sequencing Model (as depicted in Figure 1). To begin with, data preparation is to develop the datasets for training and evaluating the model. Several steps were then utilized for training the Action Sequencing Model before model evaluation was implemented.

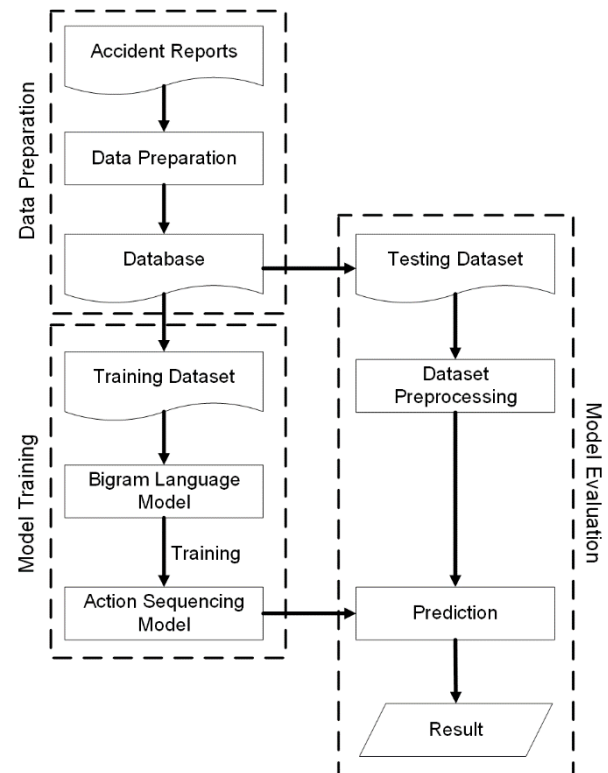


Figure 1. Methodology of Action Sequencing Model

3.1 Dataset Preparation

The accident reports are freely available online from the OSHA website [34]. In this study, accident reports, including a thorough account of construction site incidents, were picked and saved in a Microsoft Excel file. A sample size of 328 accident reports was chosen, and each report was split into separate sentences. As a result, the raw dataset of 1,689 sentences was developed. An accident report (also known as Accident Investigation Summary) is written by inspectors in free natural language to briefly describe some of the main points of an accident, so it mostly contains simple and short sentences.

Accident report	Statements	Level 1 label	Level 2 label
Employee #1 was engaged in the demolition of a structural steel amusement ride at a theme park. Employee #1 fell approximately 50 ft through a deck hole measuring approximately 2-ft by 8-ft that was created from a gear motor that had been cut and removed from the structure by the crew. Employee #1 was killed.	Employee #1 was engaged in the demolition of a structural steel amusement ride at a theme park.	Action	Demolition
	Employee #1 fell approximately 50 ft through a deck hole measuring approximately 2-ft by 8-ft that was created from a gear motor that had been cut and removed from the structure by the crew. Employee #1 was killed.	Event	Fall
	Employee #1 was killed.	Consequence	Fatality

Figure 2. Sample accident reports and labeled statements

Sentence labeling is the second step after splitting accident reports. Since accident reports are written in free natural language, it is difficult to directly identify and access vast amounts of information. The authors analyzed thoroughly and determined that, despite being unstructured text, each accident report contains three key pieces of information: sentences mention actions before the accident, sentences describe accident event, and sentences provide subsequent results. Therefore, the extracted sentences were annotated into the predefined level 1 labels, namely Action, Event, and Consequence, for grouping information. Following that, the sentences in each level 1 label were annotated into level 2 labels for the task sequences extraction (as shown in Figure 2). Aside from the summary of the incident, each accident report obtained from OSHA provides additional information such as diagnosis, cause, degree (bruise, fatality, etc.), occupation etc., which the authors referred to and reviewed for the unique labels of the statements. The authors also reused and calibrated many labels from OSHA definitions to establish the labels in the dataset. In some small number of situations, if a statement contained information that might be considered as multi-label categories and could not be aligned with one unique label based on OSHA additional information, a unique label

was assigned according to the principle of identifying the most significant contribution to the accident compared to the others. Figure 3 depicts the label diagram for the final labeling result in this study. As a result, the Action group has 30 unique labels, while Event and Consequence have the same number of unique labels of 12. Since the findings of dataset preparation occupy a large space and the paper length is limited, the authors could only show some typical labels. Among level 2 labels, “None” is a specially designed label that presents noise information in sentences without specific action, ambiguities or provides general information such as the sentence “There were no witnesses to the accident” or “The pit measured approximately 5 feet to 8 feet deep”.

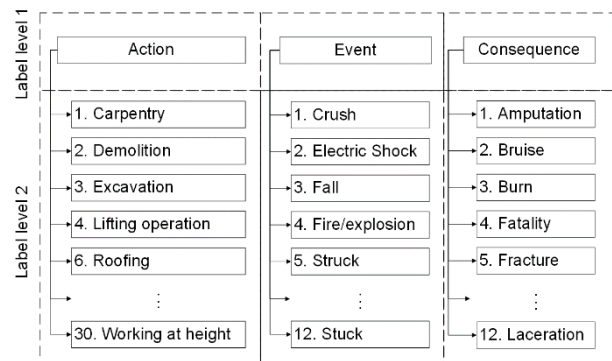


Figure 3. Labels of sentences extracted from accident reports

The next two important steps in data preparation are Dropping None label and Changing label format. As mentioned above, sentences labeled as None provided irrelevant information to extract sequences. This dropping helped filter and keep crucial extracted information. Some labels have more than one word, such as “Lifting operation” and “Working at height;” changing format step facilitated long word labels treated as one word (token) by adding underlines to link words. As a result, “Lifting operation” and “Working at height” were converted to “Lifting operation” and “Working at height” respectively. For this study, the authors read carefully and annotated the actual sequence of each accident report using level 2 labels and the order of action as a reference to develop and evaluate the model. A sequence is exhibited by a sequence of words. For example, an actual sequence with three elements is presented as “Carpentry Fall Bruise”; it can be interpreted as a sequence of actions Carpentry → Fall → Bruise. Table 1 shows the distribution of the length of the actual action sequence.

Table 1. The distribution of the length of the actual action sequence

The length of action sequence	The number of action sequence	Percentage
2 elements	15	4.57%
3 elements	147	44.82%
4 elements	118	35.98%
5 elements	35	10.67%
6 elements	12	3.66%
7 elements	0	0%
8 elements	1	0.3%
Total	328	100%

Finally, the final database was randomly split into 75% and 25% for the training and testing datasets, respectively.

3.2 Model Training

3.2.1 Bigram Language Model

In this study, the authors adopted bigram (N=2) to train the N-gram Language Model. The Bigram Language Model is the underlying Probabilistic Language Model, which has a wide application. As mentioned, the authors annotated the actual sequence of each accident report to train the model. Each element (level 2 label) of the actual sequence plays a role as a unigram, while a bigram is a sequence of two adjacent elements. Hence, the probability of an individual unigram given the bigram assumption:

$$P(w_n | w_{1:n-1}) \approx P(w_n | w_{n-1}) \quad (6)$$

The chain rule to the sequence of unigrams W :

$$P(w_{1:n}) = \prod_{k=1}^n P(w_k | w_{k-1}) \quad (7)$$

where:

$$P(w_k | w_{k-1}) = \frac{C(w_{k-1}w_k)}{C(w_{k-1})} \quad (8)$$

C is the frequency (count) of each pattern in the corpus.

Equation (8) calculates the probability of a bigram by dividing the observed frequency of this bigram by the observed frequency of the first unigram belonging to this bigram. This probability is also known as a relative frequency ratio. For example, to calculate the probability of a bigram “Struck Fall,” we need to get the counts of bigram “Struck Fall” and unigram “Struck” from the corpus level 2 labels. Afterward, we calculate the division of these two values.

$$P(\text{Fall} | \text{Struck}) = \frac{C(\text{Struck Fall})}{C(\text{Struck})}$$

$P(\text{Fall} | \text{Struck})$ denotes the probability of the unigram Fall given the unigram Struck; it is also known as $P(\text{Struck Fall})$. It can be interpreted as the probability of “Fall” occurring after “Struck.”

3.2.2 Smoothing Technique

Laplace Smoothing is used in this study to deal with the zero Bigram probability. This is the simplest and quickest technique to smooth data by adding one to all Bigram counts before normalizing them to probabilities. The probability of an individual unigram in equation (8) is expressed as:

$$P(w_k | w_{k-1}) = \frac{C(w_{k-1}w_k) + 1}{C(w_{k-1}) + V} \quad (9)$$

where V denotes the vocabulary, the set of all unigrams under consideration.

3.2.3 Training Action Sequencing Model

This study using Bigram Language Model developed the Bigram Sequence Probability Database as a root for training the Action Sequencing model:

- A Bag of Bigram was created based on a corpus of actual sequences retrieved from the training dataset.
- Adopt MLE as shown in equations (8) and (9) to estimate the probabilities of all bigrams in the Bag of Bigram. These probabilities are also known as the probabilities of the sequence of two actions. The obtained database is called the Bigram Sequence Probability Database.
- Apply the chain rule in equation (7) to calculate the probability of the action sequences. The probability of the bigrams retrieved from the Bigram Sequence Probability Database.

The obtained Bigram Sequence Probability Database contains all bigrams (sequence of two elements level 2 label) along with their probabilities that illustrate their likelihood. As a matter of fact, the resulting database is not only a component of the Action Sequencing Model but still has practical implications. For example, when considering what are immediately potential consequences following the event “Fall;” querying the Bigram Sequence Probability Database, we can achieve all results such as “Fracture” occupies the highest probability with $P(\text{Fracture}|\text{Fall}) = 0.3$, “Bruise” with $P(\text{Bruise}|\text{Fall}) = 0.05$, and “Fatality” with $P(\text{Fatality}|\text{Fall}) = 0.1$. This retrieval provides insight and enhances our understanding of all possible outcomes and what is most likely to happen for the prediction task.

3.3 Model Evaluation

Model evaluation is to evaluate the performance of the trained model on the testing dataset. The process includes preprocessing the testing dataset, prediction and evaluating results.

Firstly, data preprocessing was performed on the developed testing dataset by following steps:

- Concatenating labels (Generating preliminary sequence): The labels of sentences of each accident report were concatenated into a sequence for which the order of labels corresponds to the order of sentences. For example, an accident report contains the order of sentences corresponding to labels “Carpentry,” “Struck,” “Fracture,” “Fall”; the obtained sequence after concatenating is “Carpentry Struck Fracture Fall.”
- Generating permutations: The preliminary sequence of each accident report obtained from the Concatenating labels step was used to produce all possible sequences. For example, the sequence “Carpentry Struck Fracture” can be generated as “Carpentry Struck Fracture,” “Carpentry Fracture Struck,” “Struck Carpentry Fracture,” “Struck Fracture Carpentry,” “Fracture Struck Carpentry” and “Fracture Carpentry Struck.”

Figure 4 presents the workflow of the Action Sequence Prediction. The possible sequences got from the sequence permutation step were fed as input for the Action Sequencing Model. The output is the probability corresponding to each sequence. Eventually, Action sequence prediction was implemented by voting the permutation that had the highest probability.

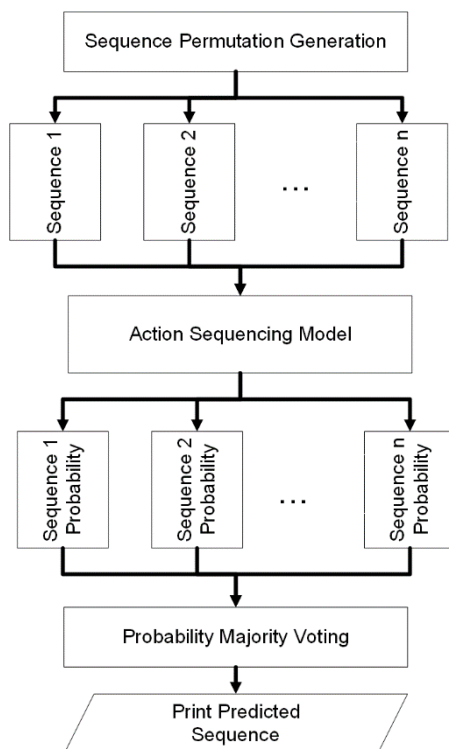


Figure 4. Workflow of the Action Sequence Prediction

Lastly, to evaluate the results (predicted orders), we

used two types of metrics: Sequence Accuracy (SA) and Pairwise Sequence Accuracy (PSA):

- **Sequence Accuracy (SA):** It measures the percentage of actions sequence that is correctly predicted.
- **Pairwise Sequence Accuracy (PSA):** this metric calculates the percentage of bigrams for which the relative order is predicted correctly. In other words, PSA is the ratio of the number of correct ordered word pairs and total possible word pairs.

4 Results and Discussion

This section presents the performance of the Action Sequencing Model. We developed two models, the first model is the baseline model without applying a smoothing technique, and the second model adopted a smoothing technique.

Our evaluation was based on sequence prediction on the testing set, and Table 2 shows the excellent performances. To begin with, the baseline model has a sequence accuracy of 74.68%, which is not too remarkable but sufficient for a successful action sequencing task. Besides, the pairwise sequence accuracy shows the ideal value of 80.83%. The second model with smoothing technique has over 10% higher than the baseline model in sequence accuracy with 84.81%. In terms of pairwise sequence accuracy, its metric is 89.12% indicating a robust value compared to the baseline model.

Table 2. Performance results of the Probabilistic Language Model

Performance Metrics	Modeling without Smoothing Technique	Modeling with Smoothing Technique
SA (%)	74.68	84.81
PSA (%)	80.83	89.12

Overall, these evaluations demonstrate that the Probabilistic Language Model is robust for developing the Action Sequencing Model and the application of the smoothing technique resulting in better performance metrics. A number of incorrect predictions are primarily due to long, complicated sequences, which made the model confusing. This error source can be observed through the difference between PSA and SA, where models performed well for judging the order of sequence pairs but operated poorer on the whole sequence. For example, the actual sequence of an accident report is “Roofing Fall Fall Struck Crush Fatality”; however, the prediction is “Roofing Fall Struck Fall Crush Fatality.” It is easy to see that there are three correct pairwise sequences, including “Roofing Fall,” “Fall Struck,” and

“Crush Fatality” out of 5 pairwise sequences. This error is able to be mitigated by expanding the dataset and applying more types of N-grams instead of bigram.

5 Conclusion

Construction accident reports are valuable documentation data, and analyzing them may give critical knowledge of prior occurrences to prevent unanticipated recurrence catastrophes. This study presents a framework that adopts a Probabilistic Language Model to sequence actions taken by workers that depict the construction scene at the time of accident from unstructured accident narrative reports. The Action Sequencing Model can deal with the problem of complicated sentences and unstructured text without any effort of reordering actions and sentences. This study produced excellent results with the highest sequence accuracy and pairwise sequence accuracy of 84.81 % and 89.12%, respectively, which illustrate the good performance for both judging the order of sequence action pair and the whole sequence actions of each record.

This research provides threefold contributions to the body of knowledge. To begin with, a reliable automated model was developed that can exploit various action relationship information from construction accident records. Secondly, a dataset was built and potentially used for further research in construction safety interest. Lastly, a sequence action database was formed in the final result that can explain the relationship between workers' actions at the time of accidents. This database can be adopted in the Sequence Mining task to provide a probabilistic forecast of likely next actions for a given action or sequence of actions. In terms of Industry implications, construction organizations can employ this automated model to analyze the sequence of action information in accident reports that generate consistent results and save time and resources. This information is used to establish safety management systems to actively forecast and prevent accidents on construction sites.

The results of this research were encouraging; however, some aspects can be further optimized in the future. A dataset size used in training is small; thus, expanding in size is needed to generalize the result. In addition, the use of trigram or more longer grams instead of bigram can potentially achieve better performance. Finally, this study introduced a simple probabilistic language model. The state-of-the-art machine learning algorithms might be incorporated into the probabilistic language model resulting in a hybrid model. The neural probabilistic language model would be a desirable objective for the action sequencing task.

References

- [1] Bureau of Labor Statistics (BLS), “National Census of fatal occupational injuries.” <https://www.bls.gov/news.release/pdf/cfoi.pdf> (accessed Feb. 20, 2022).
- [2] C. U. Ubeynarayana and Y. M. Goh, “An ensemble approach for classification of accident narratives,” in *Computing in Civil Engineering* 2017.
- [3] A. J. P. Tixier, M. R. Hallowell, B. Rajagopalan, and D. Bowman, “Automated content analysis for construction safety: A natural language processing system to extract precursors and outcomes from unstructured injury reports,” *Automation in Construction*, vol. 62, pp. 45–56, Feb. 2016, doi: 10.1016/j.autcon.2015.11.001.
- [4] Y. M. Goh and C. U. Ubeynarayana, “Construction accident narrative classification: An evaluation of text mining techniques,” *Accident Analysis and Prevention*, vol. 108, pp. 122–130, Nov. 2017, doi: 10.1016/j.aap.2017.08.026.
- [5] M. Y. Cheng, D. Kusoemo, and R. A. Gosno, “Text mining-based construction site accident classification using hybrid supervised machine learning,” *Automation in Construction*, vol. 118, Oct. 2020, doi: 10.1016/j.autcon.2020.103265.
- [6] B. Zhong, X. Pan, P. E. D. Love, L. Ding, and W. Fang, “Deep learning and network analysis: Classifying and visualizing accident narratives in construction,” *Automation in Construction*, vol. 113, May 2020, doi: 10.1016/j.autcon.2020.103089.
- [7] H. Baker, M. R. Hallowell, and A. J. P. Tixier, “AI-based prediction of independent construction safety outcomes from universal attributes,” *Automation in Construction*, vol. 118, Oct. 2020, doi: 10.1016/j.autcon.2020.103146.
- [8] B. Zhong, X. Pan, P. E. D. Love, J. Sun, and C. Tao, “Hazard analysis: A deep learning and text mining framework for accident prevention,” *Advanced Engineering Informatics*, vol. 46, Oct. 2020, doi: 10.1016/j.aei.2020.101152.
- [9] A. Chokor, H. Naganathan, W. K. Chong, and M. el Asmar, “Analyzing Arizona OSHA Injury Reports Using Unsupervised Machine Learning,” in *Procedia Engineering*, 2016, doi: 10.1016/j.proeng.2016.04.200.
- [10] D. Jurafsky, “Speech & language processing,” Pearson Education India, 2000.
- [11] R. Plamondon and S. N. Srihari, “Online and off-line handwriting recognition: a comprehensive survey,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 1, pp. 63–84, 2000.

- [12] D. Chiang, "A hierarchical phrase-based model for statistical machine translation," in *Proceedings of the 43rd annual meeting of the association for computational linguistics*, 2005.
- [13] A. Stolcke, "SRILM-an extensible language modeling toolkit," 2002.
- [14] E. Mays, F. J. Damerau, and R. L. Mercer, "Context based spelling correction," *Information Processing & Management*, vol. 27, no. 5, pp. 517–522, 1991.
- [15] B. Croft and J. Lafferty, "Language modeling for information retrieval," vol. 13. Springer Science & Business Media, 2003.
- [16] S. Missaoui, M. Viviani, R. Faiz, and G. Pasi, "A language modeling approach for the recommendation of tourism-related services," in *Proceedings of the Symposium on Applied Computing*, 2017.
- [17] J. M. Ponte and W. B. Croft, "A language modeling approach to information retrieval," in *ACM SIGIR Forum*, 2017.
- [18] K. Armeni, R. M. Willems, and S. L. Frank, "Probabilistic language models in cognitive neuroscience: Promises and pitfalls," *Neuroscience and Biobehavioral Reviews*, vol. 83. Elsevier Ltd, pp. 579–588, Dec. 01, 2017. doi: 10.1016/j.neubiorev.2017.09.001.
- [19] V. Cherian and M. S. Bindu, "Heart disease prediction using Naive Bayes algorithm and Laplace Smoothing technique," *Int. J. Comput. Sci. Trends Technol*, vol. 5, no. 2, pp. 68–73, 2017.
- [20] S. F. Chen and R. Rosenfeld, "A survey of smoothing techniques for ME models," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 1, pp. 37–50, 2000, doi: 10.1109/89.817452.
- [21] E. R. Setyaningsih and I. Listiowarni, "Categorization of Exam Questions based on Bloom Taxonomy using Naïve Bayes and Laplace Smoothing," in *3rd 2021 East Indonesia Conference on Computer and Information Technology, EIconCIT 2021*, Apr. 2021, doi: 10.1109/EIconCIT50028.2021.9431862.
- [22] Marrara, S., Pasi, G., Viviani, M., Cesarini, M., Mercorio, F., Mezzanzanica, M., & Pappagallo, M., "A language modelling approach for discovering novel labour market occupations from the web," Apr. 2017, doi: 10.1145/3106426.3109035.
- [23] T. Brants, A. C. Popat, P. Xu, F. J. Och, and J. Dean, "Large language models in machine translation," 2007.
- [24] R. Pickhardt, T. Gottron, M. Körner, P. G. Wagner, T. Speicher, and S. Staab, "A Generalized Language Model as the Combination of Skipped n-grams and Modified Kneser-Ney Smoothing," Apr. 2014.
- [25] G. Husari, X. Niu, B. Chu, and E. Al-Shaer, "Using entropy and mutual information to extract threat actions from cyber threat intelligence," in *2018 IEEE International Conference on Intelligence and Security Informatics (ISI)*, 2018.
- [26] M. Manshadi, R. Swanson, and A. S. Gordon, "Learning a Probabilistic Model of Event Sequences from Internet Weblog Stories.," in *FLAIRS Conference*, 2008.
- [27] W. Feng, H. H. Zhuo, and S. Kambhampati, "Extracting action sequences from texts based on deep reinforcement learning," 2018.
- [28] H. Mei, M. Bansal, and M. R. Walter, "Listen, attend, and walk: Neural mapping of navigational instructions to action sequences," 2016.
- [29] X. Chen, X. Qiu, and X. Huang, "Neural sentence ordering," 2016.
- [30] Y. Yin, L. Song, J. Su, J. Zeng, C. Zhou, and J. Luo, "Graph-based neural sentence ordering," 2019.
- [31] Y. Zhu, K. Zhou, J.-Y. Nie, S. Liu, and Z. Dou, "Neural Sentence Ordering Based on Constraint Graphs," 2021.
- [32] L. Logeswaran, H. Lee, and D. Radev, "Sentence ordering and coherence modeling using recurrent neural networks," 2018.
- [33] S. B. R. Chowdhury, F. Brahman, and S. Chaturvedi, "Is Everything in Order? A Simple Way to Order Sentences," 2021.
- [34] Occupational Safety and Health Administration (OSHA), "Fatality and Catastrophe Investigation Summaries." <https://www.osha.gov/pls/imis/accidentsearch.html> (accessed Dec. 11, 2021).