Smart Structures and Systems, *Vol. 33, No. 5 (2024) 349-358*
https://doi.org/10.12989/sss.2024.33.5.349

Loop closure-based high-resolution façade digital modeling technique of large-scale dams using UAV

Myung Soo Kang [1a], Keunyoung Jang [1b], Yong-Rae Yu [2c] and Yun-Kyu An*[1]

[1] *Department of Architectural Engineering, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Republic of Korea*
[2] *Department of Civil and Environmental Engineering, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Republic of Korea*

(*Received November 20, 2023, Revised April 25, 2024, Accepted May 13, 2024*)

Abstract. Structural digital models can be effectively established by spatially obtaining digital images using an unmanned aerial vehicle (UAV). One of the main purposes of the structural digital modeling is computer vision-based exterior damage detection of a target structure. To investigate micro-scale damage from the digital model, high-resolution digital images obtained with a close-up vision survey is typically required. However, serial image synthesis such as image stitching may cumulate stitching errors as the number of scanned images increases. Therefore, in this paper, a novel loop closure-based digital image stitching technique is proposed and experimentally validated using the close-up surveyed digital images acquired from an in-situ dam structure located in South Korea. The test results reveal that the proposed technique outperforms a non-loop closure-based image stitching technique, which can cause serious distortions, such as ghosting and vanishing phenomena.

Keywords: close-up UAV surveying; computer vision; dam; digital image stitching; loop closure; structural façade digital model

1. Introduction

With a rapid development of unmanned aerial vehicle (UAV), structural digital modeling recently has got limelight for effective structural reconstruction and maintenance. By employing UAVs, the structural digital models can be established much faster and safer than conventional expert-dependent measurements. UAVs facilitate the rapid acquisition of structural geometry and surface information, even in areas inaccessible to humans. This capability enables a significant reduction in economic costs while ensuring high-quality data acquisition compared to conventional methods. For example, Jiang *et al*. (2020) proposed a UAV-based 3D modeling technique of a construction site to effectively manage the construction planning and safety. Similarly, Varbla *et al*. (2020) studied 3D model reconstruction of as-built construction using real-time kinematic global navigation satellite system (RTK-GNSS)-equipped UAVs. The reconstructed geometry can be directly utilized as building information modeling input. Bang *et al*. (2017) introduced a UAV-based approach for generating large-scale panoramas of construction sites, enabling timely and frequent monitoring of construction progress. Additionally, UAV-based structural digital modeling has been employed not only for effective construction process management but also for the maintenance of existing structures such as buildings (Jhan *et al*. 2021, Drešček *et al*. 2020, Murtiyoso and Grussenmeyer 2017), bridges (Chen *et al*. 2019, Pan *et al*. 2019, Khaloo *et al*. 2018) and dams (Zhao *et al*. 2021).

In terms of structural digital modeling methodologies, various techniques based on point cloud data have been proposed. As one of the representative techniques, light detection and ranging (LiDAR)-based point cloud generation is extensively employed for constructing structural digital models. By utilizing laser pulses to measure the precise distance between the sensor and surface of the target structure, LiDAR can accurately generate a 3D point cloud model through the integration of spatially scanned distance data. For example, Vosselman and Dijkman (2001) proposed the airborne LiDAR-based 3D reconstruction of urban buildings using 3D Hough transform and ground plans. Chiang *et al*. (2017) constructed a direct geo-referenced point cloud model of terrain and buildings by integrating inertial navigation, GNSS and LiDAR-embedded UAV. Wu *et al*. (2017) applied a graph theory-based method for urban building model reconstruction, employing hierarchical structure analysis of building contours obtained through LiDAR-equipped UAV. Although LiDAR-based sparse point cloud modeling offers precise geometric information, it may lack the resolution necessary to discern structural surface damage. To address this technical constraint, photogrammetry-based 3D structural digital modeling techniques, which rely on high-resolution digital images rather than sparse laser scanning, have been studied (Cavalagli *et al*. 2020, Zollini *et al*. 2020, Pan *et al*. 2019). By incorporating with UAV-mounted vision cameras, the

*Corresponding author, Ph.D., Professor,
E-mail: yunkyuan@sejong.ac.kr
[a] Ph.D. Student, E-mail: kms35954@sju.ac.kr
[b] Ph.D. Student, E-mail: jj99137@sju.ac.kr
[c] Ph.D. Student, E-mail: pop9849@gmail.com

ISSN: 1738-1584 (Print), 1738-1991 (Online)

photogrammetry software such as MetaShape (Agisoft 2021), 4DiXplorer (4DiXplorer 2021) and ColMap (Schonberger 2021) have been considerably well commercialized to establish structural digital models. The photogrammetry-based digital modeling techniques initially build meshes from dense point cloud data, which is generated using structure-from-motion (Snavely *et al*. 2008) and a multi-view stereo algorithm (Furukawa and Ponce 2010, Shen 2013). Then, additional texturing processes are often employed to enhance the fidelity of digital models (Zollini *et al*. 2020, Dabous *et al*. 2023). However, it is important to note that surface features, including damages, may become blurred, superimposed, or even disappear during the texturing process. Consequently, directly inspecting and quantifying structural damages using photogrammetric 3D models remains a challenging task.

In this paper, a novel 2D structural façade digital modeling technique is proposed to tackle the mentioned technical issue. For planar-like structures such as dams or continuously wide buildings, a 2D high-resolution façade digital model proves to be more effective than 3D photogrammetry models for investigating exterior damage. This effectiveness stems from its capacity to reduce data size and acquire high-resolution images. Additionally, employing spatial data surveying with a long working distance for 3D digital modeling often result in insufficient surface image resolution. With these advantages of 2D digital modeling, Boller *et al*. (2015) utilized a 2D façade digital model constructed through digital image stitching to inspect structural surface integrity and assess the life cycle of ageing concrete structures. Wang *et al*. (2021) proposed a fast image stitching for characterizing cracks in large-scale structures. They noted the possibility of precise damage evaluation by combining with 2D façade digital model-based visual inspection and conventional non-destructive techniques. Similarly, Cheng *et al*. (2021) demonstrated visual inspection of structures based on 2D façade digital model constructed via digital image stitching. Then, they quantitatively assessed surface cracks on the target structure using a computer-aided method in pixel level. Jang *et al*. (2021) developed a feature control-based exterior damage mapping technique for high-rise bridge pier using continuously obtained digital images with a ring-type climbing robot. Similarly, Jang *et al*. (2024) introduces a deep learning-based digital damage mapping on 3D model for vertical-type tunnels, aiming to overcome the hazards and unreliability associated with expert-dependent visual inspection.

Additionally, Wang *et al*. (2020) introduced a 2D digital model of a steel box girder for computer vision-based automatic inspection. They synthesized spatially continuous digital images obtained through close-up surveying with a UAV-mounted vision camera. However, in reality, errors in façade digital modeling of large-scale structures might be inevitable. It is because that stitching errors inherently accumulate during the concatenative camera parameter estimation in the image stitching process, resulting in serious distortions of established façade digital model (Brown and Lowe 2006). To tackle these technical huddles, this paper proposes a novel loop closure-based technique for establishing structural façade digital models of large-scale dam structures. The proposed technique aims to minimize distortions in the established façade digital model and reduce computational costs through feature matching analysis and optimal image selection algorithms. This paper is organized as follows. First, the loop closure-based digital image stitching technique including the pre-processing of digital images is explained in Section 2. The experimental validation results are then shown in Section 3. Finally, this paper is concluded with a brief discussion in Sections 4 and 5.

2. Loop closure-based façade digital model establishment

Fig. 1 illustrates the procedure for establishing a loop closure-based façade digital model for a dam structure.

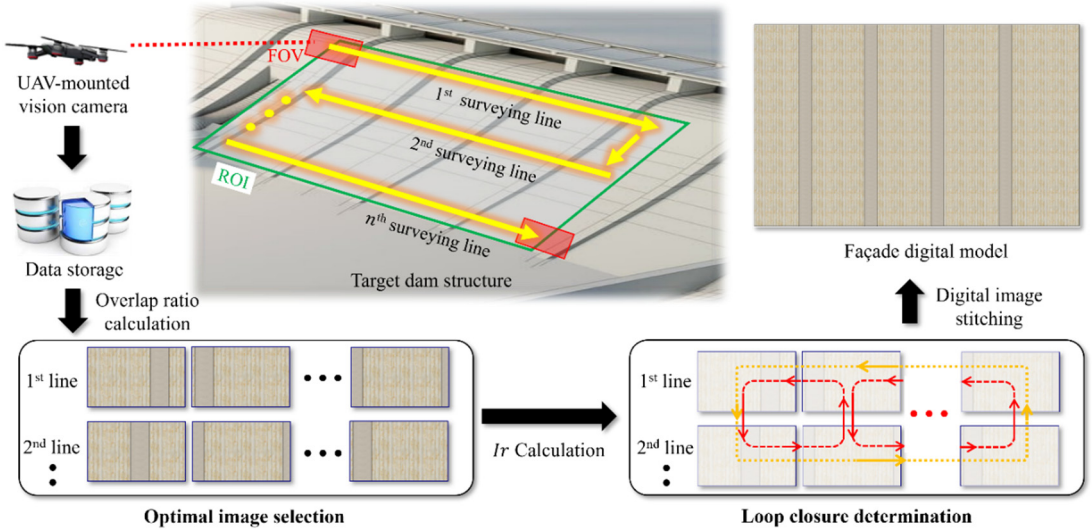


Fig. 1 Overview of loop closure-based façade digital model establishment for a dam structure: FOV, ROI and *Ir* denote the field of view, region of interest and inlier ratio, respectively

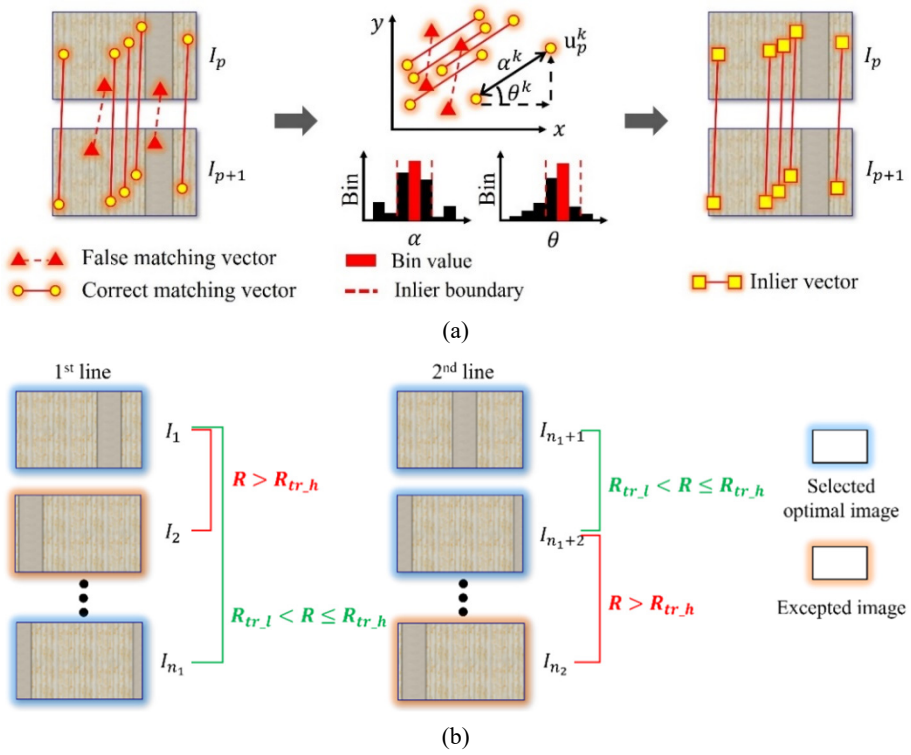


Fig. 2 Optimal image selection algorithm: (a) Feature matching analysis and (b) R calculation-based optimal image selection. I_p is the p^{th} image. $u_{p,p+1}^k$ is the k^{th} matching vector between I_p and I_{p+1}. θ^k and α^k are the angle and magnitude values of $u_{p,p+1}^k$, respectively. R is the overlap ratio, R_{tr_l} and R_{tr_h} are the lower and upper threshold values of R, respectively

First, a UAV-mounted vision camera surveys the dam surface with a certain working distance along a pre-defined surveying path. This path comprises multiple surveying lines, as shown in Fig. 1. While a closer working distance offers advantages for investigating micro-scale structural damages, it is crucial to design the appropriate working distance by considering factors such as UAV safety, camera pixel resolution, data size and computational costs.

2.1 Optimal image selection for effective façade digital modeling

Among the raw image surveyed from the target dam, the selection of optimal images is crucial for digital image stitching, as depicted in Fig. 2. Since the raw data often contain densely surveyed, blurred or abruptly camera pose changed images, it is necessary to filter out unwanted image data to minimize stitching errors. Moreover, using only optimal images can reduce computational costs.

First, image features such as points, corners or edges are extracted and preliminarily matched between adjacent images along each surveying line based on scale-invariant feature transform (SIFT) (Lowe 1999). The initial connections of the matched features are defined as the matching vectors as shown in Fig. 2(a). Here, false matching vectors may arise due to low image resolution, similar or repeated surface features within the images, or a lack of the target surface features. To eliminate these false matching vectors and retain only inlier vectors, the constraints on the magnitude (α_f) and angle (θ_f) of the matching vectors are applied. A set of inlier vectors $\{\mathcal{U}_{p,p+1}\}$ between I_p and I_{p+1} can be written as

$$\{\mathcal{U}_{p,p+1}\} = \mathrm{c}^k \mathrm{u}_{p,p+1}^k, \quad (k = 1, 2, 3, \ldots, \mathcal{K}),$$
$$\mathrm{c}^k \begin{cases} 1, & if\ \alpha^k \in [\alpha_f - \alpha_{tr}, \alpha_f + \alpha_{tr}] \text{ and} \\ & \theta^k \in [\theta_f - \theta_{tr}, \theta_f + \theta_{tr}] \\ 0, & \text{Otherwise} \end{cases} \tag{1}$$

where $u_{p,p+1}^k$ is the k^{th} matching vector between I_p and I_{p+1}. $\mathcal{K}$ is the number of initial matching vectors. c^k is the feature matching coefficient. α^k and θ^k are the magnitude and angle values of $u_{p,p+1}^k$, respectively. α_{tr} and θ_{tr} are the threshold values of α^k and θ^k. Next, R between I_p and I_{p+1} is calculated by

$$R = \frac{(W - \mathcal{X}_p) \times (H - \mathcal{Y}_p)}{W \times H} \times 100 \tag{2}$$

where W and H are the width and height of I_p, respectively. The averaged values of x and y directional components of $\{\mathcal{U}_{p,p+1}\}$ are defined as $\mathcal{X}_p$ and $\mathcal{Y}_p$. By the sequential R calculation from I_1 to I_{n_1}, the image satisfying the thresholds of R (R_{tr_l} and R_{tr_h}) is selected as the first optimal images as described in Fig. 2(b). If the selected image is $I_{p'}$, R is then repetitively calculated from $I_{p'}$ with $I_{p'+1}$, $I_{p'+2}$, …, I_{n_1} ($I_{p'} < I_{n_1}$) to determine the next optimal image. The optimal images corresponding to each surveying line can be obtained by

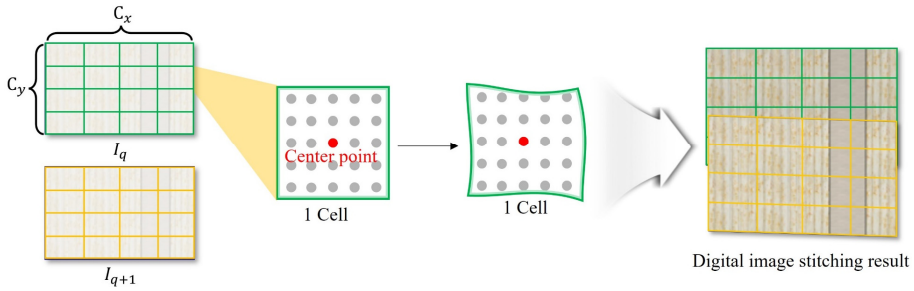


Fig. 3 Mesh-based digital image stitching: C_x, C_y are the grid mesh of selected optimal image I_q along x, y directions

repeating the above procedure.

2.2 Loop closure determination and façade digital model establishment

In order to achieve accurate stitching of the selected optimal images, precise intrinsic and extrinsic camera parameters need to be estimated. However, obtaining reliable camera parameters from surveyed digital images is challenging because of factors such as lens distortion, camera noise and low image data quality resulting from harsh outdoor conditions. To address these challenges, bundle adjustment technique is commonly employed in large-scale digital image stitching. This technique optimizes camera parameters while minimizing reprojection errors across multiple surveyed digital images (Brown and Lowe 2006). Typically, two images along the surveying path are used to estimate and optimize the camera parameters during the bundle adjustment process. However, employing more related images, such as those from adjacent surveying lines, can significantly reduce reprojection errors by leveraging multi-view information. Therefore, the loop closure, which is formed by determining the image pairs to be utilized for the bundle adjustment, is proposed. In this process, it is assumed that each surveying line is spatially close enough to have overlapping area.

The optimal images selected in Section 2.1 are stitched using the mesh-based natural image stitching with the global similarity prior (NISwGSP) is employed in this paper (Chen and Chuang 2016). As shown in Fig. 3, each optimal image is divided into grid meshes to guide the image deformation. In this algorithm, as-projective-as-possible (APAP) method is adopted for a local warp (Zaragoza *et al.* 2013), because a global warp is often not flexible enough to stitch large number of images. The mesh-based stitching algorithm attempts to find a set of deformed vertex positions $\widetilde{\mathrm{V}}$ such that an energy function $\Phi(\mathrm{V})$ is minimized for naturally stitching multiple images as much as possible. To achieve this purpose, the original perspective of each image has to be preserved locally, while it attempts to globally maintain a good structural shape by finding proper scales and rotations for images. Thus, $\Phi(\mathrm{V})$ consists of the following three terms, i.e., alignment (Φ_a), local similarity (Φ_l) and global similarity (Φ_g). First, Φ_a is the term that guarantees the quality of alignment after deformation by keeping overlapped mesh vertices aligned with their correspondences. Then, Φ_l provides regularization and propagates alignment constraints from the overlap regions to the non-overlap ones. Lastly, Φ_g requires each deformed image to undergo a similarity transform as much as possible, which is essential to achieving a naturalness of the stitched image. The optimal deformation of meshes ($\widetilde{\mathrm{V}}$) can be determined by

$$\widetilde{\mathrm{V}} = \underset{\widetilde{\mathrm{V}}}{\operatorname{argmin}}\, \Phi_\mathrm{a}(\mathrm{V}) + \lambda_l \Phi_l(\mathrm{V}) + \Phi_\mathrm{g}(\mathrm{V}) \tag{3}$$

where V denotes the set of vertices and edges in the grid mesh for the image. λ_l is the coefficient of $\Phi_l(\mathrm{V})$.

Once the mesh deformation is determined, the focal length and 3D rotation of each image are subsequently estimated by using the improved bundle adjustment technique. As a result of APAP, a homography is obtained for each quad of mesh, and the corresponding focal lengths can be calculated (Szeliski 2006). The median of these focal lengths is utilized as the initialization of the focal length, and the initial intrinsic matrix $\mathbf{K}_i$ for I_i is derived. Using $\mathbf{K}_i$, the initial guess for 3D rotation $\mathbf{R}_{ij}$ between I_i and I_j is calculated as

$$\mathbf{R}_{ij} = \underset{\mathbf{R}}{\operatorname{argmin}} \sum_{\mathrm{v}_{ij}^m \in \mathrm{M}_{ij}} \left\| \mathbf{K}_j \mathbf{R} \mathbf{K}_i^{-1} \mathrm{v}_{ij}^m - \psi\left(\mathrm{v}_{ij}^m\right) \right\|^2 \tag{4}$$

where M_{ij} is the mesh vertices of I_i in the overlapped area between I_i and I_j. v_{ij}^m and $\psi(\mathrm{v}_{ij}^m)$ denote the overlapped mesh vertices and its correspondence after warping. Subsequently, with initial guess of $\mathbf{K}_i$ and $\mathbf{R}_{ij}$, the optimal focal length $\boldsymbol{\mathcal{F}}_i$ and 3D rotation $\mathbf{R}_i$ can be obtained by bundle adjustment.

$$\begin{aligned} e &= \sum_{i=1}^{N} \sum_{j \in \mathfrak{T}(i)} \left| \mathrm{r}_{ij}^s \right|^2, \\ \mathrm{r}_{ij}^s &= \mathrm{v}_{ij}^s - p_{ij}^s, \\ p_{ij}^s &= \mathbf{K}_i \mathbf{R}_i \mathbf{R}_j^{\mathrm{T}} \mathbf{K}_j^{-1} \mathrm{r}_{ij}^s. \end{aligned} \tag{5}$$

where N is the number of optimal images, $\mathfrak{T}(i)$ is the set of images to be considered in bundle adjustment process. p_{ij}^s is the reprojection point of v_{ij}^s.

Despite the enhanced bundle adjustment of NISwGSP, misalignment of images within the panoramic domain is still remained especially on large-scale structures. To mitigate image misalignment, local correction of camera parameters for each image is necessary using loop closures.

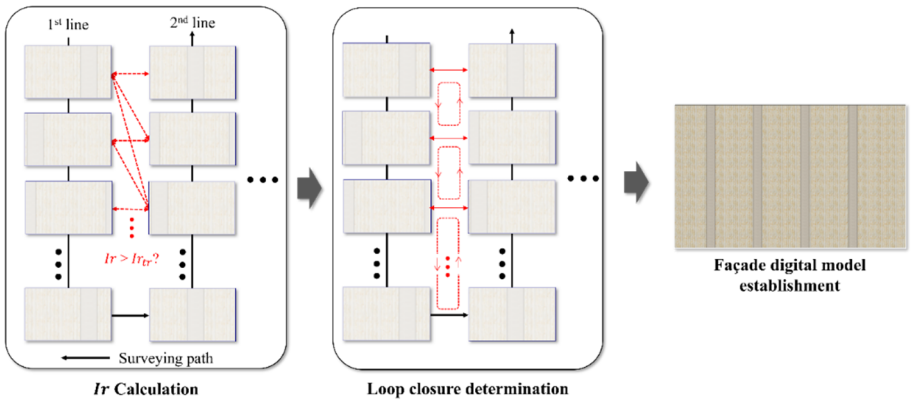


Fig. 4 Loop closure-based façade digital model establishment

(a) (b)

Fig. 5 Target in-situ dam structure: (a) Pre-defined surveying path; (b) UAV-mounted vision camera

First, Ir is calculated for all possible image pair combinations between the adjacent surveying lines using the ratio of the numbers of initial matching and inlier vectors as shown in Fig. 4.

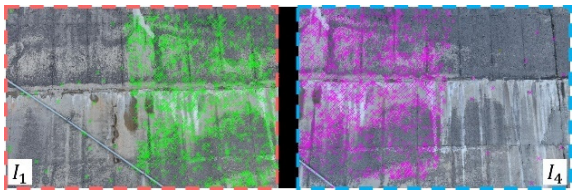
$$Ir = \frac{\mathcal{K}'}{\mathcal{K}} \times 100 \tag{6}$$

where $\mathcal{K}'$ is the number of inlier vectors. Here, a higher Ir value signifies a stronger correlation between the corresponding image pairs, physically meaning that they share more image features. The image pairs satisfying the threshold of Ir (Ir_{tr}) are then determined along the entire surveying line, and the determined image pairs will be included in $\mathfrak{T}$ of Eq. (5) to locally optimize the camera parameters of each image.

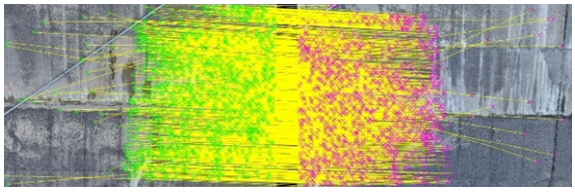
3. Experimental validation

The proposed technique was experimentally validated using a UAV-mounted vision camera at an in-situ dam located in South Korea, as described in Fig. 5. This concrete dam, standing at 28 m in height and spanning 171 m in length, possesses a total volume of 49,555 m^3. First, the target dam surface was surveyed by the UAV-mounted vision camera (EVO2 PRO of Autel Robotics) along the pre-defined surveying path to acquire spatially continuous digital images, as outlined in Fig. 5. In order to validate the proposed technique for planar-like structures, the left-non-overflow section of the dam was designated as ROI. A total of 7 horizontal lines were established as the pre-defined surveying path for the target structure. To keep the working distance of 6.4 m along the normal vector direction between the target surface and UAV, UAV was semi-autonomously controlled by a well-trained expert. Total 146 digital images with 5,472 × 3,648 of pixel resolution were acquired with respect to ROI, and the corresponding ground sample distance was 0.2 mm/pixel. Then, the optimal images were selected by the optimal image selection algorithm, and loop closures were identified between each adjacent surveying line. Finally, the dam façade digital model was established by using the mesh-based digital image stitching algorithm. To experimentally verify the superiority of the proposed technique, the façade digital model with loop closure was compared against the model established using the NISwGSP of a previous study (Chen and Chuang 2016).

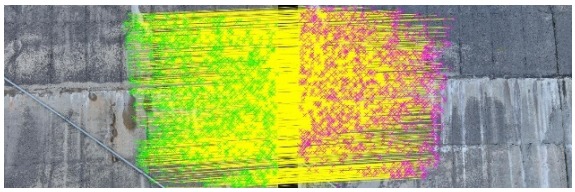
Fig. 6 shows the representative R calculation results between I_1 and I_4, for optimal image selection. As depicted in Fig. 6(a), 11,409 and 8,814 image features were extracted from I_1 and I_4, respectively. Among them, 1,845 matching vectors were initially matched between I_1 and I_4 as shown in Fig. 6(b). Here, α_f and θ_f were 1,189 pixels and -1° along the x direction, respectively. The inlier



(a)



(b)



(c)

(d)

Fig. 6 Representative R calculation results for optimal image selection using I_1 and I_4: (a) Feature extraction; (b) initial matching vector; (c) inlier vector; and (d) calculation of R, 61.58%

boundary of $\pm$ 5% was used, and the false matching vectors were removed using Eq. (1), resulting that 1,817 inlier vectors remained as displayed in Fig. 6(c). Using Eq. (2), R between I_1 and I_4 was calculated to be 61.58% as shown in Fig. 6(d), and I_4 was determined as the first optimal image. By repeating the optimal image selection procedure for each surveying line, a total of 46 optimal images were selected, resulting in a reduction of approximately 68.5% compared to the raw data. R_{th_l} and R_{th_h} were set at 55% and 65%, respectively. Here, the number of optimal images at each surveying line can be varied due to inlier feature matching sensitivity as well as UAV surveying speed variation.

Subsequently, Ir is calculated between the adjacent surveying lines using Eq. (6). Here, Ir_{tr} of 90% was used. Based on the calculation results of Ir, a total of 125 image pairs satisfies the Ir_{tr} along the entire surveying lines. The representative loop closures between the surveying lines 1 and 2 are shown in Fig. 7.

Figs. 8(a) and (b) present the comparison results of the two façade digital models established using the proposed technique and previous study, on the panoramic domain. The colors of Fig. 8 indicate the warping extent and mapping position of each stitched images. As displayed in

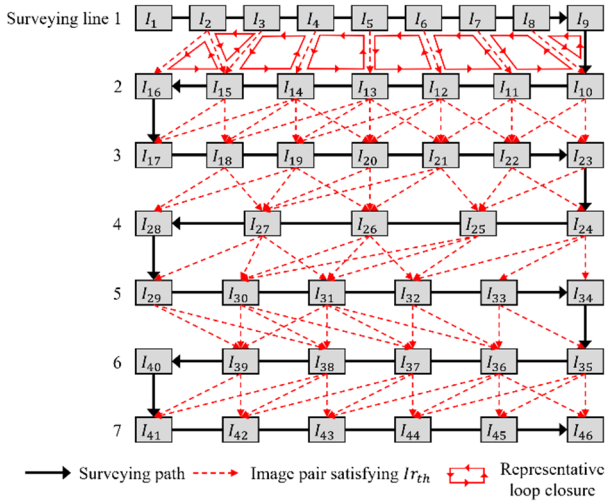
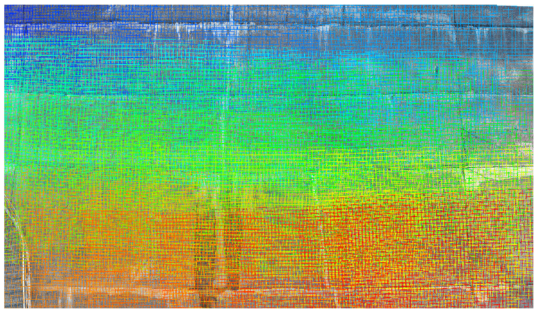
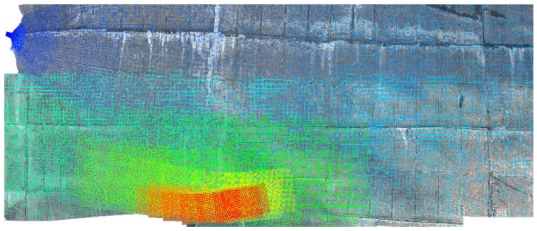


Fig. 7 The representative loop closures and image stitching order



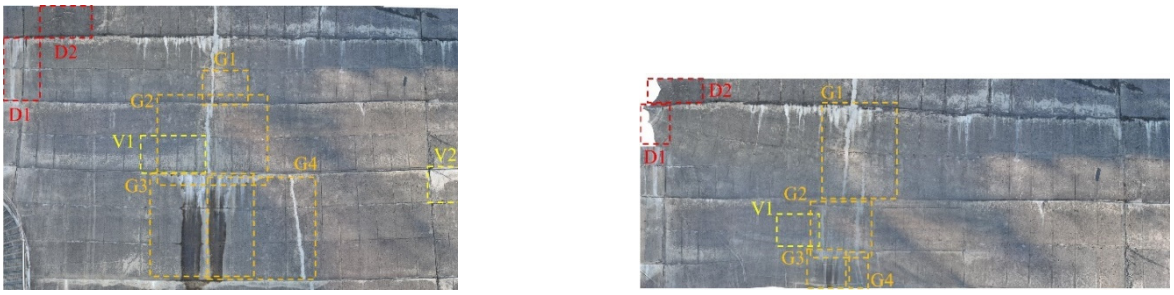
(a)



(b)

Fig. 8 Mesh-based digital image stitching results: (a) Proposed technique; (b) previous study

Fig. 8(a), each image is accurately mapped and stitched without noticeable perspective changes, distortion or scale variation. On the other hand, the previous study case exhibits spatial mis-mapping, extreme distortion as well as abnormal scale changes. In particular, as shown in Fig. 8(b), the distortions become increasingly severe towards the latter part of the surveying path due to the accumulation of errors in camera parameter estimation. In the upper-left side of Fig. 8(b), significant distortion with unnaturally warped areas can be observed. In addition, at the bottom of Fig. 8(b), most of stitched images are spatially concentrated, leading to excessive scale changes, vanishing and ghosting phenomena. Here, the vanishing phenomenon refers to the disappearance or misalignment of surface information in the established model. The ghosting phenomenon entails the blurring or overlapping of surface information, which seems



(a) (b)

Fig. 9 Dam façade digital model establishment results of (a) Proposed technique; and (b) previous study: D is the stitching distortion. V and G mean the vanishing and ghosting phenomena, respectively

(a) (b)

(d) (e) (f)

Fig. 10 Local areas stitching distortions: D1 (a) ground truth; (b) proposed technique; and (c) previous study: D2 (d) ground truth; (e) proposed technique; and (f) previous study

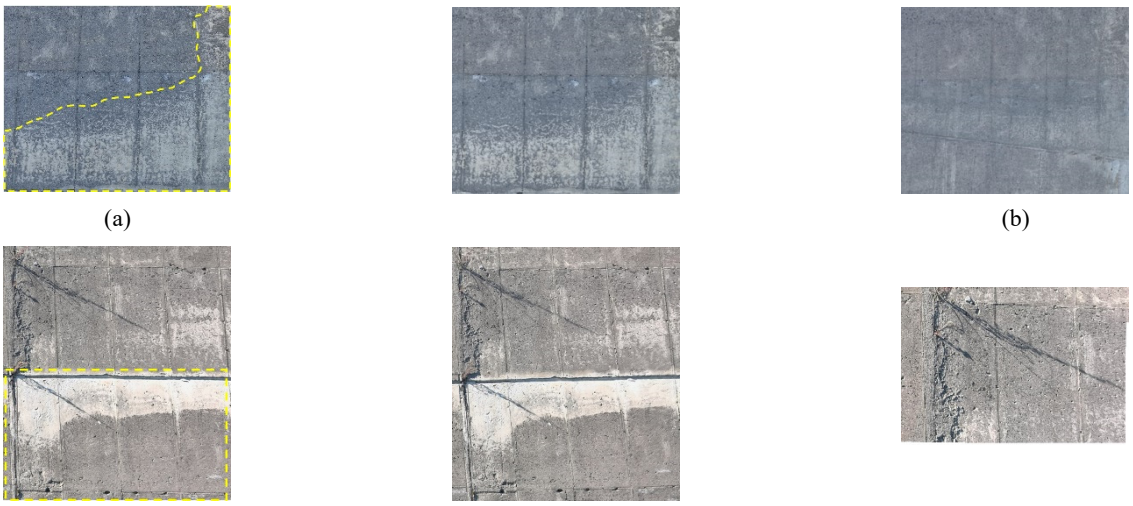
to be multiple entities.

Fig. 9 illustrates the results of dam façade digital model establishment, with representative local distortions marked with dashed boxes. In Fig. 9(a), the proposed technique case has no significant distortions, while a total of 8 critical distortions appears in Fig. 9(b), the previous study case. These representative distortions can be categorized into three types, i.e., stitching distortions (D1 and D2), vanishing (V1 and V2) and ghosting (G1 to G4) phenomena.

In order to compare the modeling accuracy between the proposed technique and previous study, local ground truth images are extracted from original surveyed images corresponding to the representative local distortion areas displayed in Fig. 9. As presented in Figs. 10(a), (b), (d) and (e), the proposed technique case shows good agreement with the ground truth, while one can easily observe not only extreme warps but also abnormal scale changes in the previous study case as shown in Figs. 10(c) and (f).

Similarly, the representative vanishing phenomena corresponding to V1 and V2 are clearly observed in the previous study case (Figs. 11(c) and (f)), compared with the ground truth (Figs. 11(a) and (d)). The surface features marked with the yellow dashed lines of Figs. 11(a) and (d) are properly reconstructed in the proposed technique case (Figs. 11(b) and (e)), while the surface features including damage information disappear in Figs. 11(c) and (f) due to image mis-mapping.

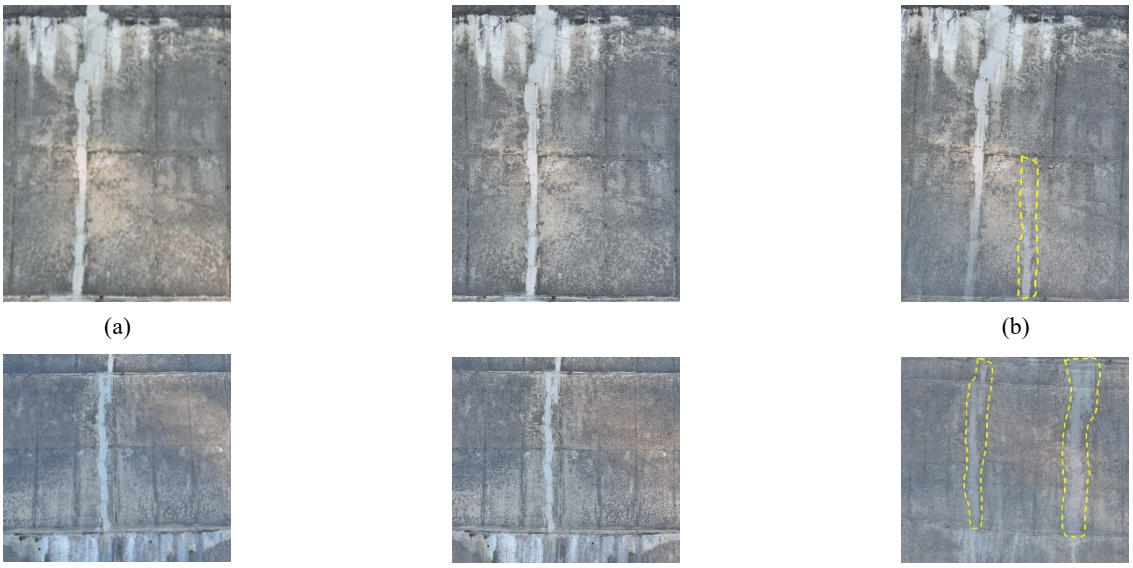
The local areas where ghosting phenomena are also compared in Figs. 12 and 13. The yellow dashed lines of Figs. 12(c), (f) and Figs. 13(c), (f) show that the surface efflorescence features are partially blurred and duplicated due to ghosting phenomenon in the previous study case. Compared to the ground truth images as depicted in Figs. 12(a) and (d) as well as Figs. 13(a) and (d), the surface features are difficult to be distinguishable or misrecognized as multiple entities in the previous study case. In particular, the unexpected white features appear on the right side of Fig. 12(c) compared to Fig. 12(a), resulting in false positive alarm increase. Similarly, in Fig. 12(f), the original surface efflorescence feature is duplicated compared to the ground truth of Fig. 12(d), meaning that the wrong feature is added on the right-hand side. In addition, the surface features are too blurred or tilted to clearly identify the exact



(a) (b)

(d) (e) (f)

Fig. 11 Local areas vanishing: V1 (a) ground truth; (b) proposed technique; and (c) previous study: V2 (d) ground truth; (e) proposed technique; and (f) previous study



(a) (b)

(d) (e) (f)

Fig. 12 Local areas ghosting: G1 (a) ground truth; (b) proposed technique; and (c) previous study: G2 (d) ground truth; (e) proposed technique; and (f) previous study

efflorescence area in Figs. 13(c) and (f). On the other hand, there is no significant distortions that can affect to structural damage inspection as shown in Figs. 12(b) and (e) as well as Figs. 13(b) and (e).

4. Discussion

As it can be observed in the established façade digital model comparison results, the proposed technique outperforms the non-loop-based previous study, NISwGSP (Chen and Chuang 2016). This superiority becomes more pronounced, particularly in spatially extensive structural areas. As explained in Section 2, increasing the number of digital images used for modeling escalates stitching errors and computational expenses, especially in large-scale structure modeling. However, stitching errors can be significantly reduced by applying the proposed technique compared to the previous stitching algorithms without loop closure. As for the computational cost, the proposed technique took 2 hours and 47 minutes with Xeon CPU E5-2630 and 64 GB RAM in this study. Such expensive computational cost can be reduced by controlling the overlap ratio during the optimal image selection process and adopting parallel computing. Nevertheless, it is the worthy noted that the proposed technique can establish high-resolution façade digital models by minimizing stitching errors.

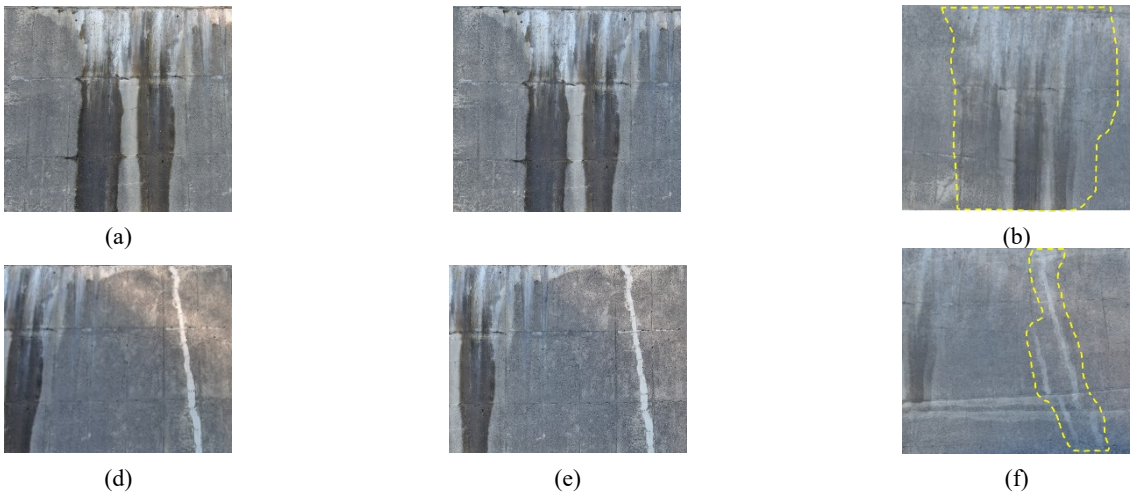


Fig. 13 Local areas ghosting: G3 (a) ground truth; (b) proposed technique; and (c) previous study: G4 (d) ground truth; (e) proposed technique; and (f) previous study

5. Conclusions

This paper proposed a loop closure-based digital image stitching technique for façade digital model establishment of a dam structure. The effectiveness of the proposed technique was experimentally validated through an in-situ dam with an unmanned aerial vehicle-mounted vision camera. The digital modeling accuracy of the proposed technique was compared with a conventional façade digital modeling technique without loop closure. The validation test results revealed that the proposed technique precisely established the façade digital model while the façade model without loop closure was extremely distorted. It means that the proposed technique enables to improve computer vision-based damage detection performance using high-resolution 2D façade digital models. As the follow-up study, time-series damage information within the façade digital model will be investigated for damage growth prediction by obtaining the digital image data periodically. Furthermore, the proposed technique will be applied to various large-scale and planar-like structures such as building, road, bridge and so on.

Acknowledgments

This research was supported by the BK21 FOUR (Fostering Outstanding Universities for Research, No.412024115147) funded by the Ministry of Education (MOE, Korea) and National Research Foundation of Korea (NRF).

References

4DiXplorer (2021), Available online: www.4dixplorer.com/software_satpp.html (accessed on 03 November 2021).

Agisoft Meta shape (2021), Available online: www.agisoft.com (accessed on 03 November 2021).

Bang, S., Kim, H. and Kim, H. (2017), "UAV-based automatic generation of high-resolution panorama at a construction site with a focus on preprocessing for image stitching", *Automat. Constr.*, **84**, 70-80. https://doi.org/10.1016/j.autcon.2017.08.031

Boller, C., Starke, P., Dobmann, G., Kuo, C.-M. and Kuo, C.H. (2015), "Approaching the assessment of ageing bridge infrastructure", *Smart Struct. Syst., Int. J.*, **15**(3), 593-608. https://doi.org/10.12989/sss.2015.15.3.593

Brown, M. and Lowe, D.G. (2006), "Automatic panoramic image stitching using invariant features", *Int. J. Comput. Vis.*, **74**(1), 59-73. https://doi.org/10.1007/s11263-006-0002-3

Cavalagli, N., Gioffrè, M., Grassi, S., Gusella, V., Pepi, C. and Volpi, G.M. (2020), "On the accuracy of UAV photogrammetric survey for the evaluation of historic masonry structural damages", *Procedia Struct. Integr.*, **29**, 165-174. https://doi.org/10.1016/j.prostr.2020.11.153

Chen, Y.S. and Chuang, Y.Y. (2016), "Natural image stitching with global similarity prior", *Proceedings of European Conference on Computer Vision* (*ECCV 2016*), Amsterdam, Netherlands, October, pp. 186-201. https://doi.org/10.1007/978-3-319-46454-1_12

Chen, S., Laefer, D.F., Mangina, E., Zolanvari, S.M.I. and Byrne, J. (2019), "UAV bridge inspection through evaluated 3D reconstructions", *J. Bridge Eng.*, **24**(4), 1-15. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001343

Cheng, K., Shan, J. and Liu, Y. (2021), "Feature-based image stitching for panorama construction and visual inspection of structures", *Smart Struct. Syst., Int. J.*, **28**(5), 661-673. https://doi.org/10.12989/sss.2021.28.5.661

Chiang, K.W., Tsai, G.-J., Li, Y.H. and El-Sheimy, N. (2017), "Development of LiDAR-based UAV system for environment reconstruction", *IEEE Geosci. Remote Sens. Lett.*, **14**(10), 1790-1794. https://doi.org/10.1109/LGRS.2017.2736013

Dabous, S.A., Al-Ruzouq, R. and Llort, D. (2023), "Three-dimensional modeling and defect quantification of existing concrete bridges based on photogrammetry and computer aided design", *Ain Shams Eng. J.*, **14**(12), 102231. https://doi.org/10.1016/j.asej.2023.102231

Drešček, U., Kosmatin Fras, M., Tekavec, J. and Lisec, A. (2020), "Spatial ETL for 3D building modelling based on unmanned aerial vehicle data in semi-urban areas", *Remote Sens.*, **12**(12), 1-24. https://doi.org/10.3390/rs12121972

Furukawa, Y. and Ponce, J. (2010), "Accurate, dense, and robust multiview stereopsis", *IEEE Trans. Pattern Anal. Mach. Intell.*, **32**(8), 1362-1376. https://doi.org/10.1109/TPAMI.2009.161

Jang, K., An, Y.-K., Kim, B. and Cho, S. (2021), "Automated crack evaluation of a high-rise bridge pier using a ring-type climbing robot", *Comput.-Aided Civ. Infrastruct. Eng.*, **36**(1), 14-29. https://doi.org/10.1111/mice.12550

Jang, K., Park, S., Jung H., Yoo H. and An, Y.-K. (2024), "Deep learning-based 3D digital damage map of vertical-type tunnels using unmanned fusion data scannning", *Automat. Constr.*, **162**, 1-13. https://doi.org/10.1016/j.autcon.2024.105397

Jhan, J.P., Kerle, N. and Rau, J.Y. (2021), "Integrating UAV and ground panoramic images for point cloud analysis of damaged building", *IEEE Geosci. and Remote Sens. Lett.*, **19**, 1-5. https://doi.org/10.1109/LGRS.2020.3048150

Jiang, W., Zhou, Y., Ding, L., Zhou, C. and Ning, X. (2020), "UAV-based 3D reconstruction for hoist mapping and layout planning in petrochemical construction", *Automat. Constr.*, **113**, 1-12. https://doi.org/10.1016/j.autcon.2020.103137

Kang, M.S. and An, Y.-K. (2021), "Deep learning-based automated background removal for structural exterior image stitching", *Appl. Sci.*, **11**(8), 1-12. https://doi.org/10.3390/app11083339

Khaloo, A., Lattanzi, D., Cunningham, K., Dell'Andrea, R. and Riley, M. (2018), "Unmanned aerial vehicle inspection of the Placer River Trail Bridge through image-based 3D modeling", *Struct. Infrastruct. Eng.*, **14**(1), 124-136. https://doi.org/10.1080/15732479.2017.1330891

Lowe, D.G. (1999), "Object recognition from local scale-invariant features", *Proceedings of the International Conference on Computer Vision* (*ICCV 1999*), Corfu, Greece, September, pp. 1150-1157. https://doi.org/10.1109/ICCV.1999.790410

Murtiyoso, A. and Grussenmeyer, P. (2017), "Documentation of heritage buildings using close-range UAV images: dense matching issues, comparison and case studies", *The Photogram. Rec.*, **32**(159), 206-229. https://doi.org/10.1111/phor.12197

Pan, Y., Dong, Y., Wang, D., Chen, A. and Ye, Z. (2019), "Three-dimensional reconstruction of structural surface model of heritage bridges using UAV-based photogrammetric point clouds", *Remote Sens.*, **11**(10), 1-20. https://doi.org/10.3390/rs11101204

Schonberger, J.L. (2021), ColMap, Available online: https://colmap.github.io (accessed on 03 November 2021).

Shen, S. (2013), "Accurate multiple view 3D reconstruction using patch-based stereo for large-scale scenes", *IEEE Trans. Image Process.*, **22**(5), 1901-1914. https://doi.org/10.1109/TIP.2013.2237921

Snavely, N., Seitz, S.M. and Szeliski, R. (2008), "Modeling the world from internet photo collections", *Int. J. Comput. Vis.*, **80**(2), 189-210. https://doi.org/10.1007/s11263-007-0107-3

Szeliski, R. (2006), "Image alignment and stitching: A tutorial", *Int. J. Comput. Vis.*, **2**(1), 1-104. http://dx.doi.org/10.1561/0600000009

Varbla, S., Puust, R. and Ellmann, A. (2020), "Accuracy assessment of RTK-GNSS equipped UAV conducted as-built surveys for construction site modelling", *Surv. Rev.*, **53**(381), 477-492. https://doi.org/10.1080/00396265.2020.1830544

Vosselman, G. and Dijkman, S. (2001), "3D building model reconstruction from point clouds and ground plans", *Int. Arch. Photogramm, Remote Sens.*, **34**(3/W4), 37-43.

Wang, D., Zhang, Y., Pan, Y., Peng, B., Liu, H. and Ma, R. (2020), "An automated inspection method for the steel box girder bottom of long-span bridges based on deep learning", *IEEE Access*, **8**, 94010-94023. https://doi.org/10.1109/ACCESS.2020.2994275

Wang, L., Spencer, B.F., Li, J. and Hu, P. (2021), "A fast image-stitching algorithm for characterization of cracks in large-scale structures", *Smart Struct. Syst., Int. J.*, **27**(4), 593-605. https://doi.org/10.12989/sss.2021.27.4.593

Wu, B., Yu, B., Wu, Q., Yao, S., Zhao, F., Mao, W. and Wu, J. (2017), "A graph-based approach for 3D building model reconstruction from airborne LiDAR point clouds", *Remote Sens.*, **9**(1), 1-16. https://doi.org/10.3390/rs9010092

Zaragoza, J., Chin, T.J., Brown, M.S. and Suter, D. (2013), "As-projective-as-possible image stitching with moving DLT", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR 2013*), Portland, OR, USA, June, pp. 2339-2346.

Zhao, S., Kang, F., Li, J. and Ma, C. (2021), "Concrete dam damage detection and localization based on YOLOv5s-HSC and photogrammetric 3D reconstruction", *Automat. Constr.*, **130**, 1-16. https://doi.org/10.1016/j.autcon.2022.104555

Zhao, S., Kang, F. and Li, J. (2022), "Concrete dam damage detection and localization based on YOLOv5s-HSC and photogrammetric 3D reconstruction", *Automat. Constr.*, **143**, 1-18. https://doi.org/10.1016/j.autcon.2022.104555

Zollini, S., Alicandro, M., Dominici, D., Quaresima, R. and Giallonardo, M. (2020), "UAV photogrammetry for concrete bridge inspection using object-based image analysis (OBIA)", *Remote Sens.*, **12**(19), 1-16. https://doi.org/10.3390/rs12193180