IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 28 July 2023, accepted 29 August 2023, date of publication 1 September 2023, date of current version 7 September 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3311017*

RESEARCH ARTICLE

# Defect Detection and Recognition of Crane Girder Structure Based on 3D Reconstruction of UAV

**ZHIPING LIU[1], YANNAN YU[1], ZHUOHUI LIANG[1,2], GUODONG HAN[1], AND YAO LU[1]**
[1]School of Transportation and Logistics Engineering, Wuhan University of Technology, Wuhan 430063, China
[2]Guangdong Special Equipment Inspection and Research Institute, Foshan 528251, China

Corresponding author: Zhuohui Liang (2859463014@qq.com)

This work was supported in part by the National Key Research and Development Program of China under Grant 2022YFB2602300 and Grant 2022YFB2602303, and in part by the Science and Technology Plan Project of Administration of National Market Supervision under Grant 2021MK043 and Grant 2022MK042.

**ABSTRACT** The traditional detection process of deformation or crack of crane main girder structure is tedious, inefficient and difficult to control safety risk. In this work, combining unmanned aerial vehicle (UAV) images with photogrammetry technology, an efficient and high-precision method for crane main girder defect detection and identification based on UAV 3D reconstruction is proposed. Firstly, the flight path, suitable for the chamfered rectangular section is selected, according to the structural characteristics of the main girder and the flight requirements of the UAV photography, and the calculation equation of flight strategy parameters is deduced. Secondly, based on 3D reconstruction methods and crucial techniques, the influence of image number, marker points and structure size effects on the accuracy of the 3D reconstruction model are investigated, separately, thus comprehensively verifying the accuracy of the 3D reconstructed model. Finally, the proposed method is tested and evaluated by the experiments for deformation detection of crane main girder structure and crack identification of scaled carton, obtaining millimeter-level accurate results, and the high precision deformation detection and crack identification using non-contact 3D reconstruction are realized. The results demonstrate that the combination of UAV and photogrammetry show potential in the deformation detection of crane main girder structure, it also has a prospect in accurately evaluating the geometry of cracks and helping structural engineers to assess the health status of structures.

**INDEX TERMS** Crane, UAV, 3D reconstruction, point cloud model, deformation, crack identification.

## I. INTRODUCTION

Crane, common special equipment in engineering construction, is employed in work scenarios such as lifting and moving heavy goods. The main girder structure, as the critical load-bearing structure, is subject to structural damage such as significant deformation and crack initiation and expansion, which can lead to catastrophic safety failures if the deformation and initial cracking that occurs in the early stages of structural performance degradation is not detected in time. The traditional crane structure detection methods, such as pulling the wire method and level method, belong to contact detection, which requires manual climbing with danger, and manual reading of the measurement error is bigger.

The associate editor coordinating the review of this manuscript and approving it for publication was Jiankang Zhang.

In addition, optical detection methods, such as theodolite [1], total stations [2] and laser rangefinders [3], can only measure a single point, which is too inefficient and cumbersome to meet the demand for convenient, efficient and automated real-time monitoring of crane safety detection. Therefore, it is critical to find a large area, non-contact and efficient detection approach for crane main girder structure health monitoring [4].

At present, camera and computer vision-based sensors have emerged as promising tools for the non-contact remote measurement of structural responses [5]. With the increase in camera resolution and computing power, vision-based technology is considered an efficient method [6] in structural health monitoring, where deep learning algorithms such as convolutional neural networks enable real-time structural damage detection of concrete [7], steel cracks [8] and

masonry historic walls [9] based on visual data including images and videos, to carry out real-time structural safety inspections and prevent catastrophic accidents [10]. Whereas detection accuracy and model robustness require substantial samples with real-world possible forms of damage to be added to the training database. This method, despite its shortcomings, has demonstrated high accuracy and detection performance and is the direction of the future.

Image-based 3D reconstruction technology has become a trend in structural inspection in the 21st century. UAVs equipped with cameras are suitable for foundation integrity assessment and damage detection [11] of large structures due to their ability to generate high-resolution structural photographs and record real-time structural surface features, and have been widely used in the structural inspection of ancient towers [12], terrain mapping [13], structural modeling [14] of civil and bridge structures, and power inspections [15]. In the aspect of structural deformation detection, Massimo et al. proposed digital photogrammetry combined with the reverse modeling method to analyze the deformation of ancient ships [16]. Singh et al. used computer vision technology to process UAV shooting data to determine the health status of the track [17]. Similarly, Zhou et al. explored image-based 3D reconstruction for damage assessment of residential buildings after hurricanes and compared the generated point clouds with moving light detection and distance data [18]. In terms of structural surface crack detection, Liu et al. proposed a method combining 2D image processing and 3D scene reconstruction to obtain 3D location information of crack edges on the surface of concrete bridge piers [19]. Ioli et al. present a method for precise assessment of crack geometry based on UAV photogrammetry to accurately estimate, through crack 3D reconstruction, the crack width [20]. Cha et al first presented a Faster Region-based Convolutional Neural Network (Faster R-CNN) for the quasi-real-time simultaneous detection of multiple damage types [8]. Furthermore, Ali et al used a modified Faster R-CNN with ResNet-101 as the base network combined with an autonomous UAV was developed for the automatic detection of multiple types of structural damage, and also offered a solution called UBS system for the autonomous navigation of a UAV in a GPS-denied area, which greatly facilitates infrastructure monitoring and inspection based on autonomous UAVs [21], and following on from this, Zhou et al applied the Faster R-CNN for crane surface crack image recognition captured by UAVs [22], but it failed accomplish the detection of structural deformation. In all of these deep learning algorithms based structural surface crack detection, hundreds or even thousands of photographs are required for training, whereas defect recognition using images to build 3D point cloud models requires fewer photographs, and the photography workload is mitigated.

From the above study, it is clear that UAV-based 3D reconstruction technology can detect not only the deformation of the structure but also the cracks on the surface of the structure. However, model accuracy is an extremely sensitive factor in the reliability of structural defect detection results. The number of images taken by the UAV, the number of pre-set control points and their distribution position, determine the accuracy of the 3D point cloud model. Among these sensitive factors, the number and location of ground control points are particularly significant [23], [24], and diverse combinations are exploited to optimize the accuracy of UAV photogrammetry. Zhao et al. studied the influence of different numbers of marker points on the accuracy of the established dam model, characterizing the model accuracy in terms of point cloud distance errors in the three directions of X, Y and Z axes to guarantee the reliability of 3D model accuracy [25]. Al-Halbounia et al. proved that setting ground control points at the edge of the measured area is favorable for acquiring more accurate topographic information [26]. Additionally, Image registration and acquisition of more redundant image information are also approaches to improve model accuracy. Li et al. proposed a non-rigid registration method, NRLI-UAV, for the registration of images collected by low-cost UAV systems, which also aims to improve the quality of LiDAR point cloud models [27]. He et al. chose to use a multi-camera system from the point of view of acquiring redundant data and producing real scale models, for which a three-point closed solution was provided to obtain a more accurate indoor 3D model [28]. Moreover, high image overlap is beneficial for model accuracy [29]. Therefore, the optimal UAV photography strategy applicable to the crane main girder structure is required to be investigated to warrant the 3D model accuracy.

Currently, few studies have been conducted to simultaneously detect both deformation and cracking defects in large mechanical structure, especially crane steel structures. In this study, a defect detection method based on UAV 3D reconstruction is proposed to achieve non-contact high-precision detection and identification of the main girder structure. Firstly, the overlap calculation formula, which applies to the chamfered rectangular section structure, is derived, and the UAV flight strategy is customized to gain sequential image sequences of the main girder structure. Secondly, on the basis of the 2D images of the main girder, the point cloud model of the main girder structure is established by combining the 3D reconstruction theory and crucial technologies. Then, the model accuracy analysis process is raised, combined with the M3C2 algorithm [30] and error evaluation indexes, to analyze the influence of image overlap, working distance, number and location of marker points, and structural size-effect on the point cloud model accuracy, respectively. Finally, the accuracy and effectiveness of the proposed method are verified with experiments.

The organizational structure of this study is shown in Figure 1, and the remainder of this study is organized as follows: Sequential image sequences acquired from UAV are conducted in Section II. Image-based 3D reconstruction approach is described in detail in Section III. Section IV

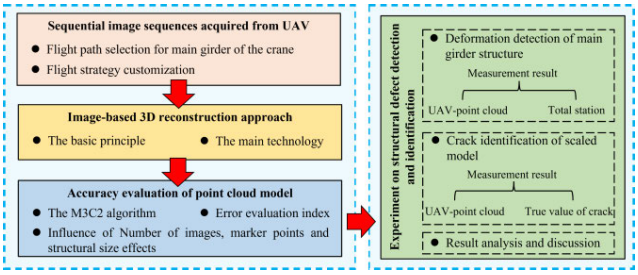



**FIGURE 1.** Methodology framework of this study.

conducts a point cloud model accuracy evaluation. Experiment on structural defect detection and identification of crane main girder in Section V. Finally, Section VI discusses the findings and concludes the study.

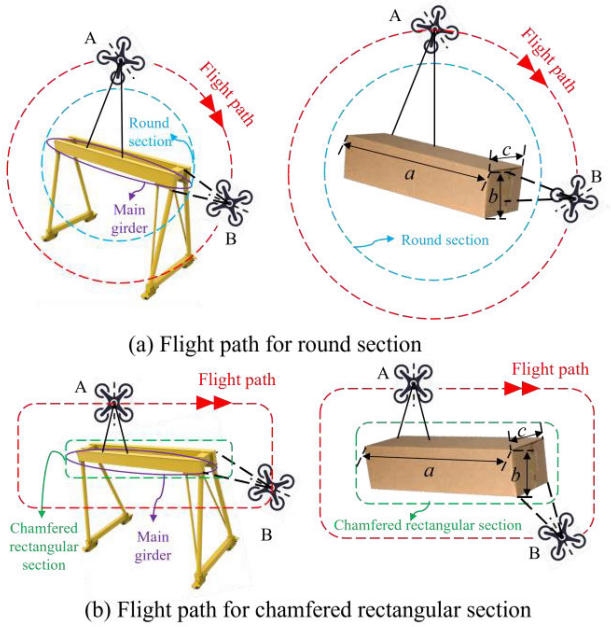



**FIGURE 2.** Flight path comparison based on structural characteristics.

## II. SEQUENTIAL IMAGE SEQUENCES ACQUIRED FROM UAV

### A. FLIGHT PATH SELECTION

There are two general types of UAV flight paths, circular and rectangular [18], and structures such as crane main girder and common carton, which have a large ratio of the long side (e.g. the $a$ side of carton) to the short side (e.g. the $b$ or $c$ side of carton), and when photographed according to the traditional circular flight path with the center of the circle as the target, as shown in Figure 2(a), the working distance of the camera from the target object varies too drastically, i.e. the working distance of the UAV at position A is much greater than that at position B, which can cause terrible accuracy of the 3D reconstructed model. In contrast, for structures such as crane main girder, where the length in the span direction is much greater than the height or width direction, a flight path applicable to the chamfered rectangular section, as shown in Figure 2 (b), is more preferable to be chosen.

### B. FLIGHT STRATEGY CUSTOMIZATION

Upon selecting the shape of the UAV flight path, a reasonable flight strategy is determined, i.e., the overlap, working distance, camera horizontal angle of view, lateral translation distance, film width and lens focal length, as well as the number of shots.

3D reconstructions based on UAV photography are required to have a minimum overlap of 50%. This study takes the overlap of 50% as an example to analyze the factors influencing the flight path applicable to chamfered rectangular section structures. As shown in Figure 3, the blue line between points S1 and S2 is the overlap area of the adjacent images, whose length d is calculated by Eq. (1), u is the working distance from the UAV to the surface of the structure under measurement, and $\alpha$ is the camera horizontal angle of view, which can be calculated from Eq. (2), v is the film width and f is the lens focal length. The overlap $\varepsilon$ can be calculated from Eq. (3), and $z$ is the lateral translation distance obtained by the UAV for each shot. The number of shots N can be calculated from Eq. (4), and L is the length of the rectangle. In addition, for the requirement of image capture at corners, three images are taken to realize full coverage of the corner region of the crane main girder.

$$d = u \cdot \tan\frac{\alpha}{2} \tag{1}$$

$$\alpha = 2\arctan\left(\frac{v}{2f}\right) \tag{2}$$

$$\varepsilon = \frac{d_{s_1 s_2}}{2d} = 1 - \frac{z}{2d} \tag{3}$$

$$N = 2 \cdot \frac{L}{z} + 10 \tag{4}$$

To meet the requirements of 3D scene reconstruction and crack detection, the flight path and the photographing strategy of UAV should follow three principles: clear image, sufficient overlap, and sufficient resolution. The proposed measures and procedures are as follows:

(1) Multiple images that should be captured in the same shooting position with the UAV to minimize blurring of the captured image.

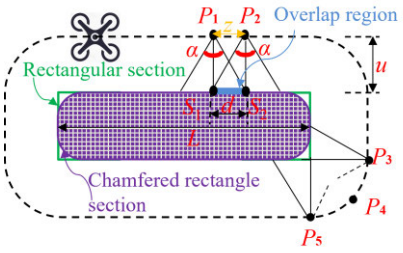



**FIGURE 3.** The flight path for the chamfered rectangular section and minimum 50% overlap.

(2) Adequate shooting points were required to be identified on the flight path to ensure at least 50% overlap. As shown in Figure 3, at 50% overlap, the edge of the camera view angle intersects the optical axis of the next image exactly. In detail, the line $P_1S_2$ is the edge of the camera view when the UAV is shooting from $P_1$, and intersects the optical axis $P_2S_2$ when shooting from $P_2$, with the intersection point located exactly at point $S_2$ on the surface of the chamfered rectangular section. The lateral translation distance z for each shot by the UAV is controlled to be equal to d, allowing a 50% overlap to be maintained. To achieve an overlap higher than 50%, a smaller lateral translation distance $z$ than d can be chosen.

(3) Given camera and lens parameters, a shorter working distance will generate a higher resolution. If the detection resolution is provided, the working distance of the UAV can be calculated via the resolution to determine the flight path and photography strategy.

## III. IMAGE-BASED 3D RECONSTRUCTION APPROACH

### A. BASIC PRINCIPLE

The original steps of 3D reconstruction include image acquisition, camera calibration, feature extraction, stereo matching, and model reconstruction.

In this study, structure from motion (SFM) [31] is added to the image-based 3D reconstruction step, and model optimization is performed after acquiring the initial reconstructed model to obtain a realistic and accurate 3D reconstruction model. While SFM, as the core technology, obtains the 3D coordinates of the scene in the image through a series of 2D images, to reconstruct the 3D scene of the unknown structure. The 3D reconstruction process of the structure adopted in this study is shown in Figure 4.

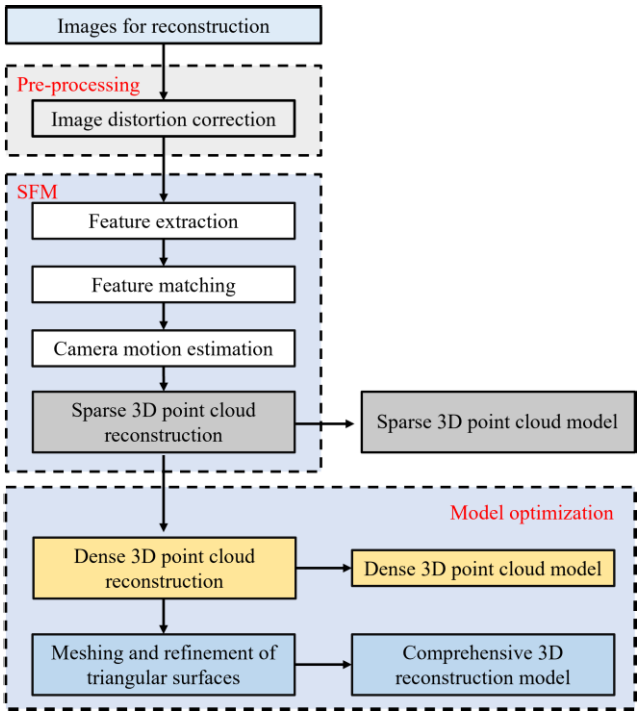



**FIGURE 4.** The 3D reconstruction process of the structure.

### B. THE CRUCIAL TECHNIQUES

After obtaining high-quality images through pre-processing, the following four main techniques are still needed for 3D reconstruction tasks, such as feature point extraction and matching, camera motion estimation, sparse 3D reconstruction, as well as model optimization, so that the real scene can be highly reconstructed.

#### 1) FEATURE POINT EXTRACTION AND MATCHING

The extraction of image feature points is a crucial part of image processing, which directly determines the quality and accuracy of the reconstruction results. Scale-invariant feature transform (SIFT) is an algorithm widely used in the field of image feature detection [32], with the invariance of scale, rotation and luminance changes, as well as excellent stability. As shown in Figure 5, the feature matching of the SIFT algorithm consists of the following three main steps: (1) detecting feature points from the image. (2) local features [33], i.e., feature vectors, are extracted from each feature point. (3) comparing each feature vector to find mutually matching pairs of feature points [34], to establish the correspondence between objects.

#### 2) CAMERA MOTION ESTIMATION

Camera motion parameter estimation aims to gain an extrinsic camera matrix that describes the position and directional relationships of the camera optical center. By shooting the target object from multiple views, the geometric relationships between the multiple images obtained are employed to recover the 3D shape and pose of the object.

Two important matrices are involved in the two-view geometry, as follows: (1) foundation matrix F, and (2) essential matrix E. These two matrices are the crux of what is needed to get the extrinsic camera.

The foundation matrix F, the mathematical description, is the epipolar geometry relationship between the imaging models of the two images covering the same scene, as shown in Figure 6. It has seven degrees of freedom (DOF) and can be estimated by at least seven points correspondences using

$$x'^{T} F x = 0 \tag{5}$$

where $x'$ and $x$ are the homogeneous pixel coordinate of a pair of feature points from the two images. Figure 6 shows the relationship of the point correspondences $o$ and $o'$ are the optical centers of two cameras, where, $x'$ and $x$ are the photographed points from point X onto the images of the two cameras.

The essential matrix $E$ is the specialization of the fundamental matrix in the case of normalized image coordinates. It could be obtained from the foundation matrix $F$ using

$$E = K'^{T} F K \tag{6}$$

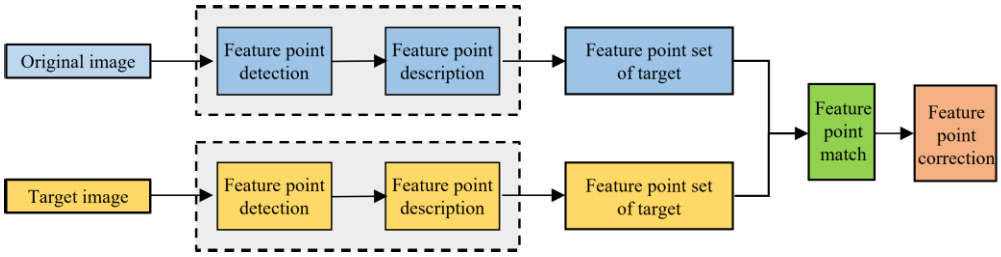



**FIGURE 5.** Flow chart of SIFT algorithm.

where, $K$ and $K'$ are the intrinsic camera matrices of the two cameras, respectively. The essential matrix $E$ is utilized to calculate the extrinsic camera matrices, which depict the camera motion. The estimated camera location is in relative positions, which can be scaled using any known absolute distance in the scene.

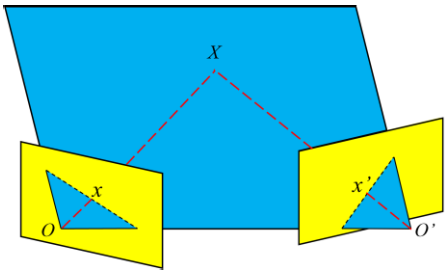



**FIGURE 6.** Epipolar geometry restriction in two-view geometry.

### 3) SPARSE 3D RECONSTRUCTION

To minimize the distance between the projected and observed image points, an optimization approach, Bundle Adjustment [35], is used over the collective set of camera and scene parameters, which is expressed as:

$$g(C, X) = \sum_{i=1}^{n} \sum_{j=1}^{m} \omega_{ij} \left\| x_{ij} - P(C_i, X_j) \right\|^2 \qquad (7)$$

where $\omega_{ij}$ is an indicator variable, $\omega_{ij} = 1$ if the camera $i$ observes a point $j$, otherwise, $\omega_{ij} = 0$. $C$ is the collection of camera parameters for a single camera $C_i = \{M, R, t, k_1, k_2\}$, $M$ is the inner parameter matrix of the camera, R and t are the rotation matrix and translation vector of the camera's position, respectively, $k_1$ and $k_2$ are the distortion coefficients. $x_{ij}$ is the observed image point, and $P(C_i, X_j)$ is the projected image point.

The objective function $g$ is the sum of squared projection errors. The purpose of bundle adjustment is to calculate the camera calibration parameters and posture of images that minimize this objective function. This process is repeated until the parameters are appropriate for all images. The result of this function can be used to generate a relatively sparse set of 3D points.

### 4) MODEL OPTIMIZATION

Sparse point cloud models cover only partial points on the surface of an object's 3D model, which is too sparse to show the real appearance of the object. Dense point cloud reconstruction is imperative for the high-precision reconstruction of the complete shape of an object in 3D space. Adjusting the pixel points in the image by rectification [27], all the epipolar lines are parallel and the searching space is reduced to one dimension. Meanwhile, the epipolar equation is no longer needed to be calculated each time. The search efficiency was enhanced and also further reduced the computational effort. Combined with triangular surface mesh and refinement operation, a comprehensive 3D reconstruction model is obtained by Agisoft Metashape Pro 1.7.4 modeling software. As an example, the carton is shown in Figure 7, which shows the sparse, dense point cloud models and comprehensive 3D reconstruction model, respectively.

(a) Sparse point cloud model (b) Dense point cloud model (c) Comprehensive 3D reconstruction model

**FIGURE 7.** Comparison of models before and after optimization.

## IV. ACCURACY EVALUATION OF POINT CLOUD MODEL

The accuracy of the 3D reconstruction model determines the precision of the defect recognition results of the measured structure, and ensuring the accuracy of the model is a prerequisite for carrying out defect recognition. For a 3D model based on UAV photogrammetry, the most critical factors affecting its accuracy are the number of images and the number and distribution of control points. For this Section, a self-made box is applied to simulate the main girder of a crane. Combined with UAV close-range photogrammetry, the box is modeled in 3D and the accuracy of the model is analyzed by changing the number of images, the number and location of control points, respectively; moreover, the influence rule of structure size effect on the accuracy is also explored. Model accuracy analyses are carried out in CloudCompare software.

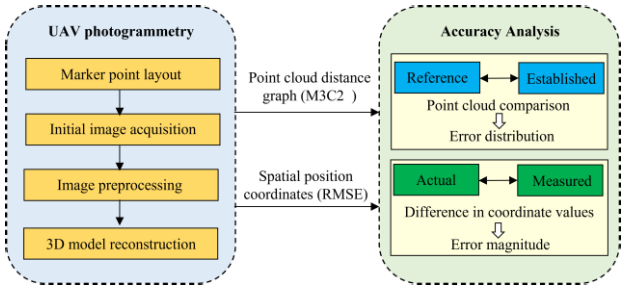



**FIGURE 8.** Accuracy analysis process of 3D reconstruction model based on UAV Image.

### A. MODEL ACCURACY ANALYSIS PROCESS

The accuracy analysis process of the 3D reconstruction model based on UAV images is shown in Figure 8. It mainly involves marker point layout, initial image acquisition, image pre-processing, 3D reconstruction, point cloud model comparison and marker point precision analysis. In which, image pre-processing is to eliminate noise interference generated by uncontrollable factors such as human operation and environment during the shooting process. Point cloud model comparison is to evaluate the error distribution by comparing the point cloud distance graph with the M3C2 algorithm, which is a comparison between the established model and the reference model, the reference model is the model with the highest accuracy under the same conditions or the actual real model. The difference in coordinate values is to evaluate the error magnitude by comparing the spatial position coordinates with the root mean squared error (*RMSE*) index, which quantifies the error magnitude between the actual coordinates and the measured coordinates of the marker points in the 3D model.

### B. M3C2 ALGORITHM

The point cloud distance is a fundamental index of the accuracy of a point cloud model, and the M3C2 algorithm is adopted to calculate the distance between the established point cloud model and the reference ones, which enables the calculation of the distance between two point cloud models [36]. The flow of the M3C2 algorithm is shown in Figure 9 and it mainly involves core point calculation, point cloud normal vector fitting, point cloud distance calculation and roughness analysis calculation. See ref [37] for details.

Point cloud model to be evaluated is established under different conditions (different number of images or marker points), which is compared with a reference model using the M3C2 algorithm to determine the distances in the horizontal and height directions, and the results are shown in different colors in a point cloud distance graph to indicate the error trend. In the point cloud distance graph, different values correspond to different colors, for example: when the distance is 0, the color is green. When the distance is greater than 0, the color is more towards blue, and conversely, the color is towards red, as shown in Figure 10.

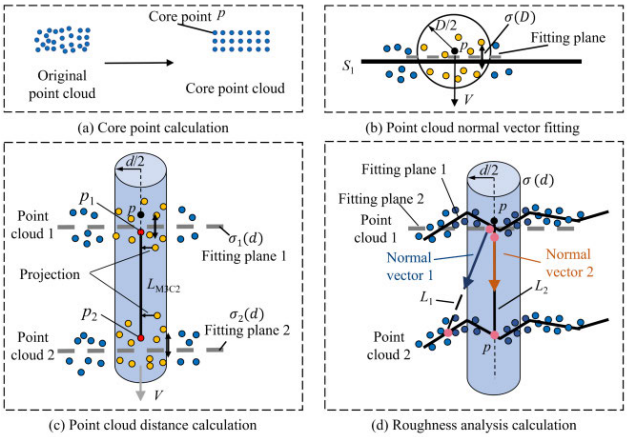



**FIGURE 9.** The flow of the M3C2 algorithm.

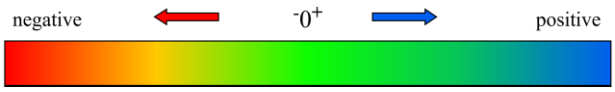



**FIGURE 10.** Color description of point cloud distance graph.

### C. EVALUATION INDEX

Root mean squared error (*RMSE*), which represents the sample standard deviation of the deviation between the predicted and actual values, reflects the degree of dispersion of the sample; the smaller the *RMSE*, the better the result. In this study, the *RMSE* was employed to assess the precision of individual marker points in the 3D model, i.e., the difference between the coordinates of the marker points in the 3D model and the actual measurements. The actual measurements are obtained from high-precision measuring instruments and are considered to be the actual values. The coordinate errors in the horizontal direction are denoted as $RMSE_X$, $RMSE_Y$ and $RMSE_{XY}$, and in the height direction as $RMSE_Z$, which are calculated by Eq. (8)-(11).

$$RMSE_X = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_{UAV}(i) - X_r(i))^2} \tag{8}$$

$$RMSE_Y = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_{UAV}(i) - Y_r(i))^2} \tag{9}$$

$$RMSE_{XY} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|(X_{UAV}(i) - X_r(i))^2 + (Y_{UAV}(i) - Y_r(i))^2\right|} \tag{10}$$

$$RMSE_Z = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Z_{UAV}(i) - Z_r(i))^2} \tag{11}$$

where $X_{UAV}$, $Y_{UAV}$ and $Z_{UAV}$ denote the two horizontal and height coordinates in the 3D model based on the UAV image, $X_r$, $Y_r$ and $Z_r$ denote the two horizontal coordinates and one height coordinate measured by the measuring instrument, and $N$ is the number of marker points.

### D. FACTORS INFLUENCING MODEL ACCURACY

#### 1) NUMBER OF IMAGES

The number of images has a significant influence on the model accuracy, while the overlap $\varepsilon$ and the working distance $u$ are the two major factors for image number. The carton 150cm * 30cm * 30cm was taken as the object of study and a 3D model was reconstructed, combined with different overlaps and working distances, by using the flight paths of section II. The 3D reconstructed model was compared with the reference model and the relationship between overlap, working distance and accuracy was analyzed separately. and this relationship may be utilized in the subsequent shooting of the original image of the crane's main girder.

##### a: IMAGE OVERLAP

The basic demand of image overlap is 50%, adjust the size of lateral translation distance z and overlap area d to set different overlap $\varepsilon$ (50%, 60%, 70%, 80%), under the condition that the UAV working distance $u$ is kept constant at 30cm, the lateral translation distance z of the UAV is obtained from Eq. (1) and (3) as 0.18m,0.144m,0.108m and 0.072m, respectively. Considering the convenience and feasibility of the experiment, two valid decimals are retained for z, and four groups of shooting schemes are acquired. See Table 1, where Test 4 was the reference model.

**TABLE 1. Image shooting schemes (different overlaps).**

| Parameters | Symbols and units | Test 1 | Test 2 | Test 3 | Test 4 |
|---|---|---|---|---|---|
| Focal length | $f$ /mm | 15 | 15 | 15 | 15 |
| Camera horizontal angle of view | $\alpha$ /° | 61.93 | 61.93 | 61.93 | 61.93 |
| Working distance | $u$ /cm | 30 | 30 | 30 | 30 |
| Lateral translation distance | $z$ /cm | 18 | 14 | 10 | 7 |
| Overlap | $\varepsilon$ /% | 50 | 60 | 70 | 80 |
| Number of images | $N$/pieces | 54 | 63 | 80 | 106 |

The 3D point cloud models corresponding to Test1 to Test4 are shown in Figure 11 (a-d). The point cloud models can all show the appearance of the carton, but the point cloud model in Figure 11 (a) has obvious missing points, while the point cloud model in Figure 11 (b) also has partially missing points, the point cloud models in Figure 11 (c) and Figure 11 (d) have no missing points, and both can show the actual box structure better, However, the point cloud model in Figure 11 (c) has a few black stray dots compared to Figure 11 (d), which is the fullest and smoothest point cloud model. Figure 11 (e-h) show the local detail of the same location (marked by red circles) on each 3D point cloud model, which shows the differences between each point cloud model more clearly, thus it can be concluded that the higher the overlap, the denser the point cloud, and the more comprehensive the model.

Test4, the point cloud model at 80% overlap, was taken as the reference point cloud model. The point cloud models of each other group (Test1, Test2, Test3) were compared with the reference model (based on the point cloud distance of the M3C2 algorithm), and the comparison results are shown in Figure 12. Figure 12(a) to Figure 12(c) show the point cloud models with 50%, 60% and 70% overlap versus the reference model. Compared to Figure 11(b), (c), it can be found that the point cloud model at 50% overlap (Figure 12(a)) has more red and green regions. However, the point cloud model at 60% overlap (Figure 12(b)), whose regions of red and green, is less than Figure 12(a) but more than Figure 12(c), The point cloud model at 70% overlap (Figure 12(c)) has only few red regions and is better overall than the first two.

Figure 12(d) to Figure 11(f) show the point cloud distance histograms of the point cloud models with 50%, 60% and 70% overlap respectively versus the reference point cloud model. From Figure 12(d)(e)(f), it can be seen that the point

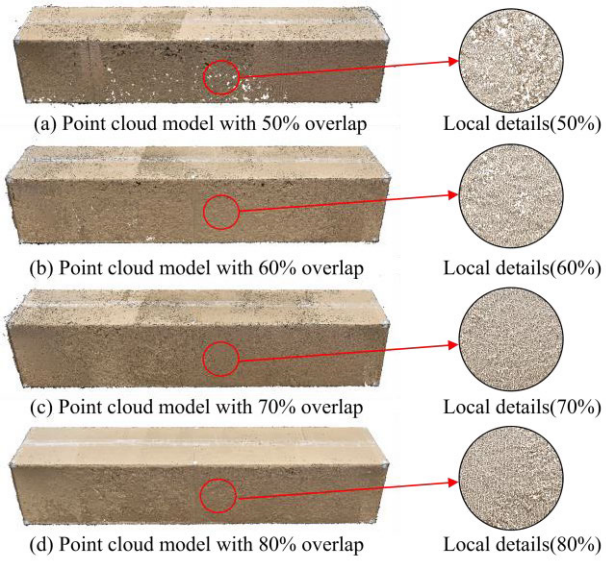



**FIGURE 11. Point cloud model and local details under different overlap.**

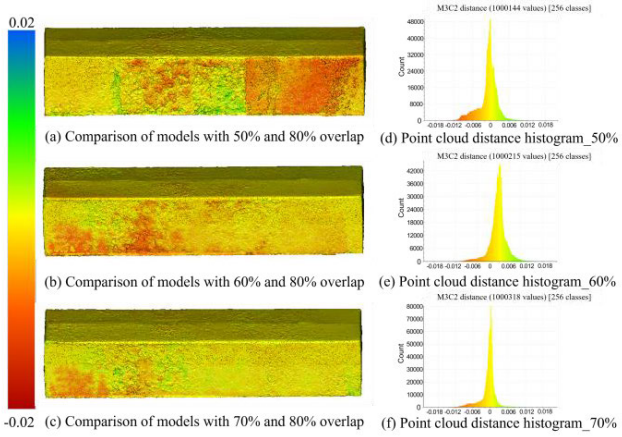



**FIGURE 12. Point cloud distances between the model to be evaluated and the reference model (under different overlap).**

cloud distances are concentrated between -0.2cm and 0.2cm, but the point cloud model with 50% overlap still exists with some point clouds in the interval of -1.2cm to -0.4cm, i.e., the red and light red areas in Figure 12(a). The point cloud model with 60% overlap shows relatively few distances in the -1.2cm to -0.4cm range, while the point cloud model with 70% overlap shows a more concentrated distribution of point cloud distances compared to the previous two. This reveals that the higher the overlap, the denser the point cloud.

### b: WORKING DISTANCE

With the same degree of overlap, the further the distance between the subject and the camera, the more content there will be per image, and for the same object, the further the distance, the fewer images there will be needed, but the images will be less sharp. Maintaining the overlap constant at 80%, the working distance $u$ from the carton surface was 30cm, 40cm, 50cm and 60cm respectively, and the correspondent UAV lateral translation distances z were calculated by Eq. (1) and (3) as 0.072m, 0.096m, 0.12m and 0.144m, and four groups of shooting schemes are acquired. See Table 2, where Test1 was the reference model.

**TABLE 2. Image shooting schemes (different working distances).**

| Parameters | Symbols and units | Test 1 | Test 2 | Test 3 | Test 4 |
|---|---|---|---|---|---|
| Focal length | $f$ / mm | 15 | 15 | 15 | 15 |
| Camera horizontal angle of view | $\alpha$ /° | 61.93 | 61.93 | 61.93 | 61.93 |
| Working distance | $u$ /cm | 30 | 40 | 50 | 60 |
| Lateral translation distance | $z$ / cm | 7 | 10 | 12 | 14 |
| Overlap | $\varepsilon$ /% | 80 | 80 | 80 | 80 |
| Number of images | N / pieces | 106 | 80 | 70 | 63 |

Test1, the point cloud model at a working distance of 30cm, was taken as the reference point cloud model. Similarly, the comparison results are shown in Figure 13. Figure 13(a) to Figure 13 (c) show the point cloud models with 40cm,50cm and 60cm working distance versus the reference model, respectively. The point cloud model with a working distance of 60 cm (Figure 12(c)) shows more dark regions compared to the other point cloud models, while the point cloud model with a working distance of 40 cm shows fewer dark regions.

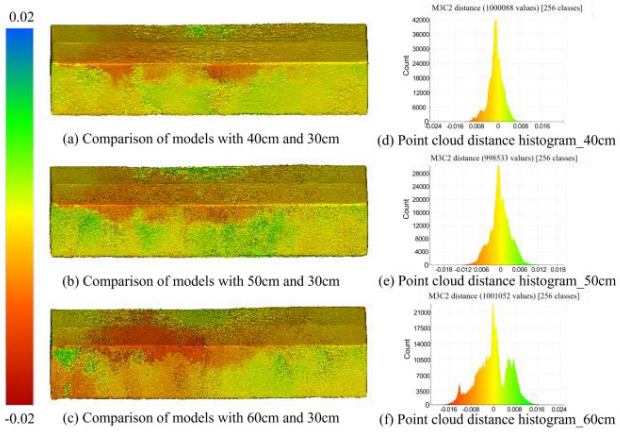



**FIGURE 13. Point cloud distances between the model to be evaluated and the reference model (under different working distances).**

Figure 13(d) to Figure 13(f) show the point cloud distance histograms of the point cloud models with 40cm,50cm and 60cm working distance respectively versus the reference point cloud model. The point cloud distances in Figure 13(d) and Figure 13(e) are concentrated between −0.4cm and 0.4cm, while Figure 13(f) is distributed between −1.6cm and 1.6cm, which is more dispersed compared to the first two. Overall, the distribution of point cloud distances gradually disperses as the working distance increases, i.e., the smaller the working distance, the denser the point cloud and the more accurate the point cloud model is.

The above analysis shows that model accuracy becomes more accurate as the number of images increases. The reason for this is that smaller numbers of images contain less feature information about the target object, which results in a lack of detail of the reconstructed 3D point cloud model and might cause the subsequent measurement results to be poorly precise. Nevertheless, over-increasing the model accuracy will also incur computational costs and inefficiencies, and the detection purpose should be taken into account to set a reasonable shooting scheme.

### 2) MARKER POINT

For UAV photogrammetry, marker points should be widely distributed and cover the whole region, however, the massive layout of marker points costs much time, especially for large-scale and medium-sized mechanical equipment like cranes, and special consideration must be made for personnel safety during the marker point layout work. Thus, how to guarantee the accuracy of the 3D reconstruction model based on fewer marker points is a fundamental issue.

There are two types of marker points, control points and check points. Control points are used to fix the coordinate positions of the 3D model and calibrate the dimensional scale of the model for reliable data of the 3D reconstructed model. Check points serve two purposes, on the one hand for assessing the accuracy of single marker points on the 3D model, on the other hand for the detection of structural defects, such as deformation, and crack parameters.

In this Section, different point cloud models are constructed by changing the number and location of control points. On the basis of the point cloud distance graph, the model errors under different shooting scenarios are evaluated, and then the *RMSE* index is utilized to evaluate the difference between the coordinates of the check points in the model and the actual coordinate values to verify the model accuracy.

### a: NUMBER OF MARKER POINTS

The rectangular carton was taken as the experimental object, and the flight path was consistent with the previous


**FIGURE 14.** The reference point cloud model constructed under full control and marker point layout.

experiment. The image Overlap was set to 80% and the working distance was 30cm. Twelve marker points were laid out in the structure to be measured (full control). As 12 marker points are dense enough to cover the whole carton uniformly, they are used as the reference model. The reference point cloud model and marker point layout constructed under full control are shown in Figure 14.

To investigate the influence of different numbers of marker points on the accuracy of the model, the numbers of control points were set as 4, 6, 8 and 12, the location distribution of which is shown in Figure 15, and the point cloud models under the corresponding control points were established respectively. The unused marker points in the model were treated as check points. Taking control point 8 as an example, marker points 5, 6, 7 and 8 of the point cloud model were the check points of the model.

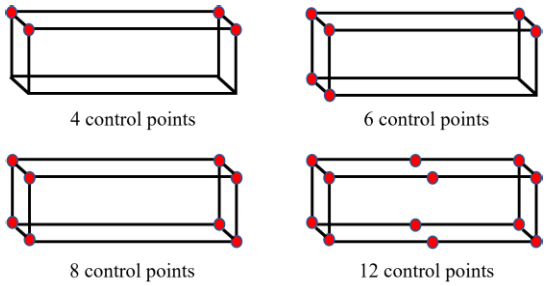



**FIGURE 15.** Different number of control points and location distribution.

Figure 16(a) to Figure 16(c) show the comparison between the point cloud model under the number of control points of 4, 6 and 8, respectively, and the reference point cloud model, and Figure 16(d) to Figure 16(f) show the corresponding point cloud distance distribution histograms. A large proportion of dark regions exist in Figure 16(a) and Figure 16(b), with the green and red regions gradually decreasing as the control points increase, which suggests that the model accuracy has been enhanced. The majority of the point cloud distances in Figure 16(d) are biased by 2mm in the positive direction (greater than 0), i.e., the overall error is around 2mm. The point cloud distance in Figure 16(f) shows a concentrated distribution between -1mm and 1mm. The point cloud distance

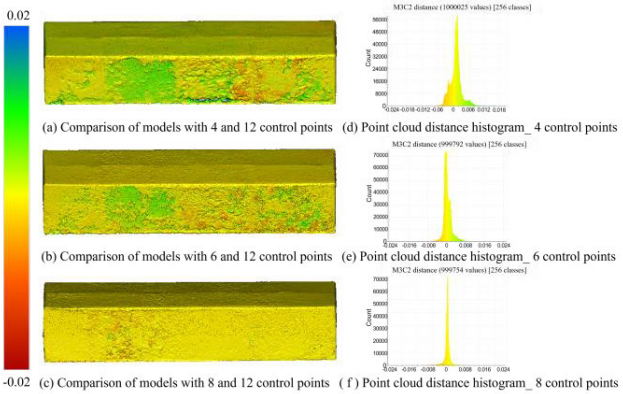



**FIGURE 16.** Histogram comparison of the distance distribution between the point cloud model and the reference model under different numbers of control points.

in Figure 16(f) is concentrated between -1mm and 1mm, the error is smaller compared to the previous two, and the accuracy of the model is better, which shows that the more the number of control points, the more concentrated the point cloud distance distribution is, and the higher the accuracy of the model. the point cloud distances created by 4 control points are mostly stable in the range of 0 to 2 mm, except for a few deviations, which proves the stability of the point cloud model with 4 control points.

The accuracy of the marker points in the model under different numbers of control points was evaluated by calculating the root mean square error of the check points in three directions, and the results are shown in Table 3. The synthetic errors *RMSEs* for each point cloud model were 4.006 mm, 2.838 mm and 1.987 mm, respectively. For the point cloud model with four control points, the *RMSEs* in the X, Y and Z directions are 2.858mm, 2.318mm and 1.585mm respectively, which are the maximum errors compared to the rest of the point cloud models. The point cloud model with 8 control points shows relatively minimal unidirectional and synthetic errors. The *RMSE* of the check points demonstrates again that the more the number of marker points, the higher the accuracy of the point cloud model.

**TABLE 3.** *RMSE* of check point in each direction (different number).

| No. of control points | $RMSE_X$ (mm) | $RMSE_Y$ (mm) | $RMSE_Z$ (mm) | *RMSE*s (mm) |
|---|---|---|---|---|
| 4 | 2.858 | 2.318 | 1.585 | 4.006 |
| 6 | 1.991 | 1.17 | 1.648 | 2.838 |
| 8 | 0.583 | 1.572 | 1.065 | 1.987 |

Note: *RMSE*x, *RMSE*y and *RMSEz* are errors in a unidirectional direction, and *RMSE*s is a synthetic error in three directions.

### b: LOCATION OF MARKER POINTS

As the 3D point cloud model under 4 control points shows the same high accuracy (2mm) and stability, and its reconstruction efficiency is significantly enhanced compared to

the 12 control points. In this Section, four vertices on the front (back), left (right) side and positive oblique and antioblique surfaces are treated as control points to establish a three-dimensional model to study the influence of the marker point location on the accuracy of the model, and the distribution of control point locations is shown in Figure 17.

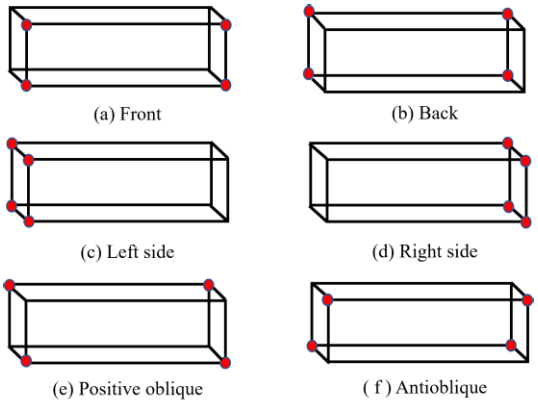



**FIGURE 17.** Histogram the distribution of control point locations.

The *RMSE* of the check points in the model at different locations of the control points are shown in Table 4. The *RMSEs* of the point cloud model at different locations are 4.006mm, 4.145mm, 14.055mm, 12.89mm, 3.044mm and 2.235mm respectively. The *RMSE* of the checkpoints under locations Figure 17(a) and Figure 17 (b) are found to be similar, and this rule is also followed between Figure 17(c) and Figure 17 (d), this is also applicable in Figures 17 (e) and 17 (f).

**TABLE 4.** *RMSE* of check point in each direction (different location).

| Type | $RMSE_X$ (mm) | $RMSE_Y$ (mm) | $RMSE_Z$ (mm) | *RMSEs* (mm) |
|---|---|---|---|---|
| a | 2.858 | 2.318 | 1.585 | 4.006 |
| b | 0.968 | 1.546 | 3.722 | 4.145 |
| c | 2.618 | 13.772 | 1.008 | 14.055 |
| d | 2.238 | 12.527 | 2.039 | 12.89 |
| e | 1.024 | 2.275 | 1.744 | 3.044 |
| f | 0.416 | 1.693 | 1.398 | 2.235 |

When the four vertices on the two oblique surfaces are treated as control points (Figure 17(e) and Figure 17 (f)), the *RMSE* of the check points is relatively small, while when the four vertices on both sides are treated as control points (Figure 17(c) and Figure 17 (d)), the *RMSE* of the check points is bigger. In addition, the areas of the two oblique surfaces are equal and are bigger than those of the sides. The following assumptions might be made: (1) When the four vertices on the faces with equal areas are treated as control points, the errors of the check points are close to each other. (2) The bigger the area of the surface on which the control points are located, the smaller the error.

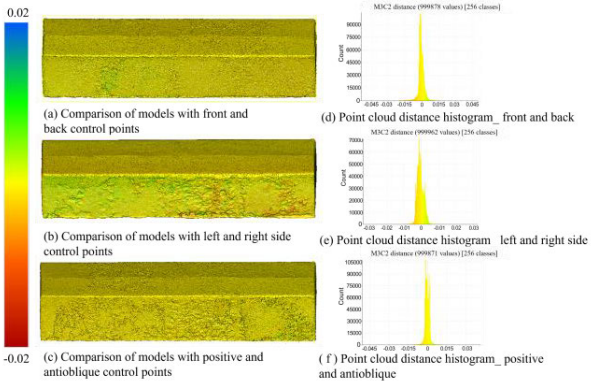



**FIGURE 18.** Comparison of model point cloud distances under different schemes.

According to the size of the area, three sets of comparisons were made, i.e., front and back, left and right sides, and oblique and antioblique sides, to establish point cloud models and compare the point cloud distances respectively. Apart from a slight fluctuation in the error of the model based on the marker points on the sides, the general error for all three sets of comparisons is small, as shown in Figure 18, As shown in Figure 18(d)(e)(f), the values of the point cloud distances are concentrated around 0, which verifies that the model errors are very close to each other when 4 vertices on a face of equal area are treated as control points.

To demonstrate that the bigger the area of the surface where the control points are located, the smaller the error of the established point cloud model, the point cloud model under 8 control points was considered to be highly accurate, so it is taken as the reference model, The point cloud model under 4 control points was compared with the reference model, i.e. the front (back) and left (right) side of the model, for the point cloud distance, and the results are shown in Figure 19.

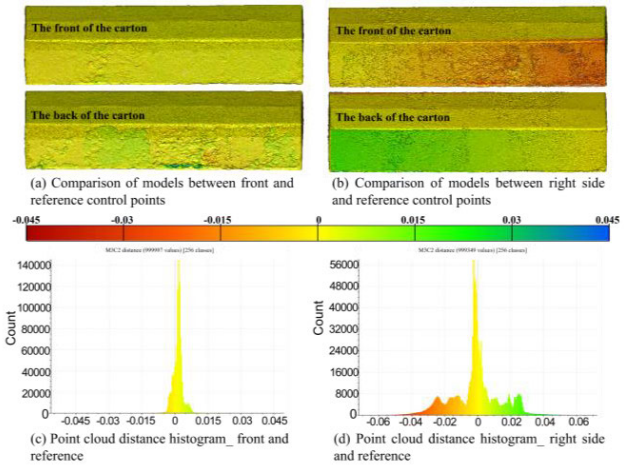



**FIGURE 19.** The point cloud distance comparison of the model under the four control points and the reference model (the front and back of the carton).

The former shows little or no red or green areas compared to Figure 18(a) and Figure 18(b), while the latter shows many red or green regions. In addition, comparing Figure 18(c) and Figure 18(d), it can be seen that the global point cloud distance of the former is concentrated around 0, while the point cloud distance of the latter is distributed in the interval from −0.04m to 0.04m. As a result, the point cloud model with four control points on the front (back) side is more accurate than the point cloud model with four control points on the left (right) side, which means that the bigger the area of the surface where the control points are located, the smaller the error of the point cloud model.

To further verify the relationship between the area of the surface on which the control points are located and the accuracy of the point cloud model, the carton in Figure 20 is taken as an example whose surfaces 1, 2 and 3 differ in area, where surface 1 is the smallest, while surfaces 3 and 4 are equal in area and these two surfaces are the largest. The experiment set up to investigate is as follows:

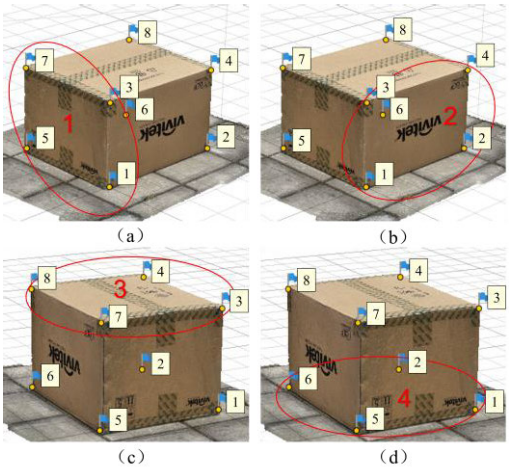



**FIGURE 20.** Comparison Model marker point and surface serial number illustration.

Step 1: Set 8 marker points, located at the 8 vertex positions of the carton.

Step 2: Establish the point cloud model of the box using the four vertices of faces 1, 2, 3 and 4 as control points and the rest of the vertices as check points, respectively.

Step 3: *RMSE* is utilized to measure the difference between the coordinates of the check points in the point cloud model and the actual measurements, which characterize the model accuracy.

Step 4: Based on the principle of control variables, compare the *RMSE* values of the point cloud models established by the control points under each of the four surfaces.

The *RMSE* of the check points of each model in each direction are shown in Table 5. The *RMSE* of the four point cloud models (a, b, c, d) are 5.473mm, 4.539mm, 2.731mm and 3.079mm respectively. The point cloud model under face 1 (the smallest area) shows the maximum *RMSE* and the worst accuracy, the point cloud model under face 3 (the biggest area) shows the minimum *RMSE* and the best accuracy, and the point cloud model under face 4 (equal in area to face 3) gives a close *RMSE* to the one under face 3, both of which are at the same level of accuracy. The above relationship between the area of the surface where the control points are located and the accuracy of the model is again verified.

**TABLE 5.** *RMSE* of check point in each direction (different areas).

| Type | $RMSE_X$ (mm) | $RMSE_Y$ (mm) | $RMSE_Z$ (mm) | $RMSE_S$ (mm) |
|---|---|---|---|---|
| a | 3.398 | 3.406 | 2.609 | 5.473 |
| b | 2.423 | 2.492 | 2.92 | 4.539 |
| c | 1.569 | 2.015 | 0.969 | 2.731 |
| d | 2.582 | 1.548 | 0.644 | 3.079 |

### E. STRUCTURAL SIZE EFFECTS

Since the carton model differs significantly in size from the main girder structure of the crane, which is the key research object of this study. The investigation experiment set up to investigate the relationship between the size of the structure to be measured and the accuracy of the model is as follows:

Step 1: Five cartons were prepared in the ratio of length, width and height, as shown in Figure 21, and 11 marker points were laid out in the same locations in each carton.

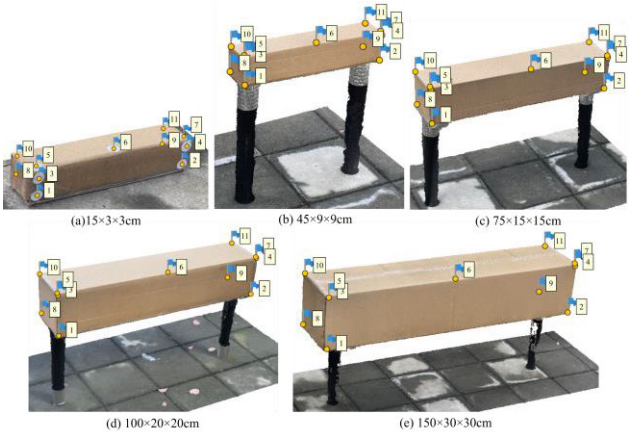



**FIGURE 21.** Five equal scale carton 3D models.

Step 2: In the above study on the factors affecting the accuracy of the model, it was found that when the overlap was 70% and the working distance was 40cm, the accuracy of the model was virtually the same as that of the reference model. Therefore, in this investigation experiment, the overlap degree was set at 70% and the shooting distance was 40cm, and image acquisition and 3D reconstruction were carried out for cartons of different sizes, and the efficiency of the reconstructed model was improved under the condition of guaranteeing accuracy.

Step 3: Establish a point cloud model of the box by using the 8 vertices of the box as control points and the marker points 5, 6 and 7 on the upper surface as check points.

Step 4: *RMSE* is utilized to measure the difference between the check point coordinates and the actual measurements in point cloud models of different size cartons, which characterizes the model accuracy.

Step 5: The *RMSE* of each 3D model check point is fit-ted linearly to the length characteristics of the carton structure, and the point cloud model error of the crane main girder structure is estimated.

In addition, as the object of study was a box, the flight path was only shot around the main body of the box, so the clarity and accuracy of the outrigger part of the point cloud model was lower.

*RMSE* of check point in each direction are shown in Table 6, The fitting results are shown in Figure 22, which is a potential relationship between the *RMSE* of the check point and the length characteristics of the carton structure, which are related as in Eq. (12):

$$y = 0.00071x + 1.4173 \tag{12}$$

**TABLE 6.** ***RMSE*of check point in each direction (different structural sizes).**

| Type | $RMSE_X$ (mm) | $RMSE_Y$ (mm) | $RMSE_Z$ (mm) | RMSEs (mm) |
|---|---|---|---|---|
| a | 0.683 | 0.925 | 0.895 | 1.457 |
| b | 0.721 | 0.946 | 1.334 | 1.787 |
| c | 0.902 | 0.84 | 1.482 | 1.928 |
| d | 1.714 | 0.942 | 1.236 | 2.227 |
| e | 1.619 | 1.634 | 0.998 | 2.412 |

R-Square indicates the overall fitting of the regression equation, and from Figure 22, the value of the fit R-Square is 0.9601, which is very close to 1, which indicates that the regression equation fits the error values extremely well.

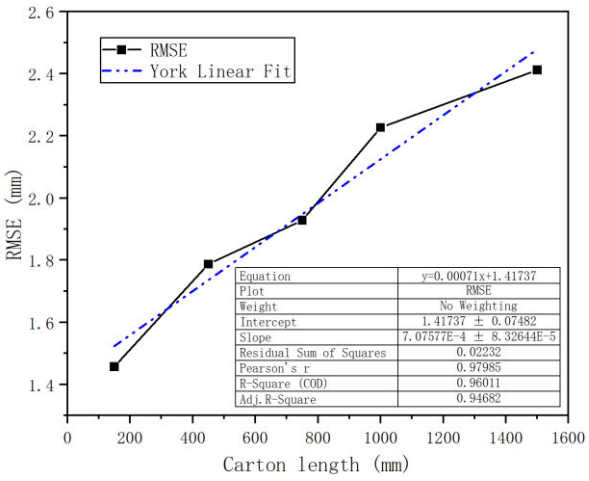



**FIGURE 22.** **Results of fitting carton length to model error.**

It can be found from the fitting curve that the model error is positively correlated with the length characteristics of the measured structure, and the model error gradually grows as the length of the structure increases, and the model error is deduced to be 22.7173mm for a crane structure size of 30*6*6m. However, according to the structural characteristics of the crane main girder, it is known that the web spacing (width) and web height (height) of a main girder of equal span (30m) is hardly possible to 6m, and the comprehensive analysis shows that the main girder model of 30m span is much less than 22.7173mm.

## V. EXPERIMENT ON STRUCTURAL DEFECT DETECTION AND IDENTIFICATION

### A. PREPARATION

The DJI M200 UAV used in this experiment, shown in Figure 23(a), was paired with a ZenmuseX5S camera with a 15mm focal length. The total station version adopted was the Leica TZ05, as shown in Figure 23(b). The main parameters of the DJI M200 UAV and the Leica TZ05 total station are presented in Tables 7 and 8 respectively.


**FIGURE 23.** **Experimental equipment.**

**TABLE 7.** **DJI M200 UAV main parameters.**

| Parameter | Value |
|---|---|
| Weight of UAV under no-load (kg) | 3.75 |
| Maximum load (kg) | 2.3 |
| Hovering accuracy (m) | Vertical: ± 0.1<br>Horizon: ± 0.3 |
| Pixel resolution | 4608×3456 |
| The focal length of the lens (mm) | 15 |
| Minimum focal length (m) | 0.2 |

**TABLE 8.** **Leica TZ05 total station main parameters.**

| Parameter | Value |
|---|---|
| Measuring range with prism (m) | 500 |
| Accuracy with prism (mm) | ±(1mm+1.5×10-6D) |
| Prism-free measuring ranges (m) | 500 |
| Prism-free accuracy (mm) | ±(2mm+2×10-6D) |
| Angle measurement accuracy (") | 2 |

Note: D is the distance measured by the total station (km)

### B. STRUCTURAL DEFORMATION DETECTION

#### 1) EXPERIMENTAL PROCESS

The process of crane main girder deformation detection based on a 3D reconstruction of UAV images to carry out experiments is shown in Figure 23. The detailed steps are as follows:

Step 1: Prepare the equipment needed for the experiment, including the UAV, total station and marker points.

Step 2: Layout marker points, as shown in Figure 25, with several uniformly laid out at the end-span of the main girder and the edge of the lower flange plate. In which, the red marker points are taken as control points to correct the coordinates of the 3D model and to determine the scale of the model based on the distance between the control points and the actual distance in the model. The blue marker points on the main girder are taken as check points for the evaluation of the main girder model accuracy. In addition, the position of the check point changes with the deformation of the main girder, so it is possible to measure the deformation of the crane main girder according to the change in coordinate data of the check point.

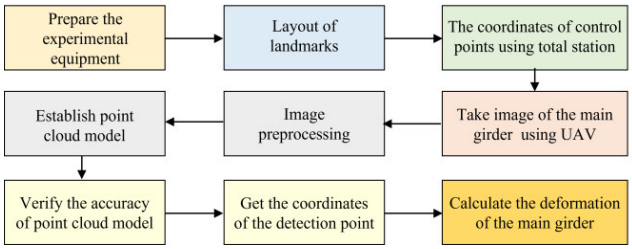



**FIGURE 24.** The process of crane main girder deformation detection.

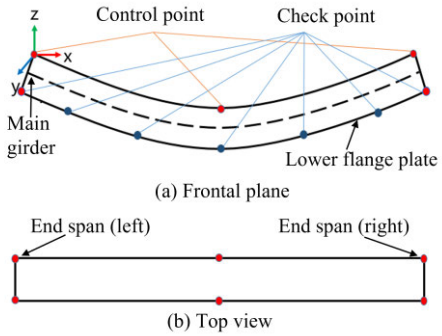



**FIGURE 25.** Schematic layout of crane main girder structure marker points.

Step 3: Acquire the coordinates of all the marker points of the main beam using the total station as the actual distance of the crane main girder.

Step 4: Operate the UAV to shoot the main girder structure according to the crane main girder UAV flight strategy in Section II. Given that the flight path is a chamfered rectangle, which is relatively simple, we used manual operation to adjust the flight status according to the working distance and other parameters displayed on the remote controller.

Step 5: pre-process the sequential image sequences shot by the UAV and reconstruct the crane main girder structure in 3D to realize the conversion from a 2D image to a 3D point cloud.

Step 6: Verify the accuracy of the crane main girder point cloud model and obtain the coordinate values of the check points.

Step 7: Compare the coordinate values of the check points before and after deformation, and evaluate the deformation of the main girder structure.

The experiment can be performed by the coordinate change of the check point in the main girder point cloud model to measure the parameters such as the upper arch, lower deflection and side bending. If the inaccuracies in the manufacture of the crane main girder are ignored and there is no side bending, then the y-directional coordinate value of the check point will not change; otherwise, the y-directional coordinate value of the check point will change.

### 2) POINT CLOUD MODEL ACCURACY VERIFICATION

The experimental object is a 41t-25m gantry crane as shown in Figure 26(a).

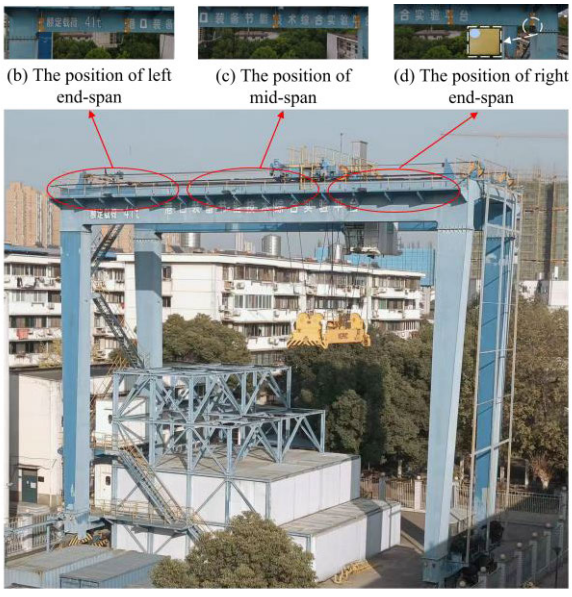



**FIGURE 26.** Gantry crane and marker point layout.

In the marker point layout, considering the safety of the experiment, the marker points are unable to be directly pasted on the lower flange plate of the crane main girder, so the white marker points are pasted on the yellow wooden board and the board is attached to the main girder. Ten marker points are laid out on the front of the main girder web, see Figure 26(b, c, d), and the enlarged details of the marker points are shown in Figure 26(d).

The Leica TZ05 total station provides a maximum measuring distance of 500m and an accuracy of $\pm(2\text{mm}+2 \times 10 - 6\text{D})$ without external interference, which adequately meets the engineering requirements of the measurement. The TZ05 is used to measure the coordinate values of the marker points on the surface of the crane main girder, and all data is

**TABLE 9.** *RMSE* for the three sets of models in each direction (different number).

| No. of control points | $RMSE_X$ (mm) | $RMSE_Y$ (mm) | $RMSE_Z$ (mm) | *RMSEs* (mm) |
|---|---|---|---|---|
| 4 | 14.1 | 21.67 | 11.21 | 27.87 |
| 6 | 10.68 | 8.22 | 9.72 | 16.61 |
| 8 | 8.22 | 5.03 | 3.42 | 10.22 |

measured three times independently and averaged to ensure the accuracy of the measurement, and the results are treated as the actual coordinates of the marker points on the crane main girder.

When shooting with the DJI M200 UAV, the image overlap was set to 70%, the working distance was 5m and the flight path was consistent with the strategy described in Section II. The point cloud of the crane main girder established is shown in Figure 27, and the comprehensive 3D reconstruction model after triangular surfaces mesh and refinement operation is given in Figure 28.

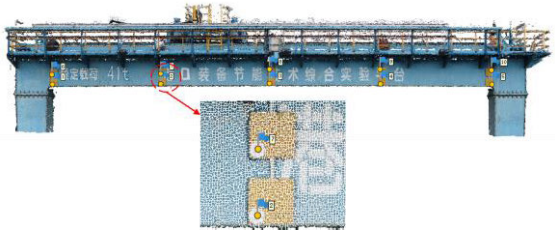


**FIGURE 27.** The point cloud and local detail of the crane main girder.

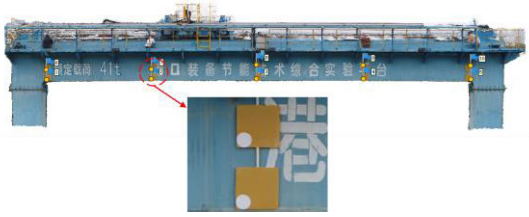


**FIGURE 28.** The comprehensive 3D reconstruction model of the crane main girder.

Four (No. 1, 5, 6, 10) marker points, six (No. 1, 5, 6, 7, 9, 10) marker points and eight (No. 1, 2, 4, 5, 6, 7, 9, 10) marker points respectively are taken as control points to establish a 3D model of the crane main girder under no load.

The balustrade part on the model was removed due to the fact that the balustrade is irrelevant to the main girder deformation detection. The point cloud models under 4 and 6 control points were compared with the reference model with 8 control points for the point cloud distances respectively, and the comparison results are shown in Figure 29.

Figure 29(a) and Figure 29(b) show the point cloud model under 4 and 6 control points respectively compared with the reference model (8 control points), and Figure 29(c) and

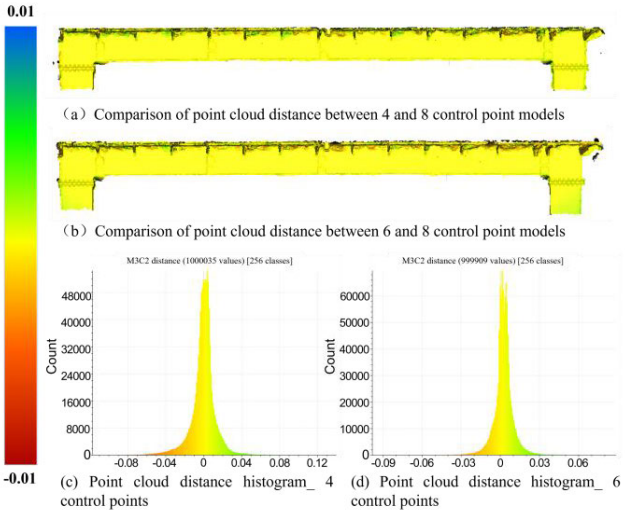



**FIGURE 29.** Point cloud distance and histogram of the model under different numbers of control points.

**FIGURE 30.** The location of 13 line segments in the point cloud model.

Figure 29(d) show the corresponding point cloud distance distribution histograms. From Figure 29(c) and Figure 29(d), one can see that the point cloud distance of the model with 4 control points is concentrated in the interval from −0.04m to 0.04m, while the model with 6 control points is mainly concentrated in the interval from −0.03m to 0.03m. Compared with the model under 4 control points, the model under 6 control points is more accurate. It can be concluded that the more control points, the higher the accuracy of the model.

For the accuracy assessment of individual marker points in the model, the *RMSE* for the three sets of models in each direction were gained as shown in Table 9, by taking the marker points except for the control points as check points.

The *RMSEs* for each model is 27.87 mm, 16.61 mm and 10.22 mm respectively. The *RMSEs* for each model are 27.87 mm, 16.61 mm and 10.22 mm respectively. For the model under 4 control points, the *RMSEs* in the X, Y and Z directions are 14.1 mm, 21.67 mm and 11.21 mm respectively, with the maximum synthetic error of the check points compared to the rest of the models, while for the point cloud model under 8 control points, the X, Y and Z directions *RMSE* are 8.22 mm, 5.03 mm and 3.42 mm respectively, with the minimum total error and much less than the maximum allowed for deformation (25 mm). The above conclusion is again verified, the more the number of marker points, the higher the accuracy of the point cloud model. It can be concluded that the model with the number of control points of 8 is chosen to achieve the highest accuracy and meets the fundamental demands of main girder deformation detection.

**FIGURE 31.** Point cloud model of the main girder after deformation.

**TABLE 10.** Comparison between actual length and measured length of 13 line segments on the main girder.

| No. of the Line segment | Actual length (m) | Measured length (m) | Error(m) |
|---|---|---|---|
| 1 | 5.49 | 5.483 | -0.007 |
| 2 | 5.5 | 5.508 | 0.008 |
| 3 | 5.58 | 5.575 | -0.005 |
| 4 | 5.63 | 5.622 | -0.008 |
| 5 | 0.502 | 0.495 | -0.007 |
| 6 | 0.502 | 0.496 | -0.006 |
| 7 | 0.725 | 0.732 | 0.007 |
| 8 | 0.725 | 0.733 | 0.008 |
| 9 | 0.725 | 0.732 | 0.007 |
| 10 | 5.49 | 5.482 | -0.008 |
| 11 | 5.49 | 5.485 | -0.005 |
| 12 | 5.6 | 5.607 | 0.007 |
| 13 | 5.63 | 5.68 | 0.005 |

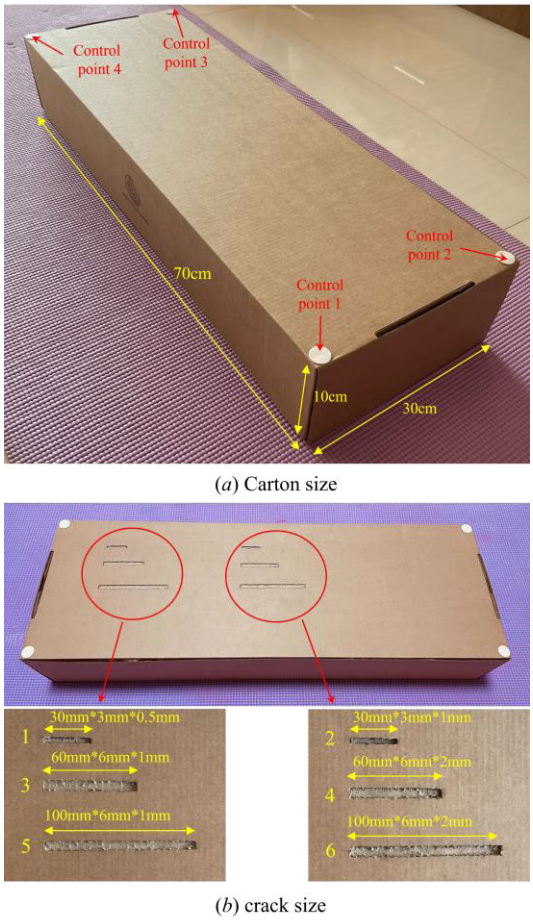
To further verify the accuracy of the point cloud model of crane main girder, the coordinate data obtained from the total station was taken as the actual value and the error analysis was carried out on the coordinate data of the check points obtained on the main girder point cloud model. For a clearer visual comparison of the gaps, the actual lengths measured by the total station were compared with the lengths in the point cloud model for the 13 line segments as shown in Figure 30, and the results of the comparison are shown in Table 10.

The actual measurement results show that the overall measurement error is within 8 mm. The point cloud model of the crane main girder based on the UAV 3D reconstruction is sufficiently accurate and feasible for crane main girder structure deformation detection.

#### 3) DEFORMATION DETECTION RESULTS

The deflection of the crane main girder is the downward displacement of the main girder span at full load compared to its position at no load, i.e., the deformation of the main girder structure which could be characterized by the change in the position of the marker point. In this experiment, the deformation of the main girder under full load is simulated by moving the marker point at the mid-span position vertically downwards by 20mm.

The same flight strategy and modeling method mentioned above were adopted, the main girder of the crane under full load was established, and the point cloud model of the main girder after deformation is shown in Figure 31. Comparing the point cloud model established under no load and rated load condition, it is found that the No. 3 marker point at the mid-span position of the main girder, whose Z coordinates are 17.315m and 17.303m respectively, the mid-span deformation is 0.012m (12mm). Compared to the actual displacement of 20mm, the error is 8mm and the accuracy reaches a millimeter level.




(*a*) Carton size

(*b*) crack size

**FIGURE 32.** Carton sizes and crack sizes.

### C. STRUCTURAL CRACK IDENTIFICATION

There is no crack defect in the crane's main girder structure in this study. There is no crack defect in the crane main girder structure itself in this study. It is difficult to manufacture cracks on the main girder structure due to the limitation of

**TABLE 11.** The specific values of the crack parameters.

| Type | Length/mm | | | Width/mm | | | Depth/mm | | |
|---|---|---|---|---|---|---|---|---|---|
| | Actual | Measured | Error | Actual | Measured | Error | Actual | Measured | Error |
| 1 | 30 | 29.9 | 0.1 | 3.0 | 2.9 | 0.1 | 0.5 | 0.4 | 0.1 |
| 2 | 30 | 29.8 | 0.2 | 3.0 | 2.9 | 0.1 | 1.0 | 1.1 | 0.1 |
| 3 | 60 | 60.2 | 0.2 | 6.0 | 5.8 | 0.2 | 1.0 | 1.0 | 0.0 |
| 4 | 60 | 60.1 | 0.1 | 6.0 | 6.1 | 0.1 | 2.0 | 1.8 | 0.2 |
| 5 | 100 | 99.8 | 0.2 | 6.0 | 5.9 | 0.1 | 1.0 | 0.9 | 0.1 |
| 6 | 100 | 99.7 | 0.3 | 6.0 | 5.9 | 0.1 | 2.0 | 2.2 | 0.2 |

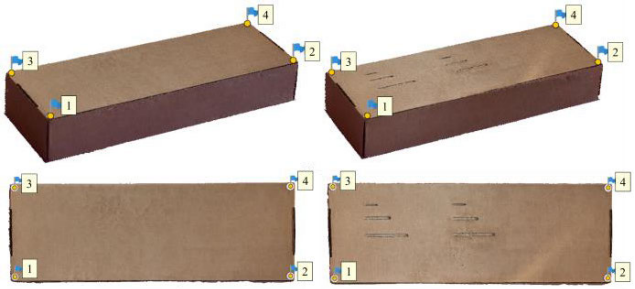



**FIGURE 33.** Carton model before and after damage.

experimental equipment and sites, for the sake of proving the effectiveness of the method in this study, a carton is taken as the research object, and cracks of different degrees are manufactured on the carton structure by human means for simulating the main girder structure crack identification.

### 1) EXPERIMENTAL PROCESS

Step 1: Prepare a carton with size 70cm*30cm*10cm, as shown in Figure 32(a). The four control points required to establish the 3D point cloud model of the carton are laid out on the four vertices of the upper surface.

Step 2: Cracks of different sizes were artificially manufactured on the carton, six types of cracks are shown in Figure 32(b), which simulate different degrees of cracks in the main girder structure.

Step 3: The UAV was utilized to shoot the carton before and after the damage, and the corresponding 3D point cloud model was established. Control points were employed to calibrate and fix the scale of the model, and the point cloud model of the carton before and after the damage under 4 control points is shown in Figure 33.

Step 4: The ICP algorithm [38] and four control points were used to calibrate the point cloud models before and after the damage. The locations of the two point cloud models before the calibration were disordered, as shown in Figure 34(a), and the point cloud model after the calibration is shown in Figure 34(b).

### 2) CRACK IDENTIFICATION RESULTS

Point cloud distance comparison maps are used to identify cracks in a structure and show the parameters and location of the cracks by means of colors and values. When cracks appear

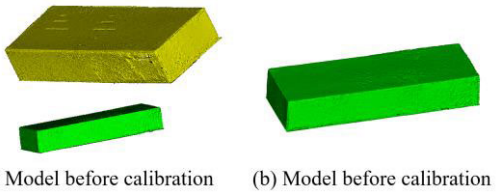



**FIGURE 34.** Point cloud model before and after calibration.

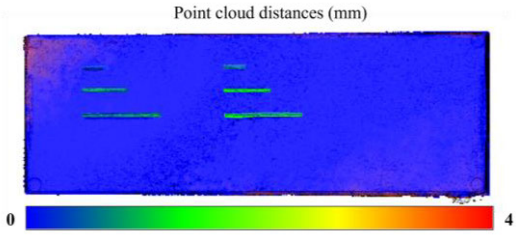



**FIGURE 35.** The point cloud distance comparison of the carton structure before and after the damage.

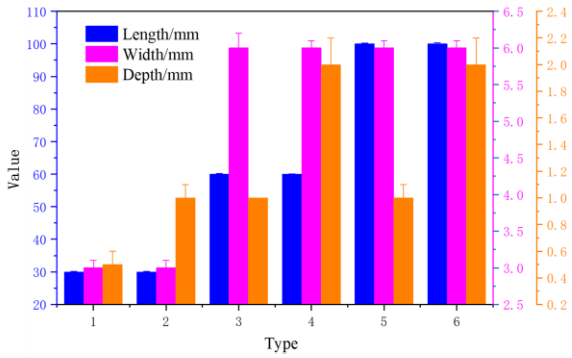



**FIGURE 36.** The results and error comparisons for the five cracks in each direction.

on the surface of a structure, the damaged area is depressed to the inside, whose color distribution will differ from the others. The red color indicates that the region contains depressed cracks and the darker the red color, the deeper the depression and the more severe the damage.

The point cloud distance comparison of the carton structure before and after the damage is shown in Figure 35. In the depth direction, for example, depths of 0.5 mm and 2 mm are clearly displayed in the point cloud distance comparison (Figure 35) in the corresponding color. The length, width and depth parameters of the cracks are obtained from the point

cloud distance. The specific values of the crack parameters are shown in Table 11, it can be seen that the measurement errors of crack depth and width are within 0.2 mm, and the maximum measurement error of length is 0.3 mm. The results and error comparisons for the six cracks in each direction (length, width and depth) are shown in Figure 36, the experiment result demonstrates that by comparing the point cloud distances of the models before and after the damage, the UAV-based 3D reconstruction method allows the identification of different degrees of structural crack damage.

## VI. CONCLUSION

A UAV-based 3D reconstruction detection and identification method for gantry crane main girder structural defects is proposed in this study, which realizes non-contact 3D reconstruction of structural deformations and cracks, thereby reducing the costs and potential unsafe factors involved in the manual inspection. The major conclusions are as follows:

(1) For the structural characteristics of the crane main girder and the flight requirements of the UAV, the UAV flight path applicable to the chamfered rectangular section was selected, and the calculation equations for each parameter in the flight strategy were derived, which provided a high-quality sequential image sequence for structural 3D reconstruction, so as to indirectly guarantee the accuracy of the point cloud model and the precision of the detection results.

(2) Based on the M3C2 algorithm, the difference between the point cloud model and the reference model was compared and analyzed in conjunction with the *RMSE* evaluation index. With the carton structure as the object, a series of possible factors affecting the accuracy of the 3D point cloud model, including the image number, the number and location of marker points, etc. were respectively analyzed, which improved the scientificity of parameter selection, such as image overlap, working distance and other parameters in the experimental scheme of crane main girder defect detection. Furthermore, the influence of the measured structural size effects on the model accuracy was analyzed, which proved that the error of the 3D point cloud model, the actual crane main girder structure, is in the millimeter range, and meets the accuracy of experimental requirements.

(3) The defect detection and identification method proposed in this study was adopted for deformation detection of the gantry crane main girder, and the detection result of mid-span deformation was 12mm with 8mm error, and the measurement accuracy reached a millimeter level. Furthermore, through scaled carton crack identification experiments, the length and depth parameters of cracks were effectively identified with an error of about 0.2mm. The two experiments jointly verify the accuracy and effectiveness of the method proposed in this study, which could serve for structural health monitoring of the crane and other large machinery equipment.

Admittedly, as a preliminary exploration of UAV-based image 3D reconstruction technology in crane structural health detection, in which there are some deficiencies and limitations, this study.

(1) The research object is only oriented to the main girder of the crane, and the rest of the entire crane structure is not considered in the scope of detection.

(2) Since the main girder structure is relatively simple compared to the entire structure, there is no consideration of route plan and obstacle avoidance in GPS-denied areas during the UAV flight process [21], and the structural health monitoring of infrastructure is an inevitable trend for the subsequent development based on the autonomous UAV flight.

(3) The 3D point cloud model in this study was built offline, so the real-time structural defect detection is poor. For future research, the modeling software should be embedded into the UAV system in the form of hardware so that the real-time performance can be improved.

## REFERENCES


[1] L. Kovanic, J. Gasinec, L. Kovanic, and P. Lechman, "Geodetic surveying of crane trail space relations," *Acta Montanistica Slovaca*, vol. 15, no. 3, pp. 188–199, 2010.

[2] P. A. Krahl, D. D. O. Martins, R. Carrazedo, I. D. Silva, and M. K. E. Debs, "Experimental and analytical studies on the lateral instability of UHPFRC beams lifted by cables," *Compos. Struct.*, vol. 209, pp. 652–667, Feb. 2019.

[3] J. M. García, J. L. Martínez, and A. J. Reina, "Bridge crane monitoring using a 3D LiDAR and deep learning," *IEEE Latin Amer. Trans.*, vol. 21, no. 2, pp. 207–216, Feb. 2023.

[4] X. Zhou, J. Wang, X. Mou, X. Li, L. Xie, and X. Feng, "Robust and high-precision vision system for deflection measurement of crane girder with camera shake reduction," *IEEE Sensors J.*, vol. 21, no. 6, pp. 7478–7489, Mar. 2021.

[5] D. Feng and M. Q. Feng, "Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection—A review," *Eng. Struct.*, vol. 156, pp. 105–117, Feb. 2018.

[6] R. S. Adhikari, O. Moselhi, and A. Bagchi, "Image-based retrieval of concrete crack properties for bridge inspection," *Autom. Construct.*, vol. 39, pp. 180–194, Apr. 2014.

[7] X. Ji, Y. Zhuang, Z. Miao, and Y. Cheng, "Vision-based seismic damage detection and residual capacity assessment for an RC shaking table test structure," *Earthq. Eng. Structural Dyn.*, vol. 52, no. 3, pp. 806–827, Mar. 2023.

[8] Y.-J. Cha, W. Choi, G. Suh, S. Mahmoudkhani, and O. Büyüköztürk, "Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 33, no. 9, pp. 731–747, Sep. 2018.

[9] N. Wang, Q. Zhao, S. Li, X. Zhao, and P. Zhao, "Damage classification for masonry historic structures using convolutional neural networks based on still images," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 33, no. 12, pp. 1073–1089, Dec. 2018.

[10] Q. Wang and Z. Zhao, "An accurate and stable pose estimation method based on geometry for port hoisting machinery," *IEEE Access*, vol. 7, pp. 39117–39128, 2019.

[11] S. Hu, H. Qiu, X. Wang, Y. Gao, N. Wang, J. Wu, D. Yang, and M. Cao, "Acquiring high-resolution topography and performing spatial analysis of loess landslides by using low-cost UAVs," *Landslides*, vol. 15, no. 3, pp. 593–612, Mar. 2018.

[12] M. R. Jahanshahi and S. F. Masri, "Adaptive vision-based crack detection using 3D scene reconstruction for condition assessment of structures," *Autom. Construct.*, vol. 22, pp. 567–576, Mar. 2012.

[13] L. Cheng, Y. Wu, Y. Wang, L. Zhong, Y. Chen, and M. Li, "Three-dimensional reconstruction of large multilayer interchange bridge using airborne LiDAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 2, pp. 691–708, Feb. 2015.

[14] X. Zheng and Y. G. Hu, "Research on the factors affecting the accuracy of three-dimensional reconstruction model of rotor UAV archaeological sites," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vols. XLII–4/W20, pp. 111–118, Nov. 2019.

[15] Z. A. Siddiqui and U. Park, "A drone based transmission line components inspection system with deep learning technique," *Energies*, vol. 13, no. 13, p. 3348, Jun. 2020.
[16] M. Martorelli, C. Pensa, and D. Speranza, "Digital photogrammetry for documentation of maritime heritage," *J. Maritime Archaeology*, vol. 9, no. 1, pp. 81–93, Jun. 2014.
[17] A. K. Singh, A. Swarup, A. Agarwal, and D. Singh, "Vision based rail track extraction and monitoring through drone imagery," *ICT Exp.*, vol. 5, no. 4, pp. 250–255, Dec. 2019.
[18] Z. Zhou, J. Gong, and M. Guo, "Image-based 3D reconstruction for posthurricane residential building damage assessment," *J. Comput. Civil Eng.*, vol. 30, no. 2, Mar. 2016, Art. no. 04015015.
[19] Y. Liu, X. Nie, J. Fan, and X. Liu, "Image-based crack assessment of bridge piers using unmanned aerial vehicles and three-dimensional scene reconstruction," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 35, no. 5, pp. 511–529, May 2020.
[20] F. Ioli, A. Pinto, and L. Pinto, "UAV photogrammetry for metric evaluation of concrete bridge cracks," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vols. XLIII–B2, pp. 1025–1032, May 2022.
[21] R. Ali, D. Kang, G. Suh, and Y.-J. Cha, "Real-time multiple damage mapping using autonomous UAV and deep faster region-based neural networks for GPS-denied structures," *Autom. Construct.*, vol. 130, Oct. 2021, Art. no. 103831.
[22] Q. Zhou, S. Ding, G. Qing, and J. Hu, "UAV vision detection method for crane surface cracks based on faster R-CNN and image segmentation," *J. Civil Structural Health Monitor.*, vol. 12, no. 4, pp. 845–855, Aug. 2022.
[23] P. Martínez-Carricondo, F. Agüera-Vega, F. Carvajal-Ramírez, F.-J. Mesas-Carrascosa, A. García-Ferrer, and F.-J. Pérez-Porras, "Assessment of UAV-photogrammetric mapping accuracy based on variation of ground control points," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 72, pp. 1–10, Oct. 2018.
[24] F. Agüera-Vega, F. Carvajal-Ramírez, and P. Martínez-Carricondo, "Accuracy of digital surface models and orthophotos derived from unmanned aerial vehicle photogrammetry," *J. Surveying Eng.*, vol. 143, no. 2, May 2017, Art. no. 04016025.
[25] S. Zhao, F. Kang, J. Li, and C. Ma, "Structural health monitoring and inspection of dams based on UAV photogrammetry with image 3D reconstruction," *Autom. Construct.*, vol. 130, Oct. 2021, Art. no. 103832.
[26] D. Al-Halbouni, E. P. Holohan, L. Saberi, H. Alrshdan, A. Sawarieh, D. Closson, T. R. Walter, and T. Dahm, "Sinkholes, subsidence and subrosion on the eastern shore of the dead sea as revealed by a close-range photogrammetric survey," *Geomorphology*, vol. 285, pp. 305–324, May 2017.
[27] J. Li, B. Yang, C. Chen, and A. Habib, "NRLI-UAV: Non-rigid registration of sequential raw laser scans and images for low-cost UAV LiDAR point cloud quality improvement," *ISPRS J. Photogramm. Remote Sens.*, vol. 158, pp. 123–145, Dec. 2019.
[28] F. He and A. Habib, "Three-point-based solution for automated motion parameter estimation of a multi-camera indoor mapping system with planar motion constraint," *ISPRS J. Photogramm. Remote Sens.*, vol. 142, pp. 278–291, Aug. 2018.
[29] O. Tziavou, S. Pytharouli, and J. Souter, "Unmanned aerial vehicle (UAV) based mapping in engineering geological surveys: Considerations for optimum results," *Eng. Geol.*, vol. 232, pp. 12–21, Jan. 2018.
[30] V. Zahs, L. Winiwarter, K. Anders, J. G. Williams, M. Rutzinger, and B. Höfle, "Correspondence-driven plane-based M3C2 for lower uncertainty in 3D topographic change quantification," *ISPRS J. Photogramm. Remote Sens.*, vol. 183, pp. 541–559, Jan. 2022.
[31] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
[32] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
[33] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "KAZE features," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2012, pp 214–227.
[34] C. Jian, L. Cong, J. Wu, H. Cui, and H. Lu, "Fast and accurate image matching with cascade hashing for 3D reconstruction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1–8.
[35] B. Triggs, P. F. Mclauchlan, R. I. Hartley, and A. W. Fitzgibbon, *Bundle Adjustment—A Modern Synthesis*. Berlin, Germany: Springer, 2000.
[36] R. Eker, "Comparative use of PPK-integrated close-range terrestrial photogrammetry and a handheld mobile laser scanner in the measurement of forest road surface deformation," *Measurement*, vol. 206, Jan. 2023, Art. no. 112322.
[37] D. Lague, N. Brodu, and J. Leroux, "Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei Canyon (N-Z)," *ISPRS J. Photogramm. Remote Sens.*, vol. 82, pp. 10–26, Aug. 2013.
[38] J. Park, P. Kim, Y. K. Cho, and J. Kang, "Framework for automated registration of UAV and UGV point clouds using local features in images," *Autom. Construct.*, vol. 98, pp. 175–182, Feb. 2019.

**ZHIPING LIU** received the B.S. and M.S. degrees from the Hubei University of Technology, in 1997 and 2000, respectively, and the Ph.D. degree from the Huazhong University of Science and Technology, in 2003. He is currently a Professor and a Ph.D. Supervisor with the Wuhan University of Technology. His main research interests include structural health monitoring and nondestructive testing and evaluation.

**YANNAN YU** received the master's degree from the Taiyuan University of Science and Technology, Taiyuan, China, in 2022. He is currently pursuing the Ph.D. degree with the Wuhan University of Technology. His main research interests include crane modern design theory and design method research, structural optimization design, and structural health monitoring.

**ZHUOHUI LIANG** received the bachelor's degree from the Shandong University of Technology, Zibo, China, in 2020, and the master's degree from the Wuhan University of Technology, Wuhan, China, in 2023. His main research interests include crane structure health monitoring and three-dimensional reconstruction.

**GUODONG HAN** received the bachelor's degree from Yangtze University, Jingzhou, China, in 2021. He is currently pursuing the master's degree with the Wuhan University of Technology. His main research interests include crane structure health monitoring and crane digital twin.

**YAO LU** received the bachelor's degree from Changsha University, Changsha, China, in 2021. He is currently pursuing the master's degree with the Wuhan University of Technology. His main research interests include crane structure health monitoring and crane failure mechanism.