Image Acquisition Planning for Image-based 3D Reconstruction Using a Robotic Arm

Rachel Hyo Son[1] and Kevin Han[1]

[1]Dept. of Civil, Construction, and Environmental Engineering, North Carolina State University, USA

hson@ncsu.edu, khan6@ncsu.edu

Abstract -

As-built modeling using advanced visual sensing technologies (i.e., photogrammetry and laser scanning) provides an opportunity for a rapid assessment of construction performance, identifying deviations from 3D CAD/Building Information Modeling (BIM) models (i.e., as-planned model). For reliable decision-making, the accuracy and quality of the as-built model are critical. In particular, data collection techniques for image-based 3D reconstruction influence the quality of as-built modeling. In addition, manual data capture is time-consuming, error-prone, and infrequent, which impedes the assessment process. To overcome the challenges, this paper proposes a model-based planning system to automate image data collection for 3D reconstruction. The data collection system uses a camera-installed robotic arm. The locations of images to be captured are planned based on a given 3D CAD/BIM model using the camera parameters, distance away from the target object, and overlap ratio. The end-effector of the robotic arm, where a camera is installed, moves along the planned locations and captures images at each location. The complete image data set uses computer vision algorithms (i.e., Structure from Motion and multiple view stereo) to create the as-built model. The preliminary experiment produces an as-built model of the object successfully. Moreover, the quality of the as-built model can be improved by the presented method on the leverage of a camera-mounted robotic arm.

Keywords -

Image-based 3D reconstruction, automation image-data collection, robotic arm, as-built modeling

1 Introduction

The advanced visual sensing technologies (e.g., photogrammetry and terrestrial laser scanning) empower efficient visual-data collection in the construction industry. Efficient methods of as-built modeling enable a rapid assessment of progress, productivity, and quality, compared with the as-planned 3D Building Information Model (BIM) or computer-aided-design (CAD) model [1, 2, 3]. Despite the advantages, there is a challenge that many practitioners face when they first try image-based 3D reconstruction: difficulty in producing a high-quality and consistent as-built models [4].

To overcome the challenge, there are research efforts on the method of creating a high-quality as-built model. Two technologies are primarily used in construction applications: terrestrial laser scanning and image-based 3D reconstruction. The terrestrial laser scanner operates based on time-of-flight or phase shift measurements. Both techniques sweep laser rays and collect 3D coordinates of the points where the laser ray hit the object surface. The collection of points becomes a point cloud model of the scanned scene. The laser scanner has the advantage of speed and precise measurement in high resolution [5]; however, the high cost of equipment and labor turns down the benefits. By comparison, image-based 3D reconstruction uses the image processing algorithms, Structure from Motion (SfM) [6], and multiple-view stereo (MVS) [7] with a set of image data taken by the camera. The image processing algorithms estimate 3D structures from pairs of images and compute 3D dense points of the scene. Due to the cost-efficient and ease of data collection, image-based 3D reconstruction gained more attention to create the as-built model in construction [8].

For this reason, the research community has focused on the development of the visualization and analysis for construction performance using a as-built model that is reconstructed from daily photographs [9, 10, 11]. A set of the daily photographs taken by site engineers often are unordered, easily led to inconsistent point cloud quality. For modular or prefabricated components, a consistent and higher level of point cloud quality is required. Since the modular components are fabricated off-site, any deviations from design can lead to unexpected costs and schedule overrun due to the reworks, including re-shipping, re-fabricating, and repair. Thus, the quality of as-built model is critical to detect the deviation to secure cost and schedule in modular construction.

Since the performance of computer vision algorithms relies on the detection and matching of the features in different images, how images are collected (i.e., the number of images and data collection positions) greatly affects the accuracy and quality of 3D reconstruction. Without proper planning (predetermining camera locations), 3D reconstruction may result in high measurement errors and incomplete and noisy point cloud.

To overcome these challenges, this paper presents a robotic arm path planning method for image-based 3D reconstruction. The presented method is designed to capture as-built models of fabricated building components.The result shows reduced errors in the 3D point cloud formation process. This paper also presents a research road map to build a reliable and practical image data acquisition method and improve as-built modeling using a robotic arm.

2 Related Work

The 3D reconstruction of real-world objects and scene has received great attention in many different applications, such as creating 3D model of heritage site [12], face recognition [13], detailed 3D spatial data for geoscience application [14], reverse engineering [15, 16], and quality inspection [17]. There are different vision sensors are explored for 3D reconstruction: stereo cameras [18], 3D depth sensors [19], monocular cameras [20], laser scanners and videos [21]. In addition to capturing technology, the image-processing algorithms are actively developed [6, 7]. Thus, the capabilities of data collection technologies and processing to create 3D models have been actively studied. Nonetheless, the quality of modeling is difficult to obtain without sensor planning for data collection.

Since image-processing techniques are robust in creating a 3D model from a set of image data, establishing the proper data collection methods and automation have a considerable potential to obtain the quality of the as-built model. The robotic system, such as unmanned ground vehicles (UGV), unmanned aerial vehicles (UAV), and robotic manipulators, is introduced to improve the data collection system. For image data collection, the UAV has been studied actively in construction applications due to the capability of extensive capturing ranges and avoiding occlusions [22, 23]. However, UAV operation is limited during construction, and the flight motion causes motion blur in captured images. The UGV has not been explored much for image-data collection since the fixed camera on UGV is not efficient in capturing multi-views in the construction site. A robotic manipulator with a high degree of freedom can be applied for data acquisition by installing the sensor on the manipulator. It allows the sensor to move to the ultimate position. Thus, the robotic arm can provide an opportunity to perform desirable motions to collect image data based on sensor planning.

The benefit of a robotic arm system is easily programmed and operates new scenarios [19]. The scenarios can be created based on sensor path planning. The current studies of sensor planning with a robotic arm focus on the small mechanical part [24] using a range sensor or laser scanner. Their planning method is finding the set of required viewpoints and planning the sequence of the viewpoints[25] [20]. The sensor path planning with a camera has not been studied much with a robotic arm system. Since the 3D reconstruction principle is different between the camera and laser scanner, the sensor path planning method cannot be the same. The image data collection method to reduce reconstruction error has not been covered enough in the research community.

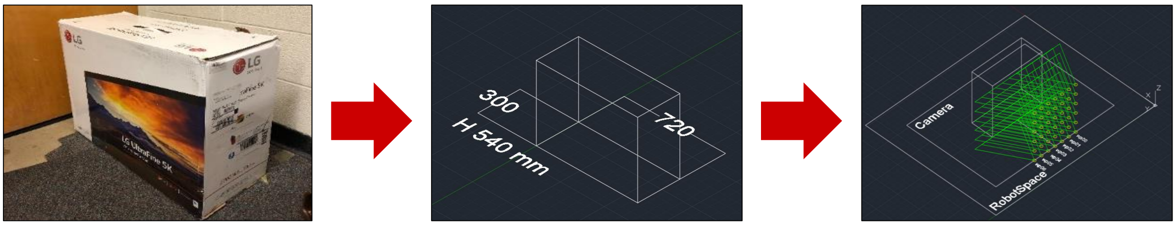
3 Method

The proposed framework aims to automate image data collection to reconstruct a complete and reliable quality as-built model. The automation of the data collection system uses an unmanned ground vehicle (UGV) integrated with a camera-mounted robotic arm. The robotics are programmable to move through the optimal path for collecting data. Thus, the overall system is associated with two motion plannings for a UGV and a camera-mounted robotic arm. In this paper, the study scope is limited to surface reconstruction to validate the robotic arm operation in image-based 3D reconstruction. Since a robotic arm has its allowable workspace, the study assumes that the camera installed on the robotic arm can cover a certain size of the surface. The study implements the model-based planning for the robotic arm to find camera positions from the given CAD model. The considerations for the image acquisition planning include capture distance and angle, field of view, overlap ratio, and robotic arm workspace to reduce a source of error for 3D reconstruction in the data collection step. The data acquisition with a robotic arm starts after a UGV stops. After a robotic arm finishes scanning one surface, the UGV moves to the other surface and starts taking image data of another surface. This procedure is repeated until collecting a complete image-data set for the object. The main focus of the paper is robotic arm operation for one surface data acquisition, not a complete object. The success of the surface reconstruction provides a promising conclusion for the 3D reconstruction of the complete building components as moving around the UGV. The overall study method is illustrated in Figure 1.

For camera sensor planning using a robotic arm, the following components, identified as a source of reconstruction error in the stage of data collection, are studied: camera distance, orientation, robotic arm workspace, and image overlap ratio.

3.1 Camera distance and orientation from the object surface

The constant distance between the image plane and capturing object surface is an important element to assure 3D point cloud quality. It minimizes the error finding corre-



(a) Target object (b) Create 3D model (c) Configure camera pose

(d) Arm trajectory design (e) Robotic arm operation (f) 3D reconstruction result

Figure 1. Research method overivew - Image data collection for 3D reconstruction

sponding 3D points from two different images. According to Yousfi et al. [26], the images used for 3D reconstruction from difference distances cause an error in matching corresponding points. Although the image processing structure-from-motion and multiview stereo (SfM-MVS) algorithms are robust to images at different scales, the images taken from significantly different distances cause to be rejected [14]. The minimum focus distance of the camera was considered to maintain the distance for image capture. Camera orientation remains constant along the surface to maintain the distance between object and image plane.

3.2 Robotic arm (end-effector) workspace

Before determining the image collecting locations, the end-effector workspace is determined. The end-effector workspace refers to the 3D space that the end-effector can reach without joint collisions. Defining a workspace preserves the quality of images since it reduces the failure of the arm trajectory and improves the accuracy of end-effector positioning.

From the Kinova robotics user manual, the maximum reachable robotic arm distance is given as 1260 mm in the z-direction and 984 mm in x,y direction when the arm is placed on the ground, as shown in Figure 2 [27]. However, the actual reachable space for this study is different from the maximum because of the constraint due to the end-effector's fixed orientation and distance. The constraint

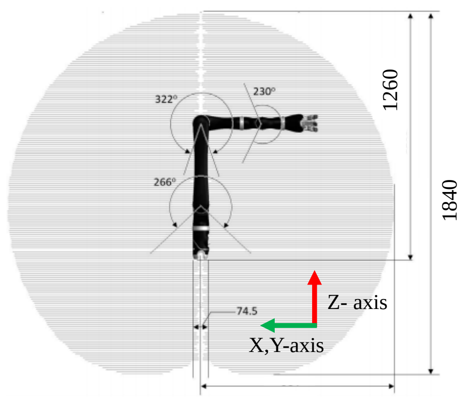


Figure 2. JACO robotic arm usable workspace in mm [27]

limits the robotic arm motion planning space on the yz-plane (Figure 3); i.e., the workspace for this data collection task only needs to be defined in the yz-plane.

Therefore, the maximum camera reachable space maintaining a fixed angle and distance of end-effector is identified. Based on the maximum, the range from the arm base is determined. Between the range, the robotic arm

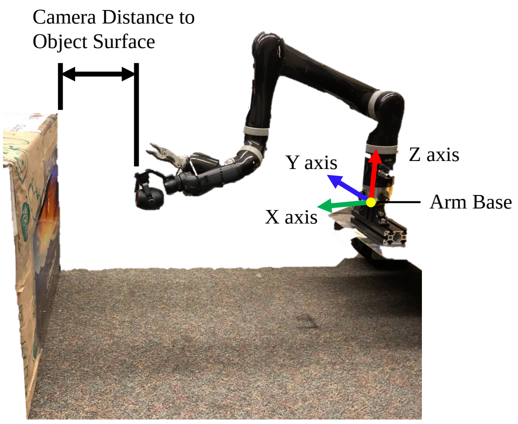


Figure 3. Robotic arm coordinate system

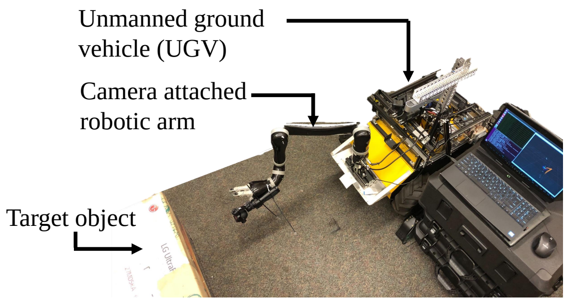


Figure 4. Experimental setup, including target object, camera attached robotic arm, and unmanned ground vehicle (UGV)

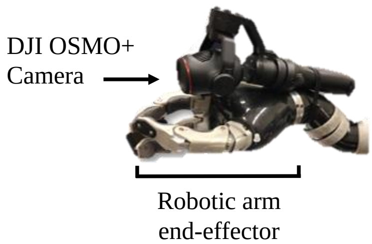


Figure 5. Camera mounted robotic arm end-effector

motion is tested at each 3D point location every 0.05 m. For example, if the identified maximum space in the y-axis is +0.3 m from the arm base, then the minimum space is -0.3 m. Each space point every 0.05 m from -0.3 m to +0.3 m was tested. The test is conducted by commanding the move distance through a python script and compares the arrival position based on the publishing data from the encoder. The position errors compared with the commend are less than 0.02 m, then include the 3D point into the workspace.

3.3 Image overlap ratio

The importance of image overlap ratio is image-based 3D reconstruction process starts with feature detection and matching in the different photographs by the Scale-Invariant Feature Transform (SIFT). Insufficient feature overlap attributes to the failure of the scene reconstruction [28]. For this reason, the different overlap ratio has been studied, changing the spacing between waypoints.

4 Experimental Setup

The experiment setup of the research is shown in Figure 4. The target object was placed in front of the robotic arm installed on UGV. The visual sensor, a camera (OSMO Plus from DJI), was mounted on a six-degree-of-freedom manipulator's (Jaco robotic arm from Kinova robotics) end-effector as shown in Figure 5.

The test object was selected and its size was 0.3 m (width) x 0.72 m (length) x 0.54 m (Height). For the model-based approach, the 3D CAD model of the object was created (Figure 1(b)). The field of view of the OSMO Plus camera is a minimum 35 ° maximum 92 ° based on the focal length. Assuming the camera operating field of view is 60 °, data acquisition positions were decided and drawn on a 3D CAD model (Figure 1(c)). The camera candidate positions are tested with the robotic arm to make sure that no joint collision occurs during the end-effector movement to the candidate positions (Figure 1(d)). The image data was taken when the robotic end-effector stops at the capturing positions (Figure 1(e)). The robotic arm was programmed to stop every waypoint and capture images. After robotic arm operation is completed, a set of images run dense cloud algorithms (i.e., SfM-MVS) to create an as-built model [29].

5 Experimental results

- **Robotic arm workspace**: The test result of the workspace was that from -0.3 m to 0.3 m in the y-axis and from -0.15 m to 0.6 m in the z-axis from the center of the arm base. These spaces are the robot arm able to operate spaces maintaining the fixed camera angle and distance. All planned waypoints are inside of the defined workspace.
- **Sensor trajectory and operation**: The sensor tra-

(a) Camera trajectory (b) Image taken locations (c) 3D reconstruction result

Figure 6. Results of the proposed framework - (a) camera trajectory, (b) image taken locations [29], (c) 3D reconstruction result

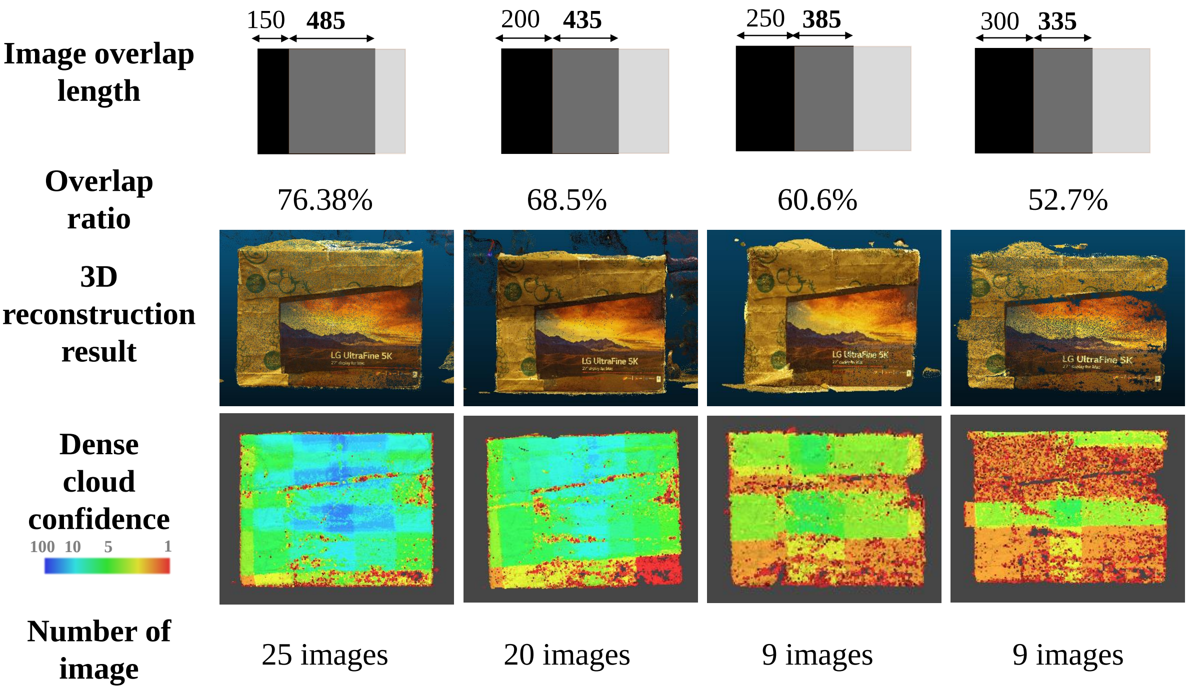


Figure 7. 3D reconstruction results based on overlap ratio

jectory is shown in Figure 6 (a). The camera moved along the designed trajectory and stopped at the planned positions to take photographs. The images (in blue boundary) created as-built models are shown with image-taken positions. Figure 6 (c) shows the surface reconstruction result.

- **Image overlap ratio**: The 3D reconstruction results with different overlap ratios are shown in Figure 7. The maximum tested overlap ratio was 76.38 %, and the minimum was 52.7%. The more overlap ratio achieves a better to complete 3D point cloud model, the least model with nine images was not able to completely reconstruct, especially on the edge of the object. The comparison of the dense cloud confidence result shows that a small number of images creates the lower confidence level of dense cloud (Figure 7) since the lower number of images is associated with creating an as-built model, which reduces the accuracy of the as-built model.
- **Construction object**: The designed camera trajec-

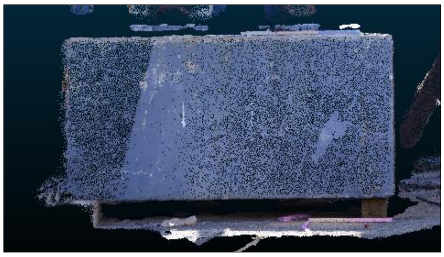


Figure 8. Image-based 3D reconstruction of concrete structure

tory was tested on the concrete object located in the Construction Facility Laboratory (CFL) at North Carolina State University (NCSU). The successful 3D reconstruction is shown in Figure 8.

6 Conclusion and future research

This paper presents a sensor planning method for image-based 3D reconstruction using a robotic arm. The planning method is to minimize the source of errors from data collection. The preliminary results show a great potential that the quality of 3D reconstruction can be improved by the presented method. Moreover, this research can be extended to complete as-built modeling, installing a robotic arm on the unmanned ground vehicle (UGV). The sensor planning method will be automated regardless of the object shape in future studies.

The automation of data collection with a camera or a laser scanner installed robotic arm can be applied to various applications in construction. The high degree of freedom operation allows placing a sensor in desirable locations depending on the object's shape. Therefore, the leverage of robotic arm data acquisition enables efficient data collection methods for the construction building components with diverse complexity of shape and a wide range of quality requirements.

7 Acknowledgement

This work was funded by the Department of Energy, under Award Number DE-AR0001155. The United States Government has a non-exclusive, paid-up, irrevocable, world-wide license to publish and reproduce the published form of this work and allow others to do so, for United States Government purposes. Any opinions, findings, conclusions and/or recommendations expressed in this paper are those of the authors and do not necessarily state or reflect those of the United States Government or any agency thereof.

References

[1] Kevin K Han and Mani Golparvar-Fard. Potential of big visual data and building information modeling for construction performance analytics: An exploratory study. *Automation in Construction*, 73: 184–198, 2017.

[2] Mani Golparvar-Fard, Jeffrey Bohn, Jochen Teizer, Silvio Savarese, and Feniosky Peña-Mora. Evaluation of image-based modeling and laser scanning accuracy for emerging automated performance monitoring techniques. *Automation in construction*, 20 (8):1143–1155, 2011.

[3] Mani Golparvar-Fard, Feniosky Peña-Mora, and Silvio Savarese. Integrated sequential as-built and as-planned representation with d 4 ar tools in support of decision-making tasks in the aec/fm industry. *Journal of Construction Engineering and Management*, 137(12):1099–1116, 2011.

[4] Habib Fathi, Fei Dai, and Manolis Lourakis. Automated as-built 3d reconstruction of civil infrastructure using computer vision: Achievements, opportunities, and challenges. *Advanced Engineering Informatics*, 29(2):149–161, 2015.

[5] Czesław Suchocki. Comparison of time-of-flight and phase-shift tls intensity data for the diagnostics measurements of buildings. *Materials*, 13(2):353, 2020.

[6] Changchang Wu et al. Visualsfm: A visual structure from motion system, 2011.

[7] Yasutaka Furukawa and Jean Ponce. Accurate, dense, and robust multiview stereopsis. *IEEE transactions on pattern analysis and machine intelligence*, 32(8): 1362–1376, 2009.

[8] Kevin K Han and Mani Golparvar-Fard. Automated monitoring of operation-level construction progress using 4d bim and daily site photologs. In *Construction Research Congress 2014: Construction in a Global Network*, pages 1033–1042, 2014.

[9] Kevin K. Han and Mani Golparvar-Fard. *BIM-Assisted Structure-from-Motion for Analyzing and Visualizing Construction Progress Deviations through Daily Site Images and BIM*, pages 596–603. doi:10.1061/9780784479247.074.

[10] Mani Golparvar-Fard, Feniosky Pena-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.

[11] Mani Golparvar-Fard, Feniosky Peña-Mora, and Silvio Savarese. D4ar–a 4-dimensional augmented reality model for automating construction progress monitoring data collection, processing and communication. *Journal of information technology in construction*, 14(13):129–153, 2009.

[12] Sabry F El-Hakim, J-A Beraldin, Michel Picard, and Guy Godin. Detailed 3d reconstruction of large-scale heritage sites with integrated techniques. *IEEE computer graphics and applications*, 24(3):21–29, 2004.

[13] Dalong Jiang, Yuxiao Hu, Shuicheng Yan, Lei Zhang, Hongjiang Zhang, and Wen Gao. Efficient 3d reconstruction for face recognition. *Pattern Recognition*, 38(6):787–798, 2005.

[14] Mike R James and Stuart Robson. Straightforward reconstruction of 3d surfaces and topography with a camera: Accuracy and geoscience application. *Journal of Geophysical Research: Earth Surface*, 117 (F3), 2012.

[15] Jun Wang, Dongxiao Gu, Zeyun Yu, Changbai Tan, and Laishui Zhou. A framework for 3d model reconstruction in reverse engineering. *Computers & Industrial Engineering*, 63(4):1189–1200, 2012.

[16] MJ Milroy, DJ Weir, C Bradley, and GW Vickers. Reverse engineering employing a 3d laser scanner: A case study. *The International Journal of Advanced Manufacturing Technology*, 12(2):111–121, 1996.

[17] P Rodríguez-Gonzálvez, M Rodríguez-Martín, Luís F Ramos, and D González-Aguilera. 3d reconstruction methods and quality assessment for visual inspection of welds. *Automation in construction*, 79: 49–58, 2017.

[18] Stefan Isler, Reza Sabzevari, Jeffrey Delmerico, and Davide Scaramuzza. An information gain formulation for active volumetric 3d reconstruction. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3477–3484. IEEE, 2016.

[19] Debdeep Banerjee, Kevin Yu, and Garima Aggarwal. Robotic arm based 3d reconstruction test automation. *IEEE Access*, 6:7206–7213, 2018.

[20] Khaled El-Rayes and Hisham Said. Dynamic site layout planning using approximate dynamic programming. *Journal of computing in civil engineering*, 23 (2):119–127, 2009.

[21] Ioannis Brilakis, Habib Fathi, and Abbas Rashidi. Progressive 3d reconstruction of infrastructure with videogrammetry. *Automation in construction*, 20(7): 884–895, 2011.

[22] Laura Inzerillo, Gaetano Di Mino, and Ronald Roberts. Image-based 3d reconstruction using traditional and uav datasets for analysis of road pavement distress. *Automation in Construction*, 96:457–469, 2018.

[23] Weiguang Jiang, Ying Zhou, Lieyun Ding, Cheng Zhou, and Xiaodi Ning. Uav-based 3d reconstruction for hoist site mapping and layout planning in petrochemical construction. *Automation in Construction*, 113:103137, 2020.

[24] Peng Wu, Hiromasa Suzuki, and Kiwamu Kase. Model-based simulation system for planning numerical controlled multi-axis 3d surface scanning machine. *JSME International Journal Series C Mechanical Systems, Machine Elements and Manufacturing*, 48(4):748–756, 2005.

[25] Li-jun Ding, Shu-guang Dai, and Ping-an Mu. Cad-based path planning for 3d laser scanning of complex surface. *Procedia Computer Science*, 92:526–535, 2016.

[26] Jezia Yousfi, Samir Lahouar, and Abdelmajid Ben Amara. Study on the strategy of 3d images based reconstruction using a camera attached to a robot arm. *Engineering Research Express*, 2(2):025006, 2020.

[27] Kinova Inc. *User guide: Kinova Jaco prosthetic robotic arm*. 2018.

[28] Jean Liénard, Andre Vogs, Demetrios Gatziolis, and Nikolay Strigul. Embedded, real-time uav control for improved, image-based 3d scene reconstruction. *Measurement*, 81:264–269, 2016.

[29] Agisoft photoscan standard (version 1.2.6). Retrieved from : `http://www.agisoft.com/downloads/installer/`, Accessed: 03/04/2019.