4D point cloud-based spatial-temporal semantic registration for monitoring mobile crane construction activities

Dong Liang[a], Sou-Han Chen[a], Zhe Chen[a], Yijie Wu[a], Louis Y.L. Chu[b], Fan Xue[a]*

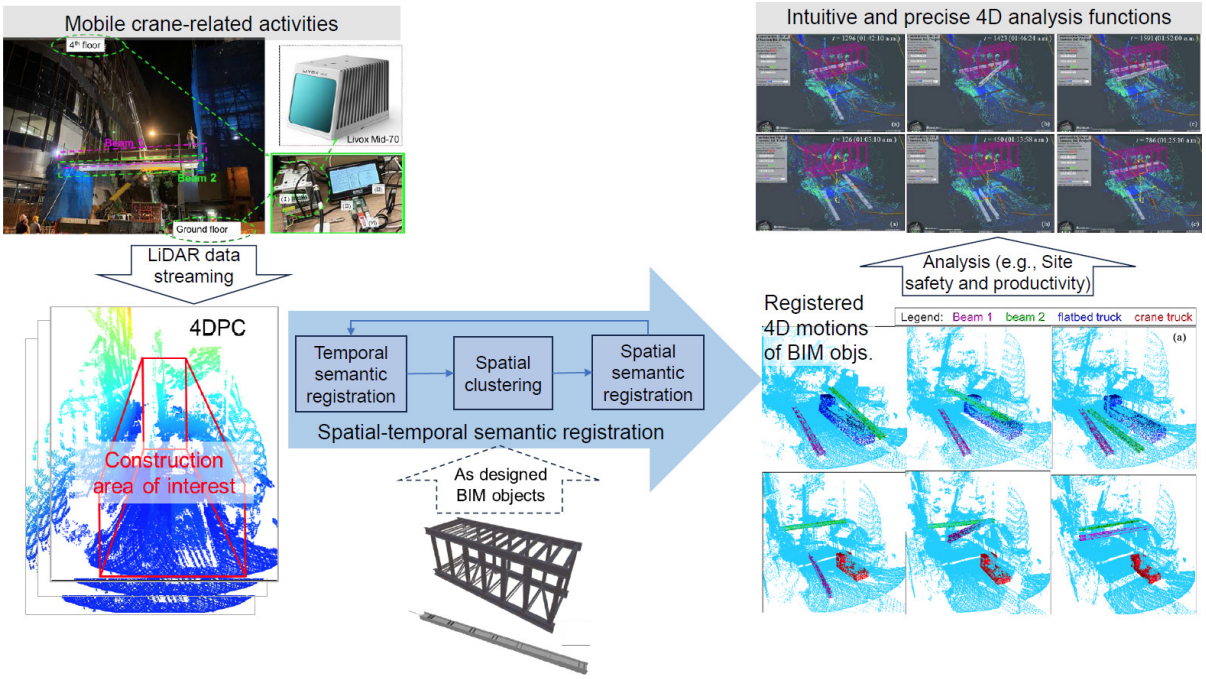
This is the peer-reviewed post-print version of the paper:
Liang, D., Chen, S.-H., Chen, Z., Wu, Y., Chu, L. Y.L. & Xue, F. (2024). 4D point cloud-based spatial-temporal semantic registration for monitoring mobile crane construction activities. *Automation in Construction*, 165, 105576. Doi: 10.1016/j.autcon.2024.105576
The final version of this paper is available at: https://doi.org/10.1016/j.autcon.2024.105576.

[a] Department of Real Estate and Construction, The University of Hong Kong, Pokfulam, Hong Kong, China

[b] Estates Office, The University of Hong Kong, Pokfulam, Hong Kong, China

* Corresponding author, Email: xuef@hku.hk, Tel: +852 3917 4174

Graphical Abstract



Highlights

- Registering spatial-temporal 4D point cloud (4DPC) with as-designed BIM objects for monitoring
- 4DPC-based spatial-temporal semantic registration method for mobile cranes
- Successful validation on one-hour 4DPC data from a night-time footbridge construction project
- Satisfactory registration accuracy (F1 = 99.95% on average) in 4DPC at object level
- Intuitive and precise analysis of mobile crane-related activities based on registration results

Abstract

Existing construction activity-monitoring technologies, such as CCTV cameras and IoT devices, have limitations, such as lack of depth information, 3D measurement errors, or wireless signal vulnerability. The limitations are particularly problematic for activities related to mobile cranes due to their high mobility and flexibility. This paper presents a 4D point cloud (4DPC)-based spatial-temporal semantic registration method to overcome the limitations. The proposed method integrates spatial-temporal semantic registration into process site 4DPC with as-designed BIM semantics. Results from a one-hour on-site experiment demonstrated that the proposed method achieved 99.93–100% F_1 accuracy in detecting BIM objects, and high resolution (centimeter-second granularity) of the trajectories of hoisting activities. This paper offers a two-fold contribution. First, spatial-temporal semantic registration represents an innovative approach to 4D point cloud (4DPC) processing. Secondly, the hoisting activities are comprehensively analyzed based on semantic registration, which can improve safety and productivity monitoring for smarter construction in the future.

Keywords: 4D point cloud (4DPC); Spatial-temporal semantic registration; Monitoring construction activities; Mobile cranes; Building information modeling (BIM).

1 Introduction

Cranes play a pivotal role in hoisting and transporting heavy objects in the construction industry (Neitzel et al. 2001). It is widely recognized that crane-related activities are closely associated with safety (Lee et al. 2012) and productivity issues (Sacks et al. 2005) in construction sites. In terms of safety, a considerable number of fatalities and injuries in construction sites were caused by crane-related accidents (Beavers et al. 2006). Meanwhile, Cheng and Teizer (2011) attributed many crane-related safety incidents, such as crashing, to restricted visibility during operations, such as sight occlusion, poor weather, or inadequate lighting conditions. Extra monitoring measures will improve operators' visibility. Regarding productivity, Lee et al. (2006) have found that effective sensing and monitoring of construction sites can improve productivity. Therefore, significant efforts have been dedicated to monitoring crane-related activities in the construction industry.

The mobile crane and the tower crane are the two primary types of crane commonly used in construction sites (Shapira et al. 2007; Görçün & Doğan 2023). The need to monitor mobile crane-related activities has been underscored in the construction industry (Teizer et al. 2022). The primary reason is that the high mobility and flexibility of mobile cranes mean more risks and uncertainties in construction sites compared to fixed tower cranes (Kan et al. 2018). These risks and uncertainties can be reduced through accurate sensing and understanding of mobile crane-related activities. Thus, there is a pressing need for a high-definition monitoring solution that fits the construction context and is specifically designed for mobile cranes.

Various sensing technologies, such as closed-circuit television (CCTV) cameras (Fang et al. 2018) and IoT devices (Mijwil et al. 2023), have been explored for monitoring mobile crane-related activities. However, several technical limitations of existing sensing methods are widely criticized for monitoring the construction site. For example, the lack of depth of the information from the CCTV camera hinders the accurate spatial positioning of objects (Chen et al. 2017). Meanwhile, the processing of images suffers from pixel resolution, lack of 3D measurements, continuous object tracking (Son et al. 2023), and triangulation errors (Liu et al. 2019). Internet of Things(IoT) devices (Chung et al. 2023)—such as Radio-Frequency Identification (RFID) (Lee et al. 2006), Global Positioning System (GPS) (Li et al. 2013), and Ultra-Wideband (UWB) (Shahi et al. 2013)—are popular in positioning and monitoring cranes but are limited to data resolution. Meanwhile, the IoT devices' wireless signals are also vulnerable to long-distance

or environmental occlusion (Udoh & Kotonya 2018). In addition, the installation of an IoT device for every object increases leading time and labor costs and decreases productivity. The RGB-D camera has been considered for sensing or monitoring crane-related activities, but its limited detection range, which is generally 0.8 to 5 meters (Alhwarin et al. 2014) restricts its applications in outdoor settings (Alphonse & Sriharsha 2021).

The emergence of 4D point cloud (4DPC) provides a novel way for sensing (Silva et al. 2022). Compared with traditional 3D point clouds which is a collection of three-dimensional data points or coordinates, 4DPC additionally incorporates temporal information for spatial data. In the construction industry, 3D point clouds representing static scenes for high-quality 3D reconstructed building information models (Jarząbek-Rychard & Maas 2023; Wu et al. 2021; Jiang et al. 2020; Xue et al. 2020; Wu et al. 2024; Yin et al. 2020); in contrast, 4DPC sensing can handle dynamic assignments, such as monitoring construction activities (Liang et al. 2023). Point cloud provide comprehensive information unaffected by lighting conditions, presenting advantages over two-dimensional pictures (Mirzaei et al. 2022; Li et al. 2022; Li et al. 2024; Li et al. 2023). In addition, 4DPC enables the observation of object movements in relation to their surroundings over time, granting it a distinct advantage over other motion monitoring technologies like various IoT devices and AI cameras (Liang & Xue 2022). Herein, the AI camera refers to systems embedded in ML/DL-driven vision detection algorithms, which help make intelligent real-time decisions.

However, it is challenging to achieve a semantic understanding of 4DPC data, and no ready-to-use method can be used to process 4DPC data for construction scenarios with a temporal nature. Existing semantic segmentation methods using training-relied deep learning were studied for 4DPC in applications, such as robotics and human actions (Liu et al. 2019; Fan et al. 2022). However, these training-relied methods can hardly be directly applied to the construction scenario with artifacts of modern construction projects and temporary project-based organizations. Meanwhile, the semantic result from these deep-learning-based methods is still at point level, from which it is hard to further develop some applications such as monitoring.

In contrast to semantic segmentation, Xue et al. (2019b) proposed a semantic registration approach to reuse the as-designed Building information modeling (BIM) semantics for automatic semantic modeling from 3D point clouds. Xue et al.'s (2019b) approach fundamentally differs from existing semantic segmentation approaches in that it is training-free, capable of processing complex scenes, and able to cleverly reuse existing information such as BIM models. From the perspective of machine learning, semantic segmentation is based on supervised learning with expensive training and labeling (Ma et al. 2020; Xia et al. 2022), while semantic registration is like unsupervised learning (Xue et al. 2019a; 2019b). Meanwhile, compared with the semantics of points provided by semantic segmentation, semantic registration could directly provide richer semantics at an object level.

This paper presents a 4DPC-based spatial-temporal semantic registration method for monitoring mobile crane-related activities. The presented method processes 4DPC data and as-designed BIM objects and is an optimization-driven semantic registration process. The most significant advantage of the proposed method is to fill the technical gap of the semantic enrichment of 4DPC in the construction scenario with an artifact-like and temporary nature. To validate the presented approach, we collected six-hour 4DPC data from a case study of a footbridge construction project and conducted experiments on registration accuracy and parameter sensitivity. Based on the semantic registration results, mobile crane-related activities can be accurately detected.

2 Literature review

2.1 Sensing and monitoring crane-related activities

In the construction industry, significant efforts have been dedicated to the development of computer-aided crane monitoring systems to enhance operational efficiency, ensure safety, and reduce operator workload. Table 1 compares related research in the literature, including the research focus, purpose, contributions, target crane type, method, and limitations.

Table 1 Summary of related studies in the literature

Research focus	Purpose and contribution	Sensing method	Crane type	Limitations	Source
crane pose estimation	Develop a UWB-based system to track crane boom movement, and estimate crane pose near real-time for collision avoidance	Imaging-based (camera) & IoT-based (UWB)	Mobile crane	· Long installation time of several sensors and multiple tags · High positioning error (0.1m) in open-space environment · Rough trajectory estimation based on linear interpolation extrapolation of two points	(Zhang et al. 2012)

	Develop a tower crane navigation system to help operators operate with blind spots	Imaging-based (camera) & IoT-based (encoder sensors)	Tower crane	· Many sensors used in the system. · High device and installation cost of several sensors and multiple tags · Virtual site environment update	(Lee et al. 2012)
	Develop a Safety Management System to monitor the status of tower crane groups and avoid collisions	IoT-based (encoder sensors)	Tower crane	· Only the main body of the crane considered	(Zhong et al. 2014)
	Understand construction activity by monitoring the pose of the tower crane	Imaging-based (camera)	Tower crane	· Low resolution with far distance and undesirable light condition	(Yang et al. 2014)
	Detect and Classify Cranes for monitoring crane-related safety hazards	Imaging-based (camera)	Tower crane	· Only the existence of crane detected	(Roberts et al. 2017)
	Recognize and 3D localize the crane activities	Imaging-based (camera)	Tower crane	· Sensitivity to the occlusion of boom · Inaccuracy of localization from some unrealistic assumption	(Wang et al. 2023)
crane pose estimation & load sway and rotation estimation & Object detection	Track load & Detect obstacle and worker	Imaging-based (camera and TLS) & IoT-based (encoder sensors and IMU)	Tower crane	sensor: · Errors from noisy encoders · Assumption failure due to large crane deflection vision: · Sensitivity to irregular lighting conditions and similar color between the load and environment	(Price et al. 2021)
load sway and rotation estimation	Develop real-time pro-active safety assistance for mobile crane hoisting operations	IoT-based (encoder sensors and IMU)	Mobile crane	· Static site reconstruction for environment sensing · High device and installation cost of several sensors for real-time sensing	(Fang et al. 2016)
	Track crane load sway	Imaging-based (camera)	Both	· Only the 2D location of the load identified. · Errors from similar color for objects and background	(Fang et al. 2018)
	Develop a novel method to detect and track the crane load fall zone	Imaging-based (camera)	Tower crane	· Overreliance on the quality of the training set	(Chian et al. 2022)
object detection in the workspace	Identify unauthorized work or entrance of personnel within a pre-defined risk zone	IoT-based (GPS and RFID)	Both	· Many tags, receivers, and other units used in the proposed system. · High cost and installation complexity of several sensors and multiple tags · Sensitivity to signal strength	(Li et al. 2013)
	· Detect object to update real-time 3D crane workspace	· Imaging-based (camera and TLS)	Mobile crane	· Inaccurate 2D bounding box. · Sensitivity to similar color · high positioning error (0.1-0.4m) in a non-occluded environment · Static environment reconstruction by single LiDAR data frame	(Chen et al. 2017)

#: UWB: Ultra-Wideband; GPS: Global Positioning System; RFID: Radio-Frequency Identification; WSN: Wireless Sensor Network; IoT: Internet of Things; IMU: Inertial Measurement Unit; UAV: Unmanned Aerial Vehicle; TLS: Terrestrial Laser Scanners.

The studies in Table 1 can generally be divided into three main categories: crane pose estimation, load sway and rotation estimation, and object detection in the workspace. The studies of crane pose estimation aimed to detect the crane's position, orientation, and motion to ensure precise and reliable operation. Early studies by Zhang et al. (2012), Lee, et al. (2012),

and Zhong, et al. (2014) mainly employed IoT sensors to achieve the aim. Then, more research (Yang et al. 2014; Roberts et al. 2017; Wang et al. 2023) used imaging-based methods due to the rapid development of computer vision and deep learning.

The second category focused on load sway and rotation estimation, both directly and indirectly, to monitor the load status. Load sway and rotation are related to crane stability and safety during hoisting operations. The indirect method including the study by Fang et al. (2016) used the physical model of mobile crane and sensing data from a series of encoder sensors and IMU to indirectly estimate the status of load. In comparison, research by Fang, et al. (2018) and Chian, et al. (2022) used the imaging-based method to directly capture the motion and rotation of load. Besides, Price et al. (2021) investigated both crane status monitoring and load detection, which were intended to assist operators in handling blind areas effectively.

The third category of related studies emphasized other construction objects, such as site workers, nearby vehicles, and construction components before hoisting, for identifying potential obstacles or hazards. Both IoT-based (Li et al. 2013) and imaging-based (Chen et al. 2017) methods were developed. However, the methods were mainly challenged by rapidly and dynamically changing poses of moving humans and vehicles (Nakanishi et al. 2022).

Overall, there were several common limitations, as listed in Table 1, of both IoT-based and imaging-based methods in terms of accuracy and real-timeliness (Liang & Xue 2022). For IoT-based methods, the most criticized aspect was the vulnerability of signals caused by long distances and obstacles (Zhang et al. 2012). Furthermore, it was always complex, labor-consuming, and time-consuming to deploy and maintain a fleet of sensors, such as RFID tags and receivers, at construction sites. In addition, a geometrically complex and time-dynamic site was reported to be too hard to cover by a limited fleet of sensors (Bohn & Teizer 2010). For imaging-based methods, images or videos from cameras were criticized for the absence of depth information and the sensitivity to lighting conditions and color variations (Chen et al. 2017). Moreover, 3D laser scanning through TLS was known to be time-consuming for one scan and contain mismatched geometries from different timestamps of scanning (Russhakim et al. 2018). Therefore, there has been an urgent need for the development of more effective and comprehensive sensing methods for monitoring mobile-crane-related activities.

2.2 4DPC in construction

A point cloud comprises three-dimensional data points or coordinates as an unstructured collection, or "cloud" (Hu & Brilakis 2024). In comparison with 2D pictures, 3D point clouds provide more comprehensive information and are not affected by lighting conditions (Bhople et al. 2021). In the construction industry, 3D point clouds have been applied for as-built BIM reconstruction and digital twin development (Hu et al. 2023; Yin et al. 2023; Zhou et al. 2019). However, existing applications of the traditional 3D point clouds were limited to capturing the features of non-moving objects.

The advent of the 4DPC (Shi et al. 2020), consisting of three dimensions and time information, has opened new opportunities in site surveying and monitoring. The 4DPC was known to be capable of recording and understanding a time-dynamic 4D world that it was challenging for 2D images or static 3D point clouds to describe (Fan et al. 2021). Recently, the 4DPC led to many innovative studies that focused on subjects such as monitoring plant growth (Li et al. 2013), tracking pedestrians (Chiu et al. 2021), and capturing high-definition human motions (Fan et al. 2021). Moreover, spatial-temporal information in the 4DPC was proven to be valuable in various identification tasks, such as calculating the acceleration of moving objects or identifying human activities (Fan et al. 2021). The 4DPC was expected to be advantageous for sensing and monitoring mobile crane-related activities thanks to the spatial-temporal data and high adaptability in poor visibility environments (Liang & Xue 2022). In addition, 4DPC data streams from multiple devices can be directly registered to cover the blind areas for each other.

In other industries, a lot of effort on deep learning has been invested in semantic enrichment for the 4DPC in recent years (Shi 2023). The deep learning-based methods can be categorized as two groups, i.e., point-based (Fan et al. 2021; Fan & Yang 2019) and voxelization-based (Choy et al. 2019; Luo et al. 2018). Large-scale training and testing datasets were required to develop and validate the methods (Luo et al. 2023).

However, understanding and enriching the semantics of 4DPC were challenging in the construction industry in comparison with conventional 3D point clouds. One reason was the limited 4DPC datasets. Another reason was that the training-relied methods are hard to apply to construction projects with a temporary nature. Thus, existing methods for semantic

enrichment of 4DPC in other industries cannot be transferred to construction scenarios. Recently, Liang et al. (2023) proposed a geometry-based semantic segmentation method for 4DPC processing, but the method suffered from low generalization ability. Meanwhile, the semantic segmentation in Liang et al. (2023) produced point-level segmentation, which was insufficient for understanding and enriching the object-level semantics on the site. In addition, point cloud sensors recently used to collect 4DPC returned lower density of points than that of traditional 3D point clouds (Shi et al. 2020). The low density of 4DPC may cause the loss of crucial spatial information, which will also increase the difficulty of semantic understanding of point clouds.

2.3 3D Point cloud registration methods

Point cloud registration is the task that aligns two or more sets of points within a shared coordinate system (Huang et al. 2021). Point cloud registration has evolved from manual alignments to sophisticated automatic algorithms by leveraging advancements in computational power and machine learning (Gu et al. 2020). These new registration methods significantly enhance accuracy and efficiency in aligning multiple point sets within a shared coordinate system. Point cloud registration is vital to the remote sensing of 3D point clouds and various applications in computer vision, robotics, and construction informatics (Pomerleau et al. 2015). Kin et al. (2018) used 3D point cloud registration for robotic mapping in the indoor environment in the construction industry. In addition, 3D point cloud registration was used to align two photographs of defects between two inspections (Bush et al. 2022). However, incomplete data was highlighted as a common pain point of point cloud registration in the construction scenarios since there are many occlusions from construction equipment, materials, and temporary structures (Kim et al. 2018).

Existing 3D point cloud registration methods can be broadly divided into coarse and fine registration. Coarse registration methods, such as Fast Global Registration (FGR) (Zhou et al. 2016) and Random Sample Consensus (RANSAC) (Raguram et al. 2008), aim at providing a rough alignment of two or more point clouds and serve as a preliminary step for fine registration (Bueno et al. 2017). FGR and RANSAC work by finding correspondences between features extracted from the point clouds, and they do not require an initial guess of the alignment.

In contrast, fine registration methods were developed to improve the coarse registration results to achieve a more precise alignment of the point clouds (Al-Rawabdeh et al. 2020). A fine registration method usually requires an initial alignment as a starting point and focuses on minimizing the distance between corresponding points or features in the point clouds. Iterative Closest Point (ICP) (Rusinkiewicz & Levoy 2001) and Coherent Point Drift (CPD) (Myronenko & Song 2010) were among the most widely used methods for fine registration. ICP iteratively optimizes the alignment between a source and target point cloud by minimizing the distances between corresponding points, while CPD models point clouds as Gaussian mixture models and uses probabilistic optimization to globally estimate the non-rigid deformation relationship between the source and target point clouds. In the literature, fine registration was considered essential for high-precision applications, such as 3D modeling, object recognition, and robot navigation (Zhao et al. 2022). In most studies, the coarse-to-fine strategy is employed to solve registration problems of the 3D point cloud (Bosche 2010).

Point cloud registration methods were also extended to volumetric 3D models (Wang et al. 2016; Xue et al. 2019b; Zhang & Arditi 2013) in the construction industry. Zhang & Arditi (Zhang & Arditi 2013) aligned scanning 3D point clouds with simulated columns to achieve automated progress control. In addition, Wang et al. (Wang et al. 2016) registered BIM models for scanning 3D point clouds based on features on precast concrete elements to achieve automated quality assessment. Xue et al. (2019b) developed a derivative-free optimization-based (DFO) semantic registration method for registering volumetric and semantic BIM objects to 3D point clouds. The DFO semantic registration is designed as a derivative-free optimization process for finding the optimal position of a BIM object in the input 3D point clouds. The results showed that the proposed DFO semantic registration method achieved a coarse-to-fine registration process.

In summary, both IoT-based and image-based sensing approaches were limited in monitoring crane-related activities at construction sites. 4DPC is a promising sensing source but also challenging in data processing to monitor the activities. Inspired by some studies on registration methods between 3D point clouds and as-designed BIMs in the construction industry, it is promising to develop a semantic registration method for 4DPC using BIM semantics. However, existing registration methods for 3D point clouds cannot be directly used in 4DPC due to its low density and spatial-temporal characteristics. Therefore, developing a spatial-temporal

semantic registration method that processes site 4DPC with as-designed BIM semantics is significant.

3 The 4DPC-based spatial-temporal semantic registration method

3.1 Overview

Fig. 1 shows an innovative 4DPC-based spatial-temporal semantic registration method for monitoring mobile crane-related activities. The inputs to the method include 4DPC and as-designed BIM objects. The core of the proposed method consists of three iterative steps after 4DPC pre-processing: (i) Temporal semantic registration (local), (ii) Clustering, and (iii) Spatial semantic registration (global). Three key parameters of the method are the registration pattern (whether to enable Step (i)), the time interval between key frames from Step (iii), and the optimization algorithm in Steps (i) and (iii). Finally, the results of spatial-temporal semantic registration are integrated and demonstrated for construction management in the post-processing.

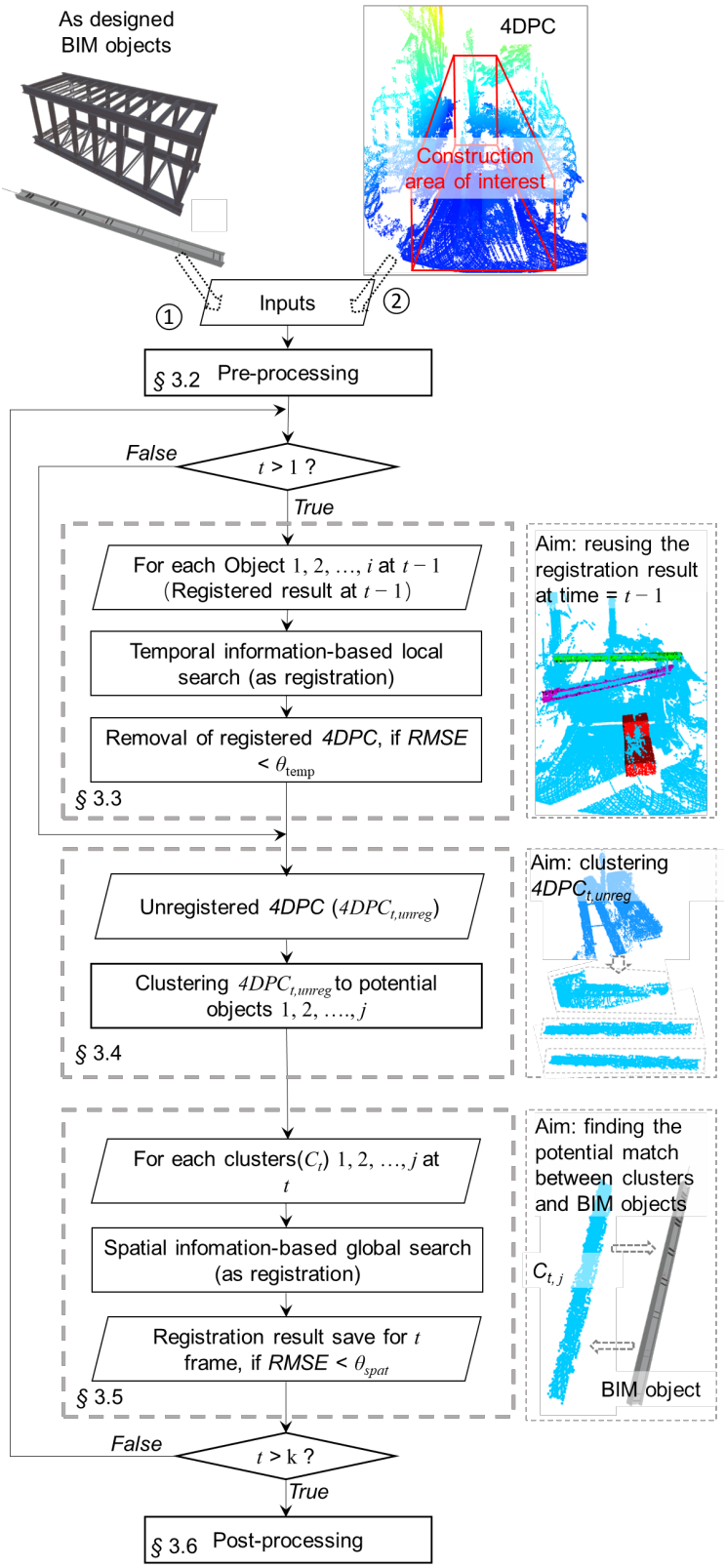


Fig. 1. Workflow of the proposed spatial-temporal semantic registration method

3.2 Pre-processing

The pre-processing part mainly includes two parts, i.e., 4DPC data registration from multiple LiDAR sensors and environment removal of integrated 4DPC data. For the first part, 4DPC data from multiple sources was registered using transformation parameters obtained by manual extraction in the first few frames of 4DPC data. The registration accuracy across frames was based on the fixed LiDAR sensors when collecting data.

For the second part, removing environmental points is as simple as cropping out the construction area of interest, as shown in Fig. 1. The environment elements within the construction area of interest, such as the ground, were removed manually using the RANSAC algorithm (Schnabel et al. 2007). The RANSAC algorithm includes three key parameters, i.e., *distance-threshold*, *ransac-n*, and *num-iterations*. In specific, *distance-threshold* indicates the distance threshold used to determine whether a data point is consistent with the estimated model. The *ransac-n* and *number-of-iterations* indicate the number of randomly sampled points used in each iteration and the number of iterations, respectively. The key parameters of the RANSAC algorithm were adopted from the suggested values from the algorithm library (open3d (Zhou et al. 2018) ver. 0.18.0) and calibrated using the initial 5 frames of 4DPC data. According to the tests, the one-off calibrated parameters effectively accomplished the task of environmental removal for the whole construction period. The successful one-off calibration was rooted in the consistent spatial information of the project's 3D environment – while the LiDAR scanners' poses were fixed throughout the entire data collection period.

3.3 Temporal semantic registration (local search)

Temporal semantic registration aims to avoid the repeated search of BIM objects in the following spatial semantic registration process. The design of this step is based on the fact of motion continuity of the construction object at site. Meanwhile, this step is activated when time $t > 1$. The temporal semantic registration of a BIM at time t is a nonlinear optimization problem:

$$\begin{aligned} &\textbf{arg min}\ \Sigma_{1\le i\le N} f_t(pos_{i,t}) \approx \textbf{arg min}\ f_t(pos_{1,t}), f_t(pos_{2,t}), \ldots, f_t(pos_{N,t}) \\ &\textbf{s.t.}\ |pos_{i,t} - pos_{i,t-1}| \le tr_{\max}, \text{ where } 1 \le i \le N \qquad \text{// continuous motion} \\ &\quad t > 1, \qquad \text{// not the first frame} \end{aligned} \tag{1}$$

where N is the number of semantic BIM objects, $i \in \{1, 2, \ldots, N\}$ is i-th BIM object, the $pos_{i,t} = \{x_{i,t}, y_{i,t}, z_{i,t}, rx_{i,t}, ry_{i,t}, rz_{i,t}\}$ is a 6-DoF pose of the i-th BIM object at time t, $|\cdot|$ is the absolute

value function. In order reduce the complexity of the problem, the minimization of a sum is approximated to the sum of the minimization of every single solution. The approximation "≈" in Eq. (1) is because all objects move independently and have no physical collisions. The first constraint in Eq. (1) indicates that any BIM object's pose change from time $t - 1$ to t is no greater than a given threshold $tr_{\max}$; that is all BIM objects move continuously.

If the DFO method is accepted to solve the Eq. (1), the objective function f_t in in Eq. (1) need to be further approximated by computing the root-mean-square error (*RMSE*) between the i_{th} BIM object's pose ($BIM(pos_{i,t})$) and input $4DPC_t$ in the current time frame (t). The approximate *RMSE* computation is presented in Eq. (2):

$$f_t(pos_{i,t}) \approx RMSE(BIM(pos_{i,t}), 4DPC_t) \quad \text{// most fitted} \quad (2)$$
$$= \mathrm{Sqrt}[\Sigma_{p_j \in BIM(pos_{i,t})}\, nndist(p_j, 4DPC_t)] \,/\, \| BIM(pos_{i,t}) \|$$

where $BIM(pos_{i,t})$ is a down-sampled point cloud from the i-th BIM, *nndist* denotes the Euclidean distance between point p_j and its nearest neighbor in 4DPC at time t, p_j is the j-th point in $BIM(pos_{i,t})$, and "$\| \cdot \|$" is the total number of points of in $BIM(pos_{i,t})$. The *nndist* metric can employ the octree data structure, which guarantees a balance between accuracy and computational time (Elseberg et al. 2013). Apart from *RMSE*, other metrics such as point cloud correspondence ratio and BIM object surface support density (number of associated points divided by surface area) are alternative metrics of f_t. The registration success is also measured by f_t in Eq. (2). If the value f_t is less than the threshold for temporal semantic registration (θ_{temp}), the successful registration result ($BIM(pos_{i,t})$) will be recorded. On the contrary, the i-th BIM object is considered out of the 4DPC.

Some modern optimization algorithms libraries, such as NLopt (ver. 2.7), offer off-the-peg DFO algorithms. Examples of update-to-update DFO algorithms are Bound Optimization BY Quadratic Approximation (BOBYQA) (Powell 2009), New Unconstrained Optimization with quadratic Approximation with bound constraints (NEWUOA_BND) (Powell 2006), Nelder–Mead (Nelder & Mead 1965), Constrained Optimization BY Linear Approximations (COBYLA) (Powell 1994), Dividing RECTangles (DIRECT) (Jones et al. 1993), PRincipal AXIS (PRAXIS) (Brent 2013). Many DFO algorithms were already confirmed to have outstanding capacity in optimizing complex problems (Siegbert et al. 2014; Wortmann et al. 2017).

In addition, some fine registration algorithms, such as ICP or CPD, can also be introduced to solve the Eq. (1), since the position in the last frame can be an excellent initial alignment for registration. After the registration of all existing objects, the registration result will be used to crop registered points out of the $4DPC_t$.

The significance of this step can be originated from two aspects. The first is that it increases the registration efficiency since the first constraint in Eq. (1) reduces the registration process's search space. In addition, the use of temporal information enhances the spatial information of point cloud by using the registration information from the previous frame, which could reduce the negative impact of incompleteness or noise of point cloud.

3.4 Density-based object clustering

This part aims to cluster unregistered 4DPC to potential objects. The input to this step is thus unregistered 4DPC ($4DPC_{t,unreg}$). Note that environment points have been removed from $4DPC_{t,unreg}$ in pre-processing. The target construction objects exhibited robust geometric connectivity and evenly distributed density. Therefore, density and connectivity-based clustering algorithms can be adopted for clustering, such as the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Ester et al. 1996). The resulting clusters (C_t) of interconnected point patches represent the surfaces of potential objects in the construction area of interest.

3.5 Spatial semantic registration (global search)

This part involves an optimization process for an optimal match between BIM objects and point cluster C_t. The registration objective f_t is to minimize the *RMSE* between the k_{th} BIM object (BIM_k) and j_{th} C_t cluster ($C_{j,t}$):

$$\textbf{arg min}\, f_t\,(pos_{k,t}) \approx \textbf{arg min}_{c_{t,j} \in C_t}\, RMSE(BIM(pos_{k,t}),\, C_{j,t}) \qquad \text{//best fitted patch} \tag{3}$$

The indicator (*RMSE*)of registration success is calculated as the minimal distance of every single point ($point_k$) in $BIM(pos_{i,t})$ to the $C_{t,j}$:

$$RMSE(BIM(pos_{i,t}),\, C_{t,j}) = \mathrm{Sqrt}[\Sigma_{p_j \in BIM(pos_{i,t})}\, nndist(p_j,\, C_{t,j})] \,/\, \|\, BIM(pos_{i,t})\| \tag{4}$$

If the minimal f_t is less than the threshold of spatial registration (θ_{spat}), it is considered to find the successful registration of $BIM(pos_{i,t})$ for k-th BIM objects and j-th cluster ($C_{j,t}$). After searching the global spatial area, the semantic registration for all clusters can be conducted

according to the optimization result. Meanwhile, the traditional coarse-to-fine registration method can also solve the Eq. (3). Compared with temporal semantic registration, spatial semantic registration is based on the random search in the global area of the point cloud frame, which is usually time-consuming. Therefore, the time interval is designed here to accelerate the registration process. Specifically, the time interval refers to the period for spatial optimization registration, i.e., conducting a global search to identify whether new objects enter the construction area of interest. A reasonable value of time intervals can achieve a good balance between the speed and accuracy of the proposed method.

3.6 Post-processing

Post-processing steps of the proposed method are also needed to visualize and apply the semantic registration results for site safety and productivity. First, a spatial-temporal load trajectory reconstruction process measures and restores the high-definition poses. The reconstructed trajectories can be visualized on the Cesium platform (ver. 1.117). Then, the construction activity analysis detects and measures the key states of construction activities. The post-processing outcomes can benefit crane operators, site managers, and other construction project stakeholders by enabling informed decision-making that will improve safety and productivity.

In conclusion, the presented method first introduces 4DPC to overcome the limitations of existing sensing methods for monitoring construction sites. Meanwhile, the three challenges of semantic enrichment for 4DPC data, that is, expensive training data, the temporary nature of construction projects, and only existing point-level results, are addressed by as-designed BIM objects, temporally inherited poses from the last frame, and the construction area of interest for mobile cranes, respectively.

3.7 Significance of the proposed method

The presented method is significant in three aspects. (i) A general spatial-temporal formulation for 4DPC semantic registration;

(ii) Efficient method with approximation in two levels, i.e., registration problem level and objective function level;

(iii) A flexible registration method compatible with various algorithms, including DFO and traditional coarse-to-fine 3D registration algorithms.

Overall, this is the first study towards the training-free semantic registration of 4DPC, as far as we are concerned.

3.8 Evaluation

To measure and compare the result among different registration methods and settings, in addition to *RMSE*, three other indicators, that is, precision, recall, and F_1, are introduced:

$$\begin{aligned} Precision &= TP / (TP + FP), \\ Recall &= TP / (TP + FN), \\ F_1 &= 2 \times Precision \times Recall / (Precision + Recall), \end{aligned} \tag{5}$$

where *TP*, *TN*, *FP*, and *FN* indicate true positives, true negatives, false positives, and false negatives for the registration result of objects, respectively. True signifies the object is precisely registered, while false indicates the opposite. Positive denotes the presence of the object within the point cloud frame, whereas false signifies the opposite.

The IoU_{3D} was used as an indicator to determine whether the BIM object is precisely registered:

$$IoU_{3D} = Volume_{overlap} / (Volume_g + Volume_{reg} - Volume_{overlap}) \tag{6}$$

where $Volume_g$ indicates the object volume of the ground truth annotated by authors, $Volume_{reg}$ indicates the volume of the registered object, $Volume_{overlap}$ indicates the overlap volume between $Volume_g$ and $Volume_{reg}$. When the IoU_{3D} is higher than 0.7, the object will be regarded as being successfully registered.

4 Experimental results

4.1 A test project

An experiment was conducted to validate the proposed method, as illustrated in Fig. 2. The test project was a footbridge project involving the installation of prefabricated steel members at midnight on June 28, 2022, in Hong Kong. The project conducted the hoisting and installation of two major steel tie-beams. Figs. 2 (a), (b), (c), and (d) show the site conditions before hoisting, the process of unloading the tie-beams from a flatbed truck, the installation of Beam 2 with a crane truck, and the completed status of the tie-beams, respectively. A mobile crane was utilized for hoisting operations in this infrastructure project. In addition, two other objects were involved: a flatbed truck, primarily used for transporting the beam, and a crane truck, utilized for further installation when the beam was hoisted into the installation position. In this

450 case, the proposed method aims to monitor hoisting load and on-site vehicles since they are the most significant source of risks (Forteza et al. 2017). The potential risks include the collision between hoisting loads and vehicles and between hoisting loads and surrounding buildings.

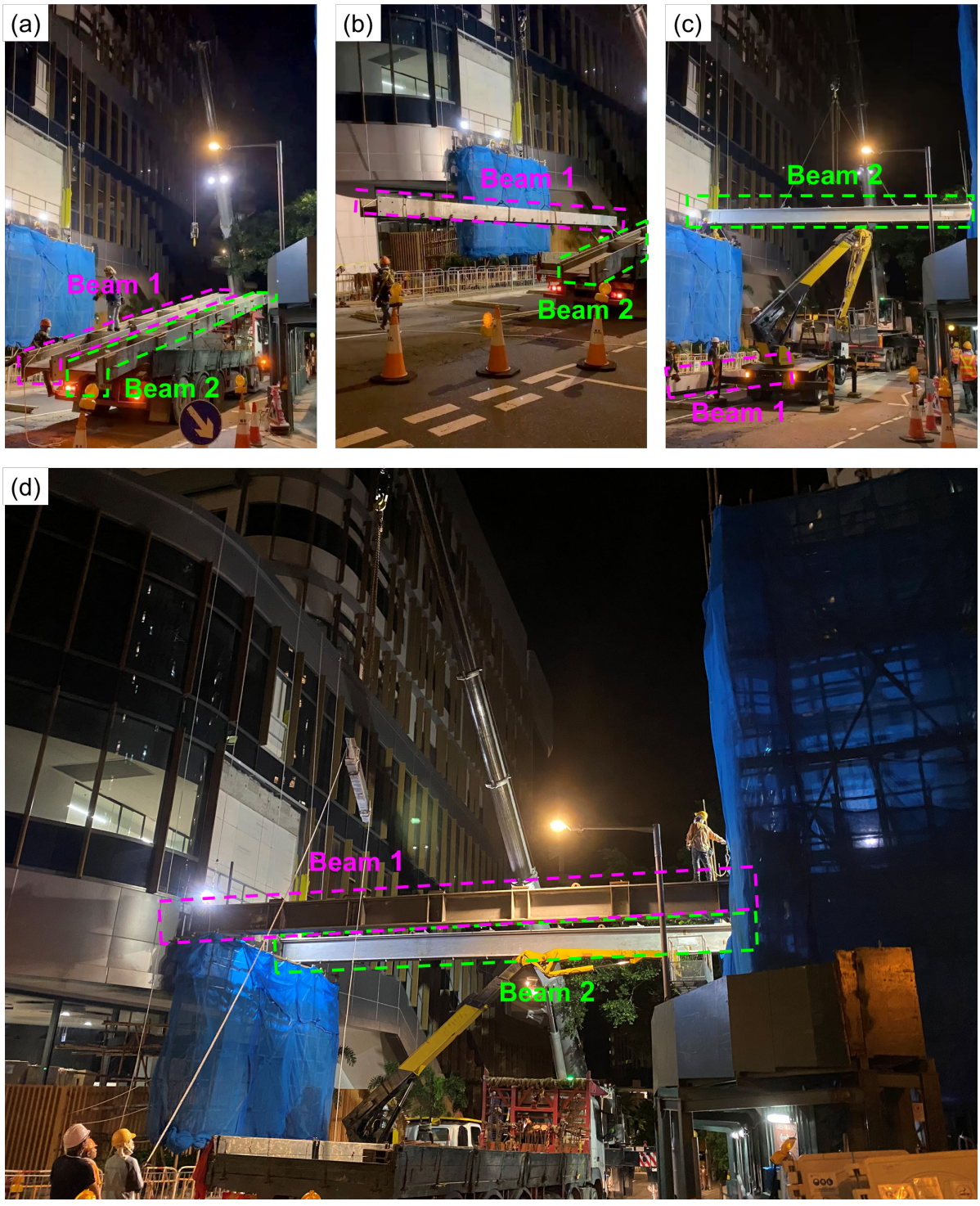


Fig. 2. Test project. (a) Site condition before hoisting; (b) hoisting process of Beam 1 from a
455 flatbed truck; (c) installation status of the Beam 1 with a crane truck; (d) completed status of two major steel tie-beams hoisting.

The case project met typical challenges regarding efficiency and safety for nighttime construction work in a congested area. First, construction activities were only allowed at midnight, and the construction work needed to be completed before dawn. This was because the site work required closing the road, which needed to resume operations in the daytime. Second, the efficiency of hoisting, as the most challenging part of the site work, considerably impacted the construction time. Third, there were buildings within close range; one building's facades were glass curtain walls. Thus, a collision between a load and the facades during the hoisting process would cause economic losses.

Thus, advanced monitoring was needed for mobile-crane-related site activities to ensure the safety and efficiency of this project. However, the traditional vision-based or sensor-based methods were challenging due to the unusual project characteristics. Specifically, the external light for the project was dim at night, so effective monitoring would be greatly impacted by common cameras that were sensitive to lighting conditions. For IoT devices, signals would be blocked or interfered with by tall buildings nearby, and the accuracy of monitoring could not be effectively guaranteed.

4.2 Data collection

A novel 4DPC sensing device, as shown in Fig. 3, was developed in-house to collect the high-definition 4D motion data of construction activities. The devices consisted of four essential modules: (I) a Livox Mid-70 sensor, (II) a controller (Raspberry Pi 4B), (III) an LED monitor, and (IV) a USB drive, as shown in Fig. 3. As an offline data transmission plan, the 4DPC data was later copied to a computer workstation via USB. Costs of parts I, II, III, and IV were USD 830, 110, 42, and 11, respectively. The total cost for a set of 4DPC sensing devices is about 993 USD.

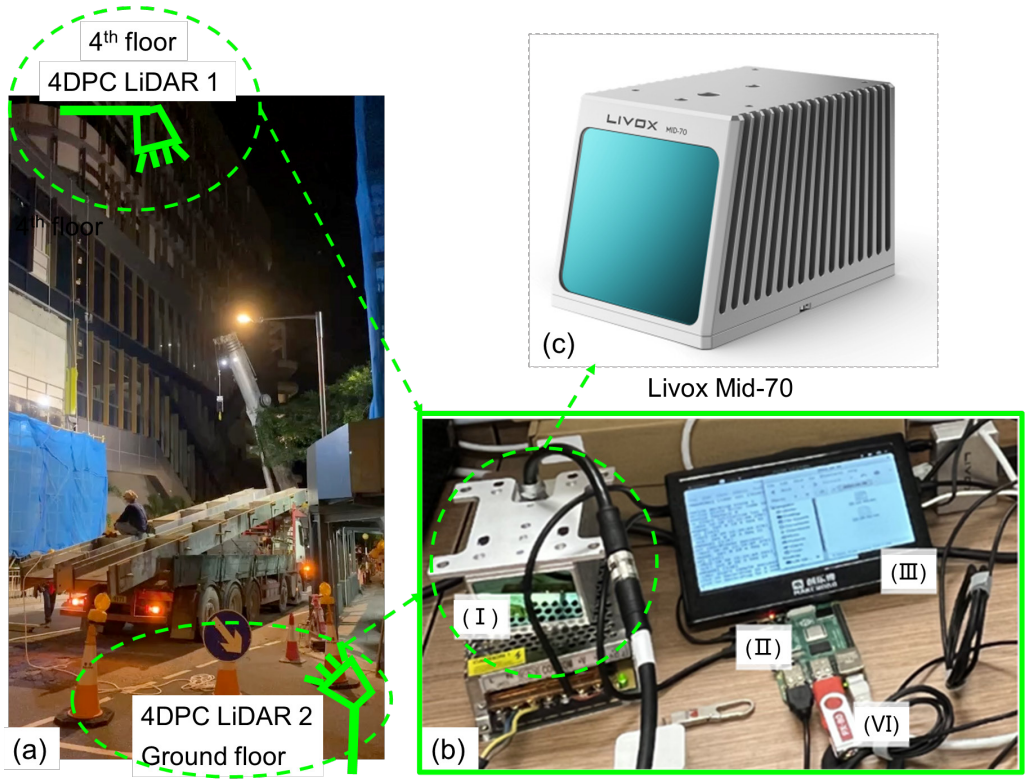


Fig. 3. 4DPC device in this study. (a) Installation illustration of 4DPC devices on site; (b) Components of our 4DPC device: (I) Livox Mid sensor, (II) controller (Raspberry Pi 4), (III) LED monitor, (IV) USB drive; (c) Livox Mid sensor (Liang et al. 2023)

To design an effective monitoring solution for mobile crane-related activities, three main factors, including the sensing range of devices, the construction area of interest, and the potential location of sensing devices, need to be considered. The basic principle is to place an appropriate number of sensors within the range where sensor equipment can be placed to achieve complete coverage of the construction area in 3D space. Therefore, the location design of sensors will vary in cases. The 4DPC sensor was Livox Mid-70, which had a 70.4° circular field of view, a minimum detection range of 5 cm, a maximum detection range of 70 m (r), and a range precision of 2 cm, as shown in Fig. 3(c). The construction area of interest was 12m × 20m × 12m (W×L×H) with a tilt angle (θ) of 4.8 degrees in the z-y plane. As shown in Fig. 4, to cover the whole 3D space of the construction area of interest, two devices were installed on the ground floor and the fourth floor, which are out of the construction area of interest.

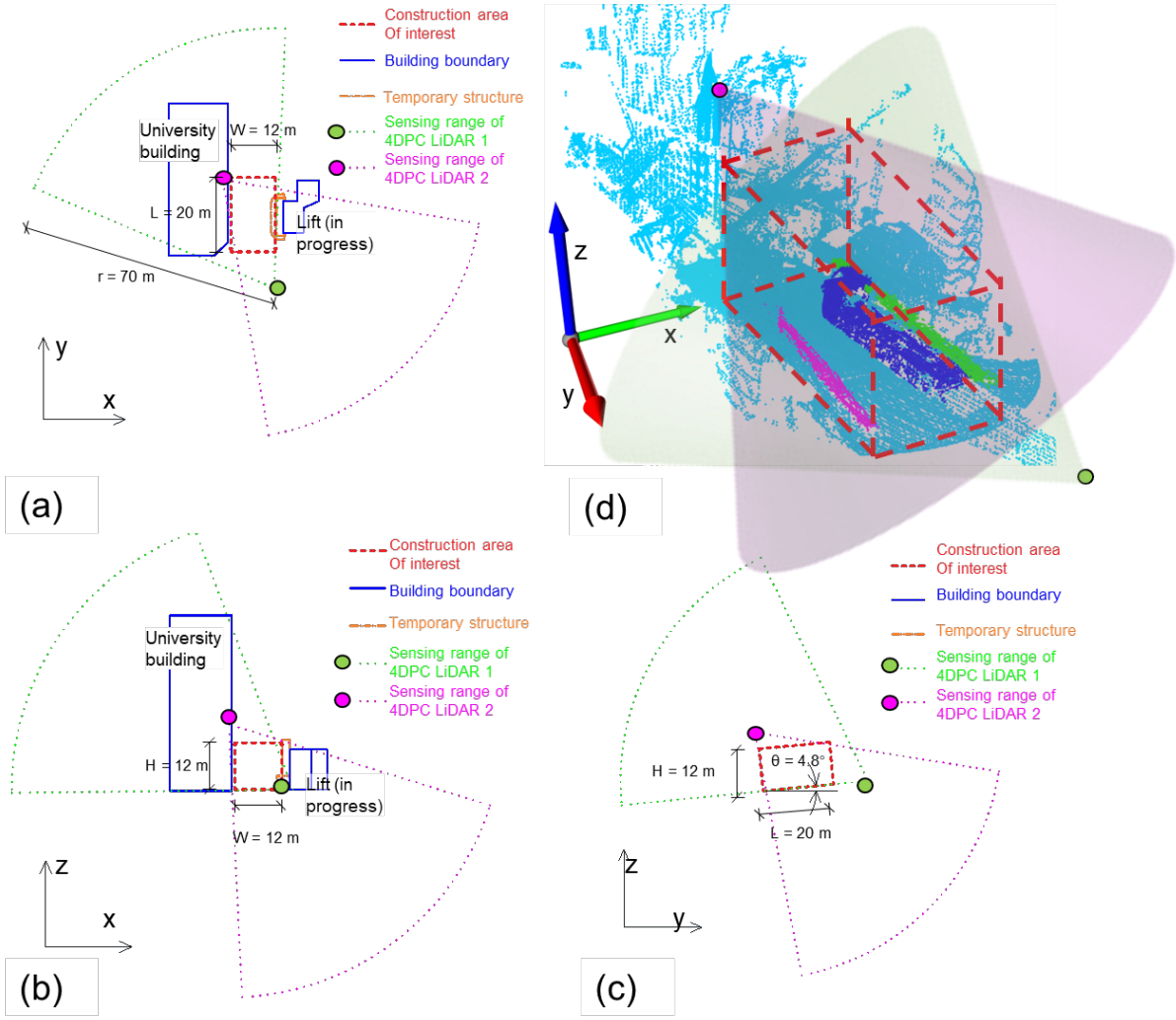


Fig. 4 Spatial relationship illustration of construction area of interest and sensing range of sensors: (a) decomposition in xy plane; (b) decomposition in xz plane; (c) decomposition in yz plane; (d) global view

505

The sensor scanned cm-accurate 4DPC at 20 frames per second, but the limited scanning duration per frame may have resulted in decreased target resolution. In total, the devices collected 6 hours of 4DPC data (size: 80 MB/min) from the construction site. One hour of 4DPC data was then selected for analysis in this paper. The selected 4DPC data contained the
510 whole hoisting activity related to the mobile crane. For a satisfying and sufficient resolution of registration, the 4DPC data was aggregated every 2 seconds, i.e., a frame rate at 0.5 fps, resulting in a total of 1800 frames.

4.3 Experimental settings

The computational experiment of the proposed method was operated in single threading mode on a desktop computer with Intel i7-10700 2.9 GHz CPU, 32 GB memory, Python 3.8, and Windows 11 64-bit. The geometric point-level error of 4DPC points was reported to be less than 2cm at 20 m. The area of interest was within 60m, and, thus, the precisions in the experiments were kept at cm level. For both BIM objects and input 4DPC, the voxel grid sizes were set to 80 mm. To remove the large planes of ground and walls in the environment, the RANSAC algorithm is selected to find the main plane of the ground and main walls. The parameters *distance-threshold*, *ransac-n*, and *number-of-iterations* were set to 150mm, 3, and 1000, respectively. DBSCAN was selected for clustering, and the maximum neighborhood distance (*eps*) was 500 mm, and the minimum number of required points for a cluster (*min_points*) was 20. According to our tests with different algorithms, a DFO algorithm BOBYQA (Powell 2009) implemented in NLopt library was selected. It should be noted that the xyz coordinates in the analysis were aligned to the site's ground plane for motion tracking and key state detection for site managers' decision-making; there existed a 4.8° angular difference from the sea-level plane, which was assumingly ignorable for site safety and productivity analyses.

First of all, one reference case with the DFO registration mode was designed; the three recommended key parameters are listed in Table 2. A total of 12 more test cases were designed for sensitivity analysis, as shown in Table 2. The first aspect examined in the sensitivity analysis was the registration algorithm. Five more cases with different DFO algorithms are designed for the DFO registration mode. Meanwhile, four coarse-to-fine registration modes combined with two coarse registration methods, i.e., FGR and RANSAC, and two fine registration methods, i.e., ICP and CPD, are considered. In addition, the time interval between the key frames varied: 8s, 16s, and 0s (every frame is a key frame). The last aspect examined in the sensitivity analysis was the spatial-only registration pattern (S8). The spatial-only pattern indicated that the Sect. 3.3 temporal registration process was disabled. In specific, the temporal information from the previous frame would not be used. The clustering and spatial registration were conducted in the current 4DPC data frame.

Table 2 Details of eight test cases and associated values of the key parameters

Category	Case code	Algorithm	Registration pattern	Time interval
Reference case	Ref. BOBYQA	BOBYQA	Spatial-Temporal	8
	NEWUOA_BD	NEWUOA_BD	Spatial-Temporal	8

Registration algorithms	NELDERMEAD	NELDERMEAD	Spatial-Temporal	8
	COBYLA	COBYLA	Spatial-Temporal	8
	DIRECT	DIRECT	Spatial-Temporal	8
	PRAXIS	PRAXIS	Spatial-Temporal	8
	FGR-ICP	FGR+ICP	Spatial-Temporal	8
	RANSAC-ICP	RANSAC+ICP	Spatial-Temporal	8
	FGR-CPD	FGR+CPD	Spatial-Temporal	8
	RANSAC-CPD	RANSAC+CPD	Spatial-Temporal	8
Time interval	ST0	DFO	Spatial-Temporal	0
	ST16	DFO	Spatial-Temporal	16
Regis. pattern	S8	DFO	Spatial-only	8

4.4 Experiment results

For the reference case, the total processing time for the one-hour 4DPC data was 3,349 seconds (0.93 hours) on the desktop computer. The average processing time for each data frame (every 2s) was 1.86s, less than 2s. Therefore, a near-real-time processing was possible.

5.4.1 Semantic registration results

Fig. 5 shows typical registration results. Fig. 5(a) shows one data frame of unloading Beam 2 from a flatbed truck, and Fig. 5(b) shows another data frame of installing Beam 2 with a crane truck. The darker points in Fig. 5 represent registered semantic BIM objects, of which the surfaces were sampled as dense 3D points. Overall, the registration results were all correct.

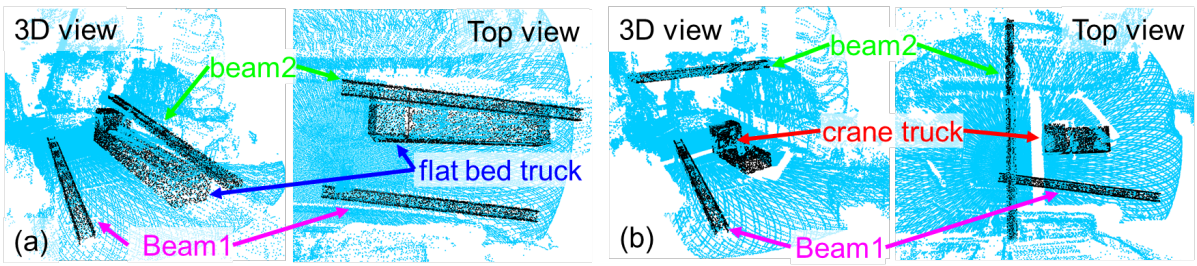


Fig. 5. Examples of registration results, where darker points represent registered semantic BIM objects. (a) Unloading Beam 2 from a flatbed truck at t =126 (01:03:10 a.m.); (b) installing Beam 2 with a crane truck at t = 786 (01:25:10 a.m.)

Table 3 illustrates the registration accuracy at the object level. Beam 1 and Beam 2 were accurately registered (Macro F_1 = 100%) on average. The average F_1 value of the flatbed truck and crane truck were 99.96% and 99.83%, respectively. The average F_1 was 99.95% for all the objects in Table 3. From our observations of the experiments, the failed registration of flatbed and crane trucks only occurred when the trucks entered or left the area of interest. Furthermore, the failed registration of the flatbed truck was attributed to data incompleteness,

where over half of the truck was outside of the area of interest when exiting. The crane truck was not timely registered because there was no key frame to detect it when it entered the area. In summary, the high accuracy demonstrated consistent and successful registration at the object level.

Table 3. List of average registration accuracy from the reference case

Object	Precision (%)	Recall (%)	Macro F_1 (%)
Beam 1	100	100	100
Beam 2	100	100	100
Flatbed truck	99.92	100	99.96
Crane truck	99.86	99.79	99.83
Average	99.95	99.95	99.95

Fig. 6 presents example hoisting motions showcasing three typical activities from the registration results. Fig. 6(a) shows the motions of the hoisting process for Beam 2 from the flatbed truck to the ground between t = 123 and t = 186. Fig. 6(b) shows hoisting for installation of Beam 2 from t = 450 to t = 786. Subsequently, Beam 1 was hoisted from the ground to the installation position between t = 1296 and t = 1591, as shown in Fig. 6(c). The motions in the registration results provided precise and near-real-time object-level information for site managers. It is important to note that the crane truck in Fig. 6(c) at t = 1591 was correctly registered, despite the sparse or discontinuous point cloud due to occlusion. The successful registration in sparse or discontinuous 4DPC was achieved due to the temporal patterns from the previous key frame.

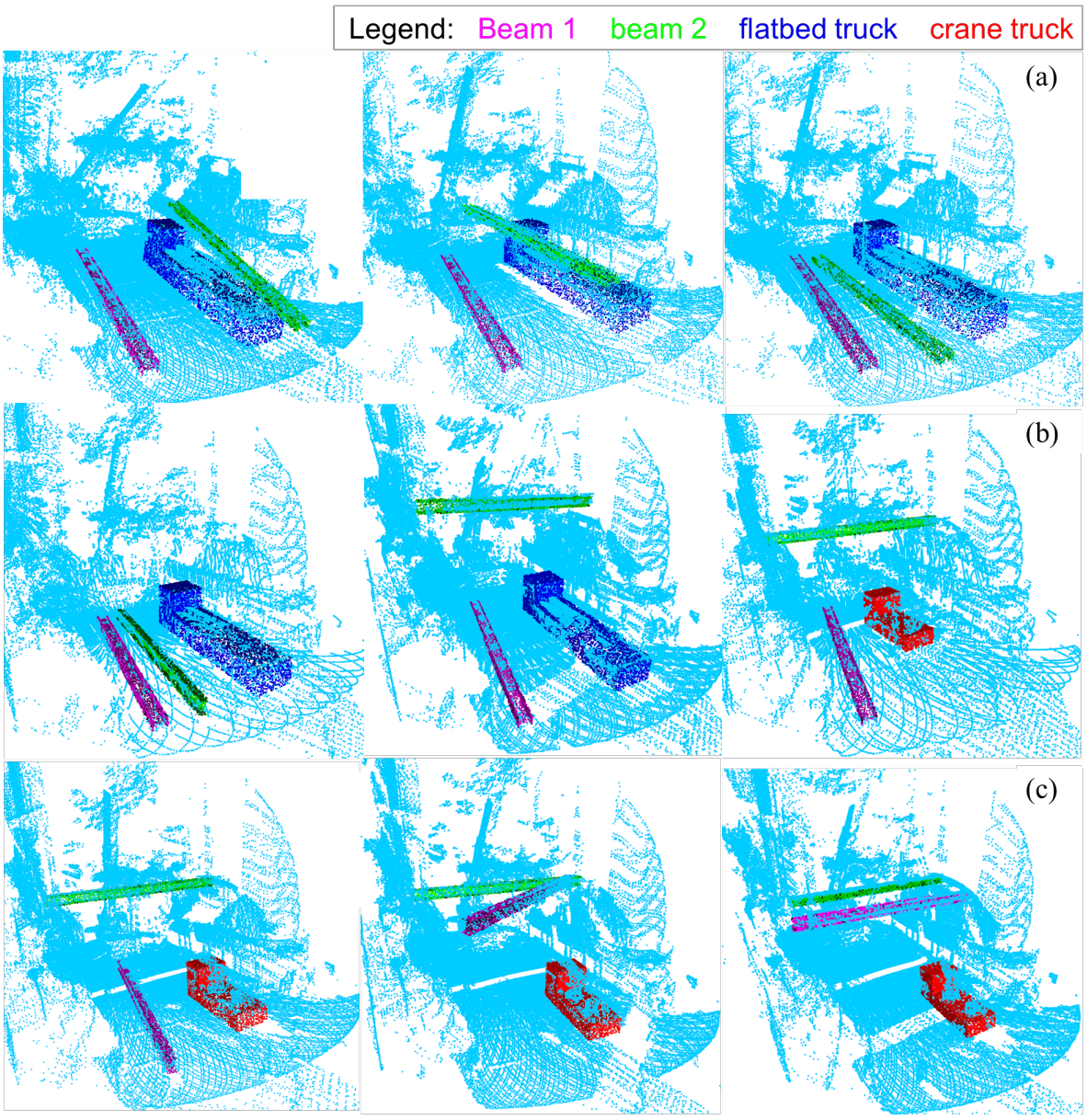


Fig. 6. Example hoisting motions represented form the registration results. (a) Hoisting Beam 2 from the flatbed truck from t = 126 to 186 (01:03:10 – 01:05:10 a.m.); (b) hoisting Beam 2 for installation from t = 450 to 786 (01:13:58 – 01:25:10 a.m.); (c) hoisting Beam 1 for installation from t = 1296 to 1591 (01:42:10 – 01:52:00 a.m.)

4.4.1 Load activity monitoring

High-definition trajectories of the tie-beams were reconstructed and visualized with the Cesium platform based on the semantic registration result of the 4DPC data. Figs. 6(a–c) visualize three typical states of hoisting Beam 1 on the Cesium platform, that is, start, process, and end. Each state was associated with a precise timestamp (2s resolution) and 3D pose. The intuitive

visualization in the main window offered site managers the graphical trajectory of Beam 1, while the timestamp at the bottom indicated the exact time of the state.

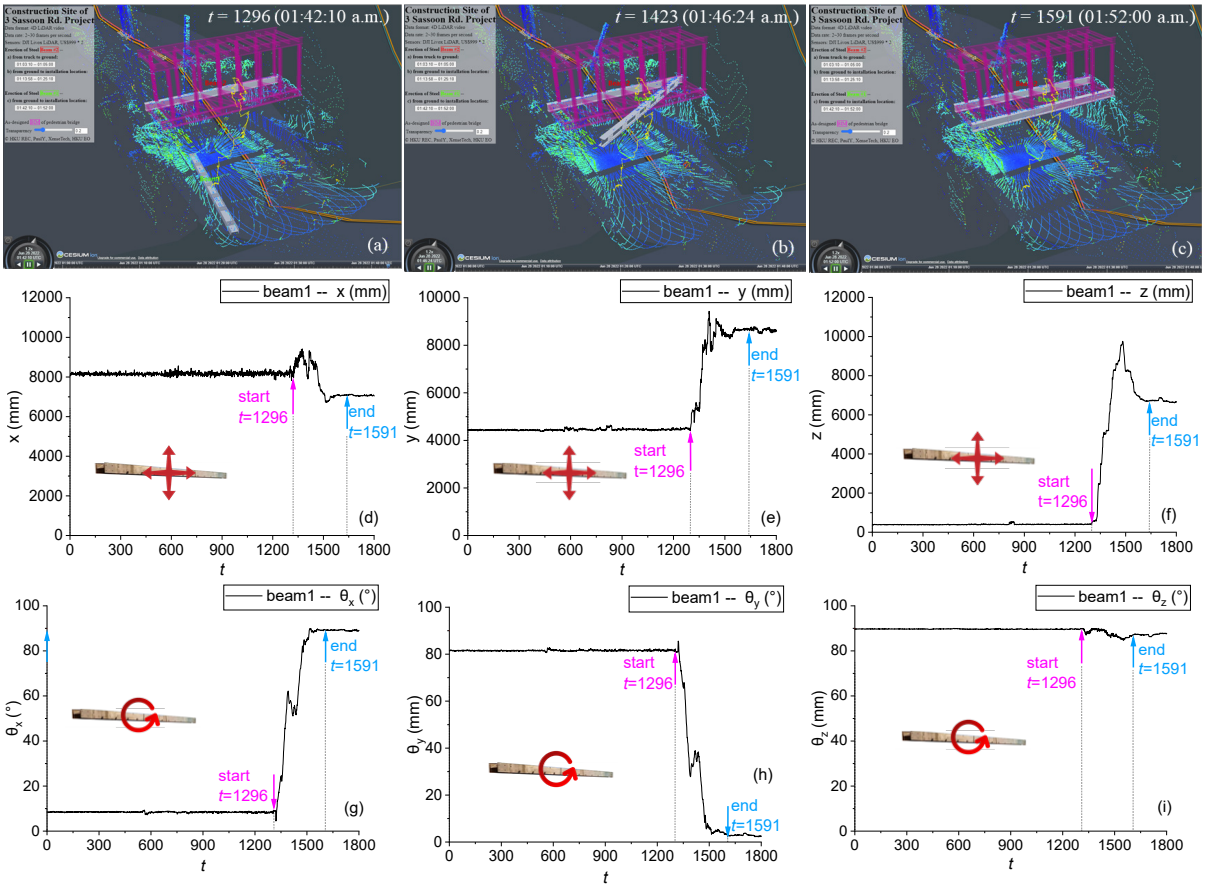


Fig. 7. Trajectory of Beam 1. (a)(b)(c) 3D illustrations; decomposed motion in (d) x-axis; (e) y-axis; (f) z-axis; decomposed angle of main axis with (g) x-axis; (h) y-axis; (i) z-axis;

To facilitate site managers' measurement of the key states, Figs. 6(d–f) show the decomposed motions of Beam 1 along the x-axis, y-axis, and z-axis, respectively. In addition, Figs. 6(g–i) depict the 3D angles of Beam 1. Figs. 6(d–f) show that during the hoisting process, the largest displacement of the center of Beam 1 was on the z-axis, with an approximately 9760 mm hoisting and subsequent lowering to the installation position at approximately 6670 mm.

In Figs. 6(g) and (h), Beam 1 was detected in a rotation of about 82°, from nearly parallel to the x-axis at t = 1296 to a near-perpendicular direction. Moreover, Fig. 7(i) indicates that the angle between Beam 1 and the z-axis remained consistent with minor fluctuations. The detected and measured rotation implied that Beam 1 maintained the angular motions in the horizontal plane during the hoisting process.

As shown in Fig. 8, Beam 2 experienced a two-stage hoisting process, which was more complex than that of Beam 1. Figs. 7(a–c) visualize the 3D trajectory of Beam 2 in two stages, which involved unloading it from the flatbed truck to the ground and hoisting it from the ground to the installation position. Figs. 7(d–i) illustrate the decomposed and angular motions of Beam 2. During the first stage from t = 126 (01:03:10 a.m.), noticeable motion changes were measured on the y-axis and z-axis, as shown in Figs. 7(e) and (f). Specifically, Fig. 8(f) reveals that the height of Beam 2 first increased from about 2550 mm to about 3890 mm and then decreased to ground level, reflecting the unloading and hoisting. Meanwhile, there was a slight rotation (about 10°) in all three axes, as shown in Figs. 7(g–i). The rotation in r_z reflected the fact that Beam 2 was initially inclined on the flatbed truck, as shown in Fig. 6(a). In the second stage, when $450 \le t \le 786$ (01:13:58 to 01:25:10 a.m.), two considerable changes were motion along the z-axis and a slight rotation in the horizontal plane. Fig. 8(f) demonstrates that Beam 2 was hoisted to a height of about 9615 mm and then lowered to 6980 mm. Furthermore, Figs. 7(g) and (h) show that Beam 2 rotated approximately 78°, from roughly parallel to the x-axis to parallel to the y-axis, during the hoisting.

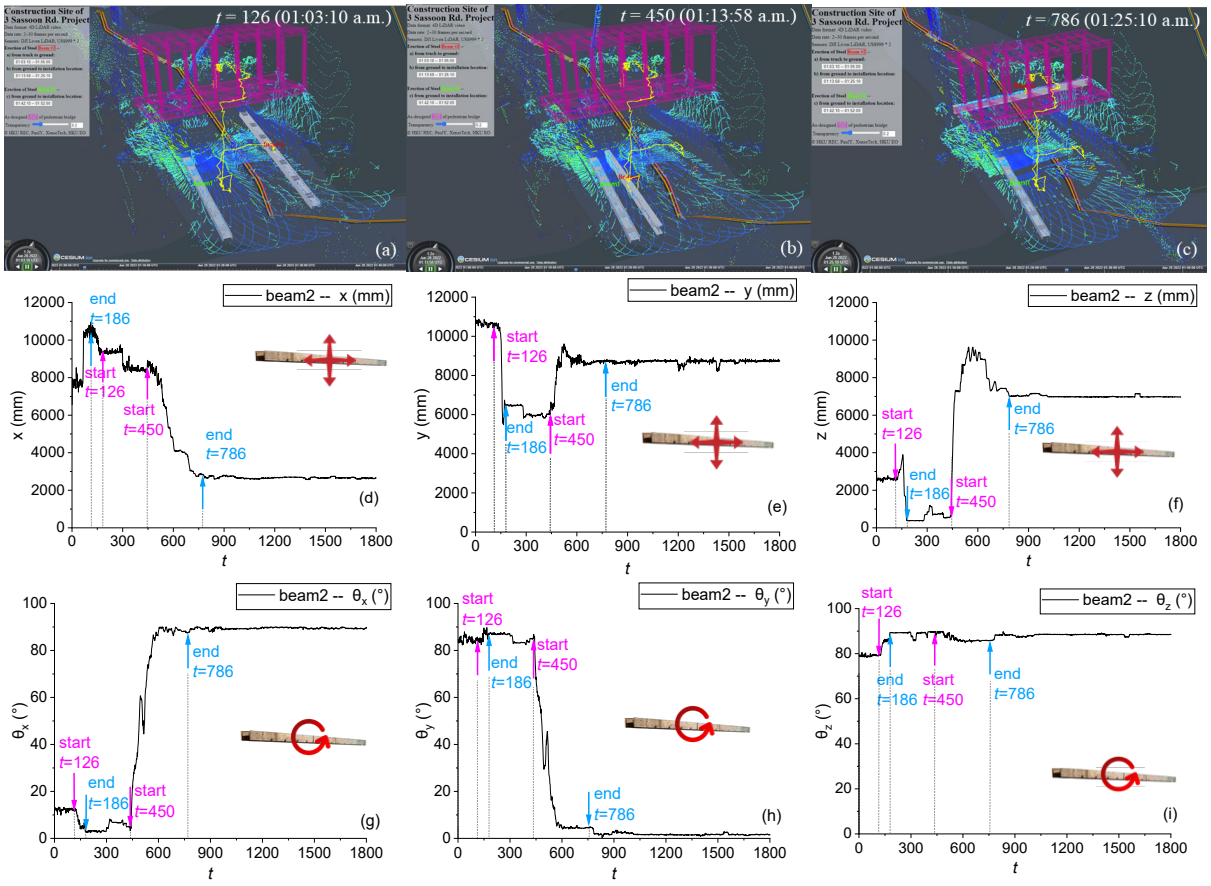


Fig. 8. Trajectory of Beam 2: (a)(b)(c) 3D illustration; decomposed motion in (d) x-axis; (e) y-axis; (f) z-axis; decomposed angle of main axis with (g) x-axis; (h) y-axis; (i) z-axis;

625

4.4.2 Monitoring truck activity

Fig. 9 illustrates the decomposed and angular motions of the flatbed truck on the horizontal xy plane. Fig. 9(a) shows that the flatbed truck initially moved along the x-axis from t = 63 (01:01:04 a.m.) and ultimately departed the area of interest at t = 633 (01:20:04 a.m.). Prior to
630 the departure, the flatbed truck moved about 7880 mm in the x-axis direction and turned 30° on the xy plane for a temporary parking place nearby.

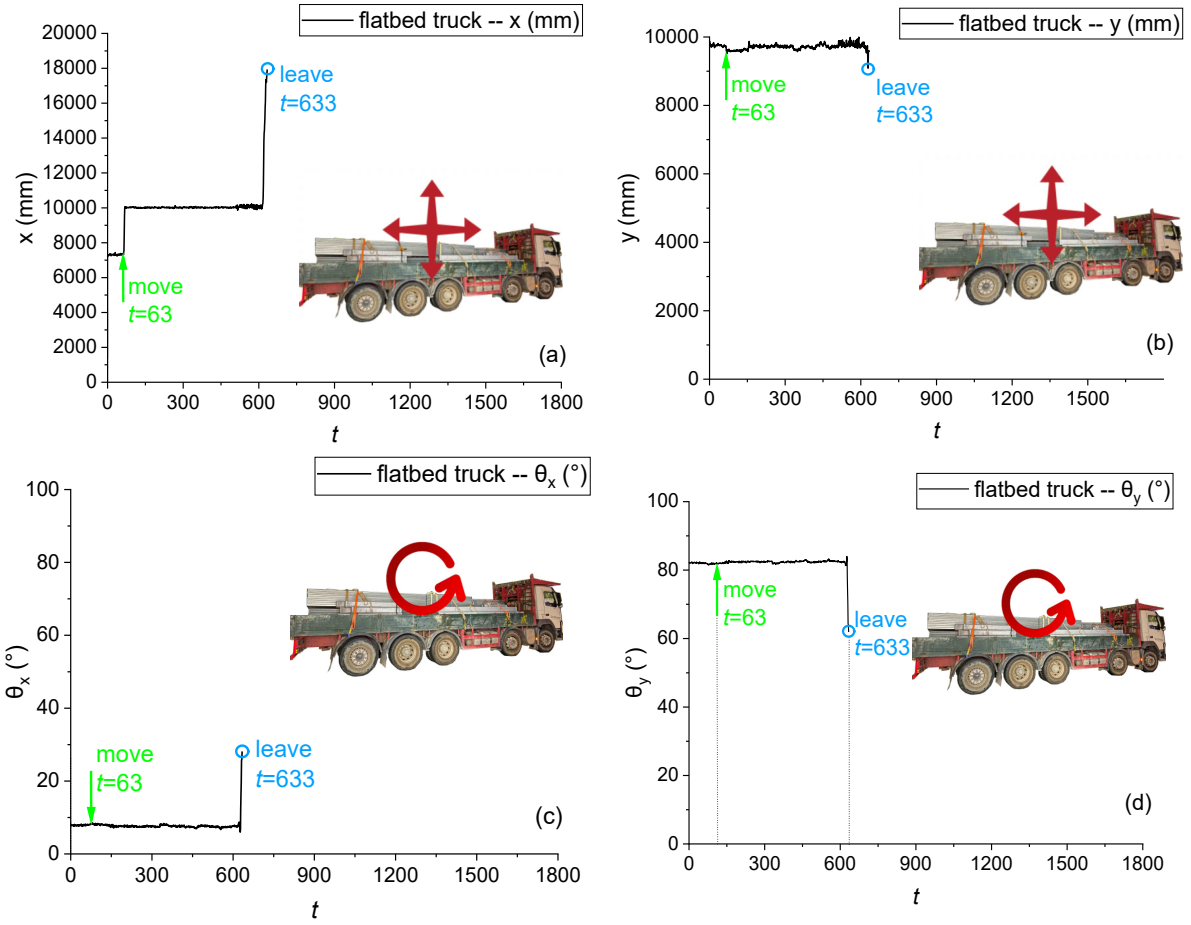


Fig. 9. Trajectory of flatbed truck: decomposed motion in (a) x-axis; (b) y-axis; decomposed angle of main axis with (c) x-axis; (d) y-axis

As shown in Fig. 10, the crane truck first entered at t = 715 (01:22:48 a.m.). Then, the crane truck changed its position from t = 1233 (01:40:04 a.m.) to install another tie-beam, as detected
635 on both the x-axis and the y-axis in Figs. 9(a) and (b), respectively. Figs. 9(c) and (d) indicate that the crane truck maintained a stable 3D angle throughout the process because only minor angular changes were measured.

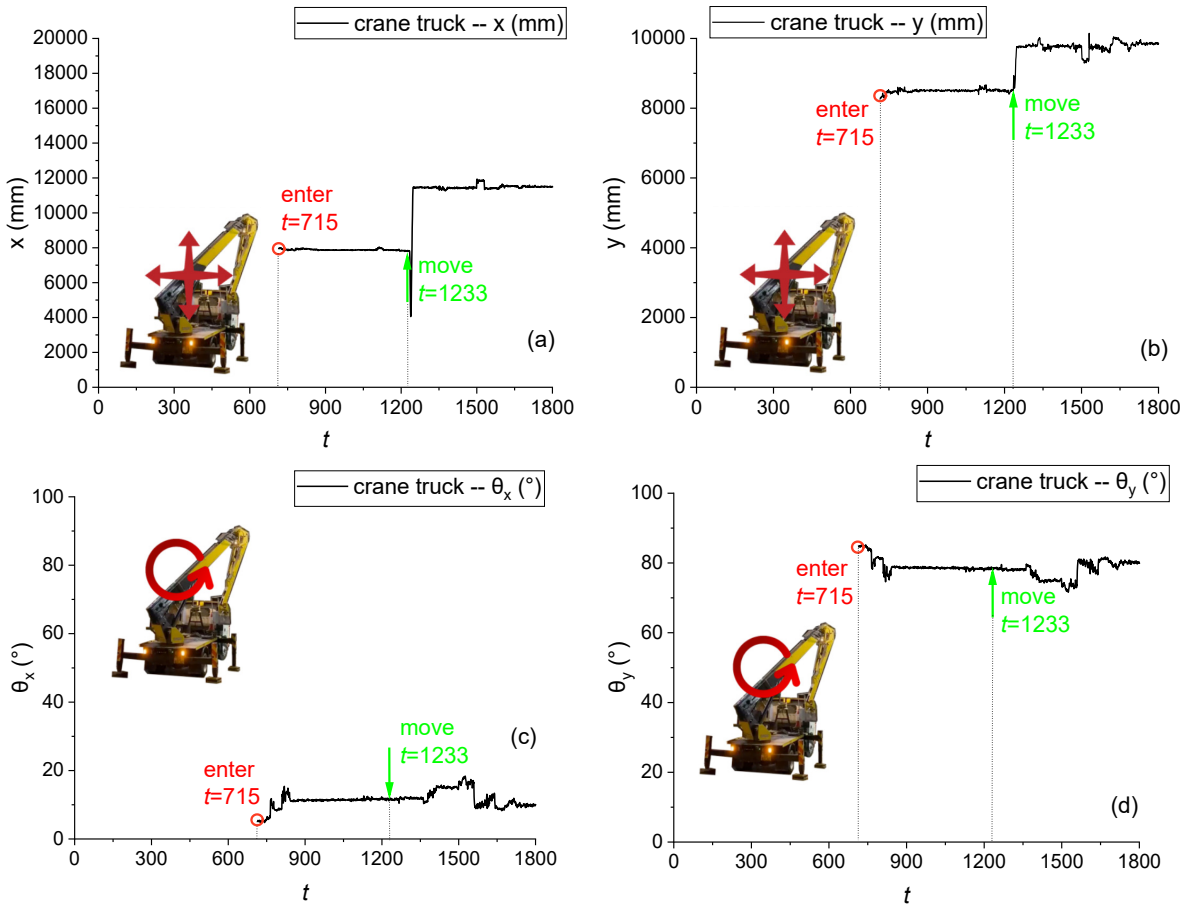


Fig. 10. Trajectory of crane truck: decomposed motion in (a) x-axis; (b) y-axis; decomposed angle of main axis with (c) x-axis; (d) y-axis

4.4.3 Sensitivity analysis

Table 4 shows two registration categories, i.e., DFO-based and coarse-to-fine registration, were first considered. Specifically, five other DFO registration algorithms and four coarse-to-fine registration methods were examined. The results showed that the performance of the BOBYQA (reference case) algorithm was better than other DFO algorithms. It can also conclude that the selection of DFO algorithms had a minor impact on the accuracy but a significant influence on the computational time.

For coarse-to-fine registration cases, the FGR-ICP had the best F1 accuracy and computational time performance than the other three cases. Compared with reference case Ref. BOBYQA, FGR-ICP significantly shortened the processing time (from 1.04s to 0.58s per frame) with a decrease of F_1 value (from 99.97% to 99.68%).

Table 4 Results for algorithm selection (default spatial-temporal parameters "ST8"; best in each category in bold)

Category	Algorithm*	F_1 accuracy (%)					Time per frame* (s)
		Beam 1	Beam 2	F-truck	C-truck	Avg*	
DFO	**Ref. BOBYQA**	100.00	100.00	99.96	99.93	**99.97**	**1.04**
	NEWUOA_BND	100.00	100.00	99.96	99.93	99.97	1.12
	NELDERMEAD	100.00	100.00	99.96	99.93	99.97	3.30
	COBYLA	100.00	100.00	99.96	98.94	99.64	1.34
	DIRECT	100.00	100.00	99.63	99.12	99.69	4.21
	PRAXIS	100.00	100.00	99.96	98.83	99.83	3.47
Coarse-to-fine	**FGR-ICP**	100.00	100.00	99.94	98.77	**99.68**	**0.58**
	RANSAC-ICP	100.00	66.46	99.94	99.35	91.44	11.65
	FGR-CPD	65.47	65.62	77.34	54.71	65.79	10.36
	RANSAC-CPD	58.73	55.15	76.38	54.34	61.15	36.51

*: Best of the column in bold

Regarding F_1 and computing time, Ref. BOBYQA and FGR-ICP were the best cases for DFO-based and coarse-to-fine registration, respectively. Furthermore, Fig. 11 shows a bi-objective comparison of the registration results between BOBYQA and FGR-ICP. The horizontal axis is the computing time per frame, and the vertical axis is *error* = 100% – F_1. In engineering practice, more minor errors and shorter computing time are expected; the scatter points are closer to the coordinate origin in Fig. R1. As shown in Fig. 11, although the computing time of FGR-ICP is about half that of BOBYQA, the error of FGR-ICP is ten times that of BOBYQA. Therefore, we need to choose a method based on actual project requirements. If the computing time requirement is high, FGR-ICP will be preferred. If accuracy (less error) is the first consideration, BOBYQA is better than FGR-ICP.

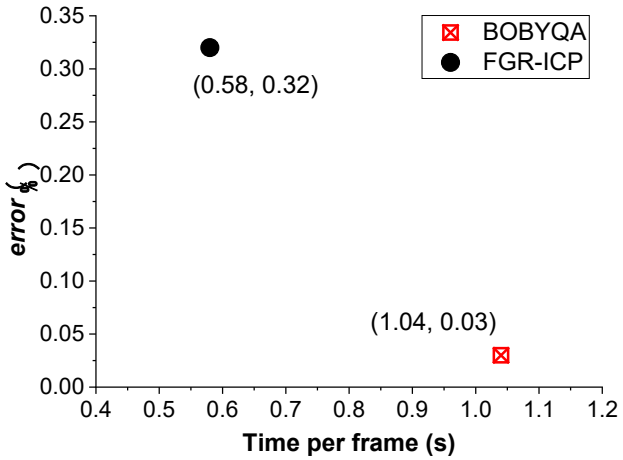


Fig. 11 Bi-objective comparison of registration results between BOBYQA and FGR-ICP

In addition, as shown in Fig. 12, the values of registration RMSE for Ref. BOBYQA and FGR-ICP were consistently close from the beginning to the end of the test project.

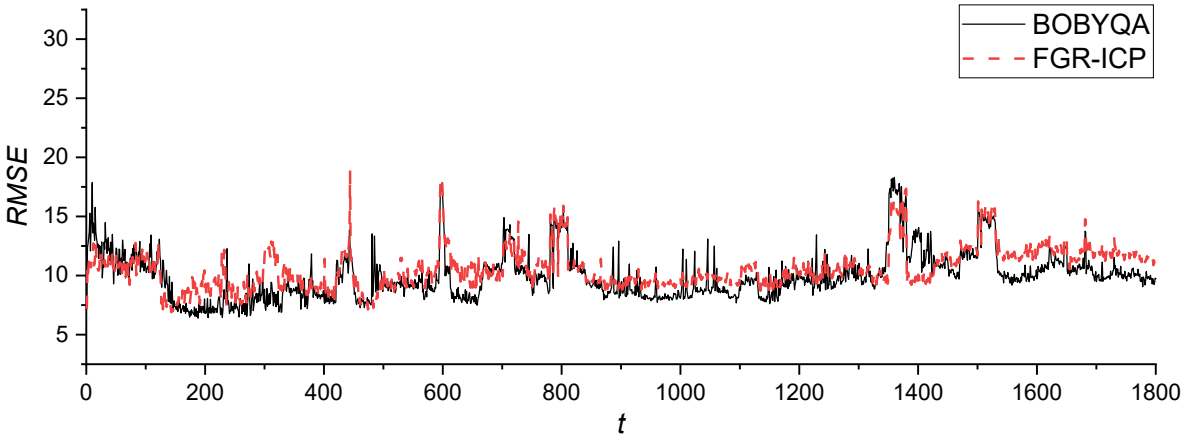


Fig. 12. Comparison of RMSE between BOBYQA and FGR-ICP (lower is better)

As shown in Table 5, the impact of the time interval between key frames on the F_1 accuracy was slight in the three test cases, i.e., ST0, Ref. BOBYQA (ST8), and ST16. However, a shorter interval led to longer computational time. The increment of computational time correctly reflected the frequency of time-consuming spatial semantic registration (global search) in Sect. 3.4. However, the savings in computational time came at the cost of registration accuracy. Overall, the parameter settings in the reference case were recommended as a 'sweet spot' for the trade-off between computational time and registration accuracy.

Table 5 Results for sensitivity analysis of spatial-temporal settings (best in F_1 and time in bold)

Case code	F_1 accuracy (%)					Time per frame (s)
	Beam 1	Beam 2	F-truck	C-truck	Avg	
Ref. BOBYQA(ST8)	100.00	100.00	99.96	99.93	99.97	1.04
ST0	100.00	100.00	99.96	99.98	**99.99**	2.21
ST16	100.00	100.00	99.96	99.84	99.95	**0.74**

Lastly, Table 6 compares the spatial-feature-relied registration pattern, i.e., spatial-only registration (S8), with the spatial-temporal registration pattern (ST8). The results showed that case S8 yielded considerably lower F_1 values in comparison with the preferred spatial-temporal registration pattern. Moreover, S8 required a longer computational time than their corresponding spatial-temporal pattern, i.e., S8. According to an in-depth analysis of the experimental results, the main barrier to the spatial-feature-relied registration method was the spatial noise in the 4DPC data, which was caused by laser occlusion, worker interferences

(e.g., guiding tie-beams with rope), contact between the ropes and the workers' arms, and point-level geometric errors.

Table 6 Results for sensitivity analysis of spatial-temporal settings (default algorithm "BOBYQA"; default time interval "8s")

Case code	F_1 accuracy (%)					Time per frame (s)
	Beam 1	Beam 2	F-truck	C-truck	Avg	
Ref. BOBYQA(ST8)	100.00	100.00	99.96	99.93	99.97	1.04
S8	32.23	55.59	91.00	72.92	62.94	2.21

In summary, the registration accuracy of the proposed method is proven sensitive to the choice of registration algorithm and the use of a spatial-temporal registration pattern. The time interval significantly impacts computational time more than registration accuracy.

5 Discussion

5.1 Pros and cons of the proposed method

The advantages of the proposed method lie in three aspects.

- The proposed spatial-temporal semantic registration method is the first of this kind for mapping the as-built BIM semantics to the 4DPC of construction activities. The resulting semantic poses and motions of the BIM objects enable digital twin site simulations, early alerts, and early interventions (Opoku et al. 2021).
- The spatial-temporal registration pattern was proven considerably better than traditional 3D spatial registration. The experimental results (the last row) in Table 5 listed significant improvement from the spatial-only patterns in terms of both accuracy and time cost.
- The proposed method is robust in employing either global or local registration algorithms, including sophisticated DFO algorithms, in the workflow. The high accuracy in experiments confirmed the rationality and effectiveness of the proposed method. Researchers and practitioners have the freedom to change the registration algorithm to meet the project site scenario.

However, this paper also had limitations:

- The environment, including the ground and facades, was removed semi-automatically using an algorithm RANSAC. An automatic method can be developed to remove the environment from 4DPC by extracting the key feature points and planes.

- The proposed method is reliant on an optimization process of as-designed BIM objects. Therefore, construction objects excluded in the as-designed BIM cannot be detected. Examples are site workers, temporary scaffoldings, and ropes to correct load positions. Future research could explore advanced computer vision-based technologies to achieve a BIM-free semantic enrichment.
- The experiment was conducted in only one type of project. More types of projects and 4DPC data regarding mobile-crane activities are recommended to validate the proposed method in the future.

5.2 Potential applications

Several promising applications in construction safety and productivity management are recommended to apply the presented 4DPC-based spatial-temporal semantic registration method based on the identified pros and cons. Examples are:

- *Load monitoring and collision alert*. The accurate tempo-spatial positioning of loads measures the deviation and potential collisions during crane operations. Early warnings of deviations and potential collisions mitigate such risks and subsequent losses.
- *Worker safety*. The proposed method for mobile cranes can facilitate worker safety management on-site. First, the 4D trajectories of all major equipment and construction objects narrow down and predict hazardous situations for workers. One example is that a worker who enters a restricted zone under a hoisting load's 10-second trajectory can be alerted by a smart helmet.
- *Component logistics management*. The outcomes of the method, such as the tracked movements of beams and components before hoisting, can facilitate just-in-time logistics from delivery to installation.
- *Productivity monitoring and optimization*. The near-real-time analysis of registration results in Sects. 4.4.2 and 4.4.3 quantitatively measures and monitors the movement and handling of components and equipment. The accumulated analysis can facilitate the site manager's identification of bottlenecks of construction processes eventually.
- *Progress monitoring and quality control*. The monitored object-level registration results reflect 3D project progress in near real-time to enable multiple progress monitoring and quality control applications such as compliance checking.

Overall, the proposed 4DPC-based optimization-based semantic registration method promises to fill the technical gap of the semantic enrichment of the 4DPC to supplement construction safety and productivity management.

6 Conclusions

In the construction industry, there is a pressing need for high-definition and construction context-aware monitoring for mobile cranes due to their high mobility and flexibility. However, conventional technologies such as CCTV cameras and IoT devices are limited in monitoring mobile crane-related activities, while new and promising data sources such as 4DPC are emerging. This paper proposes a 4DPC-based spatial-temporal semantic registration method for monitoring mobile crane-related activities. Intuitive visualizations and key state detection are designed for safety and productivity monitoring based on the results of the proposed method. From experimental results, one-hour 4DPC data showed a highly accurate semantic registration of two tie-beams and two trucks at $F_1 = 99.95\%$ on average, while the average processing time for a frame (2s) can be shortened to 1.04s. The load and truck activities were qualitatively measured and visualized for site managers' informed decision-making. In summary, this paper addresses a critical issue in high-definition and construction context-aware monitoring and opens up new avenues for research and applications in construction safety and productivity management.

This paper offers a two-fold contribution. First, the spatial-temporal semantic registration method is proposed and proved effective for tracking objects in the crane workspace. It outperformed those traditional spatial-information-relied registration methods in sparse 4DPC with noise and occlusion. Secondly, the comprehensive analysis based on the method's results offers high-definition and construction context-aware tools for construction practitioners to monitor mobile crane-related activities. Nevertheless, this paper is limited to semi-automatic environment data removal, reliance on as-designed BIM objects, and limited experimentation.

Future research directions for 4DPC construction activities are recommended in automatic 4DPC preprocessing for environment removal, BIM-free semantic enrichment processes such as deep learning algorithms, and diversified project types for validation and evaluation. Researchers also could explore cross-referencing 4DPC and other site sensing technologies.

Acknowledgment

The work presented in this paper was supported by Hong Kong Research Grants Council (RGC) (No. 27200520) and in part by Innovation and Technology Commission (ITC) (No. ITP/004/23LP) and The University of Hong Kong (No. 2307102469). We appreciate Paul Y. Engineering and XenseTech Ltd. for providing data collection support on the case project site.

References

Alhwarin, F., Ferrein, A. & Scholl, I. (2014). IR stereo kinect: improving depth images by combining structured light with IR stereo. *PRICAI 2014: Trends in Artificial Intelligence: 13th Pacific Rim International Conference on Artificial Intelligence* (pp. 409-421). Springer. doi:10.1007/978-3-319-13560-1_33

Alphonse, P. J. & Sriharsha, K. V. (2021). Depth perception in single rgb camera system using lens aperture and object size: a geometrical approach for depth estimation. *SN Applied Sciences, 3*(6), 595. doi:10.1007/s42452-021-04212-4

Al-Rawabdeh, A., He, F. & Habib, A. (2020). Automated feature-based down-sampling approaches for fine registration of irregular point clouds. *Remote Sensing, 12*(7), 1224. doi:10.3390/rs12071224

Beavers, J. E., Moore, J. R., Rinehart, R. & Schriver, W. R. (2006). Crane-related fatalities in the construction industry. *Journal of Construction Engineering and Management, 132*(9), 901-910. doi:10.1061/(ASCE)0733-9364(2006)132:9(901)

Bhople, A. R., Shrivastava, A. M. & Prakash, S. (2021). Point cloud based deep convolutional neural network for 3D face recognition. *Multimedia Tools and Applications, 80*, 30237-30259. doi:10.1007/s11042-020-09008-z

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

Bosche, F. (2010). 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. doi:10.1016/j.aei.2009.08.006

Brent, R. P. (2013). *Algorithms for minimization without derivatives.* Courier Corporation. doi:10.2307/2005713

Bueno, M., González-Jorge, H., Martínez-Sánchez, J. & Lorenzo, H. (2017). Automatic point cloud coarse registration using geometric keypoint descriptors for indoor scenes. *Automation in Construction, 81*, 134-148. doi:10.1016/j.autcon.2017.06.016

Bush, J., Ninić, J., Thermou, G., Bennetts, J., Denton, S., Tachtsi, L. & Hill, P. (2022). Point cloud registration for bridge defect tracking in as-built models. *Bridge Safety, Maintenance, Management, Life-Cycle, Resilience and Sustainability* (pp. 1053-1060). CRC Press. doi:10.1201/9781003322641-126

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, T. & Teizer, J. (2011). Crane operator visibility of ground operations. *Proceedings of the 28th International Symposium on Automation and Robotics in Construction* (pp. 699-705). IAARC. doi:10.22260/isarc2011/0131

Chian, E. Y., Goh, Y. M., Tian, J. & Guo, B. H. (2022). Dynamic identification of crane load fall zone: A computer vision approach. *Safety Science, 156*, 105904. doi:10.1016/j.ssci.2022.105904

Chiu, H.-k., Li, J., Ambruş, R. & Bohg, J. (2021). Probabilistic 3D multi-modal, multi-object tracking for autonomous driving. *International Conference on Robotics and Automation (ICRA)* (pp. 14227-14233). IEEE. doi:10.1109/ICRA48506.2021.9561754

Choy, C., Gwak, J. & Savarese, S. (2019). 4D spatio-temporal convnets: Minkowski convolutional neural networks. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 3075-3084). IEEE. doi:10.1109/cvpr.2019.00319

Chung, W. W., Tariq, S., Mohandes, S. R. & Zayed, T. (2023). IoT-based application for construction site safety monitoring. *International Journal of Construction Management, 23*(4), 58-74. doi:10.1080/15623599.2020.1847405

Elseberg, J., Borrmann, D. & Nüchter, A. (2013). One billion points in the cloud-an octree for efficient processing of 3D laser scans. *Remote Sensing, 76*, 76-88. doi:10.1016/j.isprsjprs.2012.10.004

Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining* (pp. 226-231). AAAI. Retrieved from https://cdn.aaai.org/KDD/1996/KDD96-037.pdf

Fan, H. & Yang, Y. (2019). PointRNN: Point Recurrent Neural Network for Moving Point Cloud Processing. *arXiv:1910.08287v2.*

Fan, H., Yang, Y. & Kankanhalli, M. (2021). Point 4d transformer networks for spatio-temporal modeling in point cloud videos. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2021* (pp. 14204-14213). IEEE. doi:10.1109/cvpr46437.2021.01398

Fan, H., Yang, Y. & Kankanhalli, M. (2022). Point spatio-temporal transformer networks for point cloud video modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 45*(2), 2181-92. doi:10.1109/TPAMI.2022.3161735

Fan, H., Yu, X., Ding, Y., Yang, Y. & Kankanhalli, M. (2021). Pstnet: Point spatio-temporal convolution on point cloud sequences. *International Conference on Learning Representations* . ICLR. doi:10.48550/arXiv.2205.13713

Fang, Y., Chen, J., Cho, Y. K., Kim, K., Zhang, S. & Perez, E. (2018). Vision-based load sway monitoring to improve crane safety in blind lifts. *Journal of Structural Integrity and Maintenance, 3*(4), 233-242. doi:10.1080/24705314.2018.1531348

Fang, Y., Cho, Y. K. & Chen, J. (2016). A framework for real-time pro-active safety assistance for mobile crane lifting operations. *Automation in Construction, 72*, 367-379. doi:10.1016/j.autcon.2016.08.025

Forteza, F. J., Carretero-Gómez, J. M. & Sesé, A. (2017). Effects of organizational complexity and resources on construction site risk. *Journal of Safety Research, 62*, 185-198. doi:10.1016/j.jsr.2017.06.015

Görçün, Ö. F. & Doğan, G. (2023). Mobile crane selection in project logistics operations using Best and Worst Method (BWM) and fuzzy Measurement of Alternatives and Ranking according to COmpromise Solution (MARCOS). *Automation in Construction, 147*, 104729. doi:10.1016/j.autcon.2022.104729

Gu, X., Wang, X. & Guo, Y. (2020). A review of research on point cloud registration methods. *IOP Conference Series: Materials Science and Engineering* (pp. Vol. 782, No. 2, p. 022070). IOP Publishing. doi:10.1088/1757-899x/782/2/022070

Hu, D., Gan, V. J. & Yin, C. (2023). Robot-assisted mobile scanning for automated 3D reconstruction and point cloud semantic segmentation of building interiors. *Automation in Construction, 152*, 104949. doi:10.1016/j.autcon.2023.104949

Hu, Z. & Brilakis, I. (2024). Matching design-intent planar, curved, and linear structural instances in point clouds . *Automation in Construction, 158*, 105219. doi:10.1016/j.autcon.2023.105219

Huang, X., Mei, G., Zhang, J. & Abbas, R. (2021). A comprehensive survey on point cloud registration. *arXiv preprint arXiv:2103.02690.* doi:10.48550/arXiv.2103.02690

Jarząbek-Rychard, M. & Maas, H.-G. (2023). Modeling of 3D geometry uncertainty in Scan-to-BIM automatic indoor reconstruction. *Automation in Construction, 154*, 105002. doi:10.1016/j.autcon.2023.105002

Jiang, W., Zhou, Y., Ding, L., Zhou, C. & Ning, X. (2020). UAV-based 3D reconstruction for hoist site mapping and layout planning in petrochemical construction. *Automation in Construction, 113*, 103137. doi:10.1016/j.autcon.2020.103137

Jones, D. R., Perttunen, C. D. & Stuckmann, B. E. (1993). Lipschitzian optimization without the lipschitz constant. *Journal of optimization Theory and Applications, 79*, 157. doi:10.1007/bf00941892

Kan, C., Fang, Y., Anumba, C. J. & Messner, J. I. (2018). A cyber–physical system (CPS) for planning and monitoring mobile cranes on construction sites. *Management, Procurement and Law, 171*(6), 240-250. doi:10.1680/jmapl.17.00042

Kim, P., Chen, J. & Cho, Y. K. (2018). SLAM-driven robotic mapping and registration of 3D point clouds. *Automation in Construction, 89*, 38-48. doi:10.1016/j.autcon.2018.01.009

Lee, G., Cho, J., Ham, S., Lee, T., Lee, G., Yun, S.-H. & Yang, H.-J. (2012). A BIM- and sensor-based tower crane navigation system for blind lifts. *Automation in Construction, 26*, 1-10. doi:10.1016/j.autcon.2012.05.002

Lee, U.-K., Kang, K.-I., Kim, G.-H. & Cho, H.-H. (2006). Improving tower crane productivity using wireless technology. *Computer-Aided Civil and Infrastructure Engineering, 21*(8), 594-604. doi:10.1111/j.1467-8667.2006.00459.x

Li, H., Chan, G. & Skitmore, M. (2013). Integrating real time positioning systems to improve blind lifting and loading crane operations. *Construction Management and Economics, 31*(6), 596-605. doi:10.1080/01446193.2012.756144

Li, M., Xue, F. & Yeh, A. G. (2023). Bi-objective analytics of 3D visual-physical nature exposures in high-rise high-density cities for landscape and urban planning. *Landscape and Urban Planning, 233*, 104714. doi:10.1016/j.landurbplan.2023.104714

Li, M., Xue, F., Wu, Y. & Yeh, A. G. (2022). A room with a view: Automatic assessment of window views for high-rise high-density areas using City Information Models and deep

transfer learning. *Landscape and Urban Planning, 226*, 104505. doi:10.1016/j.landurbplan.2022.104505

Li, M., Yeh, A. G. & Xue, F. (2024). CIM-WV: A 2D semantic segmentation dataset of rich window view contents in high-rise, high-density Hong Kong based on photorealistic city information models. *Urban Informatics , 3*(1), 12. doi:10.1007/s44212-024-00039-7

Li, Y., Fan, X., Mitra, N. J., Chamovitz, D., Cohen-Or, D. & Chen, B. (2013). Analyzing growing plants from 4D point cloud data. *ACM Transactions on Graphics, 32*(6), 1-10. doi:10.1145/2508363.2508368

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. *2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference.* Heraklion, Greece: European Council on Computing in Construction. doi:10.35490/EC3.2023.258

Liu, X., Yan, M. & Bohg, J. (2019). Meteornet: Deep learning on dynamic 3d point cloud sequences. *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 9246-9255). IEEE. doi:10.1109/iccv.2019.00934

Luo, W., Yang, B. & Urtasun, R. (2018). Fast and furious: Real time end-to-end 3d detection, tracking and motion forecasting with a single convolutional net. *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 3569-3577). IEEE. doi:10.1109/cvpr.2018.00376

Luo, X., Li, H. & Lee, S. (2023). Bridging the gap: Neuro-Symbolic Computing for advanced AI applications in construction. *Frontiers of Engineering Management, 10*(4), 727-735. doi:10.1007/s42524-023-0266-0

Ma, J. W., Czerniawski, T. & Leite, F. (2020). Semantic segmentation of point clouds of building interiors with deep learning: Augmenting training datasets with synthetic BIM-based point clouds. *Automation in Construction, 113*, 103144. doi:10.1016/j.autcon.2020.103144

Mijwil, M. M., Hiran, K. K., Doshi, R. & Unogwu, O. J. (2023). Advancing construction with IoT and RFID technology in civil engineering: A technology review. *Al-Salam Journal for Engineering and Technology, 2*(2), 54-62. doi:10.55145/ajest.2023.02.02.007

Mirzaei, K., Arashpour, M., Asadi, E., Masoumi, H., Bai, Y. & Behnood, A. (2022). 3D point cloud data processing with machine learning for construction and infrastructure applications: A comprehensive review. *Advanced Engineering Informatics, 51*, 101501. doi:10.1016/j.aei.2021.101501

Myronenko, A. & Song, X. (2010). Point set registration: Coherent point drift. *IEEE transactions on pattern analysis and machine intelligence, 32*(12), 2262-2275. doi:10.1109/tpami.2010.46

Nakanishi, Y., Kaneta, T. & Nishino, S. (2022). A review of monitoring construction equipment in support of construction project management. *Frontiers in Built Environment, 7*, 632593. doi:10.3389/fbuil.2021.632593

Neitzel, R. L., Seixas, N. S. & Ren, K. K. (2001). A Review of Crane Safety in the Construction Industry. *Applied Occupational and Environmental Hygiene, 16*(12), 1106-1117. doi:10.1080/10473220127411

Nelder, J. A. & Mead, R. (1965). A simplex method for function minimization. *The computer journal, 7*(4), 308-313. doi:10.1093/comjnl/7.4.308

Opoku, D. G., Perera, S., Osei-Kyei, R. & Rashidi, M. (2021). Digital twin application in the construction industry: A literature review. *Journal of Building Engineering, 40*, 102726. doi:10.1016/j.jobe.2021.102726

Pomerleau, F., Colas, F. & Siegwart, R. (2015). A review of point cloud registration algorithms for mobile robotics. *Foundations and Trends® in Robotics, 4*(1), 1-104. doi:10.1561/9781680830255

Powell, M. J. (1994). A direct search optimization method that models the objective and constraint functions by linear interpolation. *Advances in Optimization and Numerical Analysis, 275*, 51-67. doi:10.1007/978-94-015-8330-5_4

Powell, M. J. (2006). The NEWUOA software for unconstrained optimization without derivatives. In M. R. G. Pillo, *Large-Scale Nonlinear Optimization* (pp. 255-297). Springer. doi:10.1007/0-387-30065-1_16

Powell, M. J. (2009). *The BOBYQA algorithm for bound constrained optimization without derivatives.* Cambridge NA Report NA2009/06, University of Cambridge, Cambridge, 26, 26-46. Retrieved from

http://bison.ihep.su/~kachaev/books/Powell/NA2009_06_bobyqa_bound_constrained.pdf

Price, L. C., Chen, J., Park, J. & Cho, Y. K. (2021). Multisensor-driven real-time crane monitoring system for blind lift operations: Lessons learned from a case study. *Automation in Construction, 124*, 103552. doi:10.1016/j.autcon.2021.103552

Raguram, R., Frahm, J. & Pollefeys, M. (2008). A comparative analysis of RANSAC techniques leading to adaptive real-time random sample consensus. *European Conference on Computer Vision (ECCV)* (pp. 500-513). Springer. doi:10.1007/978-3-540-88688-4_37

Roberts, D., Bretl, T. & Golparvar-Fard, M. (2017). Detecting and classifying cranes using camera-equipped UAVs for monitoring crane-related safety hazards. *Computing in Civil Engineering 2017* (pp. 442-449). ASCE. doi:10.1061/9780784480847.055

Rusinkiewicz, S. & Levoy, M. (2001). Efficient variants of the ICP algorithm. *Proceedings third international conference on 3-D digital imaging and modeling*, (pp. 145-152). doi:10.1109/im.2001.924423

Russhakim, N. A., Ariff, M. F., Darwin, N., Majid, Z., Idris, K. M., Abbas, M. A., N., K. Z. & Yusoff, A. R. (2018). The suitability of terrestrial laser scanning for strata building. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42*, 67-76. doi:10.5194/isprs-archives-XLII-4-W9-67-2018

Sacks, R., Navon, R., Brodetskaia, I. & Shapira, A. (2005). Feasibility of Automated Monitoring of Lifting Equipment in Support of Project Control. *Journal of Construction Engineering and Management, 131*(5), 604-614. doi:10.1061/(ASCE)0733-9364(2005)131:5(604)

Schnabel, R., Wah, R. & Klein, R. (2007). Efficient RANSAC for point-cloud shape detection. *Computer Graphics Forum, 26*(2), 214-226. doi:10.1111/j.1467-8659.2007.01016.x

Shahi, A., West, J. & Haas, C. (2013). Onsite 3D marking for construction activity tracking. *Automation in construction, 1*(30), 136-43. doi:10.1016/j.autcon.2012.11.027

Shapira, A., Lucko, G. & Schexnayder, C. J. (2007). Cranes for Building Construction Projects. *Journal of Construction Engineering and Management, 133*(9), 690-700. doi:10.1061/(ASCE)0733-9364(2007)133:9(690)

Shi, H. (2023). *4D point cloud semantic segmentation.* Singapore: Doctoral thesis, Nanyang Technological University. doi:10.32657/10356/172100

Shi, H., Lin, G., Wang, H., Hung, T.-Y. & Wang, Z. (2020). Clouds, SpSequenceNet: Semantic Segmentation Network on 4D Point. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 4574-4583). IEEE. doi:10.1109/cvpr42600.2020

Siegbert, R., Kitschke, J., Djelassi, H., Behr, M. & Elgeti, S. (2014). Comparing Optimization Algorithms for Shape Optimization of Extrusion Dies. *85th Annual Meeting of the International Association of Applied Mathematics and Mechanics* (pp. 789-794). Wiley. doi:10.1002/pamm.201410377

Silva, M. F., Green, A., Morales, J., Meyerhofer, P., Yang, Y., Figueiredo, E., Costa, J. C. & Mascareñas, D. (2022). 3D structural vibration identification from dynamic point clouds. *Mechanical Systems and Signal Processing, 166*, 108352. doi:10.1016/j.ymssp.2021.108352

Son, G., Kim, J. & & Kim, Y. (2023). Implementation of Pedestrian Tracking in Low-Resolution Video using Multi-Camera. *Bulletin of Networking, Computing, Systems, and Software, 12*(1), 31-32. Retrieved from http://www.bncss.org/index.php/bncss/article/view/165

Teizer, J., K, W. J. & Schultz, C. (2022). The concept of digital twin for construction safety. *Construction Research Congress* (pp. 1156-1165). ASCE. doi:10.1061/9780784483961.121

Udoh, I. S. & Kotonya, G. (2018). Developing IoT applications: challenges and frameworks. *IET Cyber-Physical Systems: Theory & Applications, 3*(2), 65-72. doi:10.1049/iet-cps.2017.0068

Wang, J., Zhang, Q., Yang, B. & Zhang, B. (2023). Vision-based automated recognition and 3D localization framework for tower cranes using far-field cameras. *Sensors, 23*(10), 4851. doi:10.3390/s23104851

Wang, Q., Kim, M.-K., Yoon, S., Cheng, J. C. & Sohn, H. (2016). Automated quality assessment of precast concrete elements with geometry irregularities using terrestrial laser scanning. *Automation in Construction, 68*, 170-182. doi:10.1016/j.autcon.2016.03.014

Wortmann, T., Waibel, C., Nannicini, G. & Evins, R. (2017). Are genetic algorithms really the best choice for building energy optimization? *Proceedings of the Symposium on Simulation for Architecture and Urban Design*, (pp. 1-8). doi:10.22360/simaud.2017.simaud.006

Wu, Y., Shang, J. & Xue, F. (2021). RegARD: Symmetry-based coarse registration of smartphone's colorful point clouds with CAD drawings for low-cost Digital Twin Buildings. *Remote Sensing, 13*(10), 1882. doi:10.3390/rs13101882

Wu, Y., Xue, F., Li, M. & Chen, S.-H. (2024). A novel Building Section Skeleton for compact 3D reconstruction from point clouds: A study of high-density urban scenes. *ISPRS Journal of Photogrammetry and Remote Sensing, 209*, 85-100. doi:10.1016/j.isprsjprs.2024.01.020

Xia, T., Yang, J. & Chen, L. (2022). Automated semantic segmentation of bridge point cloud based on local descriptor and machine learning. *Automation in Construction, 133*, 103992. doi:10.1016/j.autcon.2021.103992

Xue, F., Lu, W., Chen, K. & Webster, C. J. (2019a). BIM reconstruction from 3D point clouds: A semantic registration approach based on multimodal optimization and architectural design knowledge. *Advanced engineering informatics, 42*, 100965. doi:10.1016/j.aei.2019.100965

Xue, F., Lu, W., Chen, K. & Zetkulic, A. (2019b). From Semantic Segmentation to Semantic Registration: Derivative-Free Optimization–Based Approach for Automatic Generation of Semantically Rich As-Built Building Information Models from 3D Point Clouds. *Journal of Computing in Civil Engineering, 33*(4), 04019024. doi:10.1061/(ASCE)CP.1943-5487.0000839

Xue, F., Lu, W., Chen, Z. & Webster, C. J. (2020). From LiDAR point cloud towards digital twin city: Clustering city objects based on Gestalt principles. *ISPRS Journal of Photogrammetry and Remote Sensing, 167*, 418-431. doi:10.1016/j.isprsjprs.2020.07.020

Yang, J., Vela, P., Teizer, J. & Shi, Z. (2014). Vision-based tower crane tracking for understanding construction activity. *Journal of Computing in Civil Engineering, 28*(1), 103-112. doi:10.1061/(ASCE)CP.1943-5487.0000242.

Yin, C., Yang, B., Cheng, J. C., Gan, V. J., B. W. & Yang, J. (2023). Label-efficient semantic segmentation of large-scale industrial point clouds using weakly supervised learning. *Automation in Construction, 148*, 104757. doi:10.1016/j.autcon.2023.104757

Yin, M., Tang, L., Zhou, T., Wen, Y., Xu, R. & Deng, W. (2020). Automatic layer classification method-based elevation recognition in architectural drawings for reconstruction of 3D BIM models. *Automation in Construction, 113*, 103082. doi:10.1016/j.autcon.2020.103082

Zhang, C. & Arditi, D. (2013). Automated progress control using laser scanning technology. *Automation in Construction, 36*, 108-116. doi:10.1016/j.autcon.2013.08.012

Zhang, C., Hammad, A. & Rodriguez, S. (2012). Crane pose estimation using UWB real-time location system. *Journal of Computing in Civil Engineering, 26*(5), 625-637. doi:10.1061/(asce)cp.1943-5487.0000172

Zhao, L., Xiang, Z., Chen, M., Ma, X., Zhou, Y., Zhang, S. & Hu, K. (2022). Establishment and Extension of a Fast Descriptor for Point Cloud Registration. *Remote Sensing, 14*(17), 4346. doi:10.3390/rs14174346

Zhong, D., Lv, H., Han, J. & Wei, Q. (2014). A Practical Application Combining Wireless Sensor Networks and Internet of Things: Safety Management System for Tower Crane Groups. *sensors, 14*, 13794-13814. doi:10.3390/s140813794

Zhou, Q. Y., Park, J. & Koltun, V. (2016). Fast global registration. *The European Conference on Computer Vision (ECCV)* (pp. 766-782). Springer International Publishing. doi:10.1007/978-3-319-46475-6_47

Zhou, Q.-Y., Park, J. & Koltun, V. (2018). Open3D: {A} Modern Library for 3D Data Processing. *arXiv:1801.09847*.

Zhou, Y., Wang, L., Love, P. E., Ding, L. & Zhou, C. (2019). Three-dimensional (3D) reconstruction of structures and landscapes: a new point-and-line fusion method. *Advanced Engineering Informatics, 42*, 100961. doi:10.1016/j.aei.2019.100961