Measurement 225 (2024) 114034

Contents lists available at ScienceDirect

Measurement

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

A full 3D reconstruction of rail tracks using a camera array

Yizhong Wang, Haochen Liu *, Lichao Yang, Isidro Durazo-Cardenas, Bernadin Namoano, Cheng Zhong, Yifan Zhao *

School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield MK43 0AL, UK

ARTICLE INFO

Keywords:
3D reconstruction
Rail track
Camera array
Non-destructive testing
Sensitivity analysis

ABSTRACT

This research addresses limitations found in existing 3D track reconstruction studies, which often focus solely on specific rail sections or encounter deployment challenges with rolling stock. To address this challenge, we propose an innovative solution: a rolling-stock embedded arch camera array scanning system. The system includes a semi-circumferential focusing vision array, an arch camera holder, and a Computer Numerical Control machine to simulate track traverse. We propose an optimal configuration that balances accuracy, full rail coverage, and modelling efficiency. Sensitivity analysis demonstrates a reconstruction accuracy within 0.4 mm when compared to Lidar-generated ground truth models. Two real-world experiments validate the system's effectiveness following essential data preprocessing. This integrated technique, when combined with rail rolling stocks and robotic maintenance platforms, facilitates swift, unmanned, and highly accurate track reconstruction and surveying.

1. Introduction

McKinney reported that Europe spends €25 billion annually on rail infrastructure maintenance and renewal in 2020 [1]. Although the carrying pressure has reduced due to the epidemic [2], the maintenance cost remains substantial. Fig. 1 illustrates the classification of rail defects, divided into head defects, web defects, and base defects. Head defects, such as head cracks, wall-thinning, and transverse fissures, primarily result from transverse and longitudinal fatigue induced by the rolling stocks [3]. Common web defects include head and web separation, split web, and piped rail. Base defects are categorized as broken-based and base fractures [3]. These defects pose risks of critical rail failure, such as derailment, shutdown, and casualties. According to the 2021 report by Network Rail in the UK, surface fatigue damage accounted for 51 % of the total rail damage, while deformation and wear accounted for 46 % [4]. Non-destructive testing (NDT) technologies, such as ultrasonic [5–8], eddy current testing [9–12], magnetic particle inspection [13,14], acoustic emission [15–17], and thermal imaging [18–22], have been widely employed to detect these defects without causing harm to the rail[5–7,9–12,13,14,15–17,18–21]. Results obtained through current NDT technology are typically presented as 1D signals or 2D images. In contrast, 3D models offer a more intuitive and comprehensive representation. In the realm of digital twins' development for rail maintenance, 3D reconstruction technology could create detailed rail models, incorporating surface and geometric information for remote maintenance and inspection [23]. These models provide a tight and seamless integration between the physical and virtual space.

Various 3D reconstruction technologies have found applications in rail maintenance, including structured light [24,25], time-of-flight [26,27], triangulation [28], monocular vision [29], multi-eye vision [30,31]. These technologies serve as the foundation for remote rail maintenance. Guerrieri *et al.* proposed a mathematical model for reconstructing 3D rail models and measuring transverse and longitudinal rail profiles, validating computer vision as a promising tool for rail inspection [32]. In terms of software applications, Gabara and Sawicki compared the results of capturing pictures of running rails (including double tracks, stones, sleepers, etc.) using two 3D reconstruction software, RealityCapture and PhotoScan [33]. RealityCapture was suggested more suitable for real-world applications due to its faster generation speed, superior visual quality, and higher robustness. Zhang *et al.* proposed a reconstruction process, employing structure-from-motion (SfM) and Multi-view Stereo (MVS) techniques, using five optical cameras to reconstruct the head of a single-row rail, including artificial defects [34]. Comparative analysis with MiniProf, a contact profile surface measurement equipment, showed promising results: the cross-section mean square error (MSE) of the reconstructed model is less than 0.23 mm in the laboratory environment and, outdoors, had a 95 % confidence level that the MSE of the rail head's cross-sectional profile

* Corresponding authors.
E-mail addresses: haochen.liu@cranfield.ac.uk (H. Liu), yifan.zhao@cranfield.ac.uk (Y. Zhao).

https://doi.org/10.1016/j.measurement.2023.114034
Received 27 June 2023; Received in revised form 14 November 2023; Accepted 11 December 2023
Available online 14 December 2023

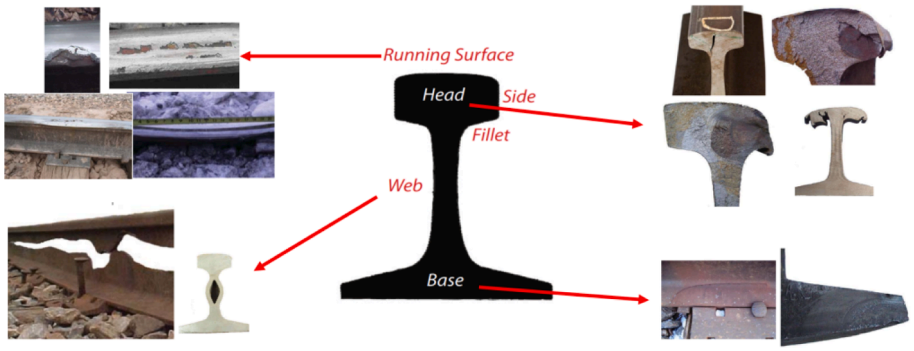


Fig. 1. Cross section and defects of rails [3].

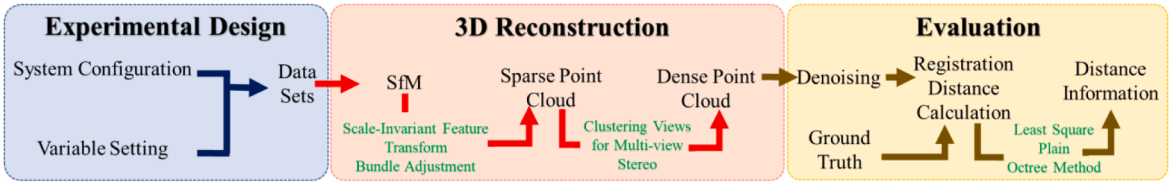


Fig. 2. The proposed workflow of this study.

was less than 0.5 mm with a root mean square error (RMSE) smaller than 0.3 mm. It should be noted that this experiment focused solely on the rail head, while rail damage also occurs on the web and base. Adham *et al.* proposed a 3D reconstruction method using feature correspondence processes and SfM based on videos of the train operating environment [35]. Their findings suggest that running speed of 10 kph, resulting in 90 % or 95 % overlap percentage, were optimal for creating 3D models of train running environments. Although their primary aim was to build 3D models for train running environments, the research output provides valuable insights for reconstructing rails using videos.

As far as we are concerned, existing related studies primarily focused on specific rail sections, such as the head, while sensitivity analysis concerning the relationship between camera layout and 3D reconstruction result remains limited. Bridging this knowledge gap is crucial for designing a lean sensing system that balances complexity and performance, a focal point addressed in this study. The key novelties of this article are as follows:

(1) Introducing a solution for reconstructing the complete 3D profile of rails, enabling a comprehensive surface evaluation rather than focusing on specific parts.
(2) Investigating the impact of camera number and layout, working distance, and moving step interval on 3D rail imaging. Optimizing these parameters lead to more accurate reconstructed results that better represent real-world rails, reducing blind zones and providing a solid foundation for future data fusion of NDT.
(3) Validating the performance of the proposed solution operating in on-track mode within an outdoor environment.

The application of such research can aid engineers in remotely and intuitively observing rails, thereby improving work efficiency, reducing the need for on-site visits, and minimising accident risks. Furthermore, the recorded 3D profile capture rail structure and surface conditions, improving the traceability of historical data and serving as a foundation for future database establishment.

2. Methods

The workflow of the proposed system can be broadly divided into three stages, as shown in Fig. 2, including experiment design, 3D reconstruction, and performance evaluation. The detail of each step is presented below.

2.1. Experiment design

2.1.1. System configuration

The 3D imaging platform is composed of four essential components: an arch sensor holder, a set of cameras, a set of connectors, and a Computer Numerical Control (CNC) scanning machine, as depicted in

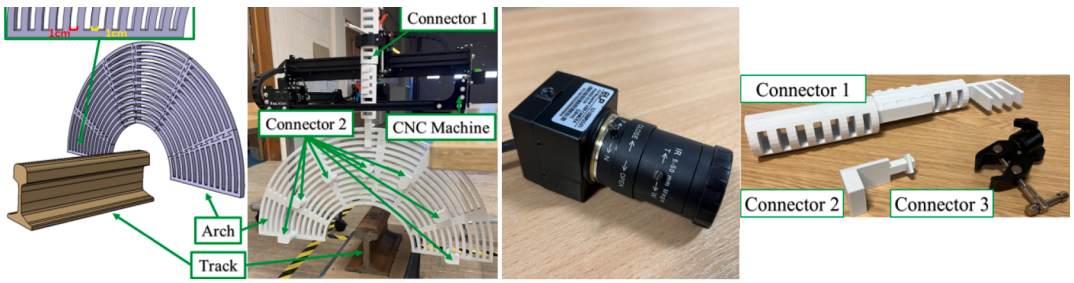


Fig. 3. Experimental setup. (a) system design; (b) a snapshot of the actual system; (c) cameras; (d) connectors (Connector 1: connecting the arch to the CNC machine; Connector 2: connecting the cameras to the arch; Connector 3: connecting the cameras to Connector 2.).

Table 1
The description of Groups.

Group ID	Description	Number of Images
1	38-7-(0-30-60-90-120-150-180)-1	217
2	38-6-(0-30-60-120-150-180)-1	186
3	38-6-(15-45-75-105-135-165)-1	186
4	38-5-(0-60-90-120-180)-1	155
5	38-5-(0-30-90-150-180)-1	155
6	38-5-(30-60-90-120-150)-1	155
7	38-4-(0-60-120-180)-1	124
8	38-4-(30-60-120-150)-1	124
9	34-7-(0-30-60-90-120-150-180)-1	217
10	30-7-(0-30-60-90-120-150-180)-1	217
11	26-7-(0-30-60-90-120-150-180)-1	217
12	38-7-(0-30-60-90-120-150-180)-2	112
13	38-7-(0-30-60-90-120-150-180)-3	77
14	38-7-(0-30-60-90-120-150-180)-4	56
15	38-7-(0-30-60-90-120-150-180)-5	49

Fig. 3. The arch sensor holder serves the purpose of securely fixing and controlling the radial distance and angles of the cameras. Its design allows for a maximum outer radius of 40 cm and a minimum inner radius of 15 cm. Constructed using polylactic acid (PLA) material through 3D printing, the arch is tailored to facilitate camera attachment. Due to space limitation within the 3D printer, the arch is divided into 6 parts, which were produced individually.

The cameras used in the system are 8.0 Megapixel USB CAMERA, operating on DC 5 V power. These cameras offer a spatial resolution of 1600 × 1200 pixels. Their positioning aligns consistently with the radial orientation of the arch, ensuring uniform imaging angles throughout the scanning process. The rail model employed in this study is the 30 cm CEN56E1, which represents the predominant rail type that accounts for approximately 70 % of usage in the UK.

The CNC machine, depicted in Fig. 3(b), plays an important role in facilitating both horizontal movement and vertical fixation of the cameras and the arch. This precision-controlled machine ensures accurate motion and positioning throughout the scanning procedure. For data collection and reconstruction, a computer with the Windows 11 x64, CPU Intel(R) Core(TM) i7-10750H CPU of 2.60 GHz with a memory of 16 GB is employed. In addition, the indoor experiment takes place in a controlled environment with stable lighting conditions and no external natural interference.

2.1.2. Variables settings

To determine the optimal system specifications, the experiment involved testing 15 groups of configurations, as detailed in Table 1. These setups were designed based on variations in working distances, shot intervals, and camera layouts. Each configuration is labelled using the "X-Y-(Z)-L" format. Specifically, "X" represents the working distance (in cm), "Y" denotes the number of employed cameras, "(Z)" indicates the angles of each camera (in°), and "L" denotes the moving step interval (in cm).

2.1.2.1. Camera quantity and layout. The angle setting for the camera positions is established with the left end of the arch serving as the reference point, marked as 0°. Subsequently, the angle between each adjacent camera position is set to 30°. Fig. 4(a) illustrates an example configuration with 7 cameras, positioned at angles of 0°, 30°, 60°, 90°, 120°, 150°, and 180°.

This study conducted tests using 4, 5, 6, and 7 cameras and evaluated their corresponding performance. To maintain a balanced distribution of inputs from the left and right sides, a symmetrical layout is adopted when the number of cameras is even. In cases of an odd number of cameras, an additional camera set at 90° is added in the layout. For example, when using 4 cameras, layout options include configurations such as 0-30-150-180, 0-60-120-180, or 30-60-120-150. In the case of 5 cameras, the above-mentioned layouts are extended to 0-30-90-150-180, 0-60-90-120-180, or 30-60-90-120-150, respectively.

2.1.2.2. Camera working distance and field of view. The working distance of the cameras refers to the distance between the cameras and the centre of the arch circle. In this study, the arch allows for a maximum working distance of 38 cm and a minimum working distance of 20 cm. To investigate the effects of varying working distance on the reconstruction, specific intervals of four centimetres were selected to enhance discrimination. Consequently, the experiment considered four working distance options: 26, 30, 34, and 38 cm.

2.1.2.3. Moving step interval. The moving step interval refers to the distance covered by the arch in the longitudinal direction during each movement, as shown in Fig. 4(b). In this experiment, various moving step interval were examined, namely 1, 2, 3, and 4 cm. It should be noted that due to the rail's length of 30 cm, different step intervals required different numbers of positions to capture data. Specifically, for the moving step interval of 1, 2, 3, and 4 cm, the corresponding numbers of positions to capture data were 31, 16, 11, 8 and 7, respectively. These positions represent discrete locations along the rail where data was collected during the scanning process.

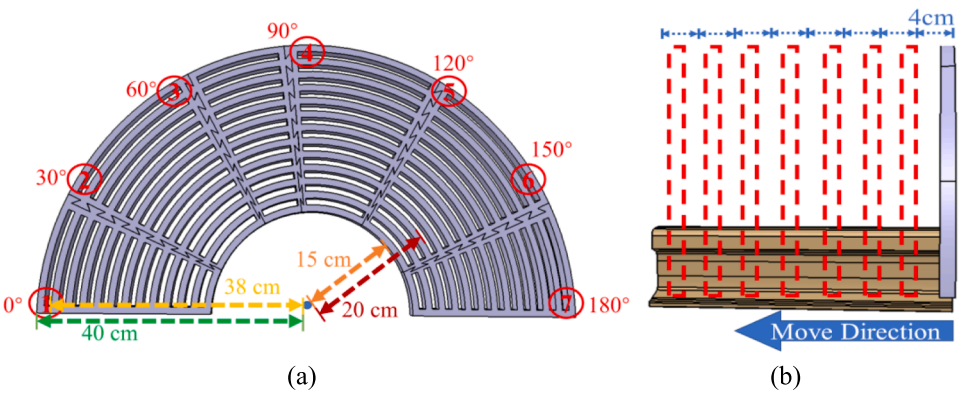


Fig. 4. (a) Group (1) layout (38–7-(0–30-60–90-120–150-180)-1). The blue polka dot represents the centre of the arch; the red circles with numbers illustrate the camera positions; the yellow dotted line indicates the working distance of 38 cm; the green dotted line indicates the maximum outer radius of 40 cm; the orange dotted line denotes the minimum inner radius of 15 cm; the deep red dotted line represents the minimum working distance of 20 cm. (b) an illustration of the moving step interval for Group 14, where the red dotted lines denote the positions of data capture with a distance of 4 cm between each adjacent position, leading to 8 data captures. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.1.3. Data collection

To maintain experimental consistency and minimize interference, each working distance underwent a single data collection process, with the exception of Group (3). The experiment involved the use of a CNC scanning machine and connector to control the movement of 7 cameras arranged in a layout of 0-30-60-90-120-150-180. Images were captured at a moving step interval of 1 cm, resulting in a total of 31 positions for data capture along the rail. Each image was taken after ensuring the CNC machine came to a complete stop, guaranteeing stability during image acquisition. Thus, a dataset consisting of 217 images (7 cameras x 31 positions) was obtained for each working distance. Table 1 illustrates the parameters corresponding to the extracted images.

For Groups 2–11, excluding Group (3), data were selected from the dataset obtained in Group (1). Whenever the working distance changed, the CNC machine returned to the starting position to initiate a repeat of the data capture process. In total, 1054 images were collected, including additional 186 pictures specifically for Group (3).

2.2. 3D models reconstruction

The experiment employed VisualSFM to accomplish the reconstruction of 3D point clouds, which involved a series of steps such as sparse point cloud reconstruction, dense point cloud reconstruction, mesh generation, and texture mapping. These stages were performed to obtain a comprehensive representation of the reconstructed object, including geometric structure and visual appearance.

2.2.1. Sparse point cloud reconstruction

The initial stage of 3D reconstruction involves sparse point cloud reconstruction. This stage utilizes the camera's intrinsic matrix (K) and extrinsic parameters to transform the scene's spatial points into an image point cloud. The extrinsic matrix, comprising a rotation matrix R and a translation matrix t, facilitates the transformations between the camera's field of view capturing the external world and the optical data received by the camera sensor. The intrinsic matrix includes parameters like scale factor, focal length, principal point, skew, and geometric distortion. These parameters enable the conversion of optical information acquired by the camera sensor into a two-dimensional pixel image. Determining these intrinsic and extrinsic parameters allows for the transformation of image pixel coordinates into camera coordinates, and subsequently into world coordinates. We primarily employed the offline structure-from-motion (SfM) approach, which integrates camera internal parameters (K) with extrinsic parameters (R and t) obtained from a collection of unordered pictures. By combining these parameters, a 3D sparse point cloud is constructed, providing an initial representation of the scene's geometry. Equation (1) corresponds the mapping of physical world coordinates in an image to their digital world.

$$x = K[R|t]X \tag{1}$$

where x stands for a pixel coordinate point in the image, X denotes its world coordinate point in the real world. Similarly, this equation implies that each pixel in the acquired images can be associated with a real-world 3D position. In summary, the extrinsic parameters facilitate the World-to-Camera transformation, while intrinsic parameters achieve the Camera-to-Pixel (including Camera-to-Image and Image-to-Pixel) transformation. Combining both transformations allows converting between pixel coordinates and global world coordinates via camera coordinate transformations.

2.2.1.1. Scale-invariant feature transform. The Scale-invariant Feature Transform (SIFT) is employed initially to extract keypoints from two pictures in the SfM method. Subsequently, the keypoints extracted from different images are linked using either the exhaustive technique or the proximity algorithm. To acquire R, t, and a sparse point cloud, the fundamental matrix F and the essential matrix E are solved using Singular Value Decomposition (SVD). The relationship between F and E is written as Eqs.(2) and (3). Thus, the essential information required for reconstruction is derived by the application of SVD.

$$x_1^T F x_2 = 0 \tag{2}$$

$$E = K_1^T F K_2 \tag{3}$$

where x_1, x_2 represent the different pixel positions of one point in different pictures, K_1, K_2 denote their corresponding intrinsic matrix. This method is iteratively executed, accumulating a substantial collection of point clouds comprising the salient features from all images and their interconnections. These feature points are critical for establishing correspondences between images. In other words, their matches imply the same real-world objects have been identified across different pixel coordinate systems. By locating common features across multiple camera views, these points enable linking images through shared scene content. These feature correspondences provide tie points between the separate image spaces, grounded by corresponding physical objects. Establishing these feature matches is thus foundational for unifying distinct image perspectives into a cohesive real-world frame of reference. In essence, the cross-view feature correspondences underpin reconciling the different image coordinate spaces in terms of a unified world coordinate system.

2.2.1.2. Bundle Adjustment. The fusion of individual point clouds is accomplished through Bundle Adjustment (BA), aiming to reduce the disparity between observed and projected image points' locations. According to [36], this study assumes that the presence of n 3D points visible from m different perspectives. In this context, x_{ij} denotes the projection of the i^{th} point on picture j. Additionally, binary variables v_{ij} are introduced, set a value of 1 when point i is visible in picture j, and 0 otherwise. Then, the camera j and each 3D point i are parameterised by vector a_j and a vector b_i, respectively. BA minimises the total reprojection error, taking into account all 3D points and the camera parameters, represented as

$$\min_{a_j,b_i} \sum_{i=1}^{n} \sum_{j=1}^{m} v_{ij} d\left(Q\left(a_j, b_i\right), x_{ij}\right)^2 \tag{4}$$

where $d(x, y)$ denotes for the Euclidean distance between image points represented by vectors x and y, while $Q(a_j, b_i)$ represents the expected projection of point i onto image j. Using accurate R and t obtained by multiple iterations, multiple small point clouds are fused iteratively until the reconstruction of the sparse point cloud is completed. The proposed solution utilizes camera poses and 3D coordinates of measured points as positional parameters, while previously extracted feature point image coordinates serve as known inputs. Through optimization, the technique determines optimal camera parameters along with the global world coordinates of points.

2.2.2. Reconstruction of dense point cloud

To gain dense point clouds, the initial sparse point clouds generated from SfM are further processed by Clustering Views for Multi-view Stereo (CVMS). According to Furukawa *et al.* [37], the CVMS method involves a series of steps, including merging of SfM points, removal of redundant images, enforcement of size constraints, and ensuring sufficient coverage. The overlapping clustering formulation is defined as follows [37]:

$$\min \sum_{k} |C_k| \text{subject to (compactness)} \tag{5}$$

$$\forall k |C_k| \leq \alpha(\text{size}) \tag{6}$$

$$\forall i \frac{\{\#\ of\ \ covered\ \ points\ \ in\ \ I_i\}}{\{\#\ of\ \ points\ \ in\ \ I_i\}} \geq \delta(\text{coverage}) \tag{7}$$

Fig. 5. Built model of (a) Group (10) and (b) the Ground Truth; Failure models of (c) Group (7) and (d) Group (11).

where $\{I_i\}$ represents input images; $\{C_k\}$ denotes overlapping image clusters; α is a constant determined by computational resources, particularly memory limitations; δ is the ratio of covered points. Equation (5) partitions the data into the minimum number of clusters based on the compactness criterion. Equation (6) constrains the size of each cluster to contain fewer points than a set value of points. Concurrently, Equation (7) ensures convergence by requiring substantial overlap between clusters and coverage of all feature points. This clustering process categorizes feature points into distinct but intersecting subsets according to compactness. By fusing multiple images with known camera intrinsic and extrinsic parameters, the approach estimates 3D information to generate a dense point cloud. In summary, the proposed method leverages compactness and overlap constraints to jointly cluster features and reconstruct 3D structure from multi-view calibrated imagery in a unified framework.

2.3. Model evaluation

Before evaluation, the model is imported into MeshLab, and any points that lie outside the model are removed manually to ensure accuracy and consistency.

2.3.1. Ground Truth reconstruction

In this experiment, Ground Truth (GT) is generated by using PolyCam software on an iPhone. A total of 76 images, captured from various angles, directions, and working distances using Lidar, serve as the basis for creating the GT. There are two reasons for choosing this method. First, although SfM has a high accuracy [38], the built models in this experiment lack multi-angle views of pictures, leading to decreased accuracy. Second, the Lidar-based method using a smart phone is recognised for its high accuracy [38].

2.3.2. Registration

For analysis, registration is performed to geometrically align the GT and the produced models (EX). The denoised point cloud and GT are imported into CloudCompare. By matching scales, matching bounding-box centres, and manual translating/rotating, the reconstructed point cloud is adjusted to match the size of the GT. After that, fine registration is performed iteratively until the error between the two models, calculated suing an average of 5×10^5 corresponding points, falls within the threshold of 1×10^{-6}.

2.3.3. Distance calculation

To calculate the distance error, the least square plane method is employed, involving fitting a plane using 12 local adjacent points. The principle is to minimise the parameter S to obtain the closest plane expression, shown in Eq. (8).

$$S = min|f(x,y,z)| \tag{8}$$

$$f(x,y,z) = \sum_{i=0}^{n-1}(a_0{}^*x_i + a_1{}^*y_i + a_2 - z_i)^2 \tag{9}$$

where $a_0 = -\frac{A}{C}, a_1 = -\frac{B}{C}, a_2 = -\frac{D}{C}$, n is the number of input points. Among them, A, B, C, and D are constant making points (x,y,z) satisfy Eq. (10).

$$A^*x + B^*y + C^*z + D = 0\,(C \neq 0) \tag{10}$$

Finally, the Octree method, as the default algorithm, is used to calculate the distance between the GT and the reconstructed model, where the Octree level is set as 10.

It is worth noting that choosing a reference model for distance calculation can yield various results. This study adopts two methods. The first one uses the GT as the reference, comparing it against the reconstructed point cloud to calculate the error. The second method uses the reconstructed point cloud as the reference, comparing it against the GT. This approach aims to identify any missing areas in the produced point cloud and assess the evenness of the point cloud distribution. The formula for calculating the distance error is as follows:

$$\mu_E = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} d_{i,j}}{N_{Compared}} \tag{11}$$

$$\sigma_E = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\left(d_{i,j} - \mu_E\right)^2}{N_{Compared}}} \tag{12}$$

where μ_E represents the mean distance of error; σ_E is the standard deviation of errors corresponding to μ_E; $d_{i,j}$ denotes the minimum distance from the point (i, j) in the compared point cloud to the reference point cloud; $N_{compared}$ is the total number of points in the compared point cloud.

3. Results

A total of 15 groups of experiments were conducted in this study, with 9 of them resulting in successful 3D reconstruction of rail models. The final rendered models exhibit a high degree of similarity to the GT, considering their shape, structure, and surface information, as shown in Fig. 5(a)-(b). However, the remaining 6 groups encountered challenges such as extensive missing areas or deformations, as demonstrated in Fig. 5(c)-(d). To determine the optimal experimental setup, this study investigates the impact of various experiment configurations, including camera layout, camera working distance, and data capture interval, on the precision and accuracy of rail reconstruction.

For the visual and quantitative evaluation of reconstruction performance, the Least Square Plane (LSP) method is proposed to fit the point cloud plane and calculate the error between the reference model and the produced model. The notation EX-GT means using the GT as the reference, which provides insights into the accuracy of the produced point clouds. Conversely, GT-EX means that the EX is used as the reference, allowing for the identification of missing areas in the produced model assuming a comprehensive reconstruction in the reference model.

In the below visualisations of reconstruction errors, a ‘jet’ colour map is used to illustrate the error. The colour scale ranges from blue for small distances to red for larger distances. To highlight the difference, the scale saturation in the compared groups is adjusted to ensure

Table 2
Reconstruction errors of different camera numbers and layouts.

Group	Description	EX-GT Mean (mm)	EX-GT Std. (mm)	GT-EX Mean (mm)	GT-EX Std. (mm)
1	38-7-(0-30-60-90-120-150-180)-1	0.96	0.66	1.29	2.11
2	38-6-(0-30-60-120-150-180)-1	1.30	0.81	1.65	2.53
6	38-5-(30-60-90-120-150)-1	0.56	0.47	1.54	3.53
8	38-4-(30-60-120-150)-1	0.49	0.40	1.70	3.84

consistency.

3.1. Camera quantity and layout

Out of the 8 groups of experiments investigating the impact of the camera number and layout (as shown in Table 2), Groups (1), 2, 6, and 8 successfully reconstructed 3D models, while others (Groups (3), 4, 5, and 7) encountered failures in reconstruction. As a result, it can be concluded that achieving successful reconstruction requires a minimum of four cameras. Among the tested layouts, Group (8), featuring four cameras with a layout of 30-60-120-150, exhibits the leanest configuration, while Group (1), consisting of seven cameras with a layout of 0-30-60-90-120-150-180 represents the most complex setup.

Within these 4 successful groups, it is observed from the EX-GT results that with additional images from other angles (more cameras), resulting in an increased number of feature points, there is an increase in the mean error and volatility of the extracted point clouds. For example, Group (8) with 4 cameras exhibits the lowest mean error of 0.49 mm and standard deviation of 0.40 mm. Conversely, Group (1) with 7 cameras displays the largest standard deviation of error at 0.66 mm and a relatively large mean error of 0.96 mm. This observation can be attributed to the introduction of redundant information when more cameras are involved, which consequently leads to errors in feature registration.

The results of GT-EX suggest a trend where an increased number of cameras contributes to a decrease in the mean and standard deviation of reconstruction error. This indicates enhanced integrity of the produced model or fewer missing areas, such as Area A depicted in Fig. 6. This observation is caused by the reduced missing information due to the increased coverage provided by more cameras. For instance, Group (8) with 4 cameras, demonstrates the largest mean error of 1.70 mm and a standard deviation of errors at 3.84 mm, while Group (1) with 7 cameras exhibits the smallest mean error of 1.29 mm and a standard deviation of 2.11 mm. In comparison to the extracted point clouds in Group (1), the small-scale overlap in Group (8) results in corresponding points with large distances, leading to a large mean error and standard deviation. In addition, the standard deviation of Group (8) is larger than its mean error, which is also caused by the far corresponding points or the low overlap, as evident in Fig. 6(d).

Additionally, experiments with three cameras were also conducted with a systematic layout, including 38-3-(0-90-180), 38-3-(30-90-150), and 38-3-(60-90-120). However, the input images from these experiments did not provide enough details to support a complete model reconstruction.

Thus, in terms of reconstruction effectiveness, it is observed that 4 cameras can build 3D models successfully and efficiently, while the 7-camera layout of Group (1) is recommended to form a complete and

Table 3
Reconstruction error of different camera working distances.

Group	Description	EX-GT Mean (mm)	EX-GT Std. (mm)	GT-EX Mean (mm)	GT-EX Std. (mm)
1	38-7-(0-30-60-90-120-150-180)-1	0.96	0.66	1.29	2.11
9	34-7-(0-30-60-90-120-150-180)-1	0.93	0.72	1.26	2.31
10	30-7-(0-30-60-90-120-150-180)-1	0.38	0.40	0.93	2.47

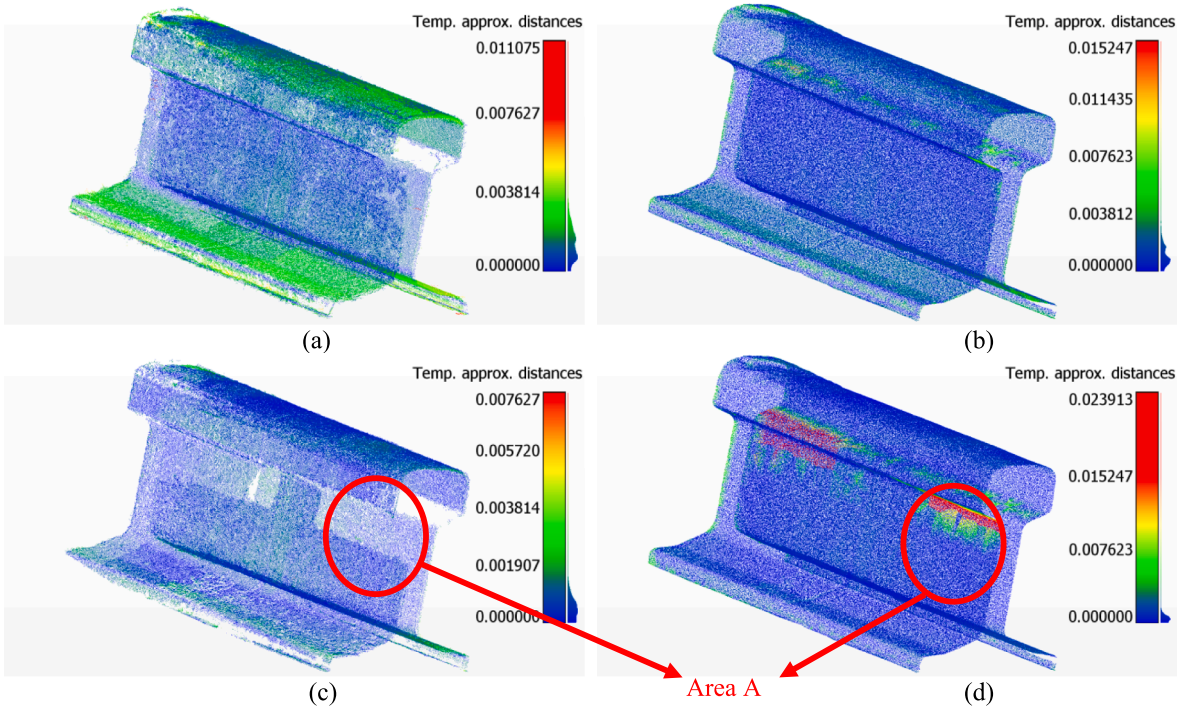


Fig. 6. Visualisation of reconstruction errors (unit: m) for different camera numbers and layouts. (a) EX-GT for Group (1); (b) GT-EX for Group (1); (c) EX-GT for Group (8); (d) GT-EX for Group (8).

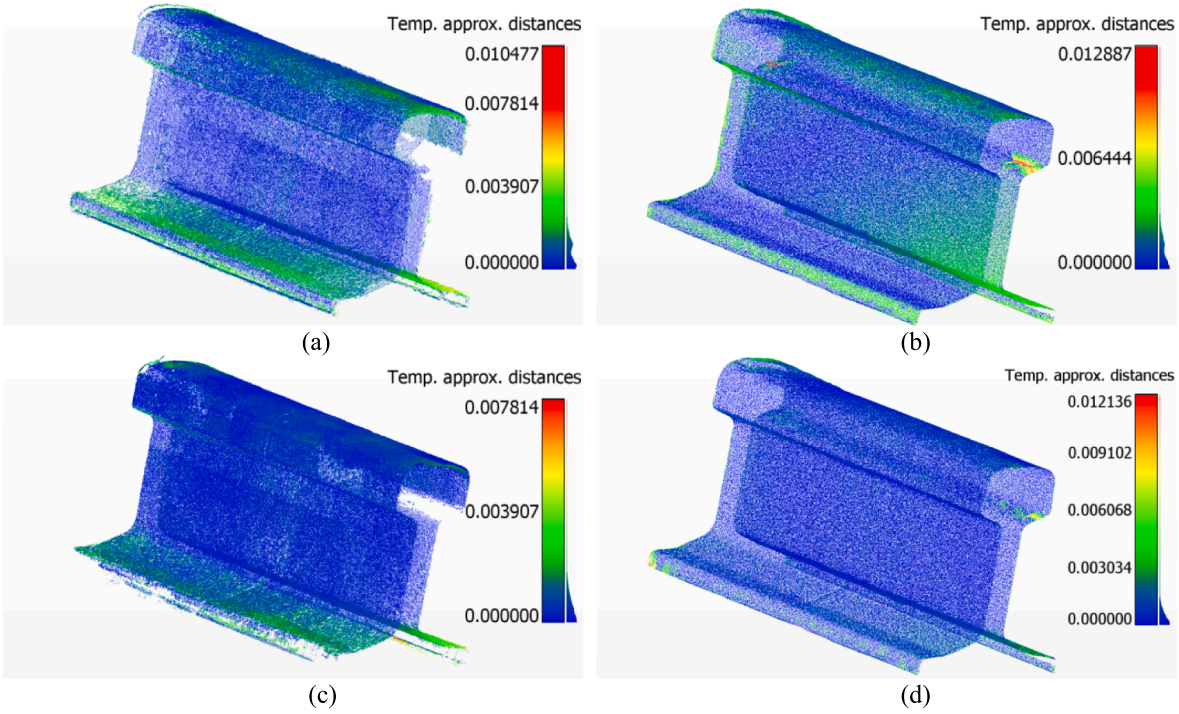


Fig. 7. Visualisation of reconstruction errors (unit: m) for different working distances. (a) EX-GT for Group (9); (b) GT-EX for Group (9); (c) EX-GT for Group (10); (d) GT-EX for Group (10).

Fig. 8. Snapshots of different working distances at the angle of 60°. (a) working distance: 38 cm; (b) working distance: 30 cm.

reliable point cloud with a mean error less than 1 mm.

3.2. Camera working distance

To investigate the effect of camera working distance, four groups of experiment were conducted, and the results are shown in Table 3 and Fig. 7. It is observed that Groups (1), 9, and 10 effectively reconstructed the 3D models, while Group (11) failed. This suggests that rails can be reconstructed within a working distance range of 30 to 38 cm. A working distance shorter than 30 cm is not suggested because the cameras cannot cover the complete view of the target object which leads to poor connections between feature points. A working distance longer than 38 cm is expected to function but may compromise spatial resolution.

The results of EX-GT for these three successful groups show that, as expected, the accuracy and precision of the extracted point clouds increase with shorter working distances due to a higher spatial resolution. It has also been observed that the performance difference between 34 cm and 38 cm is not significant while the performance at 30 cm is improved dramatically in terms of both accuracy and precision.

Additionally, the GT-EX results indicate a positive association between the mean error and the working distance, while showing a negative correlation with the standard deviation. A shorter working distance results in a smaller mean error in GT-EX due to a higher spatial resolution that results in an increased number of produced cloud points; while the standard deviation is larger caused by the unbalanced distribution of point cloud, suggesting more areas with sparse point cloud in the produced model. To further explore the impact of working distance, Figure (a) and Fig. 8(b) show snapshots at working distances of 38 cm and 30 cm, respectively, but with the same angle of 60°. Fig. 8(a) offers a larger field of view (FOV) with a compromised spatial resolution, while Fig. 8(b) has a smaller FOV with improved spatial resolution. The increasing 3D geometric linkages and restrictions are provided with more detailed information, improving the possibility of feature point identification. Numerically, the number of vertices in Group (1) is 5.7×10^5, while that in Group (10) is 9.9×10^5.

Eventually, the working distance is advised to be set between 30 and 38 cm based on the selected cameras to ensure the mean error smaller than 1 mm. Considering the lowest means of errors in both GT-EX and EX-GT for the shortest working distance, 30 cm is suggested in this study.

Table 4
Reconstruction error of different data capture intervals.

Group	Description	EX-GT		GT-EX	
		Mean (mm)	Std. (mm)	Mean (mm)	Std. (mm)
1	38-7-(0-30-60-90-120-150-180)-1	0.96	0.66	1.29	2.11
12	38-7-(0-30-60-90-120-150-180)-2	0.50	0.43	1.06	2.57
13	38-7-(0-30-60-90-120-150-180)-3	0.54	0.45	1.19	2.69
14	38-7-(0-30-60-90-120-150-180)-4	0.52	0.40	1.19	2.49

3.3. Camera scanning step interval

This study examined the effects of five different data capture intervals and the results are shown in Table 4. It is observed that Groups (1), 12, 13, and 14 were successful in reconstructing the 3D models,

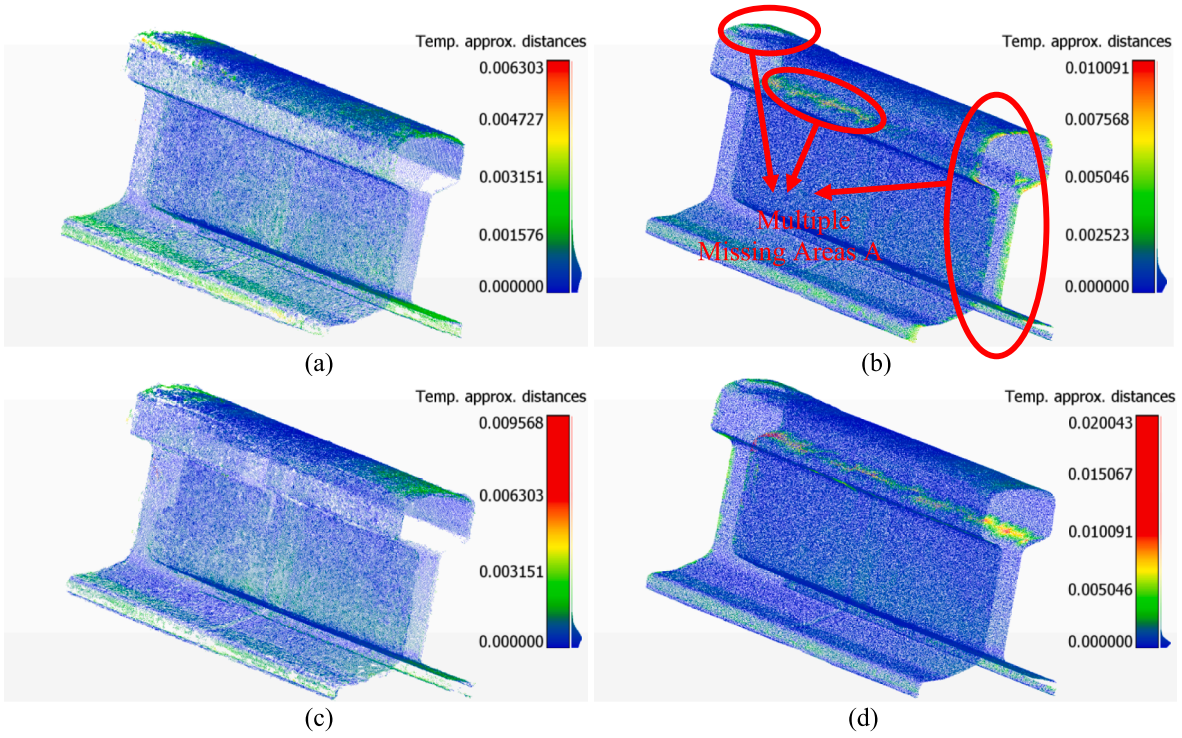


Fig. 9. Visualisation of reconstruction errors (unit: m) for different data capture intervals. (a) EX-GT for Group (12); (b) GT-EX for Group (12); (c) EX-GT for Group 14; (d) GT-EX for Group 14.

while Group 15 failed. A failed reconstruction is expected when the data capture interval is too long because of the incomplete coverage of the object between two adjacent data captures. In this experiment, a data capture interval smaller than or equal to 4 cm is suggested.

Based on these successful experiments, it is found that excessively long and short intervals lead to increases errors. Specifically, Group (1) with an interval of 1 cm exhibits a significantly larger EX-GT mean error of 0.96 mm than that of Group (12) with an interval of 2 cm resulting in a mean error of 0.50 mm. It seems that the significant amount of redundant information in Group (1) compromises the performance of image registration. Furthermore, when the interval is increased from 2 to 4 cm, the mean and standard deviation of errors increase slightly, as visualised in Fig. 9(a) and Fig. 9(c). This is because diminishing relevant feature points arise from decreased image coverage brought on by increasing intervals. In the end, failed reconstruction occurs when the interval is longer than 4 cm since there is an insufficient correlation between the images from two adjacent captures.

The mean of the reconstructed point cloud increases as noisy data decreases from 1 to 2 cm intervals, whereas the mean decreases from 2 to 4 cm due to reduced image coverage. As a result, an interval of 2 cm produced the lowest mean of GT-EX at 1.06 mm. When the interval increases from 1 to 3 cm, a smaller number of available images leads to a decrease in the number of feature points and the connection between feature points, which results in an uneven distribution between point clouds and an increase in the standard deviation. This is likely because there are positions in the point cloud that are sensitive to the interval while there are other positions that are not. For example, with the greatest number of inputs, Group (1) has the lowest standard deviation of 2.11 mm, which leads to a more consistent, dense distribution of local point clouds in the reconstruction results. Group (12) contains half as many input images (1 1 2) as Group (1). As a consequence, the vertices of Group (12) are 3.5×10^5, while that of Group (1) is 4.5×10^5. More vertices make a model comprehensive and reduce its standard deviation. Although Group (1) has more vertices, increasing noise is added, which causes the mean to increase. However, additional images provide extra information for the determination of the spatial position of feature points, resulting in a reduced standard deviation even though Group (1) has more vertices. For instance, when the number of images is further decreased, the input images of Group 14 are only half as many as those of Group (12), and its vertices are 3.2×10^5 in size. The fact that the mean and standard deviation of Group (12), 13, and 14 are not significantly different indicates that the change in point cloud vertices is not yet sensitive to the number of input images. The illustration demonstrates that whereas the blank region of Group 14 is more concentrated in one place, like Area B in Fig. 9(d), Group (12)'s blank area of the point cloud is concentrated in three areas, like Multiple Missing Areas A in Fig. 9(b). Nevertheless, the mean and standard deviation of errors between these four groups is not much different. In this study, an interval of 2 cm is suggested to achieve high integrity and accuracy of the reconstructed model and maintain the mean error within 1 mm.

3.4. Discussion & future work

3.4.1. Movement condition

In the above experiments, the current system implements the data capturing in a manner of static sampling while moving forward step-by-step. The potential influence of camera vibration, travel speed, outdoor conditions and other factors was not considered systematically at this stage. Its primary focus was to investigate how the three key parameters (camera layout, working distance and shot interval) impact 3D reconstruction performance. However, in real-world implementation, the system operates in vibration conditions, posing an inevitable challenge that must be evaluated and addressed.

To further validate the applicability of the approach, we conducted a new experiment mimicking a real-world environment where images were captured while in motion. The experiment employed the camera configuration of Group (10) (seven cameras with a 30 cm working distance) to simultaneously capture images of the targeted rail at two motion-speeds: 0.02 m/s and 0.04 m/s. These speeds were controlled by the CNC machine shown in Fig. 3(b). Under this condition, image blurring, as expected, occurred in the captured images due to the

Table 5
Reconstruction error of different moving speeds using the configuration of Group (10).

Number	Speed (m/s)	EX-GT	
		Mean (mm)	Std. (mm)
1	0.02	0.63	0.68
2	0.04	0.96	0.81

relative movement between the cameras and the object, as well as the associated vibration of cameras. To address this challenge, a motion deblurring algorithm, DeblurGAN [39–41], was applied prior to reconstruction. Repeating the previous evaluation process, the error analysis of the produced 3D models is presented in Table 5 and Fig. 10.

As shown in the Table 5 and Fig. 10, the system can successfully reconstruct the track in 3D under motion. The means errors under both dynamic conditions are between 0.5 mm and 1 mm which are higher than that in the static state (0.38 mm, as shown in Table 3), indicating a degradation in model accuracy due to motion. Additionally, with increasing speed, both the mean and standard deviation of the error rise, which is primarily caused by increased image blurring induced by motion and vibrations. Furthermore, for the high-speed scenario, there are relatively large missing areas due to significant motion blur even after applying a deblur algorithm. Therefore, developing and implementing advanced image deblurring techniques is an important research direction for further improving the performance. Additionally, considering stabilizers or high-quality cameras with high-speed shutters can further reduce the impact of operating conditions on image quality.

3.4.2. Error analysis

To further investigate the reconstruction error of different locations, the point clouds of the rail are divided into three important sections (Head, Web, and Base). Based on the result of Group (10), the average errors of Head, Web and Base are 0.22 mm, 0.21 mm and 0.35 mm respectively. It demonstrates a comparable performance in head construction to the study in [34] but offers additional information including the web and base areas. This prototype is promising to detect head damage during slow-speed travel conditions, such as corrugation on the rail head [42]. For further error reduction, a system with higher camera resolution and framerate, or considering stabilisers will be investigated.

It should be noted that the accuracy of point cloud reconstruction cannot be directly interpreted as the resolution for damage detection of abnormal wear, such as corrugation. But it has some level of potential in identifying severe damage. The 3D model incorporates surface information that captures real attributes including colour and texture. Evaluating the measurement precision of surface damage, such as corrosion or corrugation, requires additional assessment with specific Non-Destructive Testing (NDT) sensor through analysing surface information using image processing or machine learning methods. This paper primarily concentrates on the task of 3D profile reconstruction. In future vision, this rail 3D reconstruction system will be fused with high-accuracy NDT sensors to build a full digital model for rail inspection and structural health monitor. It will be capable to inspect tiny local damage characterisation like corrugation and provide a space overview of damage location and distribution.

In an initial case example shown in Fig. 11, a full reconstruction model with surface condition (Fig. 11(a)) is compared with the Ground Truth (Fig. 11(b)). It demonstrates that a surface white point can be reconstructed and identified effectively.

3.4.3. Realistic on-track experiment

To validate the feasibility under realistic operating conditions, the proposed system was integrated onto the front base of a rail track maintenance platform and tested in an outdoor environment. The robotic maintenance platform, shown in Fig. 12(a), was developed for autonomous road-rail amphibious track inspection and repair [43,44]. Fig. 12(b) depicts the environmental conditions and system setup, from which data were captured while the platform ran along the track. The rail used in the experiment was a 10-meter-long CEN56E1 track, which included other railroad components such as fastener clips, nuts and bolts, sleepers, and rail steel.

In this experiment, the arch and cameras (configured as Group (10) were integrated with the base of the robotic platform, operating at a working distance of 30 cm. The equipment maintained uninterrupted directional movement at the same maximum speed as the indoor test in Section 3.4.2, capturing a 1.5-meter section of railroad track as sample data. Following the same data preprocessing and reconstruction steps described in the previous section, the resulting point cloud and the final rendered model for the 1.5-meter rail section are shown in Fig. 13(a) and (b), respectively. This outcome demonstrates the feasibility of the proposed system operating under an outdoor on-track environment. It should be noted that in this experiment, the quality of the images used for reconstruction was affected not only by motion blur but also by challenging illumination conditions. To address this challenge, it is recommended that future studies consider implementing controllable

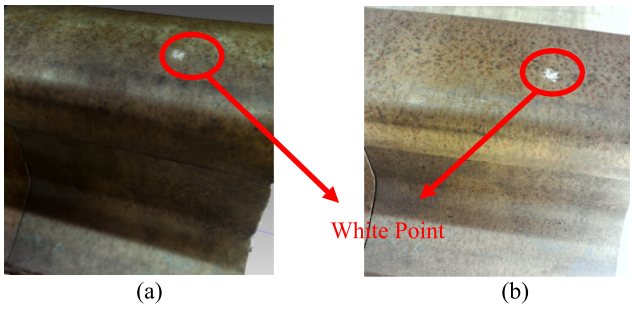


Fig. 11. Comparison of (a) Reconstructed 3D Model and (b) Actual photo at abnormal position.

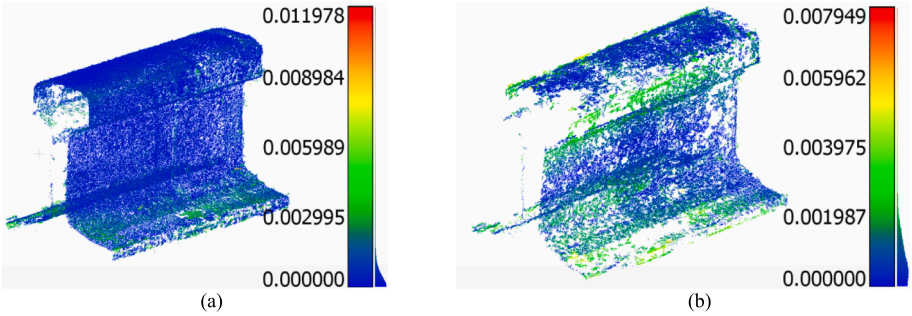
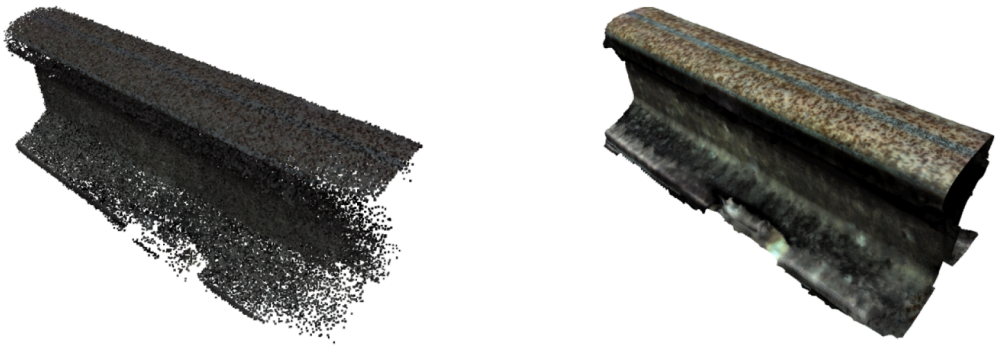


Fig. 10. Visualisation of reconstruction errors (unit: m) for different movement speeds using the configuration of Group (10). (a) EX-GT for the speed of 0.02 m/s; (b) EX-GT for the speed of 0.04 m/s.

(a) (b)

Fig. 12. (a) The robotic platform [43,44]. (b) System integration on the front base trolley.



(a) (b)

Fig. 13. Visualisation of 3D model for the outdoor on-track experiment. (a) the produced point cloud of the 1.5 m track; (b) the reconstructed model after texturing.

illumination through the use of specialised lighting.

3.4.4. Scanning

In its current state, while each group of experiments was conducted with a fixed interval, it is worth noting that photos captured at varying intervals can be combined for 3D reconstruction in practical situations. The experimental results prove that the travel distance between two shots should not exceed 4 cm. In other words, even if the speed fluctuates, the shooting frequency of cameras can be automatically adjusted based on the travel speed to satisfy this requirement. The maximum permissible travel speed will be determined by the sample rate of the cameras.

The instrument is expected to be mounted on the bottom of the specific inspection trains (e.g. NetworkRail's New Measurement Train (NMT) [45]) to achieve periodic condition inspection. Additionally, it can be integrated on the front-case of track based robotic maintenance systems (e.g., the robotic railway maintenance platform system [43]) to enable specific inspection tasks in track rail maintenance.

4. Conclusions

This paper presents a computer vision-based solution for a full 3D reconstruction of track rails. An arch layout camera array fusion scanning system is designed and optimised to achieve the structural and exterior information reconstruction of CEN56E1 standard steel rails. An experimental prototype system was established, and 15 groups of experiments were conducted to investigate the influence of the camera quantity and layout, field of view, and scanning step interval for the optimal 3D reconstruction quality. Through the sensitivity analysis using the EX-GT and GT-EX point cloud evaluation methods, the recommended specification was concluded that the camera number is 7 cameras, and the layout is 0°, 30°, 60°, 90°, 120°, 150°, and 180°; the recommended working distance and interval are 30 cm and 2 cm, respectively. The results demonstrate the proposed method can achieve a reconstruction accuracy within 0.4 mm compared to a ground truth reconstructed model by Lidar. A set of further experiments also validated that the system could operate effectively under outdoor conditions.

Future work will focus on the development of new 3D reconstruction processes, dynamic monitoring, noise processing, quantification of interaction factors, and the development of automated integrated reconstruction processes with rolling stock vehicles. Additionally, forthcoming work will take into account the real-world NDT application conditions, considering factors such as lighting, debris, and damage, etc.

CRediT authorship contribution statement

Yizhong Wang: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Haochen Liu:** Writing – original draft, Writing – review & editing. **Lichao Yang:** Validation, Writing – review & editing. **Isidro Durazo-Cardenas:** Supervision, Writing – review & editing. **Bernadin Namoano:** Supervision, Writing – review & editing. **Chen Zhong:** Formal analysis, Methodology. **Yifan Zhao:** Conceptualization, Supervision, Funding acquisition, Project administration, Resources, Writing – review & editing.

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.

References

[1] B. C. Kienzler, C. Lotz, S. Stern, "Using analytics to get European rail maintenance on track," 2020. [Online]. Available: https://www.mckinsey.com/industries/public-sector/our-insights/using-analytics-to-get-european-rail-maintenance-on-track#/.

[2] L. Downey, A. Fonzone, G. Fountas, and T. Semple, "The impact of COVID-19 on future public transport use in Scotland," *Transp Res Part A Policy Pract*, vol. 163, no. October 2021, pp. 338–352, 2022, doi: 10.1016/j.tra.2022.06.005.

[3] Nordco, Rail defects. https://www.nordco.com/Media/Assets/General-Files/NordcoRailFlawDefectsIdentificationHandbook.pdf, (Accessed 14 November 2023).

[4] B. Whitney, "PWI Lunch and Learn - March 2022 Rail Management Hot Topics and Developments Broken Rails since 2010 moving average 50 by 30," 2022 [Online]. Available: https://www.thepwi.org/knowledge/pwi-lunch-learn-rail-management-hot-topics-developments/.

[5] J. Kaščak, J. Török, M. Töröková, Utilization of the Ultrasonic Diagnostic Method in Rail Status on a Defined Railway Section, TEM Journal 10 (1) (2021) 152–157, https://doi.org/10.18421/TEM101-18.

[6] F. Wu, X. Xie, J. Guo, Q. Li, Internal Defects Detection Method of the Railway Track Based on Generalization Features Cluster Under Ultrasonic Images, Chin. J. Mech. Eng. 35 (1) (Dec. 2022) 59, https://doi.org/10.1186/s10033-022-00726-z.

[7] C.W. Yao, An ultrasonic method for 3D reconstruction of surface topography, J. Phys. Commun. 2 (5) (May 2018), 055034, https://doi.org/10.1088/2399-6528/aac691.

[8] Y. Jiang, S. Chen, K. Wang, W. Liao, H. Wang, Q. Zhang, Quantitative detection of rail head internal hole defects based on laser ultrasonic bulk wave and optimized variational mode decomposition algorithm, Measurement 218 (Aug. 2023), 113185, https://doi.org/10.1016/j.measurement.2023.113185.

[9] P. Xu, H. Zeng, T. Qian, L. Liu, Research on defect detection of high-speed rail based on multi-frequency excitation composite electromagnetic method, Measurement 187 (2022) 110351.

[10] L.U. Daura, G.Y. Tian, Characterization of Angular RCF Cracks in a Railway Using Modified Topology of WPT-Based Eddy Current Testing, IEEE Trans Industr Inform 19 (4) (Apr. 2023) 5612–5622, https://doi.org/10.1109/TII.2022.3201587.

[11] J. Wang, Q. Dai, P. Lautala, H. Yao, R. Si, Rail Sample Laboratory Evaluation of Eddy Current Rail Inspection Sustainable System, Sustainability 14 (18) (Sep. 2022) 11568, https://doi.org/10.3390/su141811568.

[12] G.Y. Dymkin, A.V. Kurkov, Y.G. Smorodinskii, A.V. Shevelev, On the Sensitivity of Eddy Current Testing of Parts of Railway Rolling Stock, Russ. J. Nondestr. Test. 55 (8) (2019) 610–616, https://doi.org/10.1134/S1061830919080059.

[13] J.-H. Ye, R.-H. Ni, Q.-C. Hsu, Image feature analysis for magnetic particle inspection of forging defects, Proc. Inst. Mech. Eng. B J. Eng. Manuf. 236 (14) (Dec. 2022) 1923–1929, https://doi.org/10.1177/09544054211014443.

[14] Y. Yang, Y. Yang, L. Li, C. Chen, Z. Min, Automatic Defect Identification Method for Magnetic Particle Inspection of Bearing Rings Based on Visual Characteristics and High-Level Features, Appl. Sci. 12 (3) (Jan. 2022) 1293, https://doi.org/10.3390/app12031293.

[15] P. Deshpande, V. Pandiyan, B. Meylan, K. Wasmer, Acoustic emission and machine learning based classification of wear generated using a pin-on-disc tribometer equipped with a digital holographic microscope, Wear 476 (2021), 203622, https://doi.org/10.1016/j.wear.2021.203622.

[16] W. Suwansin and P. Phasukkit, "Deep Learning-Based Acoustic Emission Scheme for Rail Crack Monitoring," in *2021 16th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP)*, IEEE, Dec. 2021, pp. 1–5. doi: 10.1109/iSAI-NLP54397.2021.9678162.

[17] M.S.A. Hashmi, M. Ibrahim, I.S. Bajwa, H.-U.-R. Siddiqui, F. Rustam, E. Lee, I. Ashraf, Railway Track Inspection Using Deep Learning Based on Audio to Spectrogram Conversion: An on-the-Fly Approach, Sensors 22 (5) (Mar. 2022) 1983.

[18] R. Goel, A. Sharma, and R. Kapoor, *An Efficient Object and Railway Track Recognition in Thermal Images Using Deep Learning*, vol. 841. Springer Singapore, 2022. doi: 10.1007/978-981-16-8774-7_20.

[19] X. Chen, G. Tian, S. Ding, J. Wu, Investigation of skewness feature for evaluation of defects using eddy current pulsed thermography, IEEE Sens. J. 19 (24) (2019) 12118–12125, https://doi.org/10.1109/JSEN.2019.2936221.

[20] X. Chen, G. Tian, S. Ding, J. Ahmed, W.L. Woo, Tomographic Reconstruction of Rolling Contact Fatigues in Rails Using 3D Eddy Current Pulsed Thermography, IEEE Sens. J. 21 (17) (2021) 18488–18496, https://doi.org/10.1109/JSEN.2021.3086307.

[21] J. Zhu, G. Tiany, Q. Min, J. Wu, Comparison Study of Different Features for Pocket Length Quantification of Angular Defects Using Eddy Current Pulsed Thermography, IEEE Trans. Instrum. Meas. 68 (5) (2019) 1373–1381, https://doi.org/10.1109/TIM.2018.2890053.

[22] A. Glowacz, Fault diagnosis of electric impact drills using thermal imaging, Measurement 171 (Feb. 2021), 108815, https://doi.org/10.1016/J.MEASUREMENT.2020.108815.

[23] S. Feroz and S. Abu Dabous, "UAV-Based Remote Sensing Applications for Bridge Condition Assessment," *Remote Sens (Basel)*, vol. 13, no. 9, p. 1809, May 2021, doi: 10.3390/rs13091809.

[24] S. Barone, P. Neri, A. Paoli, A.V. Razionale, 3D acquisition and stereo-camera calibration by active devices: A unique structured light encoding framework, Opt. Lasers Eng. 127 (Apr. 2020), 105989, https://doi.org/10.1016/j.optlaseng.2019.105989.

[25] S. Makhsous, H.M. Mohammad, J.M. Schenk, A.V. Mamishev, A.R. Kristal, A Novel Mobile Structured Light System in Food 3D Reconstruction and Volume Estimation, Sensors 19 (3) (Jan. 2019) 564, https://doi.org/10.3390/s19030564.

[26] C.Y. Chiu, M. Thelwell, T. Senior, S. Choppin, J. Hart, J. Wheat, Comparison of depth cameras for three-dimensional reconstruction in medicine, Proc. Inst. Mech. Eng. H 233 (9) (2019) 938–947, https://doi.org/10.1177/0954411919859922.

[27] J.H. Nam, E. Brandt, S. Bauer, X. Liu, M. Renna, A. Tosi, E. Sifakis, A. Velten, Low-latency time-of-flight non-line-of-sight imaging at 5 frames per second, Nat. Commun. 12 (1) (Nov. 2021).

[28] S. Henry, J.A. Christian, Absolute Triangulation Algorithms for Space Exploration, J. Guid. Control Dynam. 46 (1) (2023) 21–46, https://doi.org/10.2514/1.G006989.

[29] B. Xu, C. Liu, A 3D reconstruction method for buildings based on monocular vision, Comput. Aided Civ. Inf. Eng. 37 (3) (2022) 354–369, https://doi.org/10.1111/mice.12715.

[30] X. Lin, J. Wang, and C. Lin, "Research on 3D Reconstruction in Binocular Stereo Vision Based on Feature Point Matching Method," in *2020 IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE)*, IEEE, Sep. 2020, pp. 551–556. doi: 10.1109/ICISCAE51034.2020.9236889.

[31] S. Chen, Z. Xiang, N. Zou, Y. Chen, C. Qiao, Multi-stereo 3D reconstruction with a single-camera multi-mirror catadioptric system, Meas. Sci. Technol. 31 (1) (Jan. 2020), 015102, https://doi.org/10.1088/1361-6501/ab3be4.

[32] M. Guerrieri, G. Parla, and C. Celauro, "Digital image analysis technique for measuring railway track defects and ballast gradation," *Measurement (Lond)*, vol. 113, no. August 2017, pp. 137–147, 2018, doi: 10.1016/j.measurement.2017.08.040.

[33] G. Gabara, P. Sawicki, A New Approach for Inspection of Selected Geometric Parameters of a Railway Track Using Image-Based Point Clouds, Sensors 18 (3) (Mar. 2018) 791, https://doi.org/10.3390/s18030791.

[34] D. Zhang, S.N. Lingamanaik, H. Chung, Image-based 3D reconstruction for rail profile measurement, Proc. Inst. Mech. Eng. F J. Rail Rapid Transit 237 (3) (Mar. 2023) 309–321, https://doi.org/10.1177/09544097221110322.

[35] A. Mahmoud, M. Mohamed G., and E. S. Adel, "Railway Tracks Detection of Railways Based On Computer Vision Technique and GNSS Data," in *International Conference on Civil, Structural and Transportation Engineering*, Nov. 2020, pp. 269-1-269–8. doi: 10.11159/iccste20.269.

[36] R. Hartley, A. Zisserman (Eds.), Multiple View Geometry in Computer Vision, Cambridge University Press, 2004.

[37] Y. Furukawa, B. Curless, S. M. Seitz, and R. Szeliski, "Towards Internet-scale multi-view stereo," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2010, pp. 1434–1441. doi: 10.1109/CVPR.2010.5539802.

[38] Z. Alijani, J. Meloche, A. McLaren, J. Lindsay, A. Roy, A. Berg, A comparison of three surface roughness characterization techniques: photogrammetry, pin profiler, and smartphone-based LiDAR, Int. J. Digit Earth 15 (1) (Dec. 2022) 2422–2439, https://doi.org/10.1080/17538947.2022.2160842.

[39] O. Kupyn, T. Martyniuk, J. Wu, and Z. Wang, "DeblurGAN-v2 : Deblurring (Orders-of-Magnitude) Faster and Better," *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 8878–8887, 2019, doi: https://doi.org/10.48550/arXiv.1908.03826.

[40] S. Nah, T. H. Kim, and K. M. Lee, "Deep Multi-scale Convolutional Neural Network for Dynamic Scene Deblurring," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 257–265. doi: 10.1109/CVPR.2017.35.

[41] W.-Z. Shao, Y.-Y. Liu, L.-Y. Ye, L.-Q. Wang, Q.i. Ge, B.-K. Bao, H.-B. Li, DeblurGAN+: Revisiting blind motion deblurring using conditional adversarial networks, Signal Process. 168 (2020) 107338.

[42] Q. Xie, G. Tao, B. He, Z. Wen, Rail corrugation detection using one-dimensional convolution neural network and data-driven method, Measurement 200 (2022) 111624.

[43] H. Liu, M. Rahman, M. Rahimi, A. Starr, I. Durazo-Cardenas, C. Ruiz-Carcel, A. Ompusunggu, A. Hall, R. Anderson, An autonomous rail-road amphibious robotic system for railway maintenance using sensor fusion and mobile manipulator, Comput. Electr. Eng. 110 (2023) 108874.

[44] M. Rahman, H. Liu, M. Masri, I. Durazo-Cardenas, A. Starr, A railway track reconstruction method using robotic vision on a mobile manipulator: A proposed strategy, Comput Ind 148 (Jun. 2023), 103900, https://doi.org/10.1016/j.compind.2023.103900.

[45] "New Measurement Train (NMT) - Network Rail." 2022. [Online].

2023-12-14

A full 3D reconstruction of rail tracks using a camera array

Wang, Yizhong

Elsevier

Wang Y, Liu H, Yang L, et al., (2024) A full 3D reconstruction of rail tracks using a camera array. Measurement, Volume 225, February 2024, Article Number 114034
https://doi.org/10.1016/j.measurement.2023.114034