Automation in Construction 168 (2024) 105843

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Robust localization of shear connectors in accelerated bridge construction with neural radiance field

Gyumin Lee [a], Ali Turab Asad [b], Khurram Shabbir [a,c], Sung-Han Sim [a,b,*], Junhwa Lee [d,**]

[a] Department of Global Smart City, Sungkyunkwan University, Suwon, Republic of Korea
[b] Department of Civil, Architectural and Environmental System Engineering, Sungkyunkwan University, Suwon, Republic of Korea
[c] School of Engineering, Aalto University, Espoo, Finland
[d] Department of Civil Engineering, Pukyong National University, Busan, Republic of Korea

ARTICLE INFO

Keywords:
Accelerated bridge construction
Neural radiance field
Shear connector
Unmanned aerial vehicle

ABSTRACT

Accelerated bridge construction (ABC) demands precise alignment of prefabricated members to prevent assembly failure. Conventional methods struggle to localize shear connectors from point cloud data (PCD) generated by structure-from-motion due to its sparsity. This paper introduces a robust method for shear connector localization using PCD generated by a neural radiance field and a three-step narrowing-down algorithm. The PCD exhibits densely populated points for small connectors, allowing the algorithm to pinpoint their locations accurately. The method successfully identified all 72 shear connectors in a mock-up prefabricated girder, with an average error of 10 mm, demonstrating its potential for assessing constructability in ABC projects. Future research may integrate deep learning-based segmentation techniques to enhance efficiency and adaptability in complex geometries and non-standard bridge designs.

1. Introduction

Accelerated bridge construction (ABC) is a modern approach that minimizes the overall time expense by manufacturing components off-site and assembling onsite. The primary advantage of ABC is the ability to shorten project delivery time and improve constructability compared to conventional cast-in place method [1]. The full-depth prefabricated deck system that involves placing precast concrete panels onto the prefabricated girders is one of the common forms in ABC. The connections between prefabricated girders and precast concrete panels are provided by aligning shear pockets of the panels with shear connectors welded to the girders [2]. However, if the shear connectors are not co-located with shear pockets because of manufacturing errors and possible deformation during delivery [3], this misalignment could potentially lead to assembly failure and necessitate additional rework [4]. To ensure constructability and facilitate successful assembly, a minimum 1.5-in. (38.1 mm) clearance between shear connectors and shear pockets is generally advised in design specifications such as the Federal Highway Administration (FHWA) [5], the American Association of the State Highway and Transportation Officials (AASHTO) [6], and the Precast/Prestressed Concrete Institute (PCI) design specifications [7]. Thus, identifying fabrication errors or deformations in prefabricated members within design specifications is crucial to prevent assembly failures and additional rework.

The localization of shear connectors is an important aspect in the quality control of prefabricated girders. Inspectors are required to measure the position of the shear connector prior to assembling prefabricated girders with precast concrete panels to ensure the integrity and constructability of the structure. The most common method involves manual measurement using a tape measure, which is both inaccurate and ineffective. An alternative method involves using a total station, which allows for accurate measurements with a prism placed on each connector, with an error margin of 2–3 mm from a distance of approximately 100 m [8]. However, these conventional methods, while commonly used, tend to be labor-intensive, time-consuming, and prone to human error. Therefore, an efficient and accurate localization method needs to be developed for measuring the positions of the shear connectors.

Light detection and ranging (LiDAR) has been introduced for the dimensional quality assessment of prefabricated members [8–10] owing to its accurate scanning capability. A recent study has utilized LiDAR to assess the dimensional quality of shear connectors in prefabricated

* Corresponding author at: Department of Global Smart City, Sungkyunkwan University, Suwon, Republic of Korea.
** Corresponding author at: Department of Civil Engineering, Pukyong National University, Busan, Republic of Korea.
E-mail addresses: ssim@skku.edu (S.-H. Sim), lee@pknu.ac.kr (J. Lee).

https://doi.org/10.1016/j.autcon.2024.105843
Received 26 May 2024; Received in revised form 15 October 2024; Accepted 19 October 2024
Available online 24 October 2024

girders [11]. In that study, LiDAR was mounted on a crane at a height sufficient to capture the 3D point cloud data (PCD) of the entire body of the prefabricated girders. Subsequently, the locations of shear connectors were identified from the scanned PCD. However, this method requires the multiple scans for complete acquisition of PCD, which is a significant limitation for onsite applications. Each scan requires approximately 40 min to complete, further exacerbating the time constraints in construction site [12]. Because incident angle and scanning distance affect the density and quality of PCD, an insufficient number of points can be scanned for the shear connectors, depending on the position of the scanner [13]. Furthermore, challenges such as the mixed pixel phenomenon [14], limited points for the scanner installation at construction sites [15], high equipment costs, the requirement of specialized software for processing scanned results, and the demand for skilled operators, are recognized as major factors impairing the practicability of LiDAR. As such, the practical applications of LiDAR in ABC projects are intractable.

Photogrammetry, which is cost-effective and efficient, has been increasingly used to collect as-built geometric spatial information in the architecture, engineering, and construction industries [16]. Structure-from-motion (SfM) is a photogrammetric technique that can reconstruct a 3D PCD from overlapping 2D images captured from various camera angles and working distances [17]. The integration of Unmanned Aerial Vehicles (UAV) with the SfM expands its capabilities to enable the 3D reconstruction of complex structures, such as the as-built structure of bridges under construction [18]. As UAV can conveniently capture images of a full-scale bridge, the approach incorporating UAV with the SfM can efficiently acquire PCD compared with LiDAR. However, the application of the SfM in conjunction with UAV images for the localization of shear connectors faces two main issues: (1) the sparsity of the PCD reconstructed by the SfM [19], which is often insufficient to identify small shear connectors and (2) challenges in the automated search for minute shear connectors in a chaotic bridge construction environment. As such, an effective method for addressing these issues need to be developed.

Recent studies have explored advanced visualization methods for digital twin scenes during bridge construction. Wu et al. proposed a dynamic holographic modeling method that enhances the visualization of complex construction environments, addressing issues such as unstable imaging effects and poor adaptability in the current holographic visualization technology [20]. Although their approach offers valuable insights for construction progress monitoring, there remains a need for the accurate identification of specific structural components in as-built conditions.

This paper proposes a robust localization method for shear connectors using a dense PCD obtained by a neural radiance field (NeRF) [21]. The proposed method initially reconstructs a dense PCD by applying NeRF to the UAV-captured images, thereby offering detailed as-built conditions. Subsequently, the shear connectors are identified and localized using the proposed three-step narrowing-down algorithm. The proposed approach robustly localizes thin shear connectors at the construction site by densely reconstructing the PCD and effectively narrowing down the search region towards the minute shear connectors. The performance of the proposed method is demonstrated using a mock-up prefabricated girder and compared with those of a conventional SfM-based method and total station measurements.

2. Background

2.1. Structure from motion

The SfM is a widely adopted photogrammetric technique for the reconstruction of 3D PCD from UAV-captured images. The entire process is divided into two primary steps: sparse reconstruction and point densification. Sparse reconstruction begins by iteratively identifying common features across the captured images. The matched features are used to estimate the camera poses and the 3D positions of the points via bundle adjustments [22]. After sparse reconstruction, the 3D PCD are densified by a detailed examination of the images through depth map estimation and identification of additional corresponding features. Sparse PCD is used to estimate a depth map that represents the distance from the camera to the surface of the scene for each pixel. To identify additional pixel correspondences across images, densification methods such as patch matching [23] are employed on the sparse PCD. The estimated depth maps and additional pixel correspondences are then merged to generate a dense PCD. Conventionally, the application of the SfM for extracting 3D points from large structures is limited by the lengthy time required for image acquisition. However, with the recent advancements in UAV technologies and image-processing capabilities [24], 3D PCD acquisition using the SfM has become more efficient, offering advantages for capturing large scenes. As such, the SfM has gained popularity for obtaining the 3D PCD of large civil infrastructures using UAV.

The sparsity of the PCD resulting from the SfM is the primary concern when applied to the localization of shear connectors in ABC project. Given the small and thin nature of shear connectors relative to the overall structure, the PCD must encompass not only the entire construction site but also the finer details to ensure precise localization. To achieve this objective, a UAV captures images from both a far distance to encompass the entire scene and from a close distance to capture the small shear connectors [25]. As an overlap of 60–80 % between consecutive images is recommended for the SfM [26], this capturing process results in an exhaustive image acquisition time, typically lasting 1–2 h for ordinary mid-size bridges depending on the project scale and UAV capabilities. In practice, images of the construction site are acquired within a limited period, often resulting in an insufficient number of images and leading to the sparsity of the PCD. Moreover, manual UAV operation during image capture can introduce variations in lighting conditions, image blurring owing to flight instability, and object occlusions. Because the SfM heavily relies on feature extraction and matching across images, these issues can lead to incomplete reconstructions or inconsistent point densities. As such, UAV-powered SfM can efficiently reconstruct a PCD for large civil structures, whereas the point density is limitedly guaranteed for small and detailed objects, such as shear connectors.

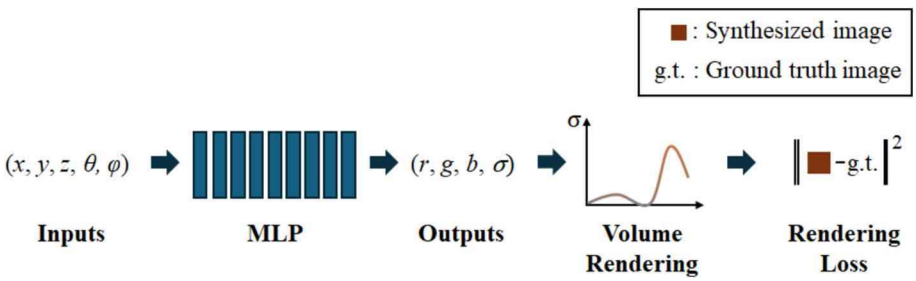


Fig. 1. Overall process of NeRF [21].

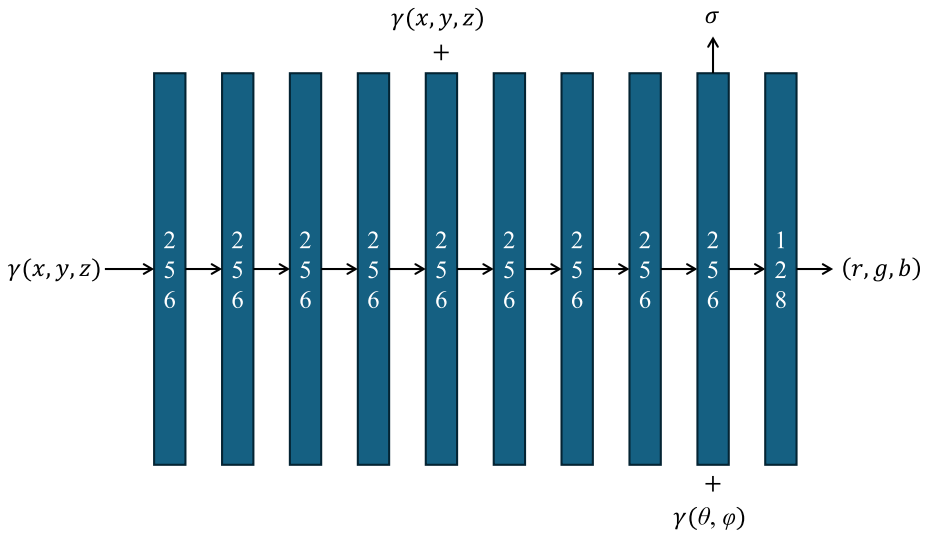


Fig. 2. MLP architecture for NeRF [21].

2.2. Neural radiance field

NeRF is a neural rendering method that can implicitly represents 3D scenes [21]. A widely adopted method for representing 3D scenes is by a collection of 3D points, each with corresponding red, green, and blue (RGB) colors. Such a conventional method can be considered as an explicit representation because geometric information is explicitly stored. On the other hand, the NeRF employs a continuous function that receives viewpoint information as input and outputs the corresponding RGB color and opacity for the given viewpoints. Herein, the function employed in the NeRF takes the form of a multilayer perceptron (MLP), thus, the name neural radiance field is appropriate. The NeRF is unable to represent the scene without inputting viewpoints into the function; hence, an implicit representation. Unlike explicit methods that discretely store points, the NeRF continuously represents the scene, capturing fine details such as texture, lighting variations, and geometric intricacies. As such, The NeRF is an implicit representation method for 3D scenes that provides an advantage over explicit representations in terms of capturing fine details.

The NeRF can synthesize photorealistic images from unseen viewpoints using three main procedures: (1) MLP configuration, (2) MLP training, and (3) view synthesis [21]. Firstly, a specialized MLP architecture was configured for the NeRF, which implicitly represents a 3D scene. One of the most well-known NeRF is illustrated in Fig. 1. The MLP takes input consisting of two components: the 3D spatial coordinates (x, y, z) of a point in space, and the 2D viewing direction (θ, φ) of the camera towards that point. Given a 5D input vector (x, y, z, θ, φ), the MLP outputs the volume density (σ) and the emitted RGB color information of the point (r, g, b). Here, higher values of the volume density indicate a more point that absorbs more light, whereas lower values suggest transparency. The outputs were used to synthesize images using a volume rendering technique, the details can be found in the literature [21]. Secondly, the MLP was trained using 2D images captured from various camera poses. The parameters of the MLP were optimized in such a way that its output closely approximated the true image. Finally, the NeRF equipped with the trained MLP was used to synthesize images of the scene from any given camera position. Details of these three procedures are discussed below.

The MLP of the NeRF was configured to represent the 3D scene implicitly. This section focuses on the original NeRF architecture as described in the initial study [21], although various architectures are also available, such as Nerfacto [27], Instant-NGP [28], and TensorRF [29]. Given the 3D coordinates of a point in space, the MLP outputs the RGB values and volume density of that point when captured from a camera with a viewing direction of (θ, φ). To do so, the MLP employs the special architecture shown in Fig. 2, the details of which can be found in an initial study [25]. This architecture utilizes positional encoding to the 5D input vector (x, y, z, θ, φ), forming a 60D vector of $\gamma(x, y, z)$ and a 24D vector of $\gamma(\theta, \varphi)$. Note that the positional encoding is used to enhance the performance of the MLP, particularly in capturing high-frequency scenes

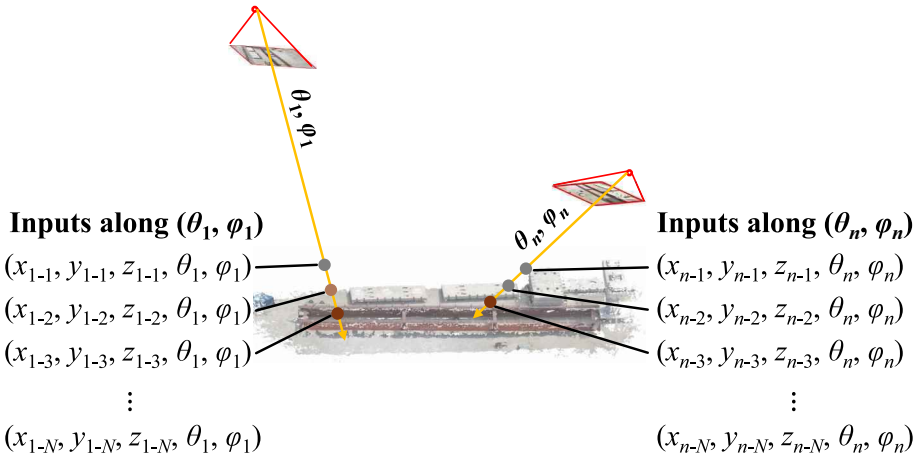


Fig. 3. Example of input vectors.

such as thin and tiny objects. These positional-encoded inputs pass through the MLP with the nine fully connected layers, and then the σ and (r, g, b) of the point are obtained as outputs. In summary, the MLP helps retrieve color information by inputting the coordinates of a desired point and its viewing direction, thereby providing an implicit representation of the 3D scene. Compared with an explicit representation of the scene, the MLP-based NeRF can continuously provide detailed color information for any object within the scene from any viewpoint. Therefore, the MLP of the NeRF was designed to represent the 3D scene as a continuous implicit function, taking a 5D input vector (x, y, z, θ, φ) and outputting the RGB values (r, g, b) and volume density (σ) of the point.

The MLP was trained by minimizing the difference between the colors estimated by the NeRF and those of the real images. Firstly, given a set of images that are captured in various camera poses, a training dataset was prepared. The viewing direction (θ, φ) of the input was prepared for each image by identifying the camera poses using the classical SfM [15]. For each image and its identified camera poses, the spatial coordinates (x, y, z) of the points were generated by sampling within the ray cast from the camera center and passing through each pixel of the given images as shown in Fig. 3. Thereby, multiple 5D input vectors (x, y, z, θ, φ) along the ray were prepared for each pixel of the images. Once the training dataset was prepared, the MLP was trained by minimizing the rendering loss, which is definition as follows. The synthesized image was prepared using a classical volume rendering technique, where the pixel color is computed by a weighted sum of (r, g, b) values over σ for each pixel to form an image. The inclusion of the viewing direction (θ, φ) in the input vectors allows the NeRF to learn how object appearances change with viewing angle. This capability is crucial for addressing complex scenarios such as occlusions caused by equipment at construction sites and varying lighting conditions. The rendering loss was then calculated by comparing the real images with the volume rendering results of the MLP, typically using a pixel-wise difference such as the mean squared error, peak signal-to-noise ratio (PSNR), and structural similarity index measure (SSIM) [30]. The MLP parameters were tuned by minimizing the rendering loss over the input images. As a result, the trained MLP could estimate the color of the pixels of an image captured at an arbitrary camera position, which is close to that of real images. This approach also enables the NeRF to handle challenges specific to UAV-based imagery in construction environments, including potential variations in lighting conditions, image blurring caused by rapid changes in viewpoints, and object occlusions.

The trained NeRF was used to render realistic images from any camera position and viewing direction. The trained NeRF outputs the RGB and volume density from a given ray that passes through the camera center and the pixel of interest. Subsequently, the pixel color corresponding to the ray is computed by applying the a volume-rendering technique to the outputs. Hence, for a given camera position with known pixel coordinates, a photorealistic image can be synthesized by estimating the color of each pixel. As such, NeRF can be used to synthesize images from novel viewpoints using a trained MLP. The capability to synthesize novel views is attributed to the continuous implicit representation of the 3D scene by the NeRF, in contrast to the reliance of the SfM discrete feature point matching. Such continuous representation enables the NeRF to synthesize realistic images from novel viewpoints, even for complex scenes with minute objects like shear connectors in ABC project. This capability remains robust in challenging conditions including lighting variations, partial occlusions, and image blur caused by UAV instability and results in dense PCD in ABC projects.

3. Proposed method for automated localization of shear connectors

This paper proposes an automated method for localizing shear connectors in ABC projects using the NeRF and a three-step narrowing-down algorithm. The overall procedure of the proposed method is shown in Fig. 4. The proposed method first prepares a dense PCD of the construction site using UAV-captured images through the NeRF. Considering the significantly smaller size of the shear connectors relative to the entire structure, a dense reconstruction of the PCD through the NeRF helps to prevent the loss of points for the shear connectors. The 3D PCD of the construction site is then processed by the three-step narrowing-down algorithm that strategically searches for thin shear connectors in a complex construction scene. The NeRF-based 3D reconstruction and the three-step narrowing-down algorithm are described in Sections 3.1 and 3.2, respectively. The summary of the proposed method is provided in Section 3.3.

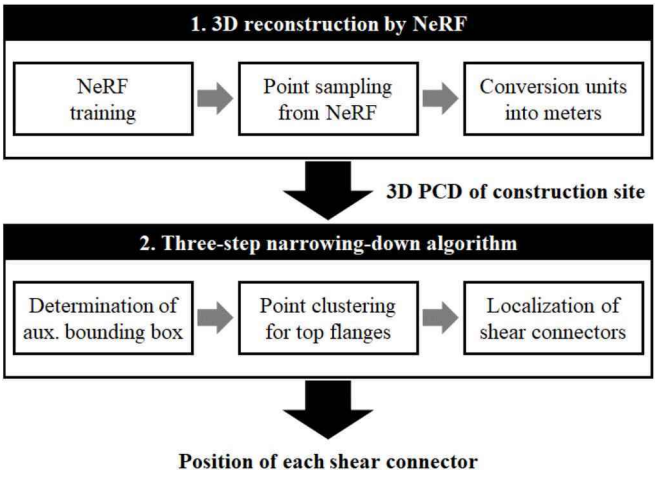


Fig. 4. Overview of the proposed method.

3.1. 3D reconstruction by NeRF

Dense 3D PCD of the construction site were generated using the NeRF. Suppose a UAV is employed to capture a construction site from various positions and angles. NeRF is then utilized to create an implicit representation of the 3D scene, as explained in Section 2.2. Subsequently, 3D points are sampled from the trained NeRF model. Finally, the points are converted into real-world metric units using the global navigation satellite system (GNSS) data recorded along with the captured images. This method results in a dense 3D PCD, with points that densely covering fine details such as shear connectors. Each step is explained in detail.

The NeRF was utilized to build an implicit representation of a construction site from the UAV-captured images. First, the NeRF model was configured. Here, not only the original NeRF architecture introduced in Section 2.2 can be employed, but also recent NeRF models can be used, such as Nerfacto [27], Instant-NGP [28], and TensorRF [29]. Once the NeRF model was configured, the next step is to train the MLP in the NeRF using the UAV-captured images. Input and output pairs, which are respectively (x, y, z, θ, φ) and (r, g, b), respectively were prepared for the training. Among the input vector, the viewing directions (θ, φ) for each pixel of each image were computed by connecting the camera center to each pixel, utilizing the camera poses obtained through the SfM. N spatial coordinates (x, y, z) along the viewing direction (θ, φ) were sampled; hence, N input vectors were prepared for each pixel. Therefore, the input–output pairs were prepared for each pixel of the UAV-captured images, consisting of the input vectors and the RGB color of the corresponding pixels. The pairs were then used to train the MLP of the NeRF by minimizing the rendering loss, as discussed in Section 2.2. Consequently, the trained NeRF could represent the 3D scene of the construction site as an implicit function.

Dense 3D PCD were generated by the sampling points from the trained NeRF. As discussed in Section 2.2, volume density (σ) refers to the probability of existence of a point ranging from 0 to 1; thus, points

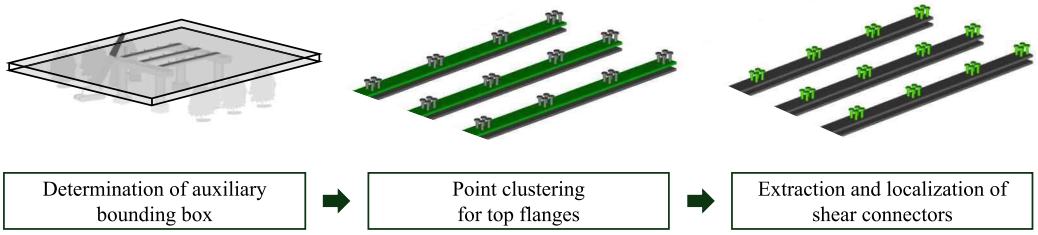
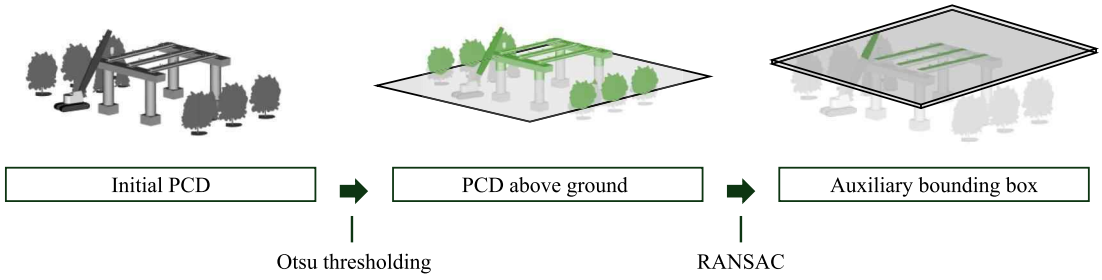
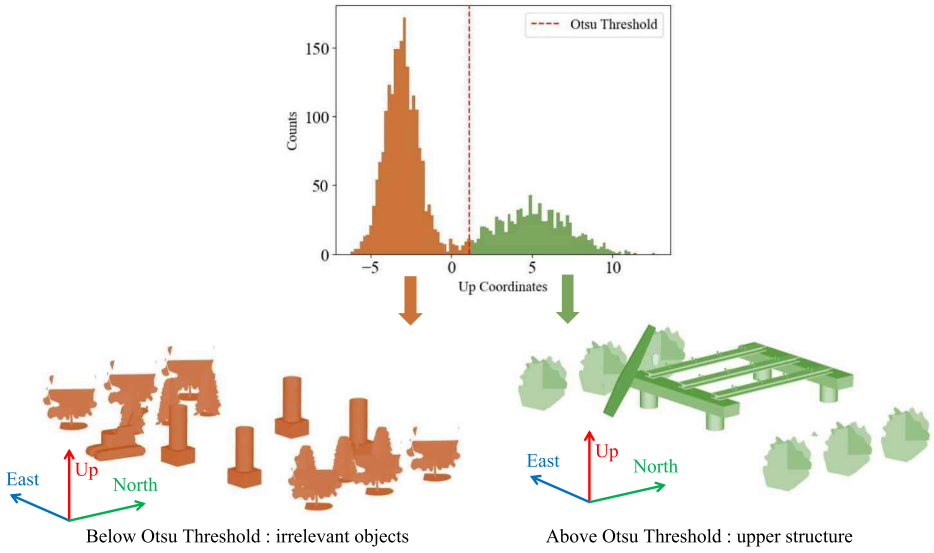


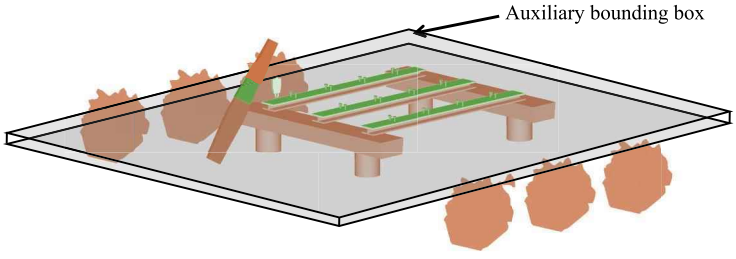
Fig. 5. Overview of the three-step narrowing-down algorithm.



(a) Overview of the extraction of auxiliary bounding box



(b) Binarization of the PCD using the Otsu threshold to separate the ground and structure



(c) Extraction of the auxiliary bounding box using RANSAC

Fig. 6. Extraction of the auxiliary bounding box.

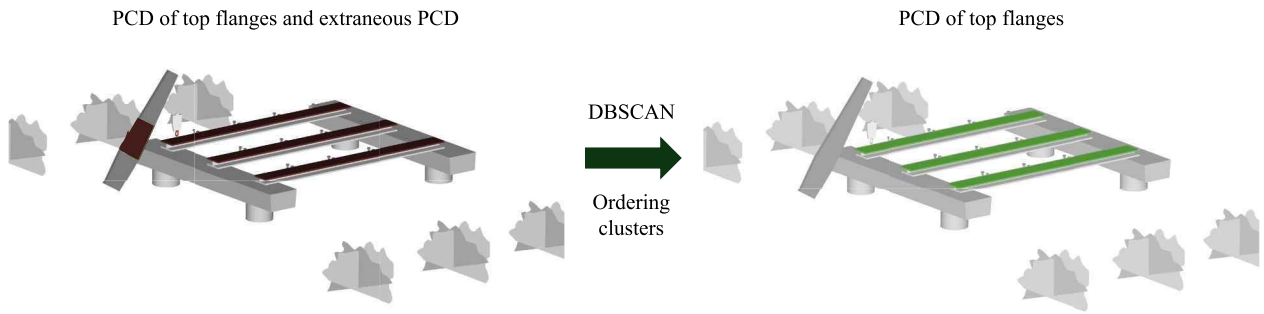


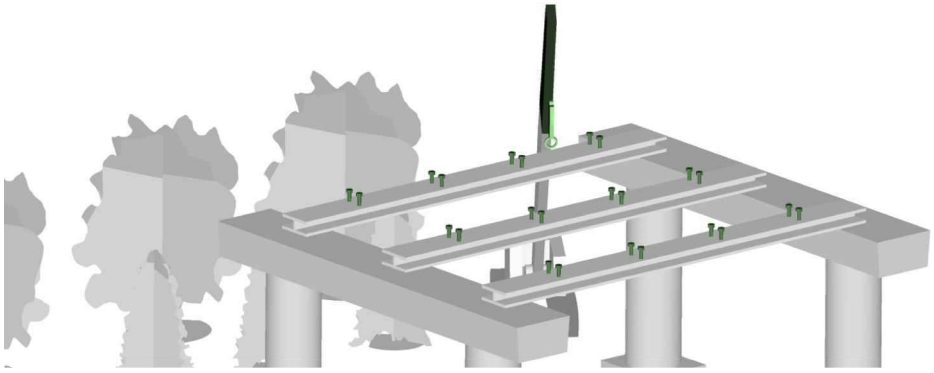
Fig. 7. Extraction of the PCD for the top flange using DBSCAN.

with low volume density were filtered to generate the 3D PCD. The volume densities were computed for each input vector by inputting the data into the trained NeRF, forming σ-augmented vectors of $(x, y, z; \sigma)$ for each input in the training data set. The σ-augmented vectors with a volume density lower than 0.5 were filtered out, and resulting in the (x, y, z) coordinates of the remaining vectors forming the desired 3D PCD. Here, the PCD density can be controlled by setting N; a higher N leads to a denser PCD and vice versa. Thus, 3D PCD could be generated from the trained NeRF, with its density controlled by the hyperparameter N.

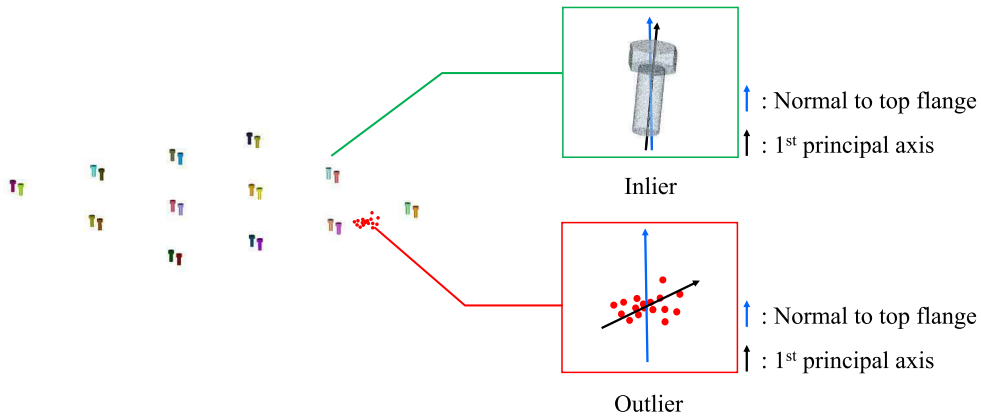
The 3D PCD extracted from the NeRF can be converted into real-world metric units by computing a 3D similarity transform using the GNSS data. The 3D PCD obtained by the trained NeRF is represented in a normalized space, indicating that the points are scaled to fit within a unit cube. The 3D similarity transformation from the normalized space to real-world metric units can be computed by comparing the GNSS data for each image with the 3D camera poses in the normalized space [17]. Once the transformation is computed, the 3D PCD are transformed into the real-world metric units. Consequently, the resulting PCD are in real-world metric units, covering the shear connectors as well as the construction site.

3.2. *Three-step narrowing-down algorithm*

The three-step narrowing-down algorithm searches for shear connectors within the PCD of a construction site, as illustrated in Fig. 5. Consider that the PCD obtained from UAV-captured images contains not



(a) Candidate shear connector PCD



(b) Identification of shear connectors based on the verticality and height

Fig. 8. Extraction of the shear connectors using the DBSCAN and design information.

only the bridge structure of interest but also extraneous objects such as construction equipment, machinery, vehicles, vegetation, and temporary structures. To automate the extraction of shear connectors while disregarding such extraneous objects, the three-step narrowing-down algorithm sequentially identifies the girder of the bridge, the top flanges of these girders, and the shear connectors on the top flanges. The first step involves narrowing the search region to an auxiliary bounding box (ABB) that encompasses the girders of the bridge. The second step is to identify the top flanges using the density-based spatial clustering of applications with noise (DBSCAN) [31] within the ABB. Finally, the shear connectors on the top flanges are extracted using the DBSCAN and a principal axis analysis. The following subsections provide the details of the proposed algorithm for each step.

The ABB was computed using Otsu's method [32] and random sample consensus (RANSAC) [33], as shown in Fig. 6 (a). First, the Otsu threshold was calculated to segment the PCD into two distinct groups: prefabricated structure and ground. Suppose that the PCD are represented in a local east-north-up coordinate system. The upper coordinates of each point in the PCD exhibit a bimodal distribution, as shown in Fig. 6 (b), with one mode representing the ground and the other representing mainly the prefabricated structure. The Otsu threshold was determined based on the upper coordinates of the PCD, effectively separating the points corresponding to the prefabricated structure from those representing the ground. Subsequently, RANSAC is applied to the points above the Otsu threshold to identify the auxiliary bounding box for the top flange, as illustrated in Fig. 6 (c). When points above the Otsu threshold were considered, RANSAC robustly detected a single plane that corresponding to the top flange. The plane determined by RANSAC is expanded along the normal direction by a specified height to encompass the entire top flanges and form the ABB. Herein, the height was manually determined based on the thickness of the top flange, with some margin added for robustness. The extracted ABB was used in the subsequent localization algorithm to narrow the search zone.

The DBSCAN [31] was employed to segment the top flanges from the PCD within the ABB, as shown in Fig. 7. The clusters resulting from the DBSCAN consist of points, where each point has at least M neighboring points within a radius of ε; hence, two hyperparameters (M and ε) need to be determined for its use. These two parameters are determined by referring to the point density in the NeRF-generated PCD and the clearance between the girders. That is, the desired point density (M/ε) was set to be lower than the average point density in the PCD reconstructed by NeRF, ensuring that the resulting clusters from the DBSCAN represent physical objects. Meanwhile, ε was set to be smaller than the distance between the girders, ensuring that each cluster contains the top flange of a single girder. Once the hyperparameters were determined based on these two rules, the PCD within the ABB were grouped into clusters corresponding to the objects, including girders and extraneous objects. Finally, given that there were G girders on the construction site, the clusters with the G highest number of points were extracted, representing the PCD for the top flange of each girder. The extracted point clusters for the top flanges were utilized in the following process to extract the shear connectors.

The shear connectors were identified based on their shapes and the PCD of the top flanges. First, the initial candidates for the shear connectors were extracted from the PCD above the top flanges, as shown in Fig. 8 (a). Despite the PCD within the candidate regions primarily consisted of shear connectors, extraneous PCD were still present, as shown in Fig. 8 (b). Subsequently, the DBSCAN was used to extract the point clusters that potentially correspond to the shear connectors. Specifically, using the DBSCAN with the same hyperparameters (M and ε) used for the extraction of the top flanges, the initial candidates were clustered into potential connectors. Third, the shear connectors were identified from the potential connectors based on their height and verticality, as depicted in Fig. 8 (b). Here, only clusters that were aligned with the normal vector of the top flange using cosine similarity and had a length similar to that of the shear connector were extracted. Finally, localization was achieved by taking the centroid of each point cluster for the shear connectors. In summary, the PCD for each shear connector can be identified using the DBSCAN and shape analysis, enabling for precise localization of their positions.

I. Reconstruction of 3D PCD by NeRF

1. NeRF configuration and training from UAV-captured images
- NeRF is configured
- NeRF is trained using UAV-captured images of construction site

2. 3D PCD generation by trained NeRF
- 3D PCD are generated by sampling points from the trained NeRF

3. Unit conversion of the 3D PCD
- PCD are transformed into real-world metric coordinate system using GNSS

II. Three-step narrowing-down algorithm

1. Determination of auxiliary bounding box (ABB)
- The PCD of upper structure are segregated using the Otsu threshold
- ABB for top flanges is computed using RANSAC

2. Point clustering for top flanges
- Hyperparameters for DBSCAN are determined
- The PCD of top flanges are identified using DBSCAN

3. Localization of shear connectors
- The shear connectors are localized by DBSCAN and shape analysis

Fig. 9. Summary of the proposed method.

3.3. Summary

The proposed method is summarized as follows (see Fig. 9). First, a dense 3D PCD of the construction site were acquired by processing the UAV-captured images using the NeRF. Subsequently, the shear connectors were localized by the three-step narrowing-down algorithm. As a result, the proposed method enabled the robust localization of shear connectors using the NeRF-aided dense 3D PCD and the automated detection algorithm.

The proposed method offers significant advantages in the localization of shear connectors, improving the accuracy after the placement of prefabricated girders and enhancing quality control. First, the NeRF can generate a denser PCD than those obtained through the SfM, facilitating a more robust detection of shear connectors. It is important to recognize that thin objects like shear connectors tend to appear significantly different depending on the viewing angle. The SfM relies on matching similar features between images for 3D reconstruction. Hence, densely extracting points corresponding to these thin 3D objects is a challenge. However, the NeRF represents a 3D scene as a continuous function in the form of an MLP, resulting in a PCD with densely generated points on the shear connectors. Second, the three-step narrowing-down algorithm automates the localization of shear connectors from the PCD of the construction site. Shear connectors are tiny compared with the construction scenes, making their identification challenging. The proposed algorithm strategically narrowed the entire construction scene to the small shear connectors. Because the NeRF-based PCD ensures the connectivity of the shear connectors in the PCD, the proposed algorithm can successfully identify the shear connectors. Thus, the proposed method provides robust localization of shear connectors from the UAV-captured images. The identified shear connectors and their precise locations can be utilized to pre-assess the feasibility of the assembly between precast concrete panels and prefabricated girders in ABC project.

4. Experimental demonstration

The proposed method was demonstrated using the prefabricated girders of a mock-up ABC project as shown in Fig. 10. These girders were composed of two steel I-beams, each of length 20 m. The blocks of the shear connectors were prepared with steel plates, each having two 0.15-

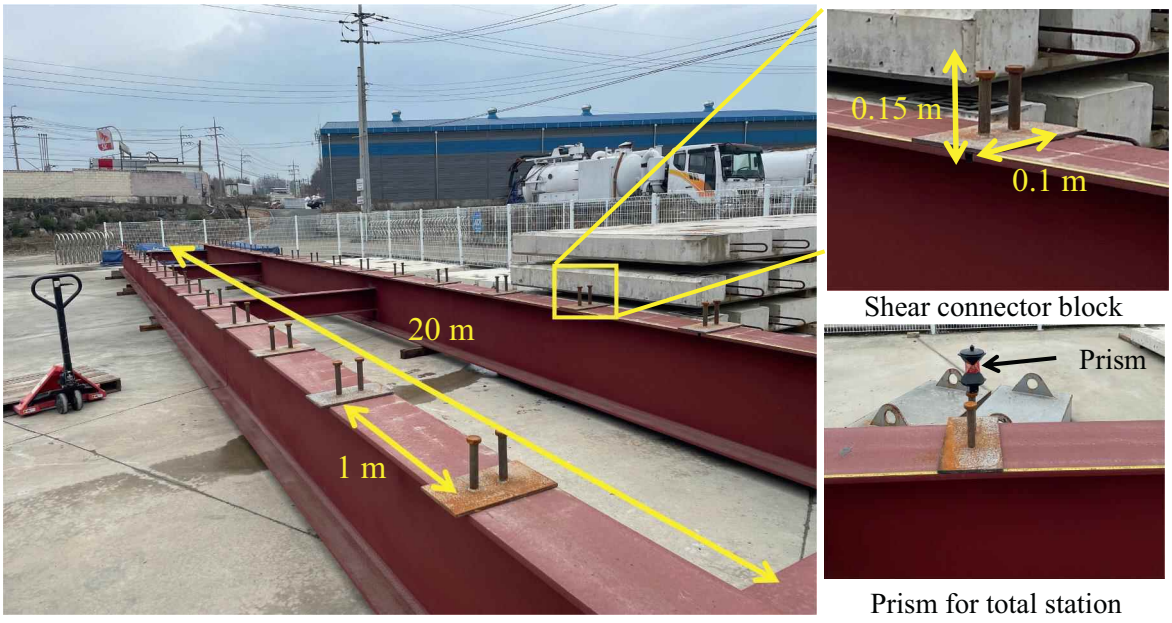


Fig. 10. Mock-up prefabricated beam for experimental validation.

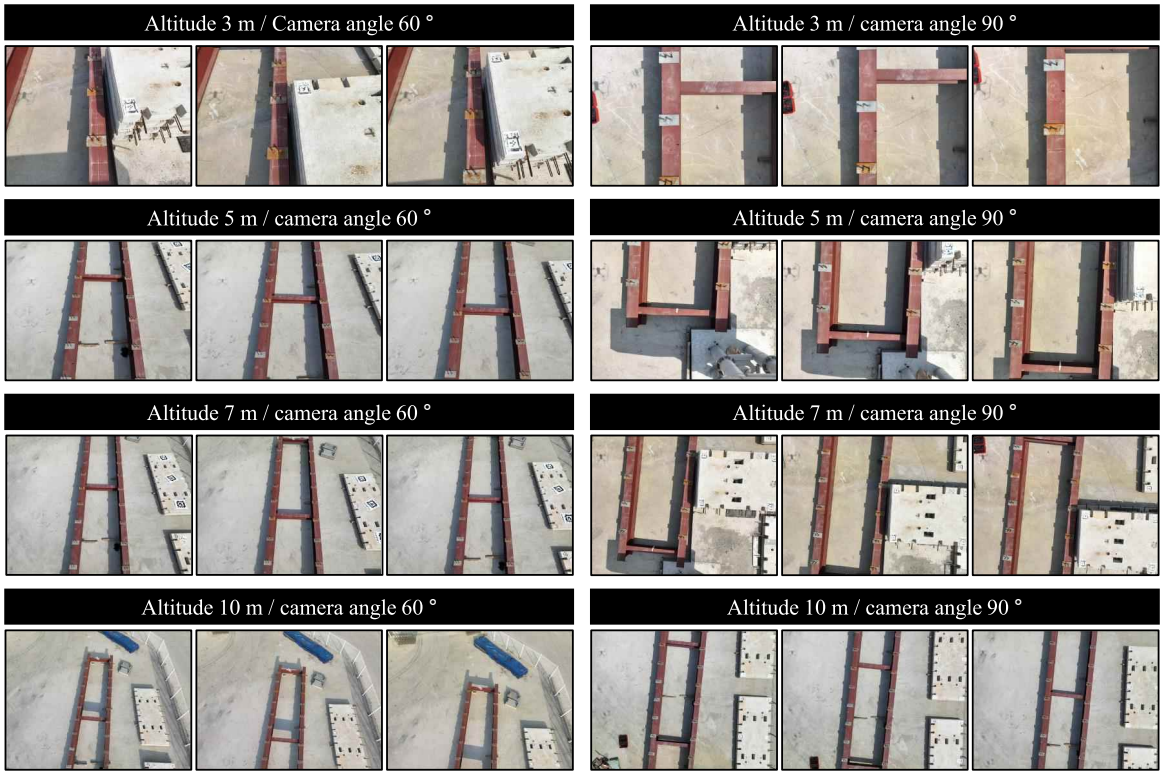
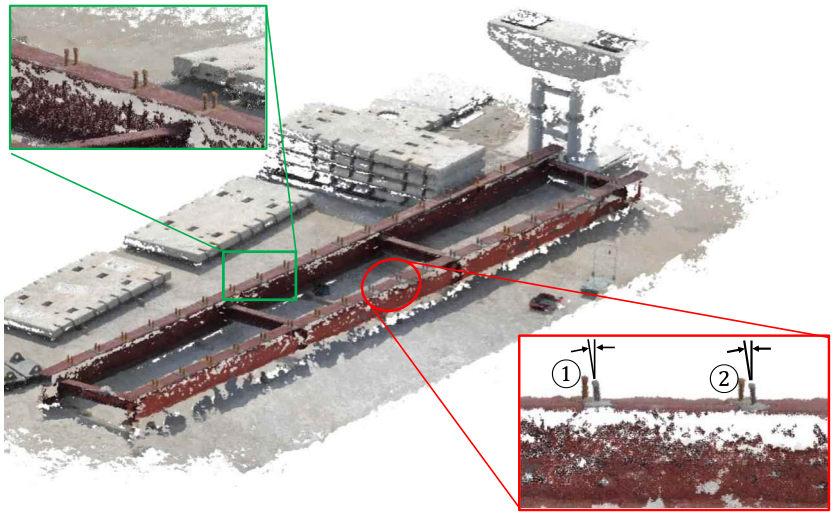


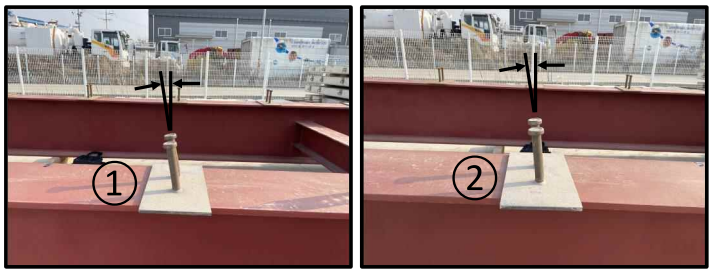
Fig. 11. UAV captured images with varying working distances and camera angles.

m-long shear studs welded with a spacing of 0.1 m, as shown in Fig. 10. Thirty-six blocks were evenly placed at intervals of 1 m, resulting in 72 shear connectors located on the girders. The 3D position of the shear connectors is a concern in this ABC project.

The positions of the shear connectors were measured using the proposed method, a conventional approach, and a total station. The conventional approach was identical to the proposed method, except that the 3D PCD were reconstructed using the SfM instead of the NeRF. A commercial grade UAV, DJI Mavic 3, was employed to capture the images of the construction site. The UAV captured 20-megapixel images with a maximum flight time of 46 min. A total of 388 images were captured by UAV flying at the altitudes of 3 m, 5 m, 7 m, 10 m with the camera angles of 60° and 90° preserving more than 50 % overlap in both the horizontal and vertical directions between images. This diverse set of capture distances and angles is recommended to enhance the quality of the NeRF reconstruction [34]. The resulting images are shown in Fig. 11. The images were processed using the proposed method to identify the 3D positions of the shear connectors. For comparison, the same images



(a) Reconstruction result from the UAV-captured images of prefabricated girders



(b) Reconstruction of shear connectors, preserving their shape

Fig. 12. PCD of the testbed reconstructed by UAV-captured images and the NeRF.

were processed using Pix4D [35], a commercial SfM software package, to obtain the shear connectors using the conventional approach. An industrial-grade total station, the Leica TS13, was installed at a fixed point to measure the 3D position of the shear connectors, which serves as a reference. Here, the total station measured the 3D position of the prism that was manually placed on the head of each connector, as depicted in Fig. 10. The onsite UAV image capture required approximately 40–50 min to complete on-site, whereas the total station measurements required approximately 90 min. The detailed procedure of the proposed method is outlined in Section 4.1. The comparison between the proposed method and the conventional approaches is presented in Section 4.2, where the total station acts as a reference.

4.1. Processing of images by proposed method

The NeRF was used to reconstruct the PCD from the images acquired using the UAV, as discussed in Section 3.1. Firstly, the NeRF model was configured and trained. Among the NeRF models, this validation adopted Nerfacto in the NeRFstudio framework [27] due to its computational efficiency. The Nerfacto model required a minimum of 6GB of GPU memory. To train the Nerfacto model, camera poses were computed for each image using the SfM. Hloc [36], an open-source SfM tool, was employed because of its robustness in determining camera poses regardless of texture similarities in the UAV-captured images of prefabricated girders. The viewing directions were computed by connecting the camera centers to each pixel in the images. Although a total of 388 × 20 million viewing directions are available, Nerfacto randomly samples the rays for computational efficiency based on the specified rays-per-batch size value. In this validation, the rays-per-batch size were set to 12,000 during training and 4096 during evaluation. For each ray, 128 spatial points were sampled to form the input vector of the MLP, which utilized a hidden dimension of 128 for both the main network and color prediction layers. Other hyperparameters included the near and far planes of 0.005 and 1000 respectively, defining the sampling range along each ray; the maximum base resolution of 16,384 for the highest spatial resolution of feature grids; the numbers of proposed samples per ray of set to 512 and 350 for two-stage sampling; and the maximum resolution of the proposed network of 512 for initial geometry estimation at a lower resolution. The training was performed on a workstation equipped with dual NVIDIA RTX 4090 GPU and Intel Core i9-10,940× CPU, and 256GB of RAM. Nerfacto is then trained for 250,000 iterations over 12 h, using the input-output pairs prepared from the 388 images. After training, the model demonstrated a PSNR of 30.43 and an SSIM of 0.96. Subsequently, a dense PCD were generated from the trained Nerfacto model, resulting in 10 million 3D points. The 3D PCD were then transformed into a real-world metric coordinate system by computing the similarity transform using the GNSS data from the captured images. The captured images were processed to generate the 3D PCD using the Nerfacto.

The resulting PCD clearly displayed high-density shear connectors. A dense coverage of the 3D points for the shear connectors was observed in the PCD, as shown in Fig. 12 (a). The inclined shear connectors were reconstructed while preserving their shape, as illustrated in Fig. 12 (b). Furthermore, because various objects were included in the UAV-

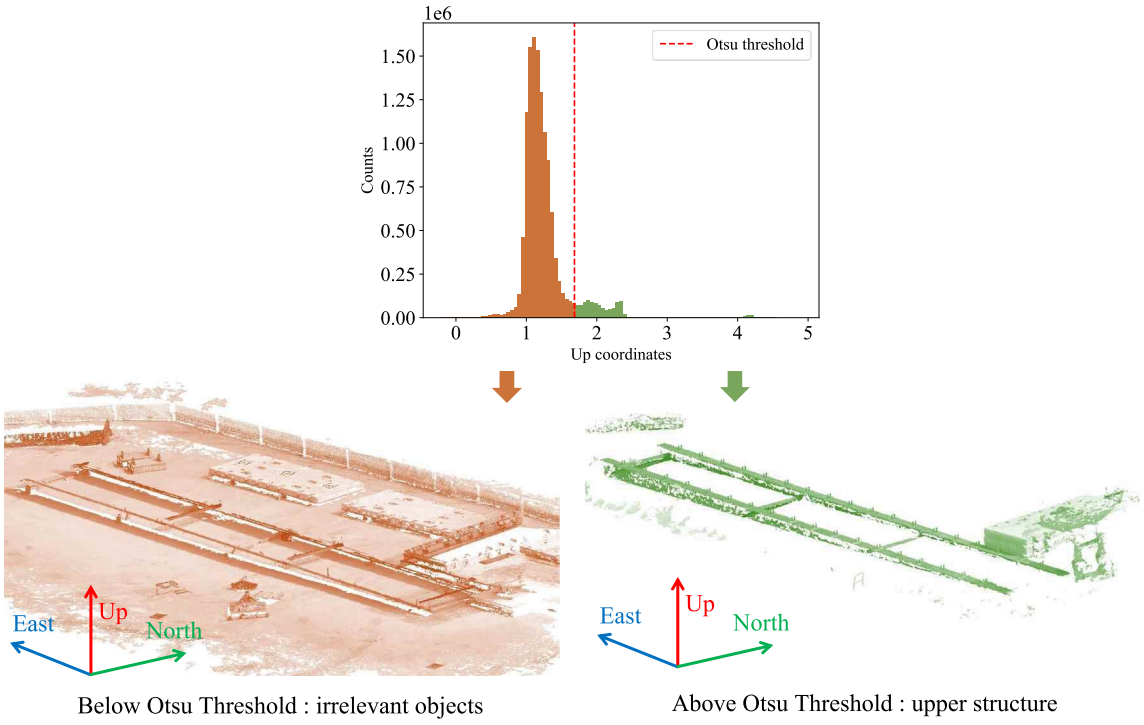


Fig. 13. Otsu binarization result of the PCD.

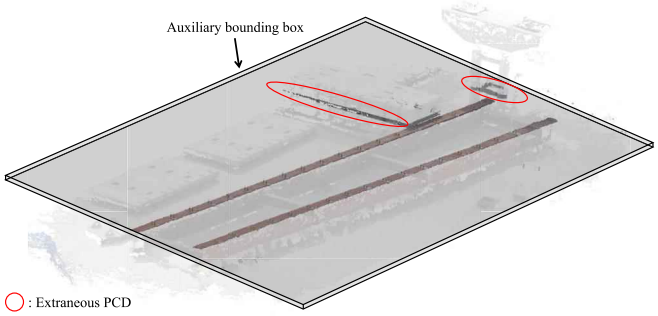


Fig. 14. ABB determined using RANSAC.

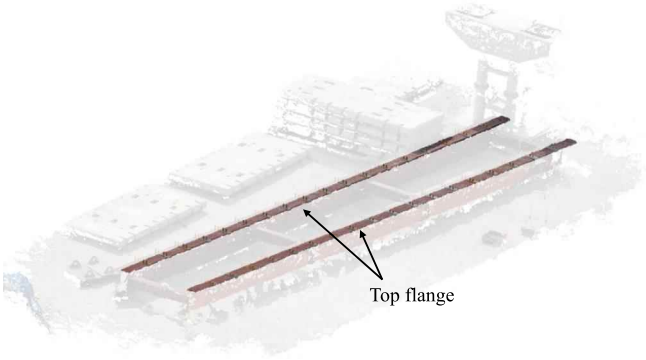


Fig. 15. PCD of the top flanges via the second step of the three-step narrowing-down algorithm.

captured images, the PCD reconstructed using the NeRF comprises both the prefabricated girders and extraneous objects on the construction site, such as construction materials and precast concrete panels. From these PCD, the 72 shear connectors are identified using the proposed three-step narrowing-down algorithm as follows.

The PCD of the construction site were processed to localize the shear connectors using the proposed three-step narrowing-down algorithm. The first step was to determine the ABB from the PCD using the Otsu threshold and RANSAC. The Otsu threshold of 1.7 m was calculated using the histogram of the "up" coordinates, as shown in Fig. 13. Points with higher"up" coordinates than the Otsu threshold were selected as shown in Fig. 13. Subsequently, RANSAC is applied to the selected PCD to search for the ABB. Note that a thickness value of 0.05 m is determined by considering that the thickness of the top flange is 0.02 m and the height of the shear connector is 0.15 m. This ABB defines the region of interest that barely includes the top flanges of the girders as shown in Fig. 14. As extraneous PCD still exists within the ABB, further processing is required to isolate the PCD of the top flanges.

The second step of the three-step narrowing-down algorithm isolated the PCD of the top flanges from the entire PCD using the ABB based on the point density in the NeRF-generated PCD and the clearance between the girders. Based on the referencing point density in the NeRF-generated PCD and the clearance between the girders as discussed in Section 3.2, the hyperparameters M and ε for the DBSCAN were calculated to be 6 and 0.012, respectively. By prioritizing the clusters based on their sizes, the PCD clusters corresponding to the top flanges can be obtained as shown in Fig. 15. Consequently, the search region was narrowed down to the areas above the top flange.

The last step of the three-step narrowing-down algorithm was applied to the PCD above the top flanges to localize the shear connectors. Firstly, the candidate regions for the shear connectors were extracted based on the PCD of the top flanges. The PCD within the candidate region of shear connectors colored in red are illustrated in Fig. 16. Subsequently, the DBSCAN was applied to cluster the PCD of the individual shear connectors. Here, the same hyperparameters used in the second step of the three-step narrowing-down algorithm can be applied because NeRF guarantees a similar point density. Clusters satisfying the criteria of having a principal axis vertical to the normal vector of the plane that were obtained from the RANSAC and having a length of 0.15 m were selected as the shear connectors. All shear

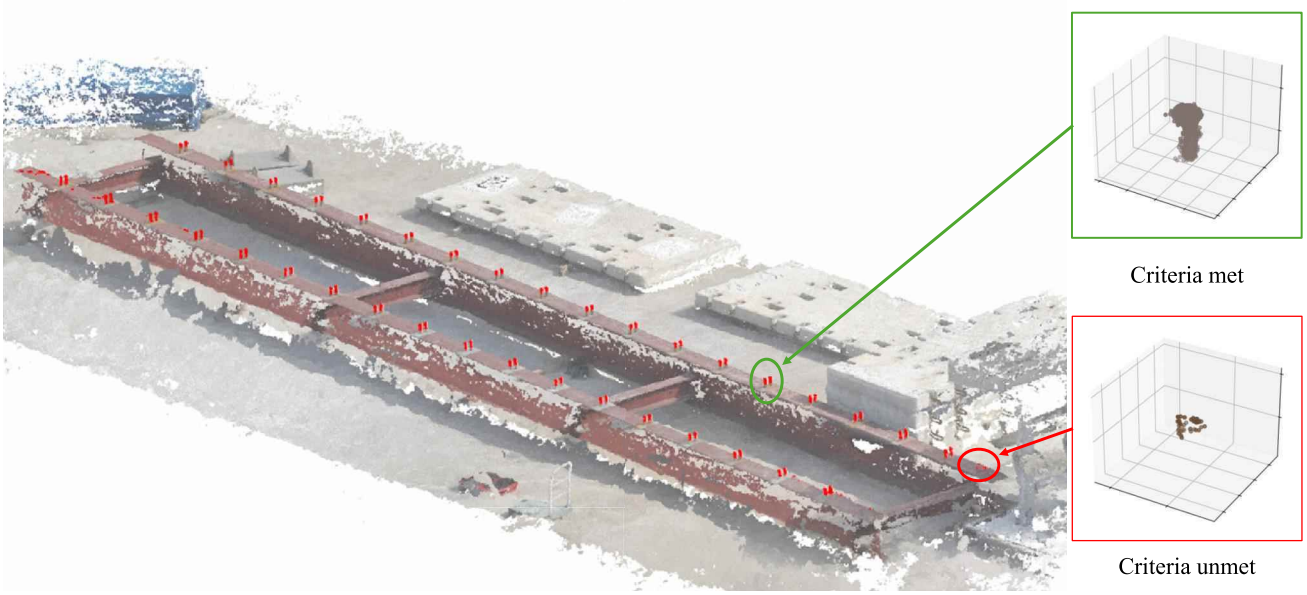


Fig. 16. Candidate PCD of the shear connectors (colored in red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

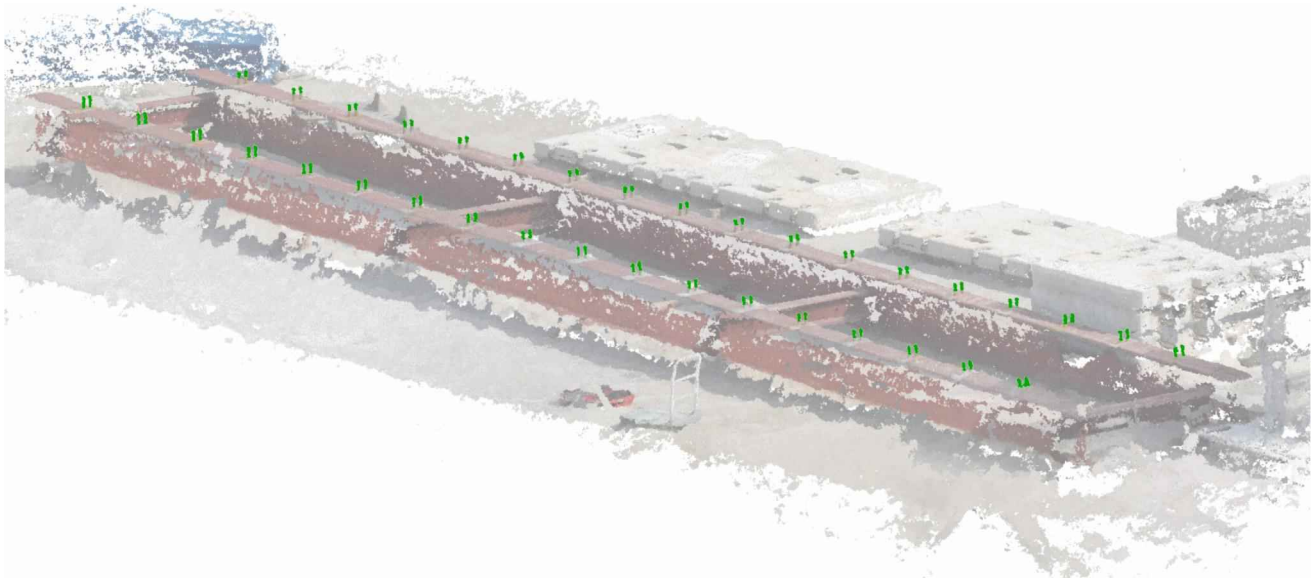


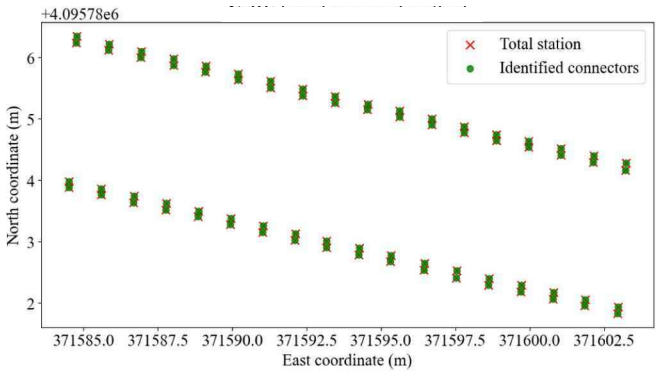
Fig. 17. Identified PCD of the shear connectors (marked as green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

connectors on the girders are successfully identified as shown in Fig. 17. The entire process of the three-step narrowing-down algorithm required approximately 3 min for the PCD containing ten million points.

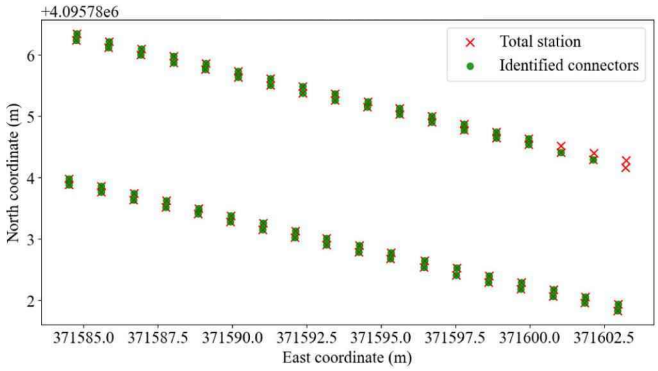
4.2. Comparative analysis between NeRF and SfM

To demonstrate the advantages of the NeRF clearly, we present a comparative analysis of the NeRF and SfM regarding the quality of the reconstructed PCD and its associated influence on the localization results. The reconstruction process required 13 h for the NeRF and 2 h for SfM respectively. The positions of the shear connectors obtained by the proposed method were compared with the total station measurements, as shown in Fig. 18 (a). Similarly, the PCD obtained by the SfM were processed using the three-step narrowing-down algorithm to localize the shear connectors, which were then compared with the total station measurements, as illustrated in Fig. 18 (b). The processing time for the three-step narrowing-down algorithm was approximately 5 min. The NeRF-based method successfully identified all 72 shear connectors. On the other hand, 68 out of the 72 shear connectors were identified using the conventional SfM-based approach. To evaluate the performance, the root mean squared error (RMSE) was computed between the total station measurements and the PCD-based approaches as shown in Fig. 19. Herein, the center of each connector in the east-north plane was extracted from both the NeRF- and SfM-based PCD. The average RMSE values for each shear connector were 0.016 m for the SfM and 0.010 m for the NeRF. The maximum deviation between the total station measurements were 0.037 m for the SfM-based method and 0.039 m for the NeRF-based method. Overall, the NeRF outperformed the SfM in terms of the localization of the shear connectors.

The reason for inaccurate localization in the SfM-based approach is



(a) NeRF-based method (Proposed approach)



(b) SfM-based method (Conventional approach)

Fig. 18. Identified location of shear connectors.

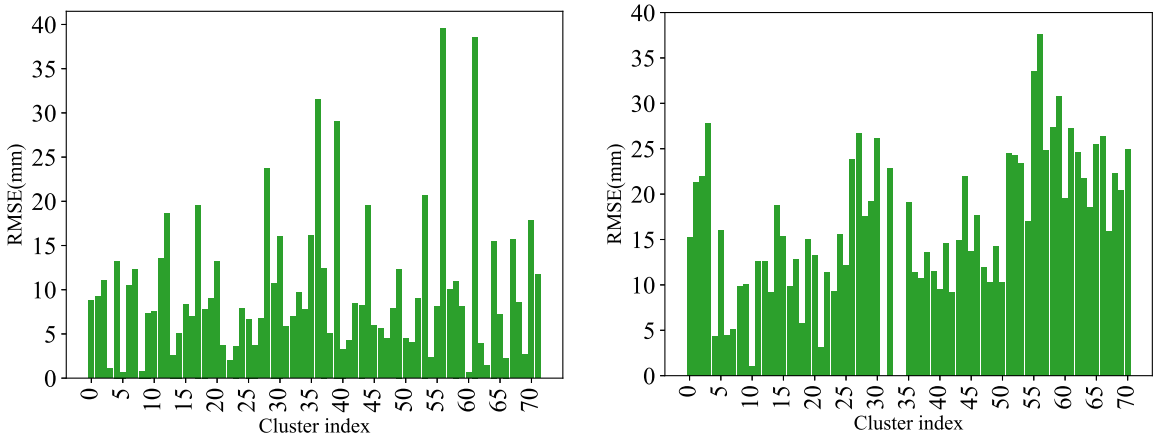
due to the heterogeneous density of the PCD. Owing to the small and thin nature of shear connectors, SfM often fails to ensure a sufficient number of points lie on them. The sparsity of points impedes the applicability of the DBSCAN, which relies on point density for clustering, in identifying point groups for the shear connectors in the three-step narrowing-down algorithm. As a result, the shear connectors are scarcely identified using SfM as shown in Fig. 20 (a). In contrast, the NeRF with the same images yields consistent number of points lying on every shear connector, robustly identifying their positions, as shown in Fig. 20 (b). Indeed, for the automated localization of shear connectors, the NeRF provides robustness when combined with the proposed three-step narrowing-down algorithm. Thus, the proposed method can serve as a foundation for dimensional quality control and collision checks prior to the assembly of girders and deck panels in ABC projects.

5. Conclusion

This paper presented a robust method for the localization of shear connectors in ABC using a UAV-powered NeRF and a three-step narrowing-down algorithm. The proposed method consists of two processes: (1) the reconstruction of a dense PCD by applying the NeRF to UAV-captured images and (2) processing the PCD using a three-step narrowing-down algorithm that automatically localizes shear connectors. The performance of the proposed method was demonstrated using a mock-up prefabricated girder and compared with those of a conventional photogrammetry-based method and total station measurements.

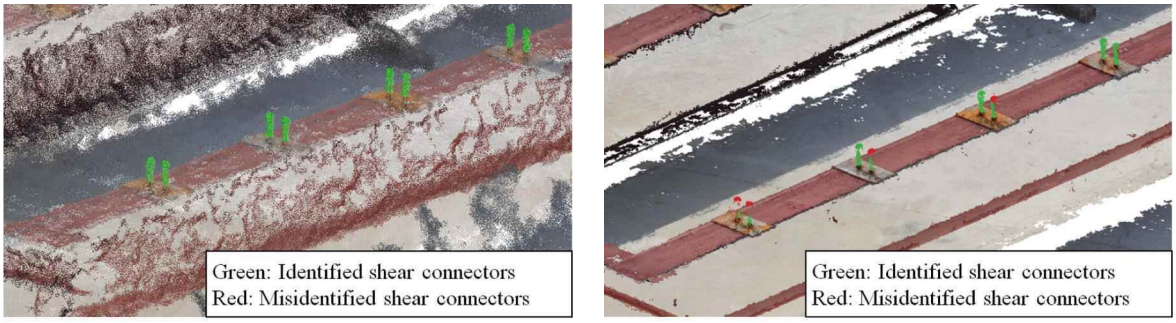
The two main advantages of the proposed method are the robustness offered by the NeRF and the automation facilitated by the proposed three-step narrowing-down algorithm. The NeRF can generate dense and evenly distributed PCD at construction sites when applied to UAV-captured images. A comparison between the results from the NeRF and SfM clearly showed that the dense PCD from the NeRF allowed successful localization, even with insufficient image information, whereas the broken PCD of the traditional SfM led to incomplete identification. Another contributing factor is the automated localization algorithm for shear connectors, which can be executed with the complicated PCD containing a wide variety of objects in at an ABC site. In the field demonstration, the proposed algorithm successfully localized 72 shear connectors, each with a height of 0.15 m, which is considerably small relative to the scale of the construction site. The proposed method offers robustness and efficiency by leveraging the NeRF for dense PCD reconstruction and an automated localization algorithm, thereby outperforming conventional approaches.

However, the proposed method has two main limitations to address. The NeRF requires substantial computational resources for training, along with the necessity to fine-tune hyperparameters such as the near and far planes, batch size, and the number of samples per ray. Thus, a computationally efficient method for 3D reconstruction is required for the wider adoption of the image-based localization of shear connectors for the quality assessment of prefabricated elements in ABC. Moreover, the localization algorithm is inapplicable to girders with arbitrary shapes or those that are tilted, such as in ramps, owing to the assumption that the top flanges of the prefabricated girders are straight and flat. The current three-step narrowing-down algorithm assumes that the top flanges of the girders form a plane, as is commonly observed in most ABC projects, and that the shear connectors are perpendicular to this plane.



(a) NeRF-based method (b) SfM-based method

Fig. 19. Comparison between the total station measurements and each proposed method.



(a) NeRF-based proposed method (b) SfM-based proposed method

Fig. 20. Comparison of resulting shear connectors between the NeRF- and SfM-based methods.

However, in some specialized ABC projects where girders may have complex shapes or shear connectors may not be vertical, the identification of shear connectors could be challenging. To address these limitations, future studies should focus on integrating deep learning-based point cloud segmentation techniques with the current algorithmic structure. This fusion approach has the potential to enhance the efficiency and robustness of shear connector identification, particularly for complex geometries, while maintaining the advantages of the current method.

CRediT authorship contribution statement

Gyumin Lee: Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Ali Turab Asad:** Writing – original draft, Validation, Data curation. **Khurram Shabbir:** Writing – original draft, Validation, Data curation. **Sung-Han Sim:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Data curation, Conceptualization. **Junhwa Lee:** Writing – review & editing, Supervision, Investigation, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This research was conducted with the support of the “National R&D Project for Smart Construction Technology (RS-2020-KA156887)” funded by the Korea Agency for Infrastructure Technology Advancement under the Ministry of Land, Infrastructure and Transport, and managed by the Korea Expressway Corporation.

References

[1] M.P. Culmo, B. Lord, M. Huie, B. Beerman, Accelerated bridge construction: Experience in design, fabrication and erection of prefabricated bridge elements and systems: Final manual (No. FHWA-HIF-12-013). United States. Federal Highway Administration. Office of Bridge Technology, 2011 http://www.fhwa.dot.gov/bridge/abc/docs/abcmanual.pdf [accessed 09 October 2024].

[2] M.P. Culmo, Connection Details for Prefabricated Bridge Elements and Systems (No. FHWA-IF-09-010), United States. Federal Highway Administration. Office of Bridge Technology, 2009. https://www.fhwa.dot.gov/bridge/prefab/if09010/report.pdf (accessed 09 October 2024).

[3] K.A. Byle, N.H. Burns, R. Carrasquillo, Time-Dependent Deformation Behavior of Prestressed High Performance Concrete Bridge Beams, 1997 https://rosap.ntl.bts.gov/view/dot/14821 [accessed 09 October 2024].

[4] R.H. Son, K. Han, Automated model-based 3D scan planning for prefabricated building components, J. Comput. Civ. Eng. 37 (2) (2023) 04022058, https://doi.org/10.1061/(ASCE)CP.1943-5487.0001055.

[5] M.P. Culmo, H. Boyle, S. Nettleton, V. Chandra, M.K. Tadros, J. Mallela, Engineering Design, Fabrication and Erection of Prefabricated Bridge Elements and Systems (No. FHWA-HIF-17-019), United States. Federal Highway Administration, 2013. https://rosap.ntl.bts.gov/view/dot/42879 (accessed 09 October 2024).

[6] American Association of State Highway and Transportation Officials (AASHTO), AASHTO LRFD Bridge Design Specifications, 9th edition, American Association of State Highway and Transportation Officials, Washington, DC, 2020. ISBN: 9781560517382. https://store.transportation.org/Item/CollectionDetail?ID=202 [accessed 09 October 2024].

[7] Precast/Prestressed Concrete Institute, PCI Design Handbook: Precast and Prestressed Concrete, Precast/Prestressed Concrete Institute, Chicago, 1992. ISBN: 9780937040515.

[8] F. Dai, A. Rashidi, I. Brilakis, P. Vela, Comparison of image-based and time-of-flight-based technologies for three-dimensional reconstruction of infrastructure, J. Constr. Eng. Manag. 139 (1) (2013) 69–79, https://doi.org/10.1061/9780784412329.094.

[9] Q. Wang, J.C. Cheng, H. Sohn, Automated estimation of reinforced precast concrete rebar positions using colored laser scan data, Comput. Aided Civ. Inf. Eng. 32 (9) (2017) 787–802, https://doi.org/10.1111/mice.12293.

[10] M.K. Kim, Q. Wang, S. Yoon, H. Sohn, A mirror-aided laser scanning system for geometric quality inspection of side surfaces of precast concrete elements, Measurement 141 (2019) 420–428, https://doi.org/10.1016/j.measurement.2019.04.060.

[11] S. Yoon, Q. Wang, H. Sohn, Optimal placement of precast bridge deck slabs with respect to precast girders using 3D laser scanning, Autom. Constr. 86 (2018) 81–98, https://doi.org/10.1016/j.autcon.2017.11.004.

[12] C.Y. Yi, F. Li, J.P.P. Thedja, S.H. Sim, Y.K. Choi, G.H. Kim, M.K. Kim, A comparison study of edge line estimation algorithms for dimensional quality assessment of precast concrete slabs, Adv. Civil Eng. 2024 (1) (2024) 4166203, https://doi.org/10.1155/2024/4166203.

[13] F. Li, H. Li, M.K. Kim, K.C. Lo, Laser scanning based surface flatness measurement using flat mirrors for enhancing scan coverage range, Remote Sens. 13 (4) (2021) 714, https://doi.org/10.3390/rs13040714.

[14] S. Kiziltas, B. Akinci, E. Ergen, P. Tang, C. Gordon, Technological assessment and process implications of field data capture technologies for construction and facility/infrastructure management, J. Inform. Technol. Const. 13 (10) (2008) 134–154. https://www.itcon.org/2008/10.

[15] F. Remondino, A. Guarnieri, A. Vettore, 3D modeling of close-range objects: photogrammetry or laser scanning?, in: Proceedings of SPIE – The International Society for Optics and Photonics, Videometrics VIII, 5665, Paper 56650M, 2005, https://doi.org/10.1117/12.586294.

[16] L. Klein, N. Li, B. Becerik-Gerber, Imaged-based verification of as-built documentation of operational buildings, Autom. Constr. 21 (2012) 161–171, https://doi.org/10.1016/j.autcon.2011.05.023.

[17] J.L. Schonberger, J.M. Frahm, Structure-from-motion revisited, in: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 4104–4113, https://doi.org/10.1109/CVPR.2016.445.

[18] H. Shakhatreh, A.H. Sawalmeh, A. Al-Fuqaha, Z. Dou, E. Almaita, I. Khalil, N. S. Othman, A. Khreishah, M. Guizani, Unmanned aerial vehicles (UAVs): a survey on civil applications and key research challenges, IEEE Access 7 (2019) 48572–48634, https://doi.org/10.1109/ACCESS.2019.2909530.

[19] M. Goesele, N. Snavely, B. Curless, H. Hoppe, S.M. Seitz, Multi-view stereo for community photo collections, in: Proceedings of 2007 IEEE 11th International Conference on Computer Vision, IEEE, 2007, pp. 1–8, https://doi.org/10.1109/ICCV.2007.4408933.

[20] J. Wu, J. Zhu, J. Zhang, P. Dang, W. Li, Y. Guo, C. Liang, A dynamic holographic modelling method of digital twin scenes for bridge construction, Int. J. Digital

Earth 16 (1) (2023) 2404–2425, https://doi.org/10.1080/17538947.2023.2229792.
[21] B. Mildenhall, P.P. Srinivasan, M. Tancik, J.T. Barron, R. Ramamoorthi, R. Ng, Nerf: representing scenes as neural radiance fields for view synthesis, Commun. ACM 65 (1) (2021) 99–106, https://doi.org/10.48550/arXiv.2003.08934.
[22] B. Triggs, P.F. McLauchlan, R.I. Hartley, A.W. Fitzgibbon, Bundle adjustment—a modern synthesis. Vision algorithms: theory and practice. IWVA 1999, Lect. Notes Comput. Sci 1883 (2000) 298–372, https://doi.org/10.1007/3-540-44480-7_21.
[23] C. Barnes, E. Shechtman, A. Finkelstein, D.B. Goldman, PatchMatch: a randomized correspondence algorithm for structural image editing, ACM Trans. Graph. 28 (3) (2009) 1–11, https://doi.org/10.1145/1531326.1531330. Article 24.
[24] N. Smith, N. Moehrle, M. Goesele, W. Heidrich, Aerial path planning for urban scene reconstruction: a continuous optimization method and benchmark, ACM Trans. Graph. 37 (6) (2018) 1–15, https://doi.org/10.1145/3272127.3275010. Article 183.
[25] H. Fathi, F. Dai, M. Lourakis, Automated as-built 3D reconstruction of civil infrastructure using computer vision: achievements, opportunities, and challenges, Adv. Eng. Inform. 29 (2) (2015) 149–161, https://doi.org/10.1016/j.aei.2015.01.012.
[26] J.A. Morgan, D.J. Brogan, P.A. Nelson, Application of structure-from-motion photogrammetry in laboratory flumes, Geomorphology 276 (2017) 125–143, https://doi.org/10.1016/j.geomorph.2016.10.021.
[27] M. Tancik, E. Weber, E. Ng, R. Li, B. Yi, T. Wang, A. Kristoffersen, J. Austin, K. Salahi, A. Ahuja, D. Mcallister, J. Kerr, Kanazawa A. Nerfstudio: A modular framework for neural radiance field development, in: Proceedings of ACM SIGGRAPH 2023 Conference, 2023, pp. 1–12, https://doi.org/10.1145/3588432.3591516.
[28] T. Müller, A. Evans, C. Schied, A. Keller, Instant neural graphics primitives with a multiresolution hash encoding, ACM Trans. Graph. 41 (4) (2022) 1–15, https://doi.org/10.1145/3528223.3530127.
[29] A. Chen, Z. Xu, A. Geiger, J. Yu, H. Su, TensoRF: Tensorial radiance fields, in: Proceedings of European Conference on Computer Vision, 2022, pp. 333–350, https://doi.org/10.48550/arXiv.2203.09517.
[30] A. Hore, D. Ziou, Image quality metrics: PSNR vs. SSIM, in: Proceedings of 2010 20th International Conference on Pattern Recognition, 2010, pp. 2366–2369, https://doi.org/10.1109/ICPR.2010.579.
[31] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (6) (1981) 381–395, doi:10.1145/358669.358692.
[32] N. Otsu, A threshold selection method from gray-level histograms, Automatica 11 (285–296) (1975) 23–27, https://doi.org/10.1109/TSMC.1979.4310076.
[33] P.E. Sarlin, C. Cadena, R. Siegwart, M. Dymczyk, From coarse to fine: Robust hierarchical localization at large scale, in: Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 12708–12717, https://doi.org/10.1109/CVPR.2019.01300.
[34] J.T. Barron, B. Mildenhall, M. Tancik, P. Hedman, R. Martin-Brualla, P. P. Srinivasan, Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields, in: Proceedings of 2021 IEEE/CVF International Conference on Computer Vision, 2021, pp. 5835–5844, https://doi.org/10.1109/ICCV48922.2021.00580.
[35] Pix4D SA, Pix4Dmapper 4.1 User Manual, Pix4D SA, Lausanne, Switzerland, 2017. https://www.pix4d.com (accessed 09 October 2024).
[36] M. Ester, H.P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, in: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, 1996, pp. 226–231, https://doi.org/10.5555/3001460.3001507.