Crack assessment using multi-sensor fusion simultaneous localization and mapping (SLAM) and image super-resolution for bridge inspection

Chu-Qiao Feng [a], Bao-Luo Li [a], Yu-Fei Liu [a,b,*], Fu Zhang [c], Yan Yue [a], Jian-Sheng Fan [a,b]

[a] *Department of Civil Engineering, Tsinghua University, Beijing 100084, China*
[b] *Key Laboratory of Civil Engineering Safety and Durability of China Ministry of Education, Tsinghua University, Beijing 100084, China*
[c] *Department of Mechanical Engineering, The University of Hong Kong, 999077, Hong Kong, China*

ARTICLE INFO

ABSTRACT

The inspection of bridges is increasingly dependent on advanced equipment and algorithms like digital cameras and SfM (Structure from Motion). However, many existing SfM-based bridge inspection methods lack efficiency due to lengthy 3D reconstruction computation times, and digital image resolution often falls short in detecting fine cracks and calculating their widths, mainly influenced by the acquisition equipment. This paper describes a fast and accurate crack assessment method that leverages multi-sensor fusion SLAM (Simultaneous Localization and Mapping) and image super-resolution. Through multi-sensor fusion SLAM, textured point clouds of the bridge structure can be obtained directly, significantly improving efficiency. Furthermore, deep learning-based image super-resolution enhances the precision of crack width calculation. Field tests demonstrate the effectiveness of the proposed methods, showcasing a 94% reduction in scene reconstruction time and a 16% improvement in crack width calculation accuracy.

1. Introduction

Given the continuous development of engineering construction in recent years, the importance of inspection, operation, and maintenance of bridges and other infrastructure has become increasingly evident. One area that has received widespread attention is the use of intelligent algorithms and multi-sensor devices for automated bridge inspection [1]. This technology utilizes unmanned aerial vehicles (UAVs), cameras, infrared sensors, and other devices to improve efficiency, reduce duplication of fieldwork, and enhance the safety of operations during bridge inspection. Furthermore, the data collected by automated bridge inspection permits quantification of the characteristics of bridge defects, circumventing the uncertainty of visual observation noted in previous human operations. Overall, the integration of automated technology in bridge inspection holds significant potential to improve current practices and promote the efficiency of infrastructure operation and maintenance.

Automated bridge inspection involves the use of a platform equipped with a range of devices positioned at a specific distance from the target structure to be inspected. As the platform moves along a particular path, it collects data, thus achieving comprehensive surveillance [2]. The primary outcomes of automated bridge inspection typically comprise 3D point clouds of bridges, digital images, or infrared images containing information on bridge defects [3]. In terms of crack assessment in bridges, the inspection process can be divided into several discrete stages, which include data acquisition, crack detection and segmentation, crack localization, and damage assessment, as illustrated in Fig. 1.

The detection and segmentation of bridge cracks is the premise of damage assessment. In existing studies, techniques such as thresholds, edges, regions, matching, and deep learning have been employed for crack detection [4]. Scholars have conducted extensive research on this topic, for example, Yeum and Dyke [5] successfully identified fatigue cracks near bridge I-beam bolts using the Frangi filter and Canny edge detection algorithm. Abdel-Qader et al. [6] compared four edge detection algorithms (FHT, FFT, Sobel, and Canny) and found that the FHT method yielded higher accuracy for bridge crack detection. Li et al. [7] utilized the neighborhood histogram difference method (NHDM) to perform crack segmentation based on the continuity of cracks and the correlation between the thresholding value and the image standard deviation. Recent advancements have seen the emergence of crack detection and segmentation methods based on deep learning. Ayele et al. [8] employed the Mask RCNN method to detect cracks from digital images collected by UAVs and developed an integrated system and interface for operation. Yu et al. [9] proposed the YOLOv4-FPM deep

learning model, which improved the efficiency and accuracy of real-time bridge crack detection and effectively addressed challenges associated with complex image backgrounds. Fu et al. [10] applied the improved Deeplabv3+ semantic segmentation algorithm to the crack segmentation task, remarkably enhancing the accuracy of crack segmentation details. Zhang et al. [11] proposed an encoder–decoder crack segmentation network that incorporates multi-scale contextual information enhancement, enabling effective differentiation between cracks and background noise.

Crack quantification mainly focuses on calculating parameters such as length and width. According to China's industry standard "Technical Specifications for Field Inspections of Existing Highway Bridges" [12], concrete crack detection should include the location, distribution, direction, width, depth and quantity of cracks. The location of the crack is a crucial factor in determining its nature (whether it is a bending crack or shearing crack). Meanwhile, the crack width serves as an important indicator for assessing the severity of the crack. Lee et al. [13] defined the crack width as the distance between the intersection points of the crack skeleton's normal line and the crack's edge line. Liu [14] proposed a stable four-step calculation method for calculating the crack width based on the above approach. Furthermore, Nishikawa et al. [15] introduced a sub-pixel interpolation method to enhance the accuracy of crack width measurement. Mirzazade et al. [16] conducted 3D reconstruction for the structure, generated an orthophoto image for the crack region, and leveraged the GSD information of the picture to measure the width of the crack.

The difficulty of UAV inspection in China lies in the need to detect cracks that measuring 0.1 mm in width [17]. As a result, the images captured by the UAV often only represent a small portion of the overall scene. This limited perspective poses a challenge to the effective identification of damage to bridge surfaces within large-scale scenes. Therefore, localizing images that contain the areas where damage has occurred becomes formidable. To tackle this issue, it is often necessary to build a 3D model of the bridge and associate the location information of the cracks with the model [18]. Moreover, the 3D reconstruction of bridges can be utilized to evaluate the service status of bridges from other aspects, such as bridge deformation detection [19]. Different inspection devices can collect various types of data for bridge damage assessment, with camera-based and LiDAR-based systems playing a crucial role in the 3D reconstruction of bridges [20]. One of the most commonly used camera-based 3D reconstruction methods is SfM (Structure from Motion). Scholars have successfully obtained 3D models of bridges or their components using the SfM method [21–25]. This method requires ensuring that the captured images have a large overlapping area and precise trajectory planning to control the working distance. Photogrammetry based on the SfM method can generate high-precision point clouds to meet the requirements of bridge 3D reconstruction.

However, using the SfM approach to generate 3D models of bridges for inspection presents several challenges. Firstly, it is often necessary to shoot at fixed points and conduct meticulous trajectory planning to avoid motion blur during data acquisition and ensure successful reconstruction. This time-consuming process leads to inefficiency. Secondly, SfM typically requires an extended computing time when dealing with a large number of images [26]. Moreover, the resulting point cloud model

cannot be directly utilized for crack detection. This is because bridge cracks are usually narrow, measured in millimeters, while the accuracy of the point cloud is at the centimeter level, which is greater than the width of the crack. Additionally, it is difficult to identify continuous cracks based on discontinuous point cloud or mesh features using existing technical means. Therefore, spending a significant amount of time to obtain surface textures utilizing SfM is not cost-effective since the crack information cannot be directly obtained through the 3D model. All of the above situations reduce the efficiency of acquiring the 3D model. Alternatively, besides SfM, SLAM(simultaneous localization and mapping) is often used for trajectory measurement of objects and 3D reconstruction of scenes. For camera-based systems, visual SLAM can be used for defect inspection. Kim and Eustice [27] implemented a monocular SLAM algorithm for automated underwater hull inspection. For LiDAR-based systems, lidar SLAM can also obtain good results in localization and mapping, such as Cartographer [28], LOAM [29], Fast-LIO [30], etc. In recent years, substantial progress has been made in the field of SLAM based on multi-sensor fusion [31–36]. The fusion of Li-DARs and RGB cameras can better improve the mapping process to simulate the real environment [20]. In this regard, Multi-sensor fusion SLAM can efficiently obtain high-precision surface textures and provide sufficient information for crack localization.

Furthermore, existing studies have rarely focused on the calculation accuracy of relatively fine cracks. As mentioned above, the size of bridge cracks is very small, which naturally increases the difficulty of crack quantification. The classical crack width calculation method requires the crack to occupy at least 3 pixels in the image so that the skeleton and edge lines can be extracted and the width can be calculated, as shown in Fig. 2(a). However, in the scene of bridge inspection, cracks usually occupy <1 pixel in the image due to the working distance and crack scale, which is manifested as a light gray sub-pixel feature, as shown in Fig. 2(b). Although there is a crack width calculation method based on gray value [37], it is very sensitive to working distance and illumination, and the calculation results are not stable. Recovering the features of fine cracks in images is of great significance to the accurate evaluation of crack parameters.

Aiming at the above problems, this paper proposes a fast and accurate crack assessment method that addresses the aforementioned issues using multi-sensor fusion SLAM and image super-resolution. Firstly, multi-sensor fusion SLAM is applied to quickly obtain 3D models and sensor trajectories. Secondly, a crack image super-resolution model is trained using deep learning network. This model helps restore crack features in the image, ensuring the accuracy of crack parameter calculation. Thirdly, image-based crack detection and segmentation based on deep learning are implemented to obtain pixel representations of cracks. Fourthly, crack width feature pair sequences can be obtained through digital image processing. Finally, the crack information is projected and located on the 3D model. Field test investigations on a real bridge pier are conducted to illustrate and validate the proposed method.

2. Framework of the proposed method

The contribution of this study is to improve the efficiency of 3D model generation and the accuracy of crack parameter calculation. The method proposed in this paper is shown in Fig. 3, which consists of 4
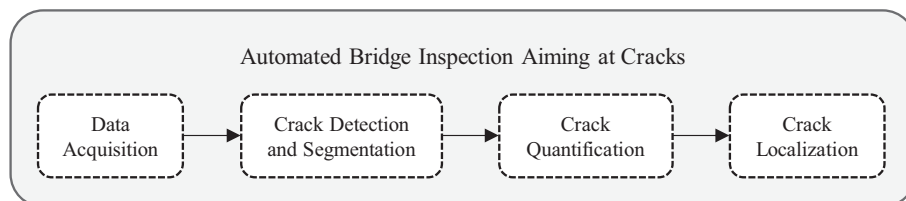

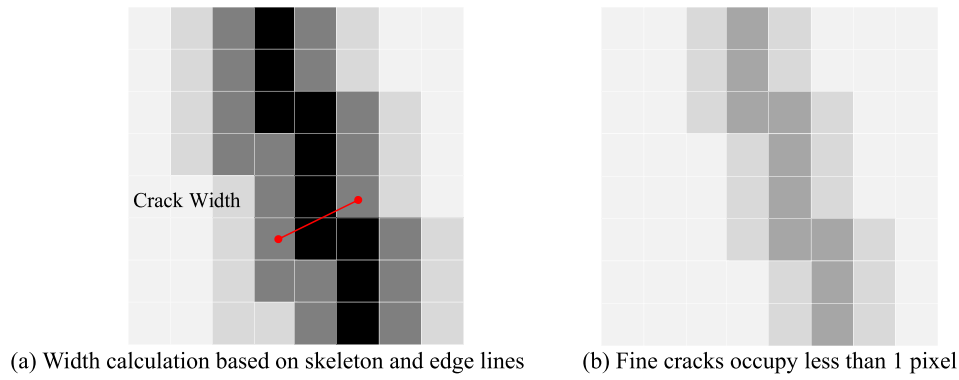**Fig. 1.** Automated bridge inspection process aiming at bridge cracks.

(a) Width calculation based on skeleton and edge lines　　(b) Fine cracks occupy less than 1 pixel

Fig. 2. Difficulty in calculating the parameters of fine cracks.
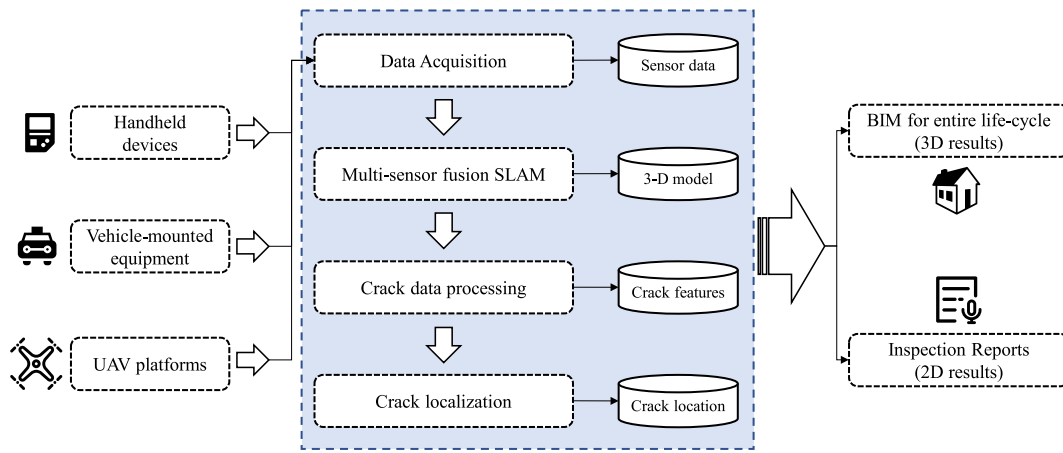

Fig. 3. Framework of the proposed method.

interrelated parts: data acquisition, multi-sensor fusion SLAM, crack data processing, and crack localization. In the data acquisition part, the determination method of maximum working distance and the principle of data collection based on multi-sensor fusion SLAM are proposed. By using multi-sensor fusion SLAM and triangular meshing, the textured point clouds and mesh models are obtained accordingly. In the crack data processing part, an object detection algorithm based on deep learning is utilized to cut the crack region. Furthermore, a crack high-definition image generation method based on super-resolution and deep learning is proposed. The information of crack feature points is then obtained through digital image processing. Finally, the information from the 3D model, camera odometry, and crack features are combined for projection to obtain the location of cracks.

The method and equipment proposed in this paper can be integrated on different platforms as a module in bridge inspection, as shown in Fig. 3. Handheld devices, vehicle-mounted equipment, and UAV platforms can be selected for data acquisition, each scheme requires corresponding additional hardware development. The data results processed by the proposed method can be used to generate or correct BIM models and issue inspection reports.

3. Data acquisition and 3D reconstruction based on multi-sensor fusion SLAM

Multi-sensor fusion SLAM is initially proposed to solve the problem of degradation of single-sensor SLAM in scenarios with less texture, drastic changes in illumination, long corridors, situations requiring strenuous movements, etc. Research and practice have demonstrated that multi-sensor fusion SLAM exhibits strong robustness and high accuracy in localization and mapping, making it suitable for bridge

inspection. Additionally, the VIO (Visual Inertial Odometry) subsystem in multi-sensor fusion SLAM allows for the acquisition of surface texture, which can be visually displayed to assist in localization. In this paper, Inertial-Visual-LiDAR fusion SLAM is used for data acquisition. The point cloud and camera odometry are obtained through state estimation, enabling the generation of a continuous surface model via triangular meshing.

3.1. Brief introduction to multi-sensor fusion SLAM

SLAM is essentially a state estimation problem that involves solving the position of the robot and the surrounding map points using measurement data from motion and observation sensors. The observation process requires various sensors, including cameras, IMU, and LiDAR, and the information from these sensors can be used simultaneously for state estimation.

Multi-sensor fusion SLAM contains subsystems such as VO(Visual odometry), LO(LiDAR odometry), VIO(Visual inertial odometry), and LIO(LiDAR inertial Odometry). Depending on whether sensor data are jointly optimized, multi-sensor fusion SLAM can be classified as loosely coupled or tightly coupled. In detail, loosely coupled SLAM involves individual subsystems performing motion estimation separately and then fusing and updating the estimation results. On the other hand, tightly coupled SLAM utilizes sensor information from different subsystems for motion estimation simultaneously and optimizes them jointly. In general, tightly coupled SLAM improves the robustness and accuracy of the system while using measurement information more efficiently [38]. For illustration, Fig. 4 shows the basic framework of multi-sensor fusion SLAM, summarized from references [35,36,38–40].
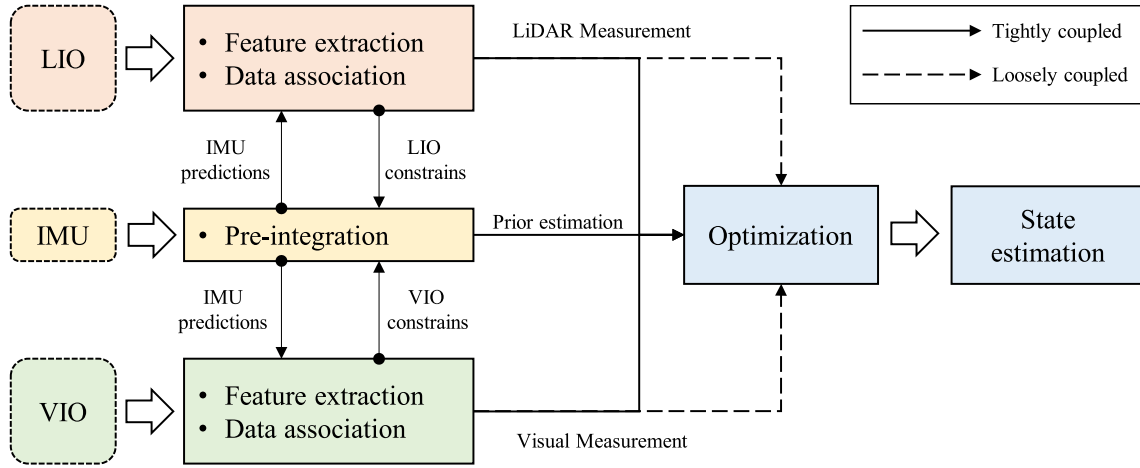
Fig. 4. Basic framework of multi-sensor fusion SLAM.

3.2. Data acquisition strategy

It is important to note that industrial cameras with relatively low resolution (e.g., 640×480, 800×600, 1280×720, etc.) are commonly used in visual or fusion SLAM. However, in this paper, a high-resolution (3072×2048) industrial camera is employed to collect visual data specifically for crack detection purposes. Although this choice may slightly affect the algorithm's real-time performance, it guarantees better crack detection outcomes.

Before acquiring data, it is necessary to determine the working distance of the system in order to obtain photographs that clearly depict cracks of a certain width. We introduce the concept of ground sampling distance (GSD) in geomatics, which represents the actual length corresponding to one pixel in the image. When employing the camera's pinhole imaging model (as shown in Fig. 5), GSD can be calculated using the following formula:

$$\text{GSD} = \frac{l \bullet u}{f \bullet N} \tag{1}$$

where l refers to the size of the camera sensor; u refers to the working distance; f refers to the focal length of the camera; and N refers to the number of image pixels corresponding to l.

The GSD of an image is directly related to the result of crack detection. When collecting images for crack assessment, it is crucial to ensure that the GSD is greater than the width of the cracks to be detected. According to Eq. (1), the maximum working distance can be determined. Furthermore, since fine cracks cannot occupy 1 pixel in an image, image super-resolution is implemented to enhance the details of the cracks. A super-resolution factor, denoted as α, is introduced to signify that the

information of 1 pixel is amplified to α pixels after super-resolution. By accounting for these considerations, the maximum working distance can be determined as follows:

$$u_{\max} = \frac{f \bullet N \bullet \alpha \bullet \text{GSD}}{l} \tag{2}$$

where $u_{\max}$ refers to the maximum working distance corresponding to a given GSD; α refers to the super-resolution factor. Note that since the working distance is less than $u_{\max}$ most of the time during data acquisition, cracks smaller than the GSD value can still be calculated, but the reliability is reduced.

Fine path planning is often required when collecting data based on SfM to ensure sufficient overlapping zones and prevent reconstruction failure. For example, for data collection on a cylindrical bridge pier, SfM necessitates locating the shooting points near the envelope circle at a distance u ($u \leq u_{\max}$) from the pier surface. The trajectory usually follows a regular pattern, depending on the surface's shape. In contrast, SLAM, dose not impose strict requirements on the shooting path. It only requires that the maximum distance from the surface to be measured during data acquisition does not exceed $u_{\max}$, as illustrated in Fig. 6. Consequently, this reduces the difficulty of data acquisition and significantly improves efficiency.

In addition, several factors should be carefully considered during the data acquisition process: (i) Avoid placing the sensor too close to the object to prevent sensor degradation caused by limited features. When conducting data collection, it is advisable to include as many structural features in the scene as possible. (ii) Reasonably control the camera's exposure time, gain, white balance, and other parameters to obtain better surface texture in 3D reconstruction. (iii) Since the collected
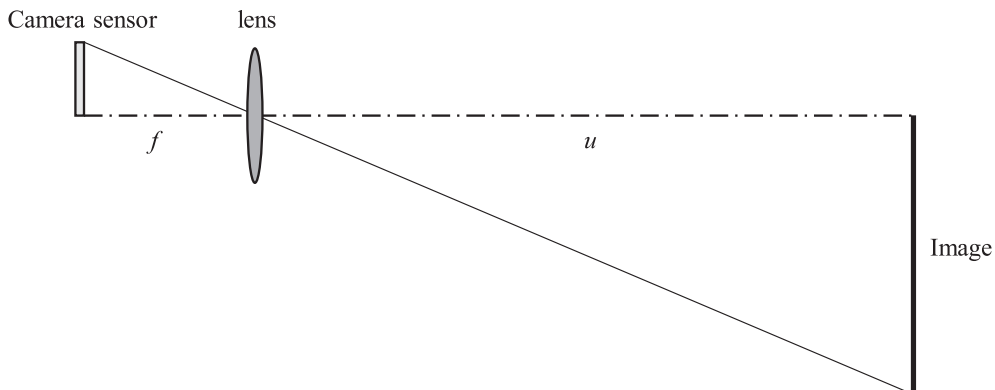

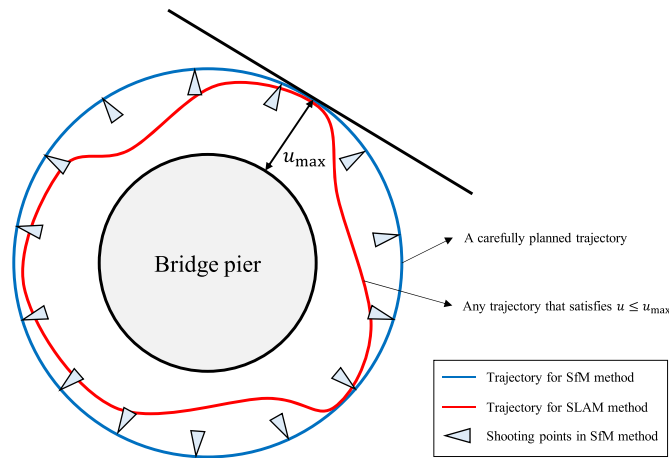**Fig. 5.** Pinhole imaging model of camera.

Fig. 6. Trajectory planning of SfM and SLAM in data acquisition.

image data will be used for various computer vision tasks such as image super-resolution, crack detection and segmentation, it is essential to acquire the data under favorable lighting conditions.

3.3. 3D reconstruction and triangular meshing

In this paper, a tightly coupled algorithm R^3IVE [36] is adopted for 3D reconstruction. R^3IVE consists of two subsystems, namely the LIO and VIO subsystems. The LIO subsystem utilizes the measurement information to construct a global map, while the VIO subsystem fuses visual data by minimizing the frame-to-map photometric error. A dense colored point cloud of the scene is obtained by the application of R^3IVE.

Since SLAM can estimate the state of the sensors, the camera odometry can be obtained accordingly. In the subsequent process of crack localization, the position of the camera's optical center can be directly determined by the camera odometry. Triangulation is performed using the offline tool provided by R^3IVE, which is based on Delaunay triangulation and graph cuts [41]. The resulting mesh is textured using vertex colors.

The applicability of the method described above is demonstrated through its successful implementation in a light rail line bridge. The data acquisition and subsequent 3D reconstruction process are presented. For the bridge scene shown in the Fig. 7, multi-sensor fusion SLAM was utilized, and the acquisition process took 149 s, resulting in a point cloud containing 3,198,550 points.

4. Crack data processing

The goal of crack data processing is to obtain crack shape and width parameters from images accurately. In this section, crack detection for acquired images is carried out. Image cropping is utilized to reduce the interference of background noise. A crack image super-resolution model

is trained using the ESRGAN-based deep learning network. Furthermore, digital image processing is implemented to extract cracks and identify their widths within the images.

4.1. Crack detection and image cropping

Since image super-resolution will generate larger images and introduce additional computation, accurately locating the crack area before image super-resolution and cropping out the parts that are not of interest can significantly improve processing efficiency. Additionally, image cropping effectively eliminates the interference of background noise and prevents unnecessary noise points from being introduced in the subsequent crack segmentation process.

In this section, we incorporate Li et al.'s [42] grid-based classification and box-based detection fusion model for bridge crack detection, which outputs the crack area in grid form. An example of a detection result is shown in Fig. 8(a). The crack detection process correctly frames the outer contour of the cracks so that most of the background noise can be excluded. Subsequently, images are cropped based on the crack detection result, eliminating the parts outside the grids as illustrated in Fig. 8(b). The above steps can significantly enhance the efficiency and accuracy of crack segmentation.

4.2. Image super-resolution

Image super-resolution is one of the classic tasks in computer vision aimed at enlarging the image and enhancing texture details. Traditional approaches to image super-resolution, such as linear interpolation, bilinear interpolation, and bicubic interpolation, are simple and fast to implement, but struggle to recover effective pixel-level features. The introduction of deep learning methods, such as SRCNN [43], treated the problem as a deep learning task that updates the parameters of the convolution kernels. However, the generated images often suffer from excessive smoothness. SRGAN [44] introduced the use of generative adversarial network(GAN) into the field of image super-resolution. It changes the input of the generator from random noise to low-resolution images and adopts residual connection. Following SRGAN, ESRGAN [45] removed the batch norm layer and replaced residual blocks with dense blocks to reduce the training difficulty of the network. Real-ESRGAN [46] further improved upon this by simulating the degradation process of high-resolution images in real life and employing pure composite data for training, which performs better in real photo super-resolution.

In this section, a deep network with residual-in-residual dense blocks (RRDB) proposed by ESRGAN is used as parts of the generator network for training. On this basis, by referring to the idea of transfer learning, the layer before the upsampling layer is frozen, and the collected dataset is used for fine-tuning, as shown in Fig. 9. The convolution layer and RRDB blocks before the upsampling layer are mainly used to extract image features. Therefore, the ability to extract image features can be better guaranteed by using the model parameters pre-trained with a

Fig. 7. 3D reconstruction result of a light rail line bridge.

(a) Crack detection   (b) Image cropping

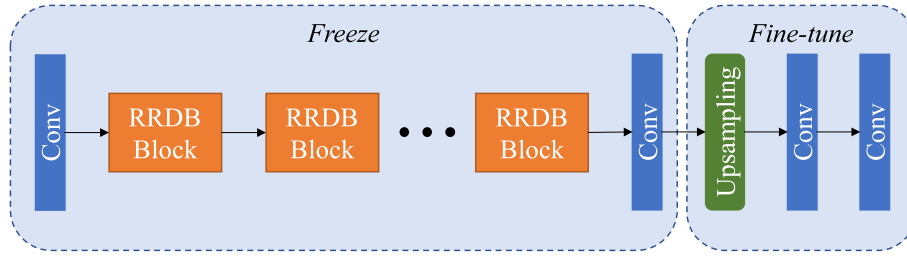**Fig. 8.** Crack detection and image cropping examples.

Fig. 9. Generator network architecture for image super-resolution.

large amount of data. The upsampling layer and the layers after it are used for image generation. In this task, fine-tuning with targeted crack data can achieve a better super-resolution effec. The pre-trained model used in this section is the X4 model for general images provided by real-ESRGAN [46], which was trained using open-source datasets including DIV2K, Flickr2K, and OutdoorSceneTraining.

Twelve hundred surface photos of 3 concrete bridges are used for the training of the image super-resolution model. These photos are processed by the second-order degradation model proposed by Real-ESRGAN and trained together with the original photos as a training set. Fig. 10 shows the crack areas in the same images before and after image super-resolution. To demonstrate the effectiveness of image super-resolution, two images representing the same area of a crack are resized using linear interpolation to the same size, as depicted in Fig. 10 (a)(b). It becomes apparent that image super-resolution is generally capable of restoring the fine details of cracks more clearly, consequently enhancing the accuracy of crack width and other parameter calculations. For relatively fine cracks, image super-resolution can effectively increase the number of pixels occupied by the cracks, enabling the calculation of previously unmeasurable crack widths (as shown in Fig. 10(c) and (d)). This improvement has significant implications for the evaluation of cracks of varying sizes.

4.3. Crack segmentation and width calculation

In this section, crack segmentation and quantification methods based on OpenCV are applied. Median filter and Niblack's local binarization are implemented to further reduce background noise. The cracks in images after cropping and super-resolution are segmented by thresholding method. Connected component analysis, PCA, and manual operation when necessary are used to further eliminate false segmentation. An example result is shown in Fig. 11(a). Subsequently, the cracks' skeleton lines and edge lines are extracted to obtain the pixel coordinates of a series of crack width feature points, as shown in Fig. 11 (b).

5. Crack localization

After obtaining the 3D model and crack information, it is necessary to project the vector information of the crack onto the model to localize the cracks. Liu et al. [23] project the crack information based on collision detection and axis aligned bounding box(AABB) tree [47]. Based on the work of Liu et al. [23], this paper proposes a crack projection method based on multi-sensor fusion SLAM, including determining the camera pose and projecting the information of cracks.

5.1. Determination of the camera's pose

The camera pose can be obtained using the camera odometry output by SLAM. The quaternion in the odometer is first converted into a rotation matrix, which represents the rotation from the world coordinate system to the camera coordinate system. Considering a crack width point p, the relationship between the world coordinate system and the camera coordinate system is given by Eq. (3), and the translation matrix from the world coordinate system to the camera coordinate system can be obtained by Eq. (4):

$$T_{WC} = -R_{WC}X_{WC} \tag{3}$$

$$R_{WC}X_{WC} + T_{WC} = X_C \tag{4}$$

where T_{WC} and R_{WC} denote the translation and rotation matrices from the world coordinate system to the camera coordinate system, respectively; X_{WC} denotes the coordinate of the camera on the world coordinate system; and X_C denotes the coordinate of point p on the camera coordinate system.

5.2. Crack projection

Crack projection is the process of mapping the crack feature points in the image to the 3D model through coordinate transformation. According to the research of Liu et al. [23], the coordinates of the crack feature points on the world coordinate system can be determined by the intersection of the projection rays and the 3D model. The transformation
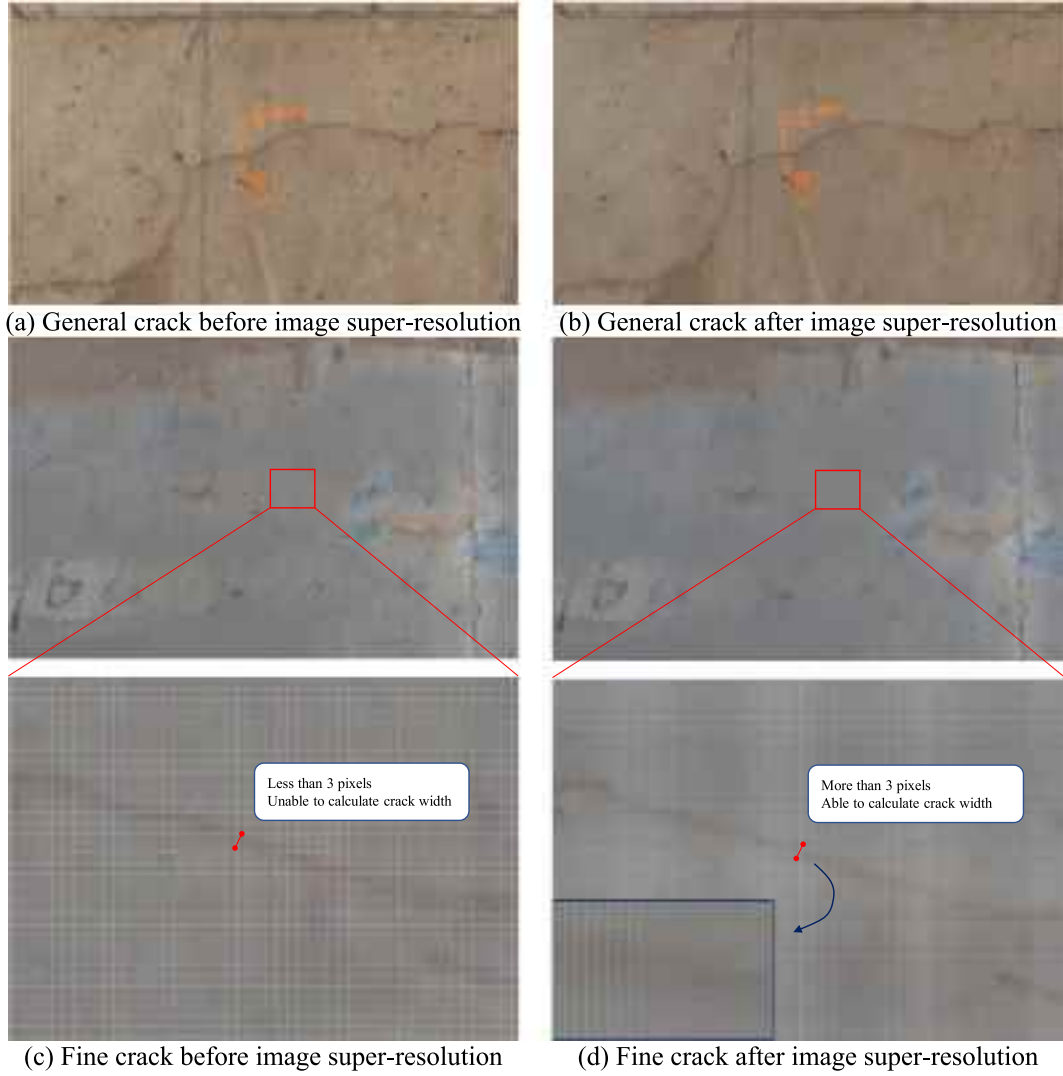
(a) General crack before image super-resolution

(b) General crack after image super-resolution

(c) Fine crack before image super-resolution

(d) Fine crack after image super-resolution

Less than 3 pixels
Unable to calculate crack width

More than 3 pixels
Able to calculate crack width

Fig. 10. Results and contributions of image super-resolution.
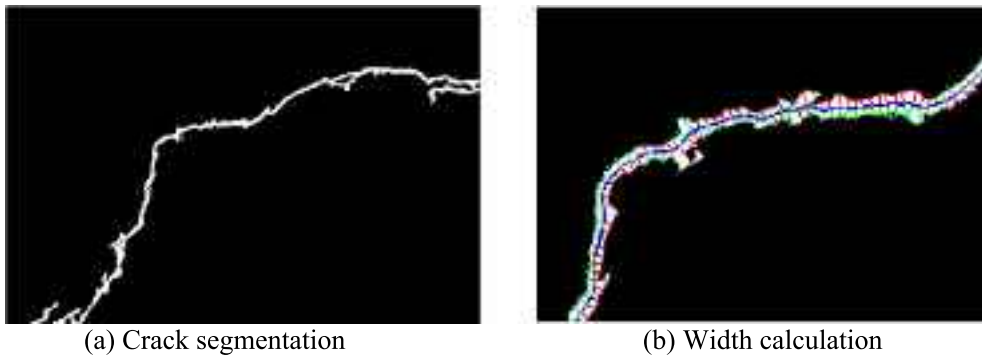

(a) Crack segmentation

(b) Width calculation

Fig. 11. Crack segmentation and width calculation examples.

of point p from the pixel coordinate system to the world coordinate system is:

$$K\left(R_{WC}X_{Wp} + T_{WC}\right) = Z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{5}$$

where K denotes the calibrated camera intrinsic parameter matrix; X_{Wp}

denotes the coordinate of point p on the world coordinate system; Z denotes the coordinate of this point in the z direction on the camera coordinate system; and u, v denote the coordinate of point p on the pixel coordinate system.

Since Z is unknown, world coordinates cannot be determined from pixel coordinates alone. Considering the normalized plane when $Z = 1$, the world coordinates of the points on the normalized plane can be obtained according to Eq. (6), and then the projection ray can be

obtained by Eq. (7):

$$X_n = R_{WC}^{-1} \bullet \left\{ K^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - T_{WC} \right\} \tag{6}$$

$$p = X_n - X_{WC} \tag{7}$$

where X_n denotes the world coordinate of point p on the normalized plane; p denotes the projection vector, which determines the orientation of the projection ray.

The projection ray intersects the 3D model to obtain the world coordinates of the crack feature points. The absolute distance between the corresponding crack width feature points can represent the real crack width at that point.

6. Field tests on a bridge pier

To illustrate and validate the proposed method, inspection tests of a bridge pier are carried out using a handheld device integrated with multi-sensor fusion SLAM.

6.1. Equipment, site, and data acquisition

The parameters of the equipment are listed in Table 1. Livox Avia is a lightweight, high-performance solid-state LIDAR with a FOV of 70.4° × 77.2° and a weight of 498 g, which is suitab le for on-site inspection. The BMI088 IMU is built into the LiDAR. Hikvision MV-CE060-10UC is a high-resolution, high-performance CMOS industrial camera that supports USB 3.0 for data transmission. The camera sensor size is 1/1.8," and the image's pixel resolution is 3072 × 2048. A lens of 6 mm focal length is used with the camera. The equipment used for field tests is shown in Fig. 12(a).

A bridge pier near Yongding River in Fengsha Second line, Beijing, was selected for field tests, as shown in Fig. 12(b). The pier in question features a truncated cone-shaped bottom half with a distribution of continuous narrow spatial cracks. Data acquisition was conducted in the field rather than in a laboratory environment, without any artificial intervention. Therefore, the results of the field tests can largely reflect the real bridge inspection situation. The field tests were conducted under good weather conditions with a cloudless sky and suitable temperature to ensure uniform lighting on the pier surface during image acquisition. Test parameters are listed in Table 2. The current bridge inspection standards and codes in China rarely specify the required accuracy for crack detection but provide information on the relationship between crack width and the degree of structural damage. Considering the current standard specification [48–50] and the working distance limitation of the LiDAR, we choose the GSD to be 0.4 mm to ensure that the quantification of cracks wider than 0.4 mm in all pictures is reliable. When the super-resolution factor α is 4.0, the maximum working distance $u_{\max}$ can be controlled as 4.11 m.

Data acquisition strategy mentioned in Section 3.2 is adopted to ensure the reliability of the results. Zhang's method [51] and livox_camera_calib [52] are adopted for intrinsic and extrinsic calibration of LiDAR and the camera. The trajectory of data acquisition is computed by

Table 1
Main parameters of equipment.

Equipment	Type	Specification
Lidar	Livox Avia	FOV:70.4° × 77.2° Range precision: 2 cm Angular precision: <0.5°
Industrial camera	Hikvision MV-CE060-10UC	Resolution: 3072 × 2048
Lens	Hikvision MVL-HF0628M-6MP	Focal length: 6 mm

SLAM and shown in Fig. 12(c), where a random trajectory is followed to control the working distance within the inspection zone less than $u_{\max}$. Cracks with relatively narrow widths (from 0.1 mm to 2.0 mm) were carefully selected on the pier for measurement (as shown in Fig. 12(e)) to verify the robustness of the proposed method in practical applications. The above cracks are marked as #1 ~ #13 for subsequent analysis.

6.2. Results of crack data processing

The crack data was processed using the technical route described in Section 4. Firstly, within the inspection zone that satisfied the maximum working distance constraint, 9 surface photos that could cover the crack regions on the pier surface were selected for processing. Following the procedure outlined in Section 4.2, the crack images underwent crack region detection and background region cropping. Subsequently, the processed images were subjected to 4× super-resolution. Afterwards, adaptive threshold segmentation was applied to transform the images into binary images. Finally, the width features of these images were identified. The entire process of crack data processing is illustrated in Fig. 13. It is important to note that due to the complexity of the real environment, some noise and misidentification are inevitably introduced during crack detection and segmentation. To mitigate this concern, certain overtly irrelevant error information is manually eliminated during data processing to emphasize the focus of the proposed methodology more effectively. At the same time, the author believes that in the engineering application and practice of this technology, some manual intervention may be necessary for crack segmentation.

6.3. Results of 3D reconstruction and crack localization

The textured point cloud calculated and output by SLAM is shown in Fig. 14(a). Furthermore, a mesh model that can be used for projection is generated after triangular meshing, as shown in Fig. 14(b). The proposed method can obtain textured 3D models efficiently and reproduce the engineering scene intuitively.

After obtaining the 3D model of the bridge pier, the pose information of the aforementioned 9 photos is obtained first according to the method described in Section 5.1. The relationship between them and the bridge pier is shown in Fig. 15(a). Then, according to the projection method in Section 5.2 and 5.3, the crack feature points are projected into the 3D model, and crack widths are calculated, as shown in Fig. 15(b)(c)(d).

6.4. Discussion of proposed methods

To demonstrate the contributions of the proposed method, 3D reconstruction based on SfM and crack width calculation based on bilinear interpolation are carried out for comparison.

6.4.1. Faster reconstruction compared with SfM

To demonstrate the efficiency of the proposed method, we compare it with other methods based on SfM for 3D reconstruction of the scene using the data collected above. We crop out the part of the bridge pier that we are interested in and show the texture details obtained by different reconstruction methods in Fig. 16. Based on the study by Chen et al. [53], a comparison of the point cloud quality for the inspection zone was conducted from three aspects: nonuniform distribution, surface deviation, and geometric accuracy, as shown in Table 3. In the nonuniform distribution comparison, the number of neighboring points within a radius of 0.01 m for each point is calculated, and its mean and standard deviation are reported. In the comparison of surface deviation, the thickness of the point cloud was measured at 5 randomly selected points on the surface of the pier, and the averaged results was obtained. In the comparison of geometric accuracy, the diameter of the bottom of the pier is selected as the reference value. The data shown in Table 3 represents: point cloud measurement value/reference true value/the

(a) Equipment used for field tests

(b) Bridge pier for field tests

(c) Trajectory of data acquisition

(c) Data acquisition with multiple sensors

(d) Crack width measurement

Fig.12　Field test scenarios and data acquisition trajectory

Fig. 12. Field test scenarios and data acquisition trajectory.

Table 2
Test parameters.

Parameters	Value
GSD	0.4 mm
Length of the image N	3072
Length of the CMOS l	7.18 mm
Focal length f	6 mm
Super-resolution factor α	4.0
Maximum working distance u_{max}	4.11 m

absolute value of the relative error.

The reconstruction results and details of the whole scene are shown in Fig. 17. During the reconstruction process, the number of points in the results obtained by the two methods is controlled to be relatively close, so that the difference in performance between the two methods can be compared scientifically. Some parameters and quality assessments during the reconstruction of the two methods are listed in Table 4.

The analysis in Table 3 shows that SfM-based reconstruction yields higher quality, better precision, and denser point clouds in the
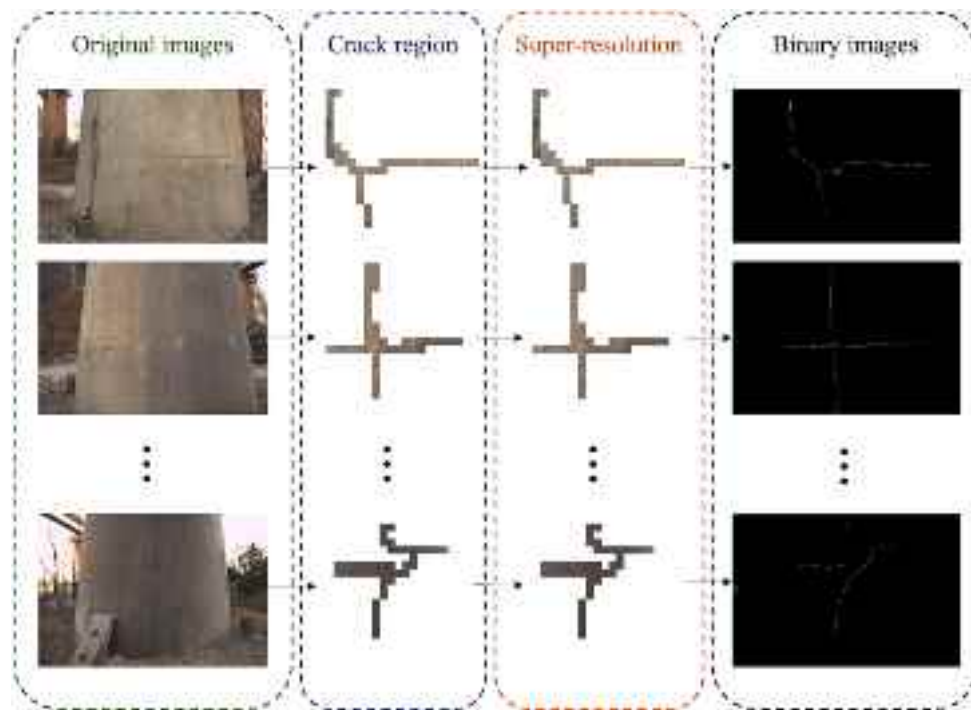
Fig. 13. Crack data processing workflow.

(a) Textured point cloud   (b) Textured mesh model

Fig. 14. Results of 3D reconstruction.

inspection zone. The point cloud quality obtained by SLAM-based reconstruction is slightly inferior, but the difference is relatively small. As can be seen from Figs. 16 and 17, for the same inspection zone, the 3D model obtained based on SfM has superior texture accuracy, but the scope of its reconstruction scene is strictly limited by the photos used in

SfM calculation. Conversely, due to its inherent characteristics, the SLAM-based method can achieve both engineering precision requirements and a boarder range of reconstruction results. Moreover, the SfM-based method necessitates meticulous consideration of photo overlap to ensure the successful reconstruction. In the example
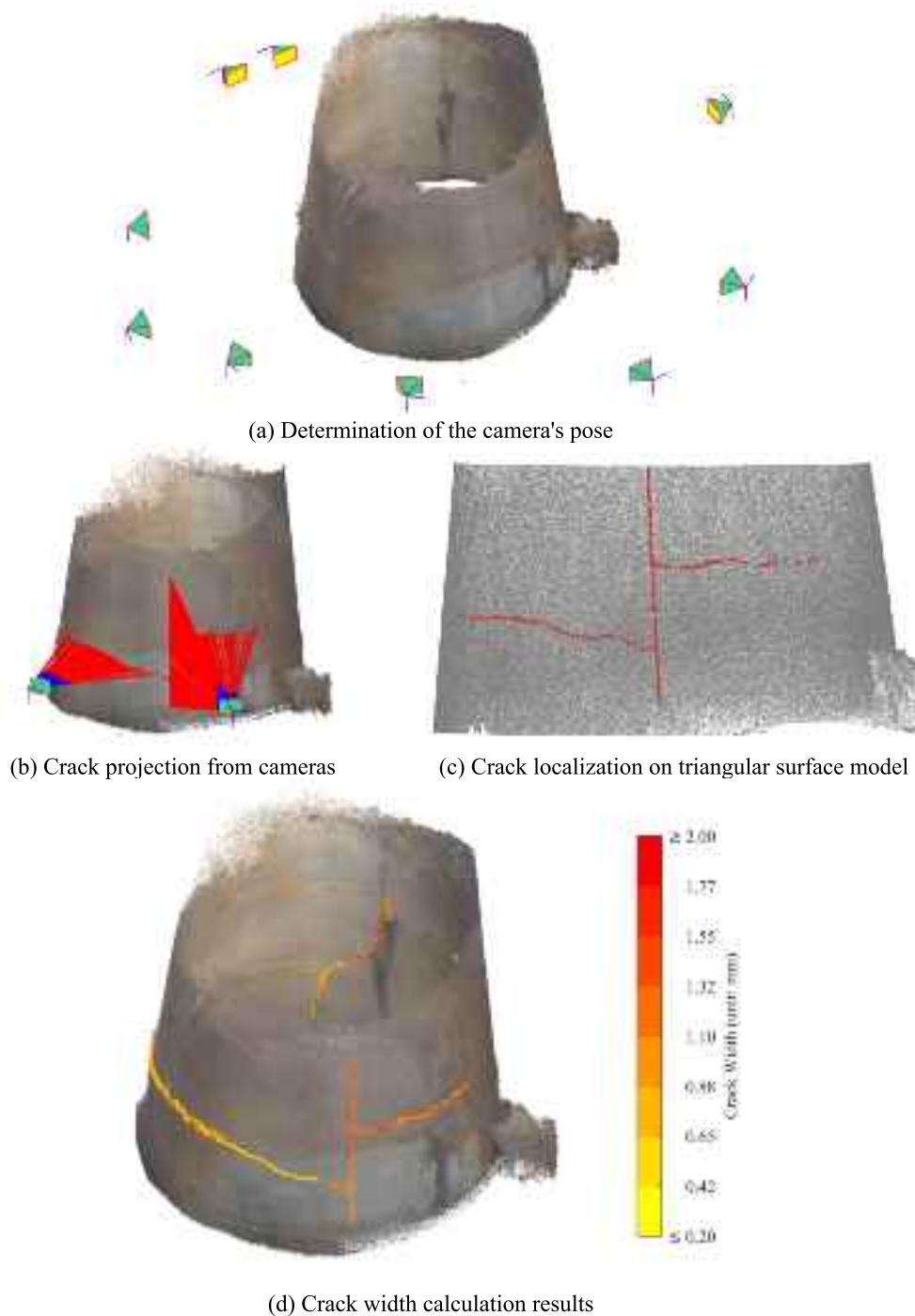
(a) Determination of the camera's pose

(b) Crack projection from cameras

(c) Crack localization on triangular surface model

(d) Crack width calculation results

Fig. 15. Results of crack projection and localization.

presented in this section, reconstruction failure occurs despite an average overlapping rate of 50% among the photos. Only when the average overlapping rate is increased to around 60% and the photos are carefully selected can the reconstruction process be successful. According to the reconstruction results in Table 4, the SLAM-based method takes approximately 94% less time than the SfM-based method to complete the same large scene reconstruction task. In engineering practice, the advantage of the SLAM-based method becomes more evident as the size of the scene increases.

From the above analysis, it can be concluded that: (i) In terms of data acquisition, the method based on multi-sensor fusion SLAM has stronger flexibility and convenience, making it more in line with the actual sit-uation of bridge inspection. (ii) SLAM significantly improves inspection efficiency by simultaneously obtaining point cloud results during mapping, unlike the time-consuming process of obtaining point cloud results through feature matching in 3D reconstruction based on SfM. (iii) Both methods can obtain the texture features of the structure surface, the texture obtained by the SfM method is more refined; however, its refinement comes at the cost of data processing speed. (iv) The method based on SfM requires precise trajectory planning before data acquisition, whereas the SLAM method only requires controlling the working distance less than u_{max}.
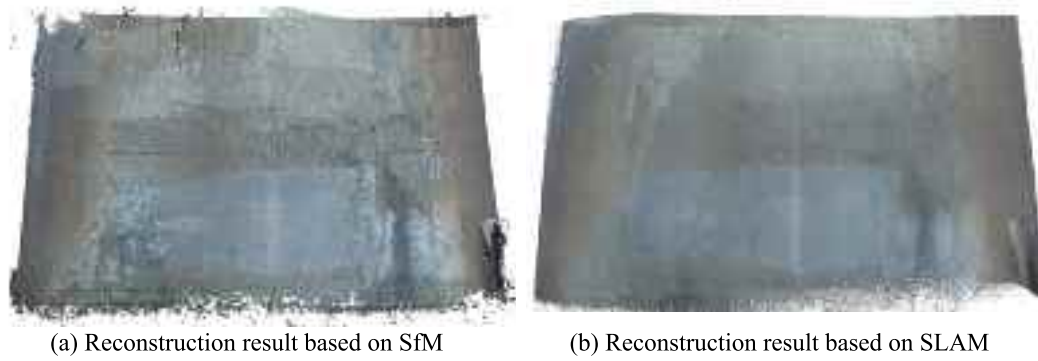
(a) Reconstruction result based on SfM   (b) Reconstruction result based on SLAM

Fig. 16. Comparison of pier reconstruction based on SfM and multi-sensor fusion SLAM.

Table 3
Comparison between point cloud quality of the inspection zone.

Evaluation metric	SfM-based	SLAM-based
Nonuniform distribution	Mean:7.67 Std.dev:2.21	Mean:2.76 Std.dev:1.41
Surface deviation	0.009 m	0.014 m
Geometric accuracy	5.14 m/5.10 m/0.78%	5.08 m/5.10 m/0.39%

6.4.2. More accurate in crack width calculation

To illustrate the contribution of image super-resolution based on ESRGAN to the accuracy of crack width calculation, this paper also utilizes bilinear interpolation method to interpolate crack images with the same magnification and calculates the width after projection according to the aforementioned process. The bilinear interpolation method, which involves simple interpolation calculations and does not introduce new features, serves as a research baseline for comparison with the ESRGAN-based crack image super-resolution method suggested in this paper. The crack width, absolute error, and relative error calculated by the two methods are listed in Table 5. Fig. 18 presents the relationship between crack width and calculation relative error.

The comparison reveals that the utilization of deep learning-based image super-resolution leads to a noteworthy enhancement in the accuracy of crack calculation. For cracks with a general width (#2 ~ #12 in this case), the calculated crack width is reduced by approximately 16% when employing deep learning-based image super-resolution, as compared to the bilinear interpolation method. For relatively fine cracks (#1 and #13 in this case), even the bilinear interpolation technique fails to effectively extract the crack features, indicating that the crack width at these locations is exceptionally small and cannot be normally
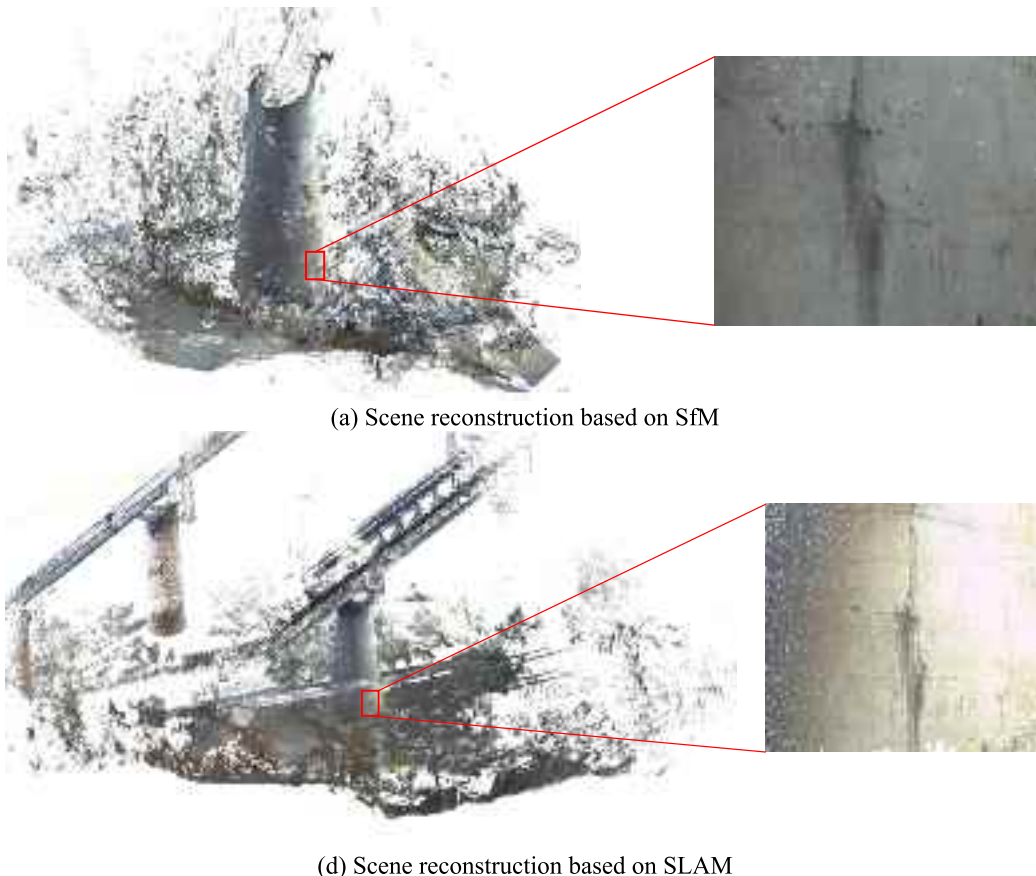

(a) Scene reconstruction based on SfM

(d) Scene reconstruction based on SLAM

Fig. 17. Results of scene reconstruction.

Table 4

Comparison between SfM-based and SLAM-based methods in scene reconstruction.

Method	SfM-based	SLAM-based
Number of points	5,766,525	7,653,839
Time taken in 3D reconstruction	SfM: 10 min 54 s MVS: 8 min 24 s Meshing: 1 min 4 s Total: 20 min 22 s	Point cloud: simultaneous Meshing: 1 min 12 s Total: 1 min 12 s
Texture quality	High	Medium
Requirements of path planning	Fine planning	Few restrictions
Scope of reconstructed scene	Medium	Large

calculated. Here, deep learning-based image super-resolution proves to be adept at restoring the intricate details of fine cracks and obtaining relatively reasonable width values.

The analysis shows that when the crack width is small, the relative error of calculation is generally large. This is because when the crack width is very small, the baseline for comparison becomes small, and measurement error, SLAM positioning error, projection error, etc. are more likely to affect the width calculation. However, in this case, the crack calculation values are higher than the actual values, indicating a bias towards safety in engineering applications. The experiment results demonstrate that for cracks with a width >0.9 mm, the relative error of calculation can be controlled within 20%. Hence, the proposed method can obtain relatively accurate crack width calculation values within this range.

According to the code "Technical specification for engineering structures inspection by digital image method" [17] promulgated by China Association for Engineering Construction Standardization, the relative accuracy of digital image method crack detection for width calculation should be controlled within 20%, so the proposed super-resolution method can meet the industry requirements. The minimum detectable crack width is 0.147 mm (not illustrated in Table 5) in the case study. This is close to 0.1 mm, which is suggested in the code [17] . However, the reliability of the calculation is reduced when the crack width is small, as demonstrated in the field tests. This is due to the accumulated errors in a series of steps such as data acquisition, crack segmentation, 3D reconstruction, and crack projection. Therefore, further research is warranted to improve the precise calculation of fine crack widths.

7. Conclusions

In this paper, a crack assessment method based on multi-sensor fusion SLAM and image super-resolution is described, which can significantly improve the efficiency and accuracy of automated bridge inspection while meeting engineering needs. It is applicable to various platforms including handheld devices, vehicle-mounted equipment, and UAVs. The paper provides guidance on data acquisition using multi-sensor fusion SLAM and explains how SLAM is used to obtain the textured point cloud and 3D model of the bridge. Moreover, the paper

Table 5

Comparison of the accuracy of crack width calculation with two interpolation methods.

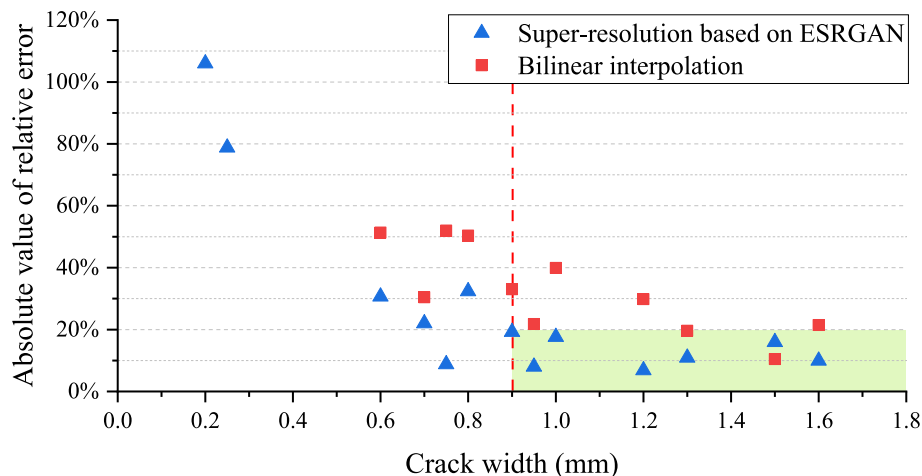| Measurement		Super-resolution based on ESRGAN			Bilinear interpolation		
Crack No.	Width	Calculation	Absolute error	Relative error	Calculation	Absolute error	Relative error
1	0.25	0.447	0.197	78.8%	–	–	–
2	0.80	1.059	0.259	32.4%	1.203	0.403	50.4%
3	1.00	1.176	0.176	17.6%	1.399	0.399	39.9%
4	1.20	1.282	0.082	6.8%	0.842	−0.358	29.8%
5	1.30	1.442	0.142	10.9%	1.045	−0.255	19.6%
6	0.75	0.816	0.066	8.8%	1.140	0.390	52.0%
7	0.60	0.784	0.184	30.7%	0.908	0.308	51.3%
8	0.70	0.854	0.154	22.0%	0.913	0.213	30.4%
9	0.90	1.073	0.173	19.2%	1.197	0.297	33.0%
10	0.95	1.026	0.076	8.0%	1.157	0.207	21.8%
11	1.60	1.441	−0.159	9.9%	1.257	−0.343	21.4%
12	1.50	1.261	−0.239	15.9%	1.343	−0.157	10.5%
13	0.20	0.412	0.212	106.0%	–	–	–
Mean relative error		#1 ~ #13:28.2% #2 ~ #12:16.6%			#2 ~ #12:32.7%		

Fig. 18. Relationship between crack width and calculation relative error.

highlights the use of deep learning-based image super-resolution to recover the details of cracks clearly and calculate the width feature accurately. In addition, the method of feature point projection is utilized to locate the cracks. Field tests on a bridge pier are conducted to showcase the effectiveness of the proposed method. The main conclusions of this paper are as follows:

1. Multi-sensor fusion SLAM can significantly reduce the acquisition time of textured point clouds in bridge inspection. In the comparative experiment of this paper, SLAM can save 94% of the time in this process compared with SfM, which dramatically improves inspection efficiency. The texture quality is slightly lower than that of the SfM method but is enough to support the needs of engineering inspection.
2. Multi-sensor fusion SLAM does not necessitate detailed path planning in the process of data acquisition. It is only necessary to ensure that the working distance does not exceed the calculated maximum distance, which reduces the complexity of inspection operation.
3. Image super-resolution based on deep learning can significantly enhance the accuracy of crack width calculation and ensure the inclusion of relatively small cracks in the width calculation. Consideration of image super-resolution can be made when it is not possible to achieve a high accuracy GSD due to limitations of the constraints of acquisition equipment and environment. By using image super-resolution, it is possible to expand the range of detectable crack widths. The improvement in results from image super-resolution depends on the availability of high-quality data acquisition, such as favorable lighting conditions and minimal motion blur.

In the future, certain technical aspects of this method can be enhanced. Firstly, it is important to address the issue of equipment selection in order to accommodate the varying scales of bridge structure inspection. Secondly, the crack detection algorithm can be further improved to acquire the width point features at the pixel level directly. Finally, there is a need to urgently study how to assess the accuracy of crack localization results in the world coordinate system, as the processes of point cloud generation and crack projection are independent.

In conclusion, the crack assessment scheme based on multi-sensor fusion SLAM and image super-resolution in bridge inspection exhibits good universality and great potential in engineering applications. The framework proposed in this paper can be used in the engineering practice of automated bridge inspection and adapted according to different crack detection and localization algorithms to achieve better results.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgment

References

[1] L. Sun, Z. Shang, Y. Xia, S. Bhowmick, S. Nagarajaiah, Review of bridge structural health monitoring aided by big data and artificial intelligence: from condition assessment to damage detection, J. Struct. Eng. 146 (2020) 04020073, https://doi.org/10.1061/(ASCE)ST.1943-541X.0002535.

[2] C. Eschmann, T. Wundsam, Web-based georeferenced 3D inspection and monitoring of bridges with unmanned aircraft systems, J. Surv. Eng. 143 (2017) 04017003, https://doi.org/10.1061/(ASCE)SU.1943-5428.0000221.

[3] Z. Ameli, Y. Aremanda, W.A. Friess, E.N. Landis, Impact of UAV hardware options on bridge inspection mission capabilities, Drones. 6 (2022) 64, https://doi.org/10.3390/drones6030064.

[4] H. Zakeri, F.M. Nejad, A. Fahimifar, Image based techniques for crack detection, classification and quantification in asphalt pavement: a review, Archiv. Comp. Methods Eng. 24 (2017) 935–977, https://doi.org/10.1007/s11831-016-9194-z.

[5] C.M. Yeum, S.J. Dyke, Vision-based automated crack detection for bridge inspection, Comp. Aid. Civ. Infrastruct. Eng. 30 (2015) 759–770, https://doi.org/10.1111/mice.12141.

[6] I. Abdel-Qader, O. Abudayyeh, M.E. Kelly, Analysis of edge-detection techniques for crack identification in bridges, J. Comput. Civ. Eng. 17 (2003) 255–263, https://doi.org/10.1061/(ASCE)0887-3801(2003)17:4(255).

[7] Q. Li, X. Liu, Novel approach to pavement image segmentation based on neighboring difference histogram method, in: Congress on Image and Signal Processing, 2008, pp. 792–796, https://doi.org/10.1109/CISP.2008.13.

[8] Y.Z. Ayele, M. Aliyari, D. Griffiths, E.L. Droguett, Automatic crack segmentation for UAV-assisted bridge inspection, Energies. 13 (2020) 6250, https://doi.org/10.3390/en13236250.

[9] Z. Yu, Y. Shen, C. Shen, A real-time detection approach for bridge cracks based on YOLOv4-FPM, Autom. Constr. 122 (2021) 103514, https://doi.org/10.1016/j.autcon.2020.103514.

[10] H. Fu, D. Meng, W. Li, Y. Wang, Bridge crack semantic segmentation based on improved Deeplabv3+, J. Mar. Sci. Eng. 9 (2021) 671, https://doi.org/10.3390/jmse9060671.

[11] L. Zhang, Y. Liao, G. Wang, J. Chen, H. Wang, A multi-scale contextual information enhancement network for crack segmentation, Appl. Sci. 12 (2022) 11135, https://doi.org/10.3390/app122111135.

[12] Ministry of Transport of the People''s Republic of China, Technical Specifications for Field Inspections of Existing Highway Bridges, JTG/T 5214–2022, People's Transportation Press, 2022.

[13] B.Y. Lee, Y.Y. Kim, S.-T. Yi, J.-K. Kim, Automated image processing technique for detecting and analysing concrete surface cracks, Struct. Infrastruct. Eng. 9 (2013) 567–577, https://doi.org/10.1080/15732479.2011.593891.

[14] Y. Liu, Multi-scale Structural Damage Assessment Based on Model Updating and Image Processing, Phd thesis, Tsinghua University, 2015, https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CDFD&dbname=CDFDLAST2022&filename=1016712246.nh&uniplatform=NZKPT&v=JULOg9yIjl3Q2sUG97jwbVZHSaOW7XTp0NCAP8DB3uC3cVLrbrh6NEyL-unvltt8 (accessed July 13, 2023).

[15] T. Nishikawa, J. Yoshida, T. Sugiyama, Y. Fujino, Concrete crack detection by multiple sequential image filtering, Comp. Aid. Civ. Infrastruct. Eng. 27 (2012) 29–47, https://doi.org/10.1111/j.1467-8667.2011.00716.x.

[16] A. Mirzazade, C. Popescu, J. Gonzalez-Libreros, T. Blanksvärd, B. Täljsten, G. Sas, Semi-autonomous inspection for concrete structures using digital models and a hybrid approach based on deep learning and photogrammetry, J. Civ. Struct. Heal. Monit. (2023) 1–20, https://doi.org/10.1007/s13349-023-00680-x.

[17] China Association for Engineering Construction Standardization, Technical Specification for Engineering Structures Inspection by Digital Image Method, T/CECS 1114–2022, China Architecture & Building Press, 2022.

[18] G. Morgenthal, N. Hallermann, J. Kersten, J. Taraben, P. Debus, M. Helmrich, V. Rodehorst, Framework for automated UAS-based structural condition assessment of bridges, Autom. Constr. 97 (2019) 77–95, https://doi.org/10.1016/j.autcon.2018.10.006.

[19] G. Cha, S. Park, T. Oh, A terrestrial LiDAR-based detection of shape deformation for maintenance of bridge structures, J. Constr. Eng. Manag. 145 (2019) 04019075, https://doi.org/10.1061/(ASCE)CO.1943-7862.0001701.

[20] S. Guan, Z. Zhu, G. Wang, A review on UAV-based remote sensing technologies for construction and civil applications, Drones. 6 (2022) 117, https://doi.org/10.3390/drones6050117.

[21] Y. Liu, S. Cho, B.F. Spencer, J.-S. Fan, Concrete crack assessment using digital image processing and 3D scene reconstruction, J. Comput. Civ. Eng. 30 (2016) 04014124, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000446.

[22] Y. Xu, J. Zhang, UAV-based bridge geometric shape measurement using automatic bridge component detection and distributed multi-view reconstruction, Autom. Constr. 140 (2022) 104376, https://doi.org/10.1016/j.autcon.2022.104376.

[23] Y. Liu, X. Nie, J. Fan, X. Liu, Image-based crack assessment of bridge piers using unmanned aerial vehicles and three-dimensional scene reconstruction, Comp. Aid. Civ. Infrastruct. Eng. 35 (2020) 511–529, https://doi.org/10.1111/mice.12501.

[24] A. Khaloo, D. Lattanzi, K. Cunningham, R. Dell'Andrea, M. Riley, Unmanned aerial vehicle inspection of the Placer River Trail Bridge through image-based 3D modelling, Struct. Infrastruct. Eng. 14 (2018) 124–136, https://doi.org/10.1080/15732479.2017.1330891.

[25] M. Pepe, D. Costantino, UAV photogrammetry and 3D modelling of complex architecture for maintenance purposes: the case study of the Masonry Bridge on the Sele River, Italy, Periodica Polytechnica, Civ. Eng. 65 (2021) 191–203, https://doi.org/10.3311/PPci.16398.

[26] J.L. Carrivick, M.W. Smith, D.J. Quincey, Structure from Motion in the Geosciences, John Wiley & Sons, 2016. ISBN: 1118895843.

[27] A. Kim, R.M. Eustice, Real-time visual SLAM for autonomous underwater hull inspection using visual saliency, IEEE Trans. Robot. 29 (2013) 719–733, https://doi.org/10.1109/TRO.2012.2235699.

[28] W. Hess, D. Kohler, H. Rapp, D. Andor, Real-time loop closure in 2D LIDAR SLAM, in: IEEE International Conference on Robotics and Automation (ICRA), 2016, pp. 1271–1278, https://doi.org/10.1109/ICRA.2016.7487258.

[29] J. Zhang, S. Singh, LOAM: Lidar odometry and mapping in real-time, in: Robotics: Science and Systems vol. 2, 2014, pp. 1–9, https://doi.org/10.15607/RSS.2014. X.007, no. 9.

[30] W. Xu, F. Zhang, FAST-LIO: a fast, robust LiDAR-inertial odometry package by tightly-coupled iterated kalman filter, in: IEEE Robotics and Automation Letters 6, 2021, pp. 3317–3324, https://doi.org/10.1109/LRA.2021.3064227.

[31] S. Jung, D. Choi, S. Song, H. Myung, Bridge inspection using unmanned aerial vehicle based on HG-SLAM: hierarchical graph-based SLAM, Remote Sens. 12 (2020) 3022, https://doi.org/10.3390/rs12183022.

[32] A. Gupta, X. Fernando, Simultaneous localization and mapping (SLAM) and data fusion in unmanned aerial vehicles: recent advances and challenges, Drones. 6 (2022) 85, https://doi.org/10.3390/drones6040085.

[33] T. Du, Y.H. Zeng, J. Yang, C.Z. Tian, P.F. Bai, Multi-sensor fusion SLAM approach for the mobile robot with a bio-inspired polarised skylight sensor, IET Radar Sonar Navigat. 14 (2020) 1950–1957, https://doi.org/10.1049/iet-rsn.2020.0260.

[34] C. Debeunne, D. Vivet, A review of Visual-LiDAR fusion based simultaneous localization and mapping, Sensors. 20 (2020) 2068, https://doi.org/10.3390/ s20072068.

[35] J. Lin, C. Zheng, W. Xu, F. Zhang, R2LIVE: a robust, real-time, LiDAR-Inertial-visual tightly-coupled state estimator and mapping, IEEE Robot. Autom. Lett. 6 (2021) 7469–7476, https://doi.org/10.1109/LRA.2021.3095515.

[36] J. Lin, F. Zhang, R3LIVE: a robust, real-time, RGB-colored, LiDAR-Inertial-Visual tightly-coupled state Estimation and mapping package, in: International Conference on Robotics and Automation (ICRA), 2022, pp. 10672–10678, https:// doi.org/10.1109/ICRA46639.2022.9811935.

[37] S. Yang, H. Chen, X. Li, Research on crack width judging by image gray scale, Highway Traffic Technol. (Appl. Technol. Ed.) 14 (2018) 71–72, https://kns.cnki. net/kcms2/article/abstract?v=3uoqIhG8C44YLTlOAiTRKibYlV5Vjs7i0-kJR0HYBJ 80QN9L51zrP6slJd751YSWU5Ue1k-U4fQC_5y7IZdi7X7PXU6DmCDM&uniplatfor m=NZKPT (accessed July 13, 2023).

[38] J. Mao, H. Fu, C. Chu, X. He, C. Chen, A review of simultaneous localization and mapping based on inertial-visual-lidar fusion, Navig. Position. Timing. 9 (2022) 17–30, https://doi.org/10.19306/j.cnki.2095-8110.2022.04.003.

[39] J. Wang, X. Zuo, X. Zhao, J. Lyu, Y. Liu, Review of multi-source fusion SLAM: current status and challenges, J. Image Graph. 27 (2022) 368–389, https://doi. org/10.11834/jig.210547.

[40] S. Zhao, H. Zhang, P. Wang, L. Nogueira, S. Scherer, Super odometry: IMU-centric LiDAR-visual-inertial estimator for challenging environments, in: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2021, pp. 8729–8736, https://doi.org/10.1109/IROS51168.2021.9635862.

[41] P. Labatut, J.-P. Pons, R. Keriven, Efficient multi-view reconstruction of large-scale scenes using interest points, delaunay triangulation and graph cuts, in: IEEE 11th International Conference on Computer Vision, 2007, pp. 1–8, https://doi.org/ 10.1109/ICCV.2007.4408892.

[42] B. Li, Y. Qi, J. Fan, Y. Liu, C. Liu, A grid-based classification and box-based detection fusion model for asphalt pavement crack, in: Computer-Aided Civil and Infrastructure Engineering, 2022, https://doi.org/10.1111/mice.12962.

[43] C. Dong, C.C. Loy, K. He, X. Tang, Learning a deep convolutional network for image super-resolution, in: Computer Vision–ECCV 2014: 13th European Conference, 2014, pp. 184–199, https://doi.org/10.1007/978-3-319-10593-2_13.

[44] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, W. Shi, Photo-realistic single image super-resolution using a generative adversarial network, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 105–114, https://doi.org/10.1109/ CVPR.2017.19.

[45] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, C. Change Loy, ESRGAN: enhanced super-resolution generative adversarial networks, in: Proceedings of the European Conference on Computer Vision (ECCV) Workshops, 2018. https://opena ccess.thecvf.com/content_eccv_2018_workshops/w25/html/Wang_ESRGAN_Enha nced_Super-Resolution_Generative_Adversarial_Networks_ECCVW_2018_paper. html (accessed July 13, 2023).

[46] X. Wang, L. Xie, C. Dong, Y. Shan, Real-ESRGAN: training real-world blind super-resolution with pure synthetic data, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 1905–1914, https://openaccess.thecvf. com/content/ICCV2021W/AIM/html/Wang_Real-ESRGAN_Training_Real-World_ Blind_Super-Resolution_With_Pure_Synthetic_Data_ICCVW_2021_paper.html (accessed July 13, 2023).

[47] C. Ericson, Real-time Collision Detection, Crc Press, 2004. ISBN:1558607323.

[48] Ministry of Transport of the People''s Republic of China, Standards for Technical Condition Evaluation for Highway Bridges: JTG/T H21–2011, People's Transportation Press, 2011.

[49] Ministry of Transport of the People''s Republic of China, Specifications for Maintenance of Highway Bridges and Culverts: JTG 5120–2021, People's Transportation Press, 2021.

[50] Ministry of Transport of the People''s Republic of China, Technical Specifications for Structural Monitoring of Highway Bridges: JT/T 1037–2022, People's Transportation Press, 2022.

[51] Z. Zhang, Flexible camera calibration by viewing a plane from unknown orientations, in: Proceedings of the Seventh IEEE International Conference on Computer Vision vol. 1, 1999, pp. 666–673, https://doi.org/10.1109/ ICCV.1999.791289.

[52] C. Yuan, X. Liu, X. Hong, F. Zhang, Pixel-level extrinsic self calibration of high resolution LiDAR and camera in targetless environments, IEEE Robot. Autom. Lett. 6 (2021) 7517–7524, https://doi.org/10.1109/LRA.2021.3098923.

[53] S. Chen, D.F. Laefer, E. Mangina, S.M.I. Zolanvari, J. Byrne, UAV bridge inspection through evaluated 3D reconstructions, J. Bridg. Eng. 24 (2019) 05019001, https:// doi.org/10.1061/(ASCE)BE.1943-5592.0001343.