Construction Data-Driven Dynamic Sound Data Training and Hardware Requirements for Autonomous Audio-based Site Monitoring

Y. Xie[a], Y.C. Lee[a], T. Huther da Costa[b], J. Park[b], J. H. Jui[c], J.W. Choi[b], Z. Zhang[b]

[a]Department of Bert S. Turner Department of Construction Management, Louisiana State University, United States
[b]School of Electrical Engineering and Computer Science, Louisiana State University, United States
[c]Division of Computer Science and Engineering, Louisiana State University, United States
[d]Louisiana Transportation Research Center, United States
E-mail: yxie21@lsu.edu, yclee@lsu.edu, thuthe1@lsu.edu, jpar132@lsu.edu, jjui1@lsu.edu, choijw@lsu.edu, doc.zhang@la.gov

Abstract –

In a dynamic construction site, sound generated by work activities and equipment operations is one of vital field data indicating construction progress, work performance, and safety issues. However, because of an enormous number of construction work types, accurate sound classification is currently limited. To address this challenge, this study proposes a schedule-based sound classification for establishing dynamic sound training data. With the schedule-based dynamic training method, this system retrieves the types of sounds of daily planned construction activities for flexibly restricting training data types of working sounds and ultimately improving the accuracy of sound classification. To reveal the implications of audio-based construction activity detection and site monitoring frameworks, this study also involves the development of a construction sound library, a hardware system, a sound classification framework, and a web-based visualization method. This proposed method is expected to play a critical role in managing a construction project by supporting site monitoring and progress analysis, and safety surveillance.

Keywords –

Schedule-based data training; Audio-based site monitoring; Audio sensor-based hardware system; Automated work activity surveillance; Machine Learning-based sound classification

1 Introduction

In the construction industry, the lack of appropriate technologies to acquire field data has been a primary obstacle to prevent real-time data collection and analyses. With the rapid development of information technology, the construction industry has been seeking for new field data acquisition and advanced analysis methods to pursue enhanced construction activity monitoring and automated field management. It is critical to allow domain professionals to manage construction activities and operations efficiently with accurate analyses and measurement of field data throughout an entire construction project, because of dynamic and complicated backgrounds of a construction project with various work activities. With the help of previous research studies of field data collection, diverse applications and monitoring technologies such as sensors- and vision-based methods have been investigated to explore their potentials and practicability in the construction industry [1, 2]. Recent studies have also identified the implications of the implementation of the audio-based site monitoring technique [3, 4]. This study states that sound data not only require less data processing weight for a data analysis than one from vision data, but also have no limitation to capture data with an unlimited level of illumination and scope [3]. Because of the diverse types of equipment, materials, and work environment, however, it is not practically feasible to establish a comprehensive sound data library covering all sound types of construction work activities in a site. To ameliorate this challenge, this study aims to establish construction schedule-based sound data training and audio-based site monitoring frameworks that dynamically develop sound training data according to daily planned work activities. This study involves the development of a construction sound library, a dynamic sound data training model, a hardware system, and a web-based sound data map visualization. To assess the proposed approach, this study has the experiment of sound classification with a real construction project schedule and illustrates improved classification accuracies with a confusion matrix.

2 Literature Review

Recent decades, diverse studies explored advanced technologies for construction field data collection and process monitoring using sensors, ultra-wide band, and computer vision-based monitoring [1, 2, 5, 6, 7]. However, vision-based monitoring is somewhat limited in tracking work activities during night time because it requires certain level of illumination and has the restricted range of a vision angle [3]. In addition, video and kinematic data are too heavy to be processed for real-time monitoring. Sound data as a new type of field data requiring less data processing weight have been explored in several studies. The execution of an audio-based approach using the support vector machine has been investigated for construction activity identification of heavy equipment [3]. The frequency domain approach was also applied to identify overlapped sounds generated by different equipment at a construction site [4]. However, these previous studies mostly focus on utilizing and improving a signal analysis, feature extraction, model training, and testing. This study has an objective to enhance sound classification accuracy by providing dynamic sound training data according to planned daily schedule.

3 Research Approach and Methodology

The primary goal of this study is to establish a schedule-based sound data library containing sound data types associated with daily planned construction activities and compare the classification result based on the schedule and that without the schedule. To identify construction activities accurately, the audio-based monitoring system requires a robust sound data library involving diverse types of work activities. Since a construction project schedule entails a detailed daily work plan including types of work activities useful for building the foundation of a site monitoring system, its data can be utilized for pre-identifying daily planned work activities and their sound types, allowing the system to flexibly retrieve sound data from an existing sound data library and establish training datasets accordingly. Figure 1 shows the research process including the establishment of a sound data library, sound classifier training, a construction activity detection framework, and a visualization interface.

To accomplish this objective, this study employs a construction schedule in the XML format extracted from a construction scheduling software. In addition, this system adopts the K-nearest neighbor (KNN) classifier for sound classification based on extracted features of selectively retrieved sound data.

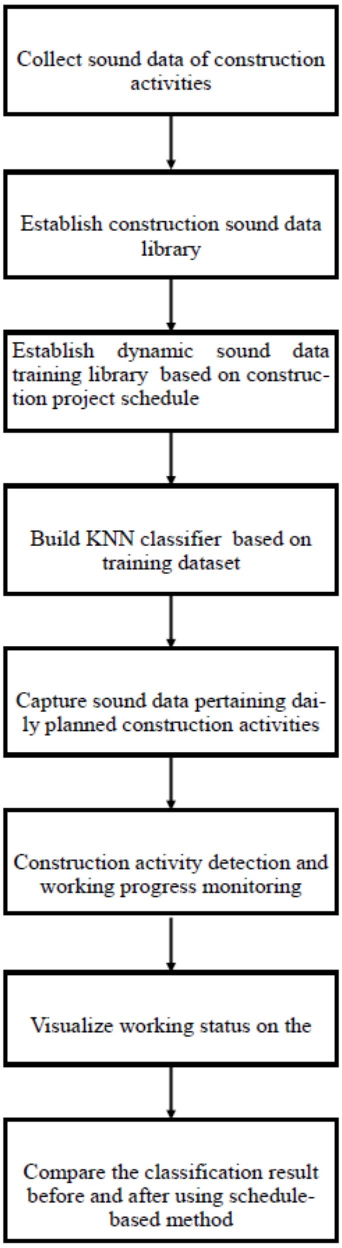

Figure 1. Framework of the methodology

With the customized XML schema, this system extracts XML-based construction schedule data including construction activities, equipment resources, and operation time from the Microsoft project, which is one of the broadly used construction scheduling software. This schedule data are imported into a framework to retrieve corresponding sound data from the established sound data library to build dynamic training datasets. The authors have already developed a sound data library that encompasses more than 100 sound data covering 15 types of construction work activities. Figure 2 shows the process of the establishment of a dynamic training data library including schedule information retrieval, sound data extraction, and training data development. The retrieved daily schedule information is referred to build a training data library including audio files (.wav) of equipment and work activities by conducting the matching process using the filenames and the semantics of audio data. With the established schedule-based library, the input file containing daily planned work types can be interpreted to extract corresponding audio data and their features that are the imperative resources analyzed by the KNN model.

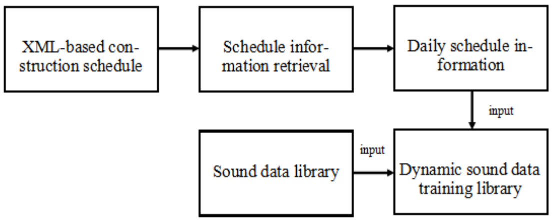

Figure 2. Establishment of dynamic sound data library

3.1 Development of a Sound Data Library

In this study, eight type sound data of construction activities including Excavating, Concrete Mixing, Compacting, Bulldozer, Hammering, Piling, Concrete Pumping and Drilling were recorded from the real bridge construction sites in East Baton Rouge, Louisiana. These sound data are divided into a short segment with bins of 2 seconds long and split into both training and testing sets. Because of the complex background and the dynamic environment of the construction site, recorded sound data of construction activities generally include noise and the classification accuracy is highly dependent on signal to noise ratio [8]. Therefore, a de-noising process is required to enhance raw sound signal data and a noise estimation algorithm [9] is adopted in this research. This algorithm is applied to reduce noise by computing the ratio of power spectrum to minimum power spectrum for each segment. If a ratio is less than one, it should be removed from original signal waves, or it can be regarded as valuable signal. The original and enhanced sound data of compacting and excavating are shown in Figure 3 and 4.

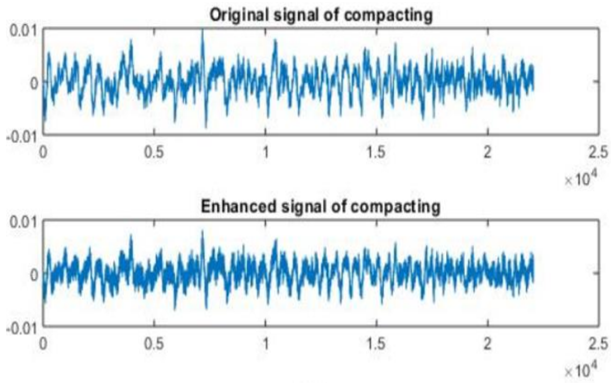

Figure 3. Original and enhanced signal of compacting

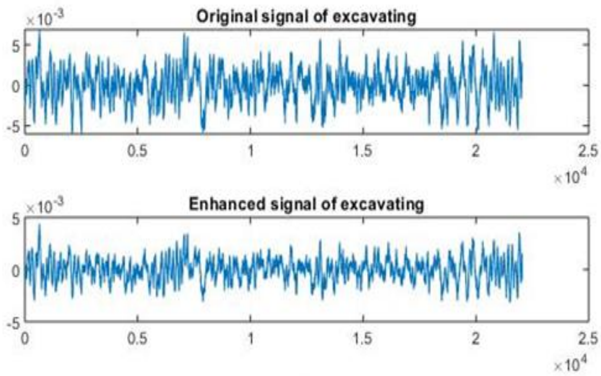

Figure 4. Original and enhanced signal of excavating

3.2 Generation of XML-based Construction Schedule Data and Training Data Library

Table 1 shows the schedule of the bridge construction project in East Baton Rouge, Louisiana, which consists of construction activities, resources, and starting and finishing time. Figure 5 illustrate the XML-based schedule data extracted from the Microsoft Project software.

Table 1. Schedule of Bridge Construction in East Baton Rouge

Construction Activities	Construction Resources	Starting Time	Finishing Time
Earth Works	Mobile excavators	2018-09-07	2018-09-20
Structural Excavation for Minor Structures	Mobile excavators	2018-09-07	2018-09-20
Fabric and Grid Reinforcing	Complete Reinforcement Cages	2018-09-11	2018-09-13
Sheeted Caissons	Pile Drivers (Rammers) and Pulling Tools	2018-09-12	2018-09-13
Sheeted Caissons	Pile Casings (linings)	2018-09-12	2018-09-13
Cast-in-Place Concrete Pile	Ready Mixed Concrete C30	2018-09-14	2018-09-17
Spreading and Compaction	Wheel loaders	2018-09-14	2018-09-20
Reinforcement Bars	Reinforcing Bars	2018-09-18	2018-09-25
Reinforcement Bars	Steel Reinforcement cutters	2018-09-18	2018-09-25
Structural Cast-in-Place Concrete Forming (steel)	Steel Forms	2018-09-26	2018-09-28
Cast-in-Place Concrete C25	Concrete truck mixer / agitator	2018-10-01	2018-10-02
Cast-in-Place Concrete C25	Concrete pumps and equipment	2018-10-01	2018-10-02
Cast-in-Place Concrete C25	Concrete vibrators	2018-10-01	2018-10-02
Cast-in-Place Concrete C25	Ready Mixed Concrete C25	2018-10-01	2018-10-02
Concrete Curing	Coatings for Concrete and Masonry	2018-10-03	2018-10-05

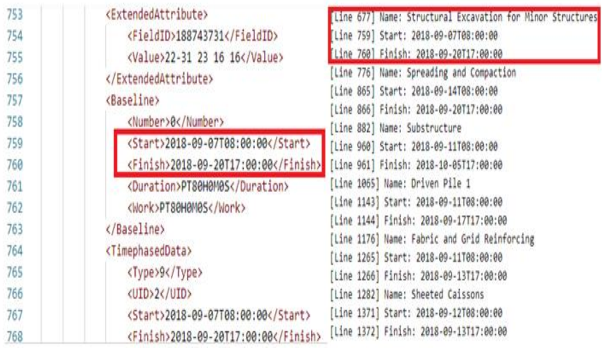

Figure 5. Generated XML based schedule

By referring the project schedule information, this system retrieves project-specific sound data from the pre-developed sound data library to build a dynamic training data library. To evaluate the proposed method, the research team conducted the case study using the following three construction activities in Table 1 scheduled from 2018-09-14 to 2018-09-17: Earth Works, Spreading and Compaction, and Cast-in-Place Concrete Pile. As shown in Figure 6, these schedule data are exported in the XML format from MS Project and imported into the framework to retrieve the associated sound data and generate a daily training library.

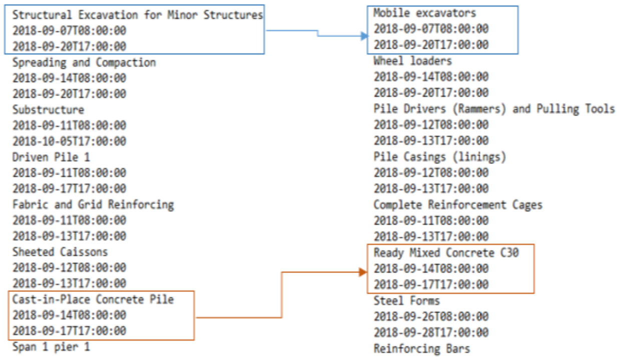

Figure 6. Schedule information retrieval

3.3 Hardware development of an audio sensor-based system

Microphones are the primary devices for capturing and recording sound data and each microphone contains a surface designed to capture audio waves. The internal circuits of a microphone should be powered to generate a sound signal and a wireless system is designed for audio signal receiving combined with the data analysis. Figure 7 shows the hardware design including a microphone, a radio frequency module (RF) including a transmitter and a receiver (nRF24L01), and a multi-track recorder (Zoom F8).

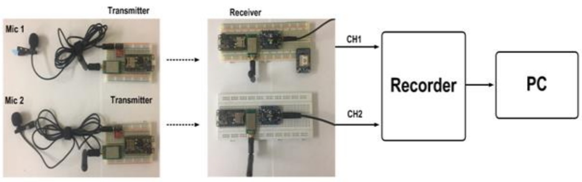

Figure 7. Hardware settings

3.4 Audio Signal Transmission and Analysis

Microphones capture surrounding sounds and transform them into electrical signals that are amplified through pre-amps making that the signals can be readable by micro-controllers. For a real-time sound signal transmission, we utilize a nRF24L01 transceiver module which is widely used in two-way radios, called walkie-talkie. In the current research, even if nRF24L01 modules work as one-way radios, the RF module can be suitable because it has an operating frequency of 2.4 GHz (ISM band) and a multichannel capability (up to 126 channels). Moreover, it can achieve a sample rate of 44 KHz with 1 Mbps data rate.

Within an available network of the RF module generally covering 1.5 miles, transmitters play a role in transferring the data sets to the receivers as shown in Figure 8. Receivers obtain the sound data in the form of digital signals. Microcontrollers select proper channels for the receivers from different channels and the digitized sound information can be transformed into the analog signal for the 8-channel recorder. The recorder collects the sound data from the multi-channel and transfers the data to a computer for analysis as shown in Figure 9.

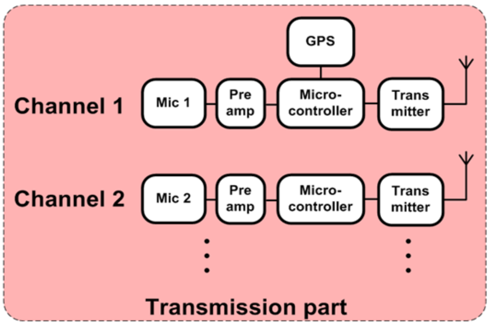

Figure 8. Sound signal transmission

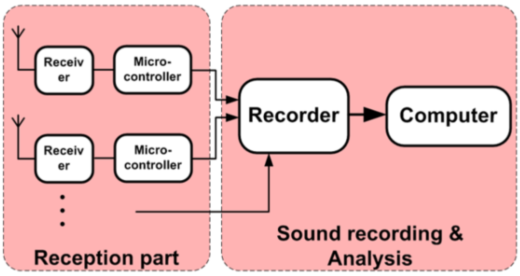

Figure 9. Sound signal receiving and analysis

3.5 Pre-processing Including De-noising, Feature Extraction, and KNN-based Sound Classification

Recorded sound data in a construction site generally contain background noise that negatively affects the accuracy of sound classification and site monitoring. Thus, this system adopts the noise estimation algorithm for de-noising, which reduces noise of a training dataset. This method computes the ratio of power spectrum to minimum power spectrum for each segment and then remove if a ratio is less than one. By using 44.1 kHz sampling frequency, this study utilized 15 distinct sets of features and in our experiments, the K nearest neighbor (KNN) algorithm achieved enough high classification accuracies, while maintaining fast execution. A set of features is required to build the KNN model. To extract the features, sound files were split in frames of 2s, from which the features shown in Table 2 were extracted to fit the KNN model. During training, each sound belongs to a defined construction activity category, as mentioned earlier. During execution, features are extracted from the recorded sound and fed to the KNN algorithm. The predicted category is found by minimizing the expected cost among the nearest neighbors of the current sample that is, by finding the nearest neighbors, shown in figure 10. When testing new audio segments, the output should be one of construction work activities or equipment operations in sound data library mentioned in section 3.1.

Table 2. Selected Sound Features

Symbol	Feature name	Number of features in the set
ZCC	Zero Crossing Count	10
VZCR	Variance of Zero Crossing Rate Ratio	1
HZCRR	High Order Zero Crossing Rate Ratio	1
RMS	Root Mean Square	10
LEF	Low Energy Frame	1
STE	Short Time Energy	1
LSTER	Low Short Time Energy Ratio	1
VLER	Variance of Low-Band Energy Ratio	1
VSFLUX	Variance Spectrum Flux	1
LFRMS	Low-Frequency Root Mean Square	1
LFZCC9	9th-order moment of LFZCC	1
SBC	Sub Band Correlation	5
HOC	High Order Crossing	8
LPC	Linear Predictive Coefficients	7
MFCC	Mel-Frequency Cepstral Coefficients	13
Total		62

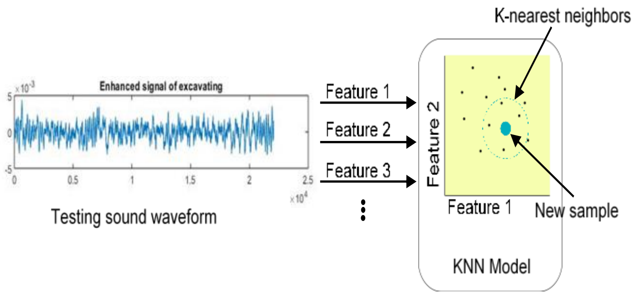

Figure 10. Predicting new sound

3.6 Construction Activity Identification and Accuracy Evaluation

By executing the fitted KNN model with daily dynamic training sound data, this study evaluates the accuracies of sound classifications before and after using the schedule-based approach. 200 short time segments (2s long) audio files of the three selected work activities are split into training and testing datasets evenly: 100 segments are the training dataset and the rest are the testing dataset. To compare the accuracies of the two scenarios without and with the integrated schedule data, the confusion matrices are adopted as illustrated in Figure 11 and 12. The classification accuracy is only 78.6% without the integration of the schedule information and 99.7% with the schedule-based dynamic training data library. These experiments clearly show that a schedule-based dynamic training framework significantly improves the accuracy of construction work sound classification.

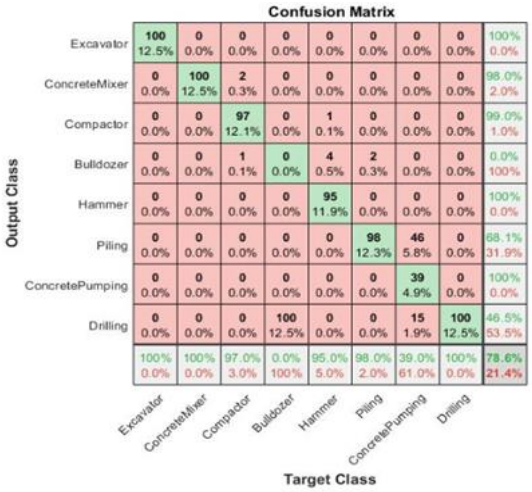

Figure 11. Classification accuracy without schedule information

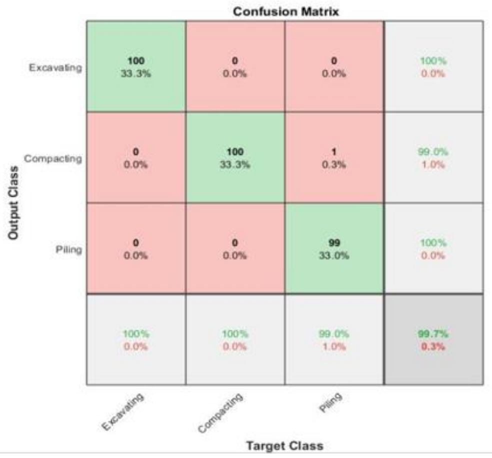

Figure 12. Classification accuracy with dynamic training data library

3.7 Web-based Visualization Map

As a visualization interface, a website has been developed to represent real-time construction work status and progress on different locations. The Web framework has been built with the Python Django framework and Leaflet is used for the base map tiles. Mapbox, ArcGIS and OSM tiles have also been integrated to the website for a better perspective of visualization data. The backend of the website is connected to a PostGreSQL database that is using the cloud Relational Database Service (RDS) from Amazon Web Services (AWS). An additional script written in Python has been used to read data from a JSON file and update the cloud database accordingly. Finally, the website has been deployed using Heroku. The website map representing site locations, job types, and working dates is shown in Figure 13.

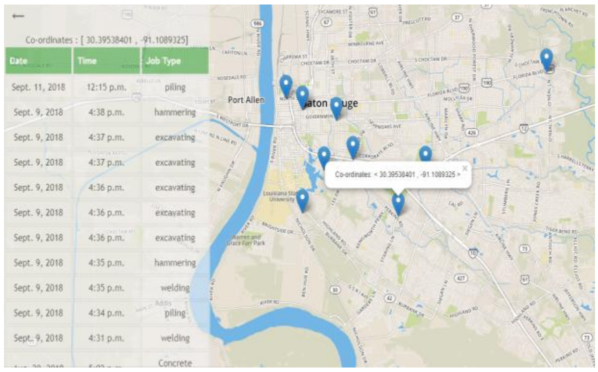

Figure 13. Construction Activity Visualization

4 Conclusion

The sound recognition technology has been adopted in diverse disciplines, but not received much attention in the construction industry. This paper proposes a schedule-based approach to improve the sound classification accuracy for providing robust construction event detection and site monitoring frameworks. As the outcomes show the significantly improved accuracy, this method is expected to be an imperative asset for diverse disciplines to develop an innovative technology that captures, analyses, and visualizes sounds of construction work and equipment in a dynamic and complicated construction site. The research results also involve new knowledge on the implications of sound recognition in the construction industry for accurately identifying working status and progress monitoring without any human effort. However, there are a few limitations identified in this research such as the insufficient size of sound data. The size of the data is not huge enough to validate the result and this research team will keep collecting sound data including various types of work activities for future research. In addition, the identification of heterogeneous sounds occurring simultaneously by using neural network is necessarily to be discussed in the next study and the recognition of sounds that do not belong to any of the categories such as the recognition of idle time will be also investigated. It is expected that the developed theoretical framework will add the innovative scientific knowledge and the new logical theory which pose significant impacts on the workflow optimization, the accurate task-performance measurement, and civil infrastructure construction projects for roads, highways, bridges, and tunnels.

5 References

[1] Teizer, J., Lao, D., & Sofer, M. (2007, September). 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* (pp. 19-21).

[2] Cheng, T., & Teizer, J. (2013). Real-time resource location data collection and visualization technology for construction safety and activity monitoring applications. *Automation in Construction*, *34*, 3-15.

[3] Cheng, C. F., Rashidi, A., Davenport, M. A., & Anderson, D. (2016, January). Audio Signal Processing for Activity Recognition of Construction Heavy Equipment. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction* (Vol. 33, p. 1). Vilnius Gediminas Technical University, Department of Construction Economics & Property.

[4] Cho, C., Lee, Y. C., & Zhang, T. (2017). Sound recognition techniques for multi-layered construction activities and events. In *Computing in Civil Engineering 2017* (pp. 326-334).

[5] Brilakis, I., Fathi, H., & Rashidi, A. (2011). Progressive 3D reconstruction of infrastructure with videogrammetry. *Automation in Construction*, *20*(7), 884-895.

[6] Golparvar-Fard, M., Peña-Mora, F., & Savarese, S. (2009). Application of D4AR–A 4-Dimensional augmented reality model for automating construction progress monitoring data collection, processing and communication, ITcon Vol. 14, Special Issue Next Generation Construction IT: Technology Foresight, Future Studies, Roadmapping, and Scenario Planning, pg. 129-153.

[7] Turkan, Y., Bosche, F., Haas, C. T., & Haas, R. (2012). Automated progress tracking using 4D schedule and 3D sensing technologies. *Automation in Construction*, *22*, 414-421.

[8] Hirsch, H. G., & Pearce, D. (2000). The Aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions. In *ASR2000-Automatic Speech Recognition: Challenges for the new Millenium ISCA Tutorial and Research Workshop (ITRW)*.

[9] Rangachari, S., & Loizou, P. C. (2006). A noise-estimation algorithm for highly non-stationary environments. *Speech communication*, *48*(2), 220-231.