Three-dimensional reconstruction and crack measurement of cultural monuments using UAV-based photogrammetry

Wei-Che Huang [1], Wen-Cheng Liu [2], Yi-Shan Luo [3], Po-Yu Chen [4], Kuei-Luo Lin [5]

[1] Department of Civil and Disaster Prevention Engineering, National United University, Miaoli 360302, Taiwan - e11856824@gmail.com.tw
[2] Department of Civil and Disaster Prevention Engineering, National United University, Miaoli 360302, Taiwan - wcliu@nuu.edu.tw
[3] Department of Civil and Disaster Prevention Engineering, National United University, Miaoli 360302, Taiwan - s13saysomething@gmail.com
[4] Shin-Mag Industrial Co., Ltd., Changhua 50444, Taiwan - jason0823183@gmail.com
[5] Fullai Construction Co., Ltd., Taoyuan 333001, Taiwan - a90302155031@gmail.com

Keywords: Unmanned Aerial Vehicle (UAV), SfM Photogrammetry, Cultural Monument, 3D Reconstruction, Crack Measurement.

Abstract

Three-dimensional (3D) modeling for the documentation, preservation, and management of cultural heritage is indispensable. To achieve this goal, a low-cost unmanned aerial vehicle (UAV) combined with the Structure from Motion (SfM) photogrammetric technique was utilized to build a 3D model and conduct surface crack measurements of cultural monuments. The results showed that, under simple conditions, non-specialists can easily generate accurate 3D models from UAV-acquired imagery. In this study, the statistical errors of checkpoints between 3D reconstruction and field measurements, expressed as total RMSE, ranged from 0.103 m to 0.848 m. However, the mean absolute errors of surface crack measurements between tape-based methods and 3D reconstruction ranged from 0.002 m to 0.099 m. Furthermore, UAV-SfM was applied to measure surface crack lengths on an inaccessible cultural monument. The findings demonstrated that employing the UAV-SfM photogrammetric technique for 3D reconstruction of cultural monuments is both feasible and reliable.

1. Introduction

Cultural heritage refers to assets of historical, artistic, scientific, or other cultural value that are designated or registered by governmental authorities or international organizations. Tangible cultural heritage is highly vulnerable to natural hazards such as fires, floods, earthquakes, landslides, and typhoons, as well as anthropogenic threats including arson and war. These events can lead to structural instability, increased fragility, or even severe deterioration and collapse (Frodella et al., 2020; Dasari et al., 2021; Konstantinidis et al., 2021). In recent years, unmanned aerial vehicle (UAV) have been widely employed for the three-dimensional reconstruction of historical and cultural assets. Their popularity stems from advantages such as low equipment costs, high operational flexibility, high-resolution imagery, reduced personnel risk, minimal environmental impact, high dimensional accuracy, and compatibility with diverse UAV platforms (Stek, 2016; Martin-Bejar et al., 2020; Block-Berlitz et al., 2022). When integrated with Structure from Motion (SfM), UAV can be applied to object modeling, three-dimensional (3D) surface generation, point cloud construction, and the digital documentation of cultural heritage sites and buildings (Bakirman et al., 2020). The combination of UAV- and SfM-based photogrammetry with conventional terrestrial techniques enables comprehensive coverage of elevated or inaccessible structures, producing complete 3D models of cultural heritage sites and their surroundings without omissions (Hidayat and Cahyono, 2016; Sang et al., 2022; Partama et al., 2025).

In this study, UAV was used to acquire imagery and data for three-dimensional reconstruction, while crack measurements were conducted by non-specialists. Case studies were carried out on cultural heritage monuments in Miaoli County, Taiwan. UAV-based photogrammetry combined with SfM was employed to identify appropriate flight altitudes, camera orientations, and image quantities to achieve accurate 3D modeling and reconstruction. Model accuracy was validated using ground checkpoints and crack measurements from three historic buildings. The proposed approach was further applied to the Neishechuan Iron Bridge, a site that is difficult to access using conventional survey methods. By integrating UAV photogrammetry and SfM techniques, the bridge was comprehensively documented, and the locations and dimensions of structural cracks were accurately identified.

2. Materials and Methods

This study adopts a UAV-based photogrammetry approach to acquire images of historical buildings, reconstruct 3D models, and subsequently measure surface cracks using the reconstructed models. Compared with conventional methods, the proposed measurement approach significantly reduces inspection time and eliminates safety risks associated with climbing by field personnel. The overall workflow consists of five main steps, as illustrated in Figure 1. (1) Layout and measurement of ground control points (GCPs): GCPs were surveyed using a total station to enhance the accuracy of the 3D reconstruction, establish the scale conversion between model pixels and physical dimensions, and validate the modeling accuracy. (2) Image acquisition: Following the deployment and measurement of GCPs, a UAV was used to capture images of the historical buildings. (3) Optimization of the UAV imaging conditions: The effects of different numbers of images, flight altitudes, and camera viewing angles on the quality of the 3D reconstruction were investigated. (4) 3D model reconstruction of cultural heritage structures: The acquired images were processed using Agisoft PhotoScan to generate dense point clouds and reconstruct 3D models of the historical buildings. (5) Surface crack measurement and validation: Surface cracks were

https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-841-2026

measured directly from the reconstructed 3D models, and the results were compared with in situ measurements obtained using a tape measure.

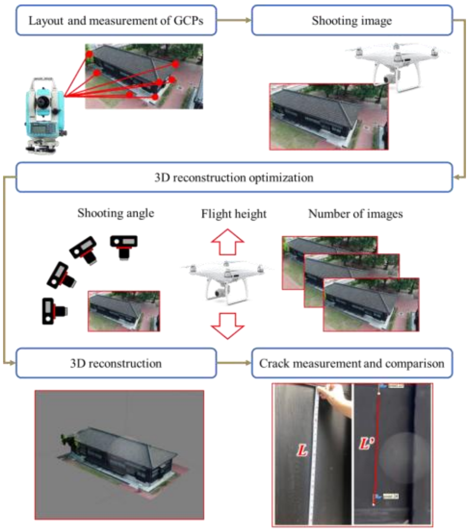

Figure 1. Flowchart followed in this research

The UAV flies by manual operation. To obtain the images, the UAV is operated to the specified height, and adjust the shooting angle of the camera with 180° in a straight line or a circular line with shooting angle of the camera with 135° (see Figure 2). The shooting angle of 180° means that the camera shoots vertically downwards, and the captured image is the roof of the building (a1-a4 in Figure 2a). The shooting angle of 135° is to shoot the facade of the building. When flying in a loop, the camera shooting direction is facing the facade of the building. Therefore the facade of the building can be photographed (b1-b4 in Figure 2b). The shooting angle of 135° was chosen because it can simultaneously capture both the roof and the facade of the building. The ground control points can be divided into two parts: one part is used as a reference point and the other part is served as a check point (for the number and location of control points and check points, refer to Description of Study Sites). After producing the 3D point cloud, select the point cloud representing the reference points, input the coordinates measured by the total station, and after the software recalculates again, the 3D point cloud with the coordinate unit in meters can be produced. Then click on the point cloud representing the check point, read the point cloud coordinates calculated by the software, and compare it with the coordinates of the check point measured by the total station to validate the accuracy of the 3D modeling reconstruction.

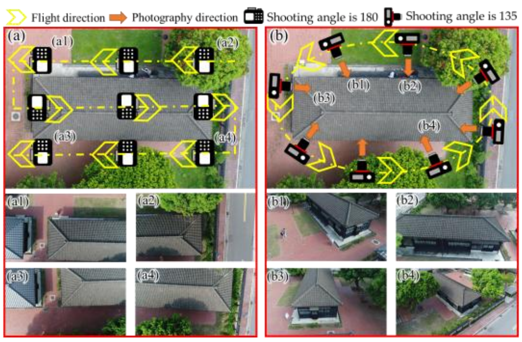

Figure 2. Flight path of UAV with (a) straight line and (b) circular line

2.1 Description of Study Sites

The research sites of this study include four cultural monuments in Miaoli County, Taiwan, as shown in Figure 3a, which are Lai's Festival Filial Piety, Shanjiao Elementary School Japanese-style Dormitory, Tasi's Jiyang Hall, and Neishechuan Iron Bridge. The Shanjiao Elementary School Japanese-style Dormitory was built in 1937 and is located in Yuanli town, Miaoli County. It is a group of traditional dormitories in the late Japanese occupation constructed with wooden materials. It was announced as a historical building in 2003. A total of ten control points (i.e., points 1 to 10) and five check points (i.e., points 11 to 15) were used for calibration (Figure 3b). Lai's Festival Filial Piety was built in 1883 and is located in Miaoli City, Miaoli County. It was originally announced as a third-level historic site on August 19, 1985, and later became a Miaoli County-designated historic monument. A total of eight control points (i.e., points 1 to 8) and four check points (i.e., points 9 to 12) were set up, as shown in Figure 3c. Tasi's Jiyang Hall built in 1840 is an ancestral house located at the foot of the mountain in Yuanli town. It is adjacent to the Japanese dormitory group of the elementary school. A total of eight control points (i.e., points 1 to 8) and six check points (i.e., points 9 to 14) were set up in this study (Figure 3d). The Neishechuan Iron Bridge was constructed in 1908 and decommissioned in 1998. It is located in Sanyi Township, Miaoli County. The bridge has a total length of 173.73 m, spans the Jingshan River, and is part of the old mountain line of the Taiwan Railway. The bridge was built with oval hollow reinforced concrete piers and employed a truss girder design, which, at the time of its construction, represented a unique case of the novel sheet-girder bridge structure. It is also the only top-loaded truss bridge built on the Taiwan Railways before World War II. A total of three control points (i.e., points 1 to 3) were established at the site, as shown in Figure 3e. The numbers of control points, check points, and the corresponding RMSE values for the four cultural monuments are summarized in Table 1.

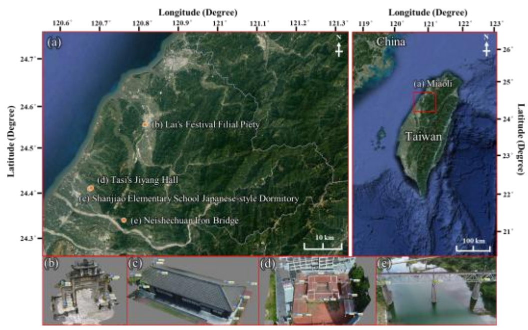

Figure 3. Location of the study sites (a) Shanjiao Elementary School Japanese-style Dormitory, (b) Lai's Festival Filial Piety, (c) Tasi's Jiyang Hall, and (d) Neishechuan Iron Bridge

Name	Quantity		RMSE (mm)
	Control points	Check points	
Shanjiao Elementary School Japanese-style Dormitory	10	5	7.6
Lai's Festival Filial Piety	8	4	2.2
Tasi's Jiyang Hall	8	6	11.6
Neishechuan Iron Bridge	3	0	3.1

Table 1. Summary of control points and check points data

2.2 UAV and Total Station

The total station and UAV were employed to measure the ground control points (GCPs) and acquire images, respectively. The total station used in this study was a Nikon DTM 332, with a measurement accuracy of $1 \pm (2 + 2 \times \text{distance} \times 10^{-6})$ mm and an angular accuracy of 2″. The instrument is equipped with a horizontal bubble device, weighs 4.9 kg, and has a maximum measuring distance of 2000 m.

To avoid potential damage to historic structures, no GCP markers were painted or affixed to the buildings. Instead, distinct architectural features, such as wall corners or construction endpoints, were selected as GCPs. Owing to the relatively small size of the investigated buildings, a custom coordinate system was adopted, with length units expressed in meters, and the coordinates were not converted to a geodetic system.

Each GCP was measured three times, and errors caused by human factors were excluded by analyzing the differences among repeated measurements. Based on total station data, the root mean square errors (RMSE) of the GCPs for Lai’s Festival Filial Piety, the Shanjiao Elementary School Japanese-style Dormitory, Tasi’s Jiyang Hall, and the Neishechuan Iron Bridge were 2.2 mm, 7.6 mm, 11.6 mm, and 3.1 mm, respectively.

The UAV employed was a DJI Phantom 4 Pro, weighing 1388 g including batteries, propellers, and the onboard camera. The camera features an 84° field of view (FOV) with autofocus capability, a 1-inch CMOS sensor with a resolution of 20 megapixels, and a three-axis gimbal stabilization system to ensure image stability. The gimbal has a controllable pitch range from −90° to +30°, with a maximum pitch control speed of 90°/s and an angular control accuracy of approximately ±0.02°.

2.3 Structure from Motion Photogrammetry

SfM is a restore continuous two-dimensional images into three-dimensional structures through the principle of motion parallax. Motion parallax means that when the observer (i.e., camera) is moving, objects at different distances from the observer move at different distances in the observer's vision. Figure 4 illustrates that three cameras shoot the objects at different positions. The vertical distances between the object and the camera are L1, L2, and L3, and L3 is greater than L2, which is greater than L1. The camera in the center shoots the image of the object. The object is in the middle of the horizontal direction of the image. When the camera moves to the left and right, the horizontal position of the object in the image is also different. The positions of the objects at distance L1 in the images captured by the left and right cameras are shifted by DL1 and DR1, respectively. Object at distance L2 is shifted by DL2 and DR2, and the object at distance is shifted by DL3 and DR3. Among the distance relationships, DL3 is smaller than DL2, which is smaller than DL1, while DR3 is smaller than DR2, which is smaller than DR1.

To establish the correspondence between continuous images, when shooting images, the overlap rate of the left and right images (i.e., forward lateral) should be at least 60%, and the overlap rate of the front and rear images (i.e., frontal side) should reach more than 80% to obtain better results (Erenoglua et al., 2017; Seifert et al., 2019). Therefore, when discussing the influence of flight height and shooting angle on the accuracy of the 3D model, the image overlap rate is maintained above 80%. In the section discussing the number of images, the number of images significantly affects the overlap rate; thus, when the number of images is 30, 60, and 120, the overlap rates are 55%, 69%, and 83%, respectively.

This study uses Agisoft PhotoScan software (version 1.4.5) for 3D reconstruction, which can perform photogrammetry processing on digital images, generate 3D spatial data, and build 3D models. The software imports ground control points and improves the accuracy of the 3D model, which requires the measurement of spatial data such as the co-ordinates of ground points and object length, area, or volume. For the parameters adopted to generate the 3D model, we used medium for quality, aggressive for depth filtering, 4096 for texture, 40000 for key point limit, and 4000 for tie point limit.

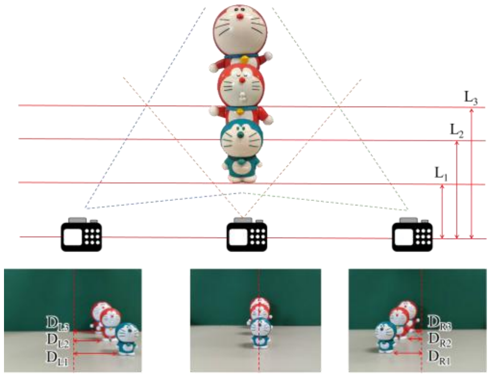

Figure 4. Schematic diagram of motion parallax

2.4 Accuracy Evaluation

To assess the 3D point cloud reconstruction, the accuracy evaluation in Easting (X), Northing (Y), and height (Z) was conducted on the check points and their coordinates were compared with the check point coordinates measured by the total station. The equations for RMSE X, RMSE Y, RMSE Z, and total RMSE can be expressed as:

$$RMSE\ \ X = \sqrt{\frac{\sum_{i=1}^{n}(X_{SfM,i} - X_{TS,i})^2}{n}} \quad (1)$$

$$RMSE\ \ Y = \sqrt{\frac{\sum_{i=1}^{n}(Y_{SfM,i} - Y_{TS,i})^2}{n}} \quad (2)$$

$$RMSE\ \ Z = \sqrt{\frac{\sum_{i=1}^{n}(Z_{SfM,i} - Z_{TS,i})^2}{n}} \quad (3)$$

$$\text{total } RMSE = \sqrt{\frac{(RMSE\ \ X)^2 + (RMSE\ \ Y)^2 + RMSE\ \ Z)^2}{3}} \quad (4)$$

where n = denotes the number of check/georeferencing points
$X_{SfM,i}$, $Y_{SfM,i}$, $Z_{SfM,i}$ = the coordinates measured for georeferencing points in the X, Y, and Z directions
$X_{TS,i}$, $Y_{TS,i}$, $Z_{TS,i}$ = the coordinates measured for check points in the X, Y, and Z directions

3. Result

3.1 Difference Flight Height of UAV

To obtain the optimal UAV shooting conditions, this study used the Shanjiao Elementary School Japanese-style Dormitory as the case site for testing. During the 3D modeling reconstruction process, different numbers of shooting images, flight heights, and shooting angles were applied, and the statistical errors of the coordinates of the check points measured by the total station and by 3D modeling were compared to determine the optimal conditions.

This study examined the effects of UAV flight heights of 10 m, 20 m, and 30 m, with a flight speed of 1 m/s, on 3D modeling accuracy. Flight height was calculated from the highest point of the building. The camera angle was 180°, i.e., vertically downward. The UAV flew in straight lines with three flight strips to ensure optimal image capture at the edges of the building. Images were extracted from 24 fps video, with one frame selected from every 12 frames. The overlap rate of images at different heights ranged from 80% to 90%.

The 3D reconstruction results showed that as UAV flight height increased, the errors of the check points gradually decreased (see Figure 5). At a flight height of 10 m, the major error occurred along the X-axis, ranging from −1.578 m to 8.765 m, indicating instability of the model in the X-axis. The Y-axis error ranged from −0.932 m to −0.895 m, with a median of −0.911 m, suggesting a consistent offset of about 0.9 m. The Z-axis error showed a similar pattern, ranging from 0.765 m to 1.456 m, with a median of 1.157 m, indicating a difference of about 0.7 m and an overall offset of about 1.1 m along the Z-axis.

At a flight height of 20 m, the X-axis error was slightly reduced compared with 10 m, ranging from −1.325 m to 7.66 m, with an overall offset of about 2.6 m. At 30 m, the errors along the X- and Z-axes were significantly reduced. The X-axis error ranged from −0.095 m to 1.4 m, while the Z-axis error ranged from −0.157 m to 0.168 m. The significant reduction in X-axis error was due to an increased overlap of images covering check points 13 and 14 near the edge of the building. At lower flight heights, the images from the second flight strip did not cover these check points; however, at 30 m, the second strip captured them, thereby improving the 3D model accuracy in this area.

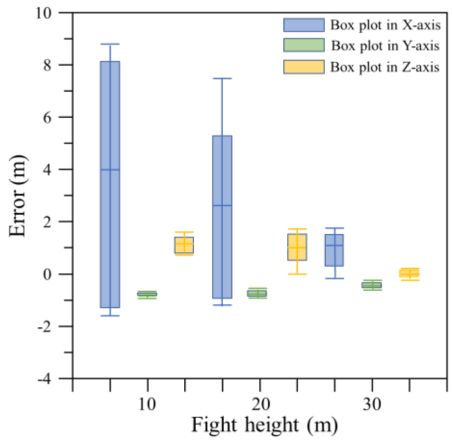

Figure 5. Statistical errors for box plot in X-axis, Y-axis, and Z-axis with different UAV flight heights

3.2 Difference Shooting Angle of UAV

To comprehend the influence of UAV shooting angle on 3D modeling accuracy, the angles were set to 135° and 180°. The UAV was flown at a height of 30 m, with a flight speed of 1 m/s and a shooting frame rate of 2 fps. A total of 120 images were used for 3D reconstruction. With a shooting angle of 180°, the UAV flew in a straight path with three flight strips, whereas at 135°, the UAV followed a circular path along the perimeter of the façade (see Figure 2).

The errors of the 3D reconstructions with different shooting angles are shown in Figure 6. The results indicated that the 3D model constructed at a shooting angle of 135° was more accurate than that at 180°. At 180°, the X-axis error ranged from −0.095 m to 1.532 m, with an overall offset of about 0.86 m. The Y-axis error ranged from −0.632 m to −0.322 m, with a median of −0.512 m. The Z-axis error ranged from −0.157 m to 0.168 m, with a median of 0.086 m. At 135°, the errors along all three axes were significantly reduced: the X-axis error ranged from −0.247 m to 0.089 m (median −0.05 m), the Y-axis error ranged from −0.124 m to 0.105 m (median 0.009 m), and the Z-axis error ranged from −0.033 m to 0.076 m (median 0.009 m).

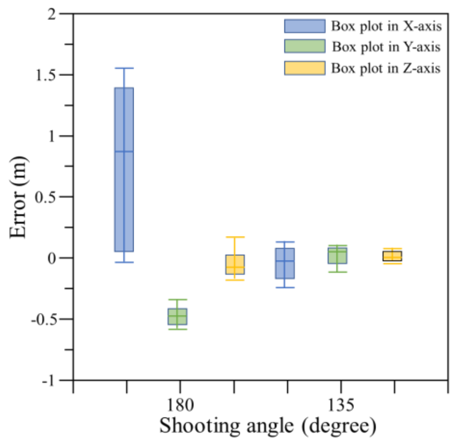

Figure 6. Statistical errors for box plot in X-axis, Y-axis, and Z-axis with different shooting angles from the UAV

3.3 Difference Number of Image

The effect of the number of images on the accuracy of 3D modeling was also investigated. The numbers of images adopted were 30, 60, and 120, with image overlap rates of 55%, 69%, and 83%, respectively. As confirmed in the previous section, a shooting angle of 135° produced better results; therefore, this section used a 135° shooting angle, and the UAV followed a circular flight path. At least two images captured the edges of the historic building for 3D modeling. The UAV flight conditions were a height of 30 m, a speed of 1 m/s, and a shooting angle of 135°.

Figure 7 shows the errors of the check points in 3D reconstructions built with different numbers of images. The results indicated that the greater the number of images used, the higher the accuracy of the 3D model. Conversely, when the number of images was too small and the overlap insufficient, the accuracy of the 3D model was unstable. When 30 images were used, the main error occurred along the Z-axis, ranging from −0.752 m to 0.269 m. With 60 images, the main error shifted to the X-axis, ranging from −0.601 m to 0.243 m, while the Z-axis error ranged only from −0.009 m to 0.279 m. Overall, the errors for the 3D model constructed with 120 images were much smaller than those using 30 or 60 images. Most check points were located in the central area of the images, and the connection strength among consecutive images was sufficient. Therefore, under a flight altitude of 30 m and a shooting angle of 135°, the median errors along the X-, Y-, and Z-axes when using 120 images were only −0.05 m, 0.09 m, and −0.004 m, respectively.

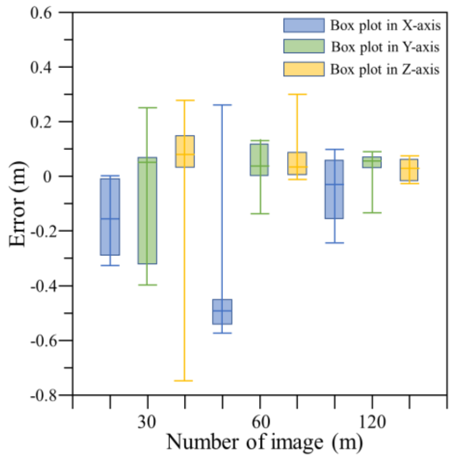

Figure 7. Statistical errors for box plot in X-axis, Y-axis, and Z-axis with different numbers of images

3.4 Validation of 3D Reconstruction and Crack Measurement

After conducting the optimization test of the 3D model at the Shanjiao Elementary School Japanese-style Dormitory, 3D reconstructions of three cultural monuments-Shanjiao Elementary School Japanese-style Dormitory, Lai's Festival Filial Piety, and Tasi's Jiyang Hall-were carried out under the optimized conditions. To verify the measurement results, three surface crack measurements were performed for each cultural monument.

Figure 8 presents the comparison between crack length measurements obtained using 3D reconstruction and a tape measure. Figure 8a shows the crack length measured with a tape, while Figure 8b illustrates the measurement using 3D reconstruction. By reading the coordinates of the two ends of the crack in the 3D model and calculating the distance, the crack length was obtained. The UAV flight conditions were a height of 30 m, a shooting angle of 135°, and 120 images, which were used to reconstruct the 3D models of the Shanjiao Elementary School Japanese-style Dormitory, Lai's Festival Filial Piety, Tasi's Jiyang Hall, and the Neishechuan Iron Bridge (Figure 9). The statistical errors for RMSE X, RMSE Y, RMSE Z, and total RMSE were 0.145 m, 0.096 m, 0.04 m, and 0.103 m, respectively (Figure 10). The surface crack lengths measured with tape were 0.605 m, 0.69 m, and 0.67 m, while the results using the 3D model were 0.614 m, 0.666 m, and 0.649 m (S1, S2, and S3 in Figure 11). The mean absolute error between both methods was 0.018 m, indicating that the 3D model accurately measured crack lengths.

Under the same conditions, the 3D model of Lai's Festival Filial Piety was reconstructed. The statistical errors for RMSE X, RMSE Y, RMSE Z, and total RMSE were 0.022 m, 0.035 m, 0.211 m, and 0.124 m, respectively (Figure 10). The crack lengths measured with tape at three locations (L1, L2, and L3) were 0.295 m, 0.185 m, and 0.097 m, while the 3D model results were 0.247 m, 0.224 m, and 0.115 m (Figure 11). The differences ranged from 0.018 m to 0.048 m, all within 5 cm.

For Tasi's Jiyang Hall, the statistical errors of the check points were 0.806 m, 0.498 m, 1.123 m, and 0.848 m for RMSE X, RMSE Y, RMSE Z, and total RMSE, respectively (Figure 10). The surface crack lengths measured with tape at three locations (T1, T2, and T3) were 0.90 m, 0.63 m, and 0.95 m, while the results from the 3D model were 0.999 m, 0.664 m, and 0.948 m (Figure 11). The differences were 0.099 m, 0.034 m, and 0.002 m at T1, T2, and T3, respectively. Across the three historical buildings, the crack length error between the 3D model and tape measurements was less than 0.1 m, with an average error of 0.033 m. These results demonstrate that crack measurements using 3D reconstruction are feasible, with centimetre-level accuracy, and suggest that this technique can be applied to inaccessible railway bridges where direct georeferencing is difficult.

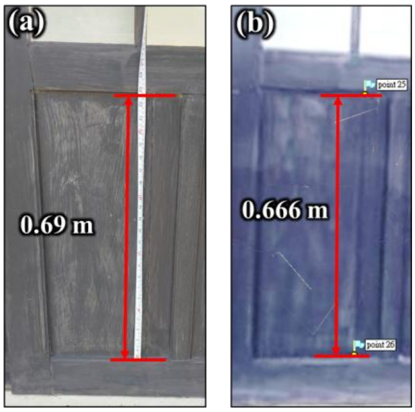

Figure 8. Measurement of crack length using (a) tape and (b) 3D reconstruction

Figure 9. 3D reconstruction using UAV-SfM in the (a) Japanese-style dormitory of Shanjiao Elementary School, (b) Lai's Festival Filial Piety, (c) Tasi's Jiyang Hall, and (d) Neishechuan Iron Bridge

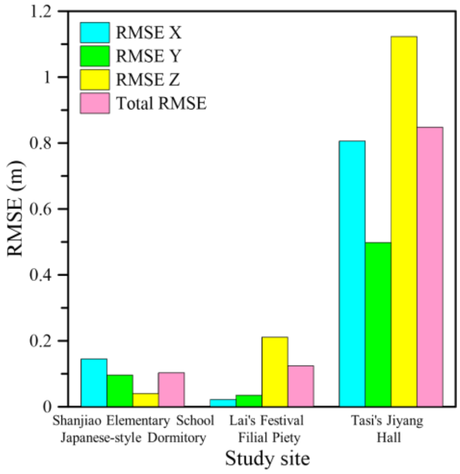

Figure 10. Statistical errors of RMSE X, RMSE Y, RMSE Z, and total RMSE for three cultural monuments

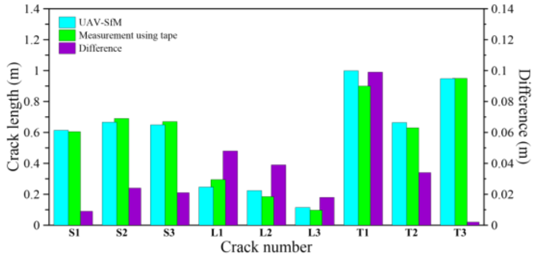

Figure 11. Measurement of surface crack length and difference using UAV-SfM and tape. Note that S1, S2, and S3 denote three locations in Shanjiao Elementary School Japanese-style dormitory; L1, L2, and L3 represent three locations in Lai's Festival Filial Piety; T1, T2, and T3 indicate three locations in Tasi's Jiyang Hall

3.5 Application to Measure Crack at Inaccessible Cultural Monument

The 3D reconstruction and crack measurements were validated to demonstrate the feasibility of applying UAV-SfM to measure surface crack length in three cultural monuments. UAV-SfM for 3D modeling and crack measurement was further applied to the Neishechuan Iron Bridge, as shown in Figure 12. Because the Neishechuan Iron Bridge spans the Jingshan River, it is currently idle and closed; therefore, it is difficult to directly access the bridge piers to measure surface cracks. The total station can perform long-distance measurements, but the surveyor must still place a target or prism on the object. By reflecting the infrared rays of the total station through the target or prism, the coordinates can be accurately measured. Only three control points were established on the Neishechuan Iron Bridge, which were used as input for the 3D model to improve reconstruction accuracy and convert the scale of the model into real-world units.

The 3D model identified three surface cracks on the Neishechuan Iron Bridge, with lengths of 1.205 m (Figure 12b), 1.447 m (Figure 12c), and 2.891 m (Figure 12d). The lengths measured by the total station were 1.195 m, 1.519 m, and 2.856 m, respectively, with differences of 0.01 m, 0.072 m, and 0.035 m compared with the 3D model results. Although the total station can also measure crack length, the limited field of view of its telescope requires significant time to locate and aim at the endpoints of the crack. This further highlights the convenience and importance of using UAVs to capture images, reconstruct 3D models through SfM, and measure surface cracks from the models.

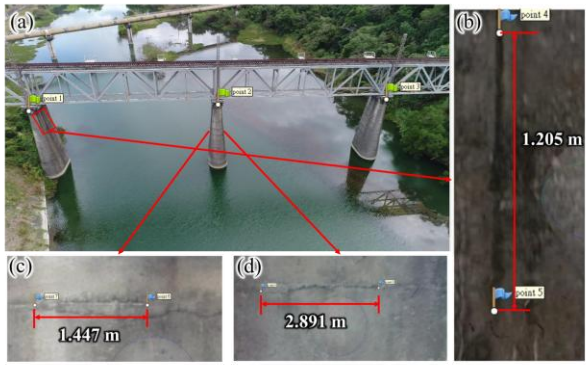

Figure 12. (a) Application of UAV-SfM to build a 3D model for the Neishechuan Iron Bridge and surface crack measurements at (b) the first location (1.205 m), (c) the second location (1.447 m), and (d) the third location (2.891 m)

4. Discussion

4.1 Flight Plan

From the section on different UAV flight heights, it was observed that the higher the flight altitude, the greater the image overlap rate. In the section on different numbers of images, it was found that more images also led to higher overlap. Several studies (Nony et al., 2012; Rakha and Gorodetsky, 2018; Chen et al., 2021) reported that when the image overlap rate exceeds 80%, 3D modeling achieves good results. Furthermore, when more details of the object are captured in the images, the resulting 3D reconstruction is also more detailed, which explains why the modeling results obtained at a shooting angle of 135° (Section 4.2) were better than those at 180°. The 135° angle provided superior reconstructions because it captured more details of the building facade.

To achieve a high overlap rate and capture sufficient detail, several factors must be considered beyond flight height, shooting angle, and image number. These include whether obstacles exist along the planned UAV flight path, whether wind causes image blurring, whether the shooting time is too long, or whether images are taken at different times, leading to inconsistent shadows and difficulties in matching feature points. All these factors may increase 3D modeling errors and even lead to reconstruction failure.

Another factor affecting 3D modeling is the computing platform. In Section 4.3, after recording at 24 fps in a circular flight, about one out of every twelve frames was selected, yielding a total of 120 images for 3D reconstruction. Processing took about 2 hours using an Intel i7-3.0 GHz CPU with 32 GB of memory and no graphics card. However, while more images increase overlap and improve modeling accuracy, they also substantially increase computation time. In this study, acquiring UAV images and performing 3D reconstruction of a historical building required roughly one full day.

4.2 Crack Measurement

Photogrammetry has great potential as a noninvasive, contactless survey technique for detecting surface cracks and improving measurement accuracy in structures (Valenca et al., 2013; Jahanshahi et al., 2013; Lins and Givigi, 2016). Therefore, the UAV-SfM technique was applied to crack measurements in cultural monuments, even when the buildings were inaccessible.

In the example shown in Figure 9a, there were trees around the historic building, but the ground control points and cracks analyzed in this study were located outside areas affected by the trees. Only manual operation was used, and no auxiliary flight software was employed to obtain images for 3D reconstruction and crack measurement. For the three historical buildings, the average error in crack length was 0.033 m. The results indicated that as long as surveyors were trained in UAV operation, no additional tools or software were required to obtain qualified images. Consequently, 3D models could be generated and used for further crack measurements in historical buildings.

It was observed that the statistical error of the check points at Tasi's Jiyang Hall was significantly higher than that at Shanjiao Elementary School Japanese-style Dormitory and Lai's Festival Filial Piety. This was likely due to the more complex structure of Tasi's Jiyang Hall compared to the other two monuments. In 3D modeling with SfM, feature points are tracked across continuous images. When the continuity of these points was disrupted, subsequent images exhibited larger absolute coordinate errors in the 3D point cloud. However, crack length measurement relies on the relative distance between two point clouds. Thus, for the T3 crack at Tasi’s Jiyang Hall, although the absolute coordinates of the surrounding point cloud had large errors, the relative distance between the crack endpoints remained stable.

Compared with tape measurements of surface crack length, the 3D model showed a maximum error of 0.099 m (Figure 11), which was larger than that of Shanjiao Elementary School Japanese-style Dormitory and Lai's Festival Filial Piety. However, the minimum error reached 0.002 m, smaller than the corresponding errors at the other two sites (Figure 11). These results demonstrate that UAV-SfM is a feasible technique for building 3D models to measure crack lengths in cultural monuments. Even when the total RMSE of the check points was as high as 0.848 m, the measurement error of surface cracks remained below 0.01 m.

4.3 Limitations and Future Work

This study employed UAV to acquire imagery of cultural monuments, constructed 3D models using SfM, and subsequently measured surface cracks. The primary advantage of this approach lies in its ability to substantially reduce the time and labor required for crack inspection. However, the method has several limitations. Adequate open space is required to ensure safe UAV operation and image acquisition. For monuments with complex geometries, UAV is unable to capture close-range imagery, resulting in insufficient representation of interior and fine-scale structural details in the reconstructed 3D models. Consequently, the proposed method is suitable for preliminary identification of crack locations and lengths, whereas crack width, depth, and potential structural risks still require in situ inspection by trained professionals.

The 3D models were generated from point clouds derived from image matching. As a result, crack length estimation depends on the selection of target points within the point cloud, with crack length calculated from the coordinates of these points (Figure 8). Owing to the limited density of the point cloud, points rarely coincide exactly with crack endpoints. Instead, the nearest available points were selected, which constitutes a primary source of uncertainty in crack length estimation based on 3D models.

Furthermore, for cultural monuments with complex architectural forms, image acquisition under a single set of flight conditions may be insufficient to ensure model accuracy. For example, in the case of Tasi’s Jiyang Hall, images were collected using only one flight configuration (flight height of 30 m, shooting angle of 135°, and 120 images). As a result, the checkpoint errors of the reconstructed model were larger than those observed for the other monuments, and discrepancies between crack lengths measured from the 3D model and those obtained by tape measurements exhibited greater variability. Future studies will investigate 3D reconstruction strategies that integrate images acquired under multiple flight conditions to improve model accuracy.

In the present study, surface cracks on cultural monuments were identified manually prior to measurement. Future work will focus on integrating machine vision and deep learning techniques (Flah et al., 2020; Yang et al., 2022; Zhu et al., 2022; Huang et al., 2025) with UAV-based photogrammetry to enable automated detection and measurement of surface cracks. In addition, point cloud data acquisition using sensors with higher relative accuracy, such as terrestrial laser scanning (TLS) (Xu et al., 2014; Manajitprasert et al., 2019; Jo and Hong, 2019), aerial LiDAR, and GeoSLAM systems, will be considered to further enhance the reliability of 3D modeling and crack assessment in future research.

5. Conclusions

Cultural heritage serves as a tangible record of history and a critical link between the past and the present; therefore, systematic documentation is essential for historical understanding and future preservation planning. In this study, a low-cost and time-efficient UAV-based photogrammetric approach combined with SfM was employed to perform 3D reconstruction and surface crack measurements for four cultural monuments in Miaoli County, Taiwan. To determine optimal UAV data acquisition parameters, the Japanese-style Dormitory at Shanjiao Elementary School was selected as a case study to evaluate the effects of flight altitude, camera orientation, and image quantity on 3D modeling performance. The results indicate that a flight altitude of 30 m, a shooting angle of 135°, and a total of 120 images provide optimal conditions for accurate 3D reconstruction of cultural monuments.

Using these optimal acquisition parameters, the accuracy of 3D reconstruction and crack measurement was validated for the Shanjiao Elementary School Japanese-style Dormitory, Lai’s Festival Filial Piety, and Tasi’s Jiyang Hall. For the Shanjiao Elementary School Japanese-style Dormitory, the RMSE of the checkpoints were 0.145 m, 0.096 m, and 0.040 m in the X, Y, and Z directions, respectively, with a total RMSE of 0.103 m. For Lai’s Festival Filial Piety, the RMSE in the X, Y, and Z directions were 0.022 m, 0.035 m, and 0.211 m, respectively, yielding a total RMSE of 0.124 m. For Tasi’s Jiyang Hall, the corresponding RMSE were 0.806 m (X), 0.498 m (Y), and 1.123 m (Z), with a total RMSE of 0.848 m, which is slightly larger than those observed for the other monuments. Nevertheless, discrepancies between tape-based and 3D model crack measurements ranged from 0.002 m to 0.099 m. In addition, the proposed UAV-SfM approach was applied to inaccessible bridges to measure surface crack lengths. These

https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-841-2026

results demonstrate that UAV-SfM photogrammetry provides a feasible and reliable method for 3D documentation and surface crack measurement of cultural heritage structures.

References

Bakirman, T., Bayram, B., Akpinar, B., Karabulut, F., Bayrak, O. C., Yigitoglu, A., Seker, D. Z., 2020. Implementation of ultra-light UAV systems for cultural heritage documentation. *Journal of Cultural Heritage*, 44, 174-184.

Block-Berlitz, M., Ducke, B., Rohland, H., Franken, C., Suchowska, P., Batbayar, T., Erdenebat, U., 2022. Area-optimized, rapid UAV-Brone recording of medieval heritage in Central Asia. *Journal of Field Archaeology*, 47(2), 90-104.

Chen, K., Reichard, G., Akanmu, A., Xu, X., 2021. Geo-registering UAV-captured close-range images to GIS-based spatial model for building façade inspections. *Automation in Construction*, 122, 103503.

Dasari, S., Mesapam, S., Kumarapu, K., Mandla, V.R., 2021. UAV in development of 3D heritage monument model: A case study of Koto Gullu, Warangal, India. *Journal of the Indian Society of Remote Sensing*, 49(7), 1733-1737.

Erenoglua, R.C., Ozgun, A., Oya, E., 2017. An UAS-assisted multi-sensor approach for 3D modeling and reconstruction of cultural heritage site. *Journal of Cultural Heritage*, 26, 79-90.

Flah, M., Suleiman, A.R., Nehdi, M.L., 2020. Classifications and quantification of cracks in concrete structures using deep learning image-based techniques. *Cement and Concrete Composites*, 114, 103781.

Frodella, W., Elashvili, M., Spizzichino, D., Gigli, G., Adikasvili, L., Vacheishvili, N., Kirkitadze, G., Nadaraia, A., Margotti-ni, C., Casagli, N., 2020. Combining infrared thermography and UAV digital photogrammetry for the protection and conservation of Rupestrian cultural heritage sites in Georgia: A methodological application. *Remote Sensing*, 12(5), 892.

Hidayat, H., Cahyono, A.B., 2016. Combined aerial and terrestrial images for complete 3D documentation of Singosari Temple based on structure from motion algorithm. In IOP Conference Series: Earth Environmental Science, 47(1), 012004.

Huang, W.C., Luo, Y.S., Liu, W.C., Liu, H.M., 2025. Deep learning-based detection on cultural heritage surfaces. *Applied Sciences*, 15(14), 7898.

Jahanshahi, M.R., Masri, S.R., Padgett, C.W., Sukhatme, G.S., 2013. An innovative methodology for detection and quantification of cracks through incorporation of depth perception. *Machine Vision and Applications*, 24, 227-241.

Jo, Y.H., Hong, S., 2019. Three-dimensional digital documentation of cultural heritage site based on the convergence of terrestrial laser scanning and unmanned aerial vehicle photogrammetry. *ISPRS International Journal of Geo-Information*, 8, 53.

Konstantinidis, I., Marinos, V., Papathanassiou, G., 2021. UAV-based evaluation of rockfall hazard in the cultural heritage area of Kipinas Monastery, Greece. *Applied Sciences*, 11, 8946.

Lins, R.G., Givigi, S.N., 2016. Automatic crack detection and measurement based on image analysis. *IEEE Transactions on Instrumentation Measurement*, 65, 583-590.

Manajitprasert, S., Tripathi, N.K., Arunplod, S., 2019. Three-dimensional (3D) modeling of cultural heritages site using UAV imagery: A case study of the pagodas in Wat Maha That, Thailand. *Applied Sciences*, 9(18), 3640.

Martin-Bejar, S., Claver, J., Sebastian, M.A., Sevilla, L., 2020. Graphic applications of unmanned aerial vehicles (UAVs) in the study of industrial heritage assets. *Applied Sciences*, 10(24), 8821.

Nony, N., Luca, L.D., Godet, A., Deseilligny, M.P., Remondino, F., 2012. Protocols and assisted tools for effective image-based modeling of architectural elements. 4th International Conference, EuroMed 2012, Progress in Cultural Heritage Preservation, Lecture Notes in Computer Science, vol. 7616. Springer, Berlin, Heidelberg.

Partama, I.G.Y., Yastika, P.E., Wijaya, I.M.W., 2025. 3D Modeling Using UAV-Photogrammetry Technique for Digital Documentation of Cultural Heritage Buildings. International *Journal of GEOMATE*, 28(126), 61-70.

Rakha, T., Gorodetsky, A., 2018. Review of Unmanned Aerial System (UAS) applications in the built environment: Towards automated building inspection procedures using drones. *Automation Construction*, 93, 252-264.

Sang, X., Leeng, X., Ran, X., Li, X., Xue, L., 2022. A virtual 3D geological library based on UAV and SFM: Application for promoting teaching and research on geological specimen and heritage online. *Geoheritage*, 14, 43.

Seifert, E., Seifert, S., Vogt, H., Drew, D., van Aardt, J., Kunneke, A., Seifert, T., 2019. Influence of drone altitude, image overlap, and optical sensor resolution on multi-view reconstruction of forest images. *Remote Sensing*, 11(10), 1252.

Stek, T.D., 2016. Drones over Mediterranean landscapes: The potential of small UAV’s (drones) for site detection and heritage management in archaeological survey projects: A case study from Le Pianelle in the Tappino Valley, Molise (Italy). *Journal of Cultural Heritage*, 22, 1066-1071.

Valenca, J., Dias-da-costa, D., Julio, E., Araujo, H., Costa, H., 2013. Automatic crack monitoring using photogrammetry and image processing. *Measurement*, 46, 433-441.

Xu, Z., Wu, L., Shen, Y., Li, F., Wang, Q., Wang, R. 2014. Tridimensional reconstruction applied to cultural heritage with the use of camera-equipped UAV and terrestrial laser scanner. *Remote Sensing*, 6(11), 10413-10434.

Yang, J., Li, H., Zou, J., Jiang, S., Li, R., Liu, X., 2022. Concrete crack segmentation based on UAV-enabled edge computing. *Neurocomputing*, 485, 233-241.

Zhu, J., Zhong, J., Ma, T., Huang, X., Zhang, W., Zhou, Y., 2022. Pavement distress detection using convolutional neural networks with images captured via UAV. *Automation Construction*, 133, 103991.