A lossy 4D point cloud (4DPC) data compression method for LiDAR surveillance streaming in construction management

Yibo Wang[1*], Fan Xue[1], Zhe Chen[1], Lingming Kong[1]

Abstract: Traditional 2D surveillance video has been widely adopted in construction project management, yet facing several challenges. An emerging surveillance technology is the latest Light Detection and Ranging (LiDAR) that streams time-dynamic 4D point cloud (4DPC), which gained wide attention in the automobile and engineering sectors. However, one of the major defects of LiDAR surveillance streaming is the huge data volume. This paper presents a novel 4DPC data compression method to ease the problem. First, the 4DPC data frame is transformed into a depth image; then, the depth and laser intensity information are coded in the RGB color spectrum. As a result, a stationed 4DPC data stream is converted into a 3D surveillance video coded in H.264. A pilot test was conducted on the 4DPC data collected in a real-world project. The results confirmed that the presented method effectively encodes the 4DPC data stream with a high (1:0.08) compression ratio at a very low (0.5%) loss ratio at 5-cm resolution to common point cloud formats such as .ply and .pcd. The result of 4DPC LiDAR surveillance video can be played in 3D view mode by a designated player. The contribution of this paper lies in a novel video-based 4DPC data compression for LiDAR surveillance video streaming in construction project management.

Keywords: 4D point cloud; point cloud data compression; LiDAR streaming; surveillance video; construction project management.

[1] Department of Real Estate and Construction, The University of Hong Kong
[*] Corresponding author, E-mail: yibo@hku.hk

1 Introduction

Diverse construction projects worldwide require the surveillance or monitoring of activities, progress, and safety problems, such as risky worker activities. Effective surveillance videos can provide contractors and stakeholders with rich information about construction sites and worker states to conduct important control decisions. With the rapid development of information technology and computer vision in recent years, high-resolution automated cameras play a major role in construction projects to collect on-site static images or dynamic videos (Bohn & Teizer 2010). By analyzing the collected data, supervisors can facilitate the progress of construction projects and find the potential risks in time. Therefore, these types of cameras have been widely installed and used in construction projects for monitoring construction activities such as safety management to reduce high-risk accidents (Luo et al. 2020).

However, a single high-resolution camera typically meets several limitations: 1) lack of depth information (Chen et al. 2017), and 2) heavy reliance on visible light (Pathak et al. 2015). The lack of depth measurement brings the inconvenience of monitoring some specific construction activities like excavation. With the growth of sensing technologies, some emerging low-cost LiDAR sensors that collect spatial and temporal information have been utilized in the construction industry to address the above issues (Liang & Xue 2022). Benefiting from the time-dynamic 4D point cloud (4DPC) data from these LiDAR sensors, the surveillance or monitoring of construction activities can be performed in a 3D view and conducted all the time.

Regardless of cameras or LiDAR sensors, the collected data is stored in the local processors and transmitted to other devices or servers in need via the network. Therefore, a large storage volume and bandwidth are required to record and process a huge amount of data with rich information. During the past decades, image and video compression techniques have been widely explored in the research field for reducing storage space and transmission bandwidth, and have matured with the development of MPEG standards (Gall 1991). Additionally, point cloud compression (PCC) methods are also developed with the growing usages of point cloud data especially in the construction industry.

4DPC data collected from novel LiDAR sensors is suitable for generating 3D surveillance video for construction project management. However, it also encounters the issues of heavy requirement of storage and bandwidth because of the simultaneous collection of spatial geometric information and temporal sequences. Therefore, to better leverage the convenience and efficiency of 4DPC, we focus on these research questions: 1) how to eliminate the spatial and temporal redundancies of 4DPC data, and 2) how can the 4DPC compression technique be applied in practical real-world projects?

To address these issues, the main contributions of this work are summarized as follows: 1) we propose a novel 4DPC compression method that encodes data collected from time-dynamic LiDAR sensors into a compact video format. 2) we conduct a series of experiments to verify the proposed method in terms of two aspects the compression effectiveness and the quality of the reconstructed 4DPC. By relying on the proposed 4DPC compression method, LiDAR surveillance streaming can be smoothly generated by 4DPC data in construction project management.

2 Related Work

Surveillance videos have been increasingly developed and applied to the construction field for better project progress and safety management. Based on the intelligent camera, unsafe behaviors of workers are detected by project managers for critical control (Guo et al. 2018). To further monitor multiple workers at the construction site by cameras, a tracking method based on machine learning was developed and tested by Yang et al. (2010). Additionally, a graph-assisted hierarchical reinforcement learning algorithm was applied to address the limitations of existing surveillance video for smart construction (Ming et al. 2022). For safety assessment in the construction industry, a case study on metro tunnel construction was conducted to test the efficiency of the safety checklist analysis approach using intelligent video surveillance (Guo et al. 2021). To address the labor-intensive and time-consuming observation of construction productivity, researchers proposed a novel framework using on-site surveillance videos with three convolutional neural networks to detect, track, and recognize excavation activities (Chen et al. 2020). However, the traditional surveillance video lacks 3D geometric information like depth.

Point cloud data is the fundamental data for 3D information, which has been increasingly developed for various applications in construction project management. To overcome the limited human resources, health, safety, and time of the traditional in-person site inspection, a smart construction monitoring system supported by photogrammetry and LiDAR-derived 4D digital model was proposed (Cheng et al. 2022). Compared with the conventional 3D point cloud, 4DPC has more potential for construction project management. For instance, 4DPC data can be used to conduct semantic registration for monitoring crane-related activities on the construction site (Liang et al. 2023).

However, the large volume of 4DPC data collected by LiDAR sensors hinders its application. For example, the large size of it leads to poor performance in data transmission and storage. Therefore, compressing 4DPC data, namely identifying and eliminating spatial and temporal redundancies, is important for high-speed storing, transmitting, and processing (Abdelwahab et al. 2019). For different LiDAR sensors, an Ethernet interface solution for data packet decoding and reconstruction utilizes hardware-assisted algorithms to improve data transmission (Cunha et al. 2022).

Different from hardware-accelerated methods, the performance of point cloud compression can also be improved by analyzing the data structure. Typical compression algorithms can be classified into four groups (Roriz et al. 2024): 1) Coding-based compression; 2) Format-based compression; 3) 2D compression; 4) 3D compression. With the rapid development of computer vision technologies and mature applications of image or video compression algorithms, 2D-based point cloud compression methods have played a major role in recent years. To leverage the advantages of 2D representations, 3D/4DPC data is converted into a 2D structure as range or reflectivity images. Based on the projected image, several compression algorithms are proposed. For instance, combined with range image-based segmentation and novel prediction methods, an algorithm tailored for eliminating spatial redundant information was proposed (Sun et al. 2019). Additionally, a baseline for range image-based point cloud compression was proposed which can reconstruct the point cloud with uniform or non-uniform accuracy loss (Wang et al. 2022).

While the approaches can identify and eliminate the spatial redundancies of point cloud data with a lossy or lossless method, few of them are tailored for the novel 4DPC data compression consisting of spatial and temporal information. Leveraging the continuous characteristic of temporal information is a great way to compress 4DPC in depth but it still needs to be researched. In this paper, we propose a novel compression method combined with the temporal information of 4DPC data collected from novel LiDAR sensors and video compression techniques tailored for conducting 3D video surveillance in construction project management. Compared with the aforementioned literature, our method can effectively encode the temporal information of 4DPC data except the spatial information.

3 4DPC Compression

The architecture of the proposed lossy 4DPC compression algorithm is presented in Figure 1, where the input time-dynamic 4DPC is transformed into the Spherical coordinate system and projected on the converted range image with RGB information by its 2D index. Range images are eventually encoded into a video and stored as a data stream combined with the corresponding index file.

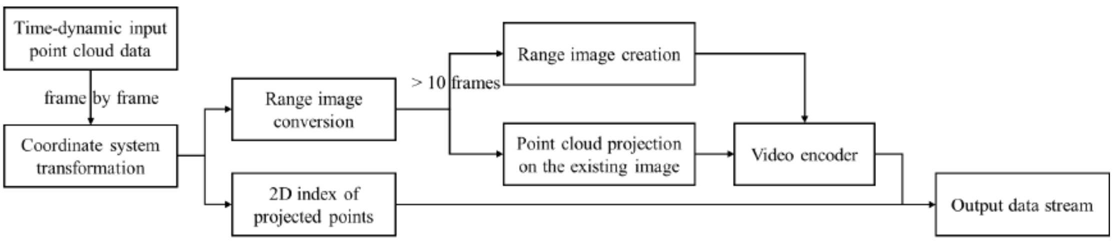

Figure 1. The architecture of the 4DPC compression algorithm.

3.1 Coordinate system transformation

To comprehensively leverage the temporal information of novel LiDAR sensors, the characteristics of 4DPC, and the advantages of image or video compression methods, we first transform the 4DPC data from a Cartesian coordinate system to a Spherical coordinate system. Each Cartesian point <*x, y, z*> in a 4DPC frame can be represented as <*r, ρ, θ*> in a Spherical coordinate system illustrated in Figure 2. In this way, the value <*ρ, θ*> of a point can determine its 2D index on the range image and is saved in a specific file during this process. These files will be utilized combined with the range and intensity value recorded on the range image to decode and reconstruct the original 4DPC data frame-by-frame.

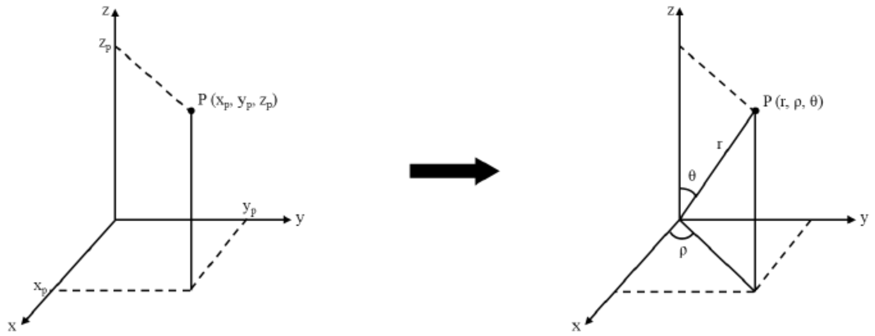

Figure 2. Transformation from Cartesian to Spherical coordinates.

3.2 Range image conversion

The details of the range image conversion can be divided into two main parts: geometric projection and attribute allocation. For geometric projection, since each point in a LiDAR frame is represented by its spherical coordinates <r, ρ, θ>, the angle value <ρ, θ> divided by their resolution can determine the location on the range image of this point and the 2D index is recorded individually. When focusing on the attribute allocation, the scope of the LiDAR projected range is typically lower than 200, the integer and decimal parts of range value <r> can be respectively saved in two channels of an RGB image with a maximum integer value of 255. In addition, the other channel of the RGB image is suitable for storing the intensity float value between 0 and 255. For instance, the integer part of the range value of the projected point can be saved in the red channel, the decimal part is saved in the green channel, and the intensity value of the point is stored in the blue channel. The process of range image conversion is illustrated in Figure 3.

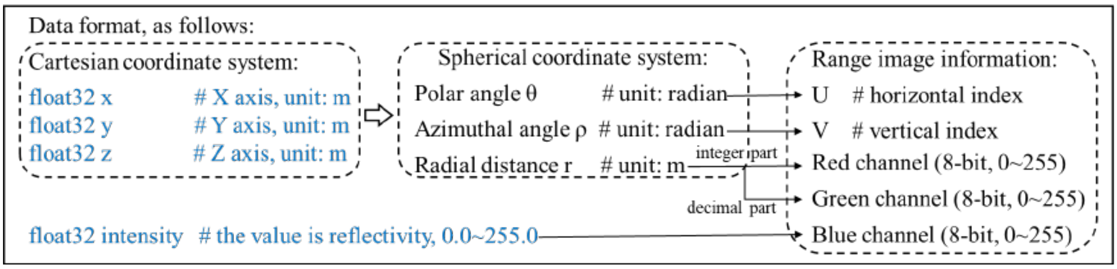

Figure 3. Projection principle from one 4DPC frame to RGB image.

Although the intensity value of a point also contains the decimal part, it is not important to be recorded and neglected in the process of data encoder. This will be qualitatively illustrated in the experimental evaluation part.

3.3 Video encoder

Following the above conversion rules, each point in a 4DPC frame can be projected on the range image. However, the range image produced by only one single 4DPC frame is less rich, thereby leading to the worse performance of the classical video compression algorithms in our case. To address this issue and fully leverage the advantages of video encoders, 4DPC data is frame-by-frame cumulated by its continuous timestamp information. Since novel LiDAR sensors produce the point cloud in a non-repetitive method, continuous 4DPC frames can therefore be projected on different locations of the range image. In this study, ten 4DPC data frames were projected and stored in the same image.

The results of single-frame projection and cumulative-frame projection are shown in Figure 4. Although continuous 4DPC data is produced by novel LiDAR sensors in a non-repetitive way, more than one point is still possible to be projected at the same pixel of the range image. To address this limitation, we calculated the average value of points located at the same pixel to represent its value.

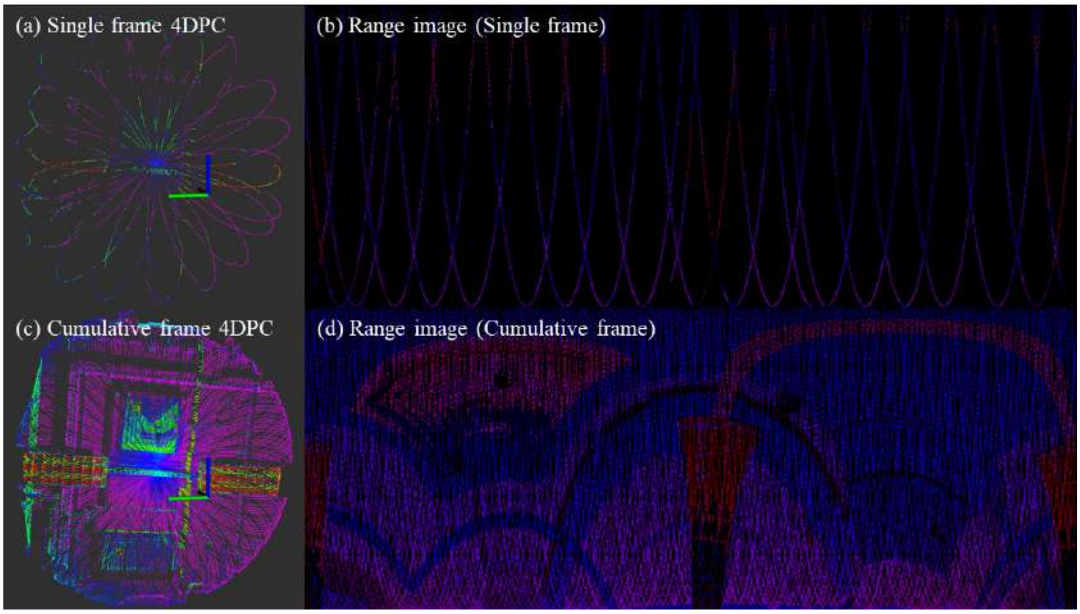

Figure 4. Range images projected by a single 4DPC frame and cumulative 4DPC frame.

After obtaining the range image with image richness projected by the cumulative 4DPC frame, we applied the general video compression method AVC H.264 (Wiegand et al. 2003) to process these images for further storage and replay in the future. It is worth mentioning that continuous 4DPC data is encoded in the compact format of a video through this method. Additionally, with the growth of time for collecting 4DPC data, the performance of our proposed compression method based on video encoders is also evident. This is because an increasing number of 4DPC data is converted to the range image and encoded as a video by the video encoders which are good at capturing and recording the difference between each image.

3.4 4DPC data reconstruction

Based on the previous encoding method, a video and some index files are stored locally. When conducting the reconstruction process, the video will first be divided into multiple image frames. Each frame contains ten 4DPC frames projected on it. Therefore, we can simply leverage the corresponding index files for inverse projection. By means of this, the original 4DPC data will be reconstructed frame-by-frame. It is worth mentioning that using the average geometric and attribute information value of points projected onto the same pixel, to represent that of all points, may result in a minor distortion in the reconstructed point cloud as compared to the original point cloud.

4 Experimental Evaluation

This paper employs a variety of performance metrics to assess point cloud compression, including the bits per frame that calculate the average storage space of each 4DPC frame and bits per point. The bits per 4DPC frame can be leveraged to compare the performance of different 4DPC compression methods. In this study, we respectively recorded and saved 4DPC data in the format of the bin, las, pcd, ply, and our proposed video. The bits per frame and bits per point of each format are listed in Table 1.

Table 1. Compression efficiency of various data formats.

Group	Data format	Bits per frame (KB)	Bits per point (KB)
Common format	.bin	175.6	0.0176
	.las	331.6	0.0333
	.pcd	341.2	0.0342
	.ply	341.3	0.0342
Ours	.avi	**27.5**	**0.00276**

Encoding 4DPC data into a video file is the most effective method for storing it with 27.5 KB per frame and 0.00276KB per point, about 1:0.08 compression ratio to common point cloud formats such as .ply and .pcd. Therefore, the proposed 4DPC compression method can achieve a high efficiency compared with other data encoders.

To assess the performance of 4DPC reconstruction, we conducted a qualitative analysis by decoding the 4DPC data compressed in a video to a point cloud sequence and visualizing that in the same platform. The front and side views of the reconstructed 4DPC compared with the original 4DPC are illustrated in Figure 5.

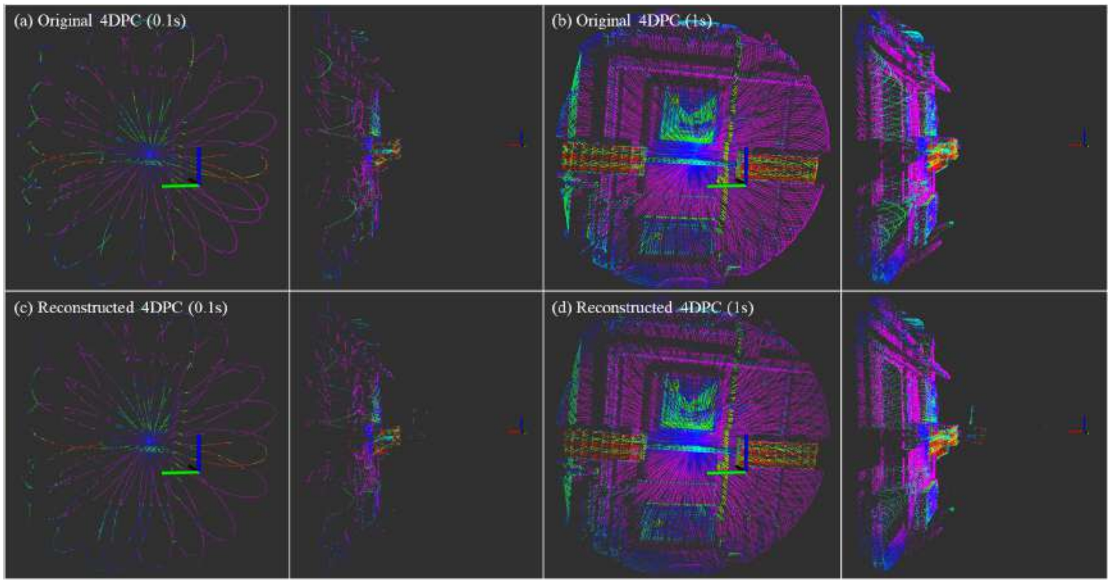

Figure 5. Front and side views of original and reconstructed 4DPC.

As shown in Figure 5, the 4DPC data is frame-by-frame decoded and details of the original 4DPC are successfully reconstructed. It is worth mentioning that the change in intensity value cannot be directly found in Figure 5 because its difference is not evident in the process of 4DPC encoding and decoding. Therefore, the float type of intensity information for each point in a 4DPC frame can be simply encoded in a channel of RGB range image. However, there is still some distortion between the reconstructed 4DPC and the original one. This is because we calculated the average pixel value and directly used this value to represent the geometric and attribute information of each point located on this pixel. Therefore, the decoded 4DPC is relatively distorted compared to the original 4DPC.

In addition to a qualitative analysis of the reconstruction process, we quantitatively assess the Precision, Recall, and F1-score values at 1 cm, 2 cm, and 5 cm tolerance thresholds between the reconstructed 4DPC and the original one as listed in Table 2.

Table 2. Precision, Recall, and F1-score values at 1 cm, 2 cm, and 5 cm tolerance thresholds

Tolerance	Precision (%)	Recall (%)	F1-score (%)
1 cm	87.07	99.08	92.69
2 cm	98.70	99.19	98.95
5 cm	99.78	99.20	99.49

Except for the extreme tolerance condition of 1 cm, the precision rate of the reconstruction process of our proposed 4DPC compression method exceeds 90 percent. It is noteworthy that the recall rate approaches 100 percent across all tolerance thresholds. Furthermore, the F1-score demonstrates an upward trend and is increasingly near 100 percent as the tolerance threshold increases, indicating that the proposed method achieves a well-balanced performance, given an appropriately set tolerance threshold.

5 Discussion and Conclusion

In the construction management field, the application of 4DPC data is important such as surveillance streaming for monitoring construction activities. However, the large volume of 4DPC data challenges the surveillance video storage and transmission, thereby compressing it is needed to be explored.

In this paper, a novel and efficient compression method for 4DPC data, which is suitable for encoding and storing data with large volumes in construction project management, has been proposed. The proposed method adapts the advantages of image and video compression algorithms to record 4DPC data as a format of video. Experimental results showed that the 4DPC data sequences can be frame-by-frame compressed in a compact format of video with about 1:0.08 compression ratio. Additionally, significant geometric and attribute information is recorded when encoding and reconstructed with a high F1-score (99.49%) at a 5-cm tolerance threshold during the decoding process.

Therefore, based on the proposed method, 4DPC data can be encoded into a video file that contains rich information on 4DPC data that can be easily stored locally or transmitted with a high speed and a low bandwidth between various construction departments. It will increase the efficiency of the construction project using 4DPC data and create a better surveillance video for construction contractors and stakeholders to conduct management and make important decisions immediately. This kind of surveillance video can also provide construction managers with a novel digital management method that the efficiency of production processes can be evidently improved.

However, there are still some limitations in this study that need more cases for feasibility verification and the algorithm needs to be improved to adapt to more different scenarios. Further study will focus on the improvement of encoding efficiency and reconstruction quality of the proposed method to produce a better 3D surveillance video, thereby creating convenience for construction management. In addition, the proposed method will be further verified with more data formats and other compression algorithms.

Acknowledgment

The work presented in this paper was supported by The University of Hong Kong (No. 2307102469).

References

Abdelwahab, M. M., El-Deeb, W. S. & Youssif, A. A. (2019). LiDAR data compression challenges and difficulties. *5th International Conference on Frontiers of Signal Processing (ICFSP)* (pp. 111-116). IEEE. doi:10.1109/ICFSP48124.2019.8938066

Bohn, J. S. & Teizer, J. (2010). Benefits and barriers of construction project monitoring using high-resolution automated cameras. *Journal of construction engineering and management, 136*(6), 632-640. doi:10.1061/(ASCE)CO.1943-7862.0000164

Chen, C., Zhu, Z. & Hammad, A. (2020). Automated excavators activity recognition and productivity analysis from construction site surveillance videos. *Automation in construction, 110*, 103045. doi:10.1016/j.autcon.2019.103045

Chen, J., Fang, Y. & Cho, Y. K. (2017). Real-time 3D crane workspace update using a hybrid visualization approach. *Journal of Computing in Civil Engineering, 31*(5), 04017049. doi:10.1061/(ASCE)CP.1943-5487.0000698

Cheng, S. Y., Liu, L., Hou, W., Hart, J. R. & Yong, Y. M. (2022). Smart Construction Monitoring Using Photogrammetry and LiDAR-derived 4D Digital Model: A Case Study from the

Tung Chung New Town Development of Hong Kong. *AIJR Proceedings* (pp. 129-140). AIJR. doi:10.21467/proceedings.133.11

Cunha, L., Roriz, R., Pinto, S. & Gomes, T. (2022). Hardware-accelerated data decoding and reconstruction for automotive LiDAR sensors. *IEEE Transactions on Vehicular Technology, 72*(4), 4267-4276. doi:10.1109/TVT.2022.3223231

Gall, D. L. (1991). MPEG: A video compression standard for multimedia applications. *Communications of the ACM, 34*(4), 46-58. doi:10.1145/103085.103090

Guo, S., Li, J., Liang, K. & Tang, B. (2021). Improved safety checklist analysis approach using intelligent video surveillance in the construction industry: a case study. *International journal of occupational safety and ergonomics, 27*(4), 1064-1075. doi:10.1080/10803548.2019.1685781

Guo, S., Xiong, C. & Gong, P. (2018). A real-time control approach based on intelligent video surveillance for violations by construction workers. *Journal of Civil Engineering and Management, 24*(1), 67-78. doi:10.3846/jcem.2018.301

Liang, D. & Xue, F. (2022). Applications of 4D Point Clouds (4DPC) in Digital Twin Construction: A SWOT Analysis. *Proceedings of the 27th International Symposium on Advancement of Construction Management and Real Estate* (pp. 1231-1238). Singapore: Springer. doi:10.1007/978-981-99-3626-7_95

Liang, D., Chen, Z., Kong, L., Wu, Y., Chen, S.-H. & Xue, F. (2023). 4D Point Cloud (4DPC)-driven real-time monitoring of construction mobile cranes. *European Conference on Computing in Construction* (pp. 0-0). Flemish Region, Belgium: European Council on Computing in Construction (EC3). doi:10.35490/EC3.2023.258

Luo, H., Liu, J., Fang, W., Love, P. E., Yu, Q. & Lu, Z. (2020). Real-time smart video surveillance to manage safety: A case study of a transport mega-project. *Advanced Engineering Informatics, 45*, 101100. doi:10.1016/j.aei.2020.101100

Ming, Z., Chen, J., Cui, L., Yang, S., Pan, Y., Xiao, W. & Zhou, L. (2022). Edge-Based Video Surveillance With Graph-Assisted Reinforcement Learning in Smart Construction. *IEEE Internet of Things Journal, 9*(12), 9249 - 9265. doi:10.1109/JIOT.2021.3090513

Pathak, P. H., Feng, X., Hu, P. & Mohapatra, P. (2015). Visible light communication, networking, and sensing: A survey, potential and challenges. *IEEE communications surveys & tutorials, 17*(4), 2047-2077. doi:10.1109/COMST.2015.2476474

Roriz, R., Silva, H., Dias, F. & Gomes, T. (2024). A Survey on Data Compression Techniques for Automotive LiDAR Point Clouds. *Sensors, 24*(10), 3185. doi:10.3390/s24103185

Sun, X., Ma, H., Sun, Y. & Liu, M. (2019). A novel point cloud compression algorithm based on clustering. *IEEE Robotics and Automation Letters, 4*(2), 2132-2139. doi:10.1109/LRA.2019.2900747

Wang, S., Jiao, J., Cai, P. & Wang, L. (2022). R-pcc: A baseline for range image-based point cloud compression. *International Conference on Robotics and Automation (ICRA).* IEEE. doi:10.1109/ICRA46639.2022.9811880

Wiegand, T., Sullivan, G., Bjontegaard, G. & Luthra, A. (2003). Overview of the H. 264/AVC video coding standard. *IEEE Transactions on circuits and systems for video technology, 13*(7), 560-576. doi:10.1109/TCSVT.2003.815165

Yang, J., Arif, O., Vela, P., Teizer, J. & Shi, Z. (2010). Tracking multiple workers on construction sites using video cameras. *Advanced Engineering Informatics, 24*(4), 428-434. doi:10.1016/j.aei.2010.06.008