Concrete Crack Assessment Using Digital Image Processing and 3D Scene Reconstruction

Yu-Fei Liu[1]; Soojin Cho[2]; B. F. Spencer Jr., F.ASCE[3]; and Jian-Sheng Fan[4]

Abstract: Traditional crack assessment methods for concrete structures are time consuming and produce subjective results. The development of a means for automated assessment employing digital image processing offers high potential for practical implementation. However, two problems in two-dimensional (2D) image processing hinder direct application for crack assessment, as follows: (1) the image used for the digital image processing has to be taken perpendicular to the surface of the concrete structure, and (2) the working distance used in retrieving the imaging model has to be measured each time. To address these problems, this paper proposes a combination of 2D image processing and three-dimensional (3D) scene reconstruction to locate the 3D position of crack edges. In the proposed algorithm, first the precise crack information is obtained from the 2D images after noise elimination and crack detection using image processing techniques. Then, 3D reconstruction is conducted employing several crack images to build the 3D scene, and the surfaces in the scene are estimated by plane fitting using the 3D point cloud. Subsequently, the crack is projected from the 2D image onto the 3D concrete surface with a crack so that the precise 3D coordinates of the crack edges are found. The final crack assessment results are given using the scaled 3D crack information. Field tests were conducted on a concrete wall including single and multiple concrete surface tests, and a concrete flange with complex crack shape, respectively. The results indicate that the proposed approach overcomes existing hurdles to offer a new tool for monitoring the health of concrete structures. **DOI: 10.1061/(ASCE)CP.1943-5487.0000446.** *© 2014 American Society of Civil Engineers.*

Author keywords: Crack assessment; Crack width; Digital image processing; Three-dimensional (3D) scene reconstruction; Projection.

Introduction

Determining the presence, location, and geometry (e.g., length and width) of cracks in a concrete structure is critical to assessing its overall health and performance. Concrete surface cracks are usually caused by various mechanisms, e.g., shrinkage and creep, excessive stress, corrosion of reinforcement bars, and severe loading events such as earthquakes. The presence of cracks provides a phenomenological indication of the stress undergone inside the structures. Currently, surface crack assessment is conducted through manual inspection carried out by trained personnel. However, manual inspection for surface cracks is time consuming, labor-intensive, and highly subjective, depending on the inspector's skill and experience. Moreover, for large-scale structures manual inspection may

be difficult and sometimes dangerous to carry out due to the sheer size of the structural members.

Due to these shortcomings, recent research activity regarding crack detection concentrates primarily on automated crack assessment techniques. Interdisciplinary studies of digital image processing (DIP) techniques have been conducted to enhance the efficiency and autonomous nature of crack assessment. By mimicking the human vision system using DIP techniques, noncontacting crack detection and characterization are made possible. The operators can remotely capture the crack on the concrete surface using cameras and DIP techniques will help obtain the crack information. Related DIP techniques include the following: (1) spatial or frequency domain filters (e.g., Fujita and Hamamoto 2011), (2) wavelet analysis (e.g., Sonka et al. 2008; Jahanshahi and Masri 2013), (3) mathematical methods such as morphological operation (e.g., Tanaka and Uematsu 1998; Sinha and Fieguth 2006), (4) edge detectors (e.g., Abdel-Qader et al. 2003), and so on.

The previously mentioned literature is based on two-dimensional (2D) images. In these approaches, the pin-hole model [Fig. 1(a)] is employed to calculate the absolute scale of the crack width. Under such circumstances, the crack image should be taken with the camera oriented perpendicular to the concrete surface; in this way, all points on the image will share a unique working distance (i.e., the distance from the optical center to the object). This situation is illustrated in Fig. 1(b). In Fig. 1(b), the camera locations are marked with triangles. For Camera A, the working distance for all the cracks on the wall is unique, i.e., the length of AA′. However if the image is taken aslant, the working distances will not be unique when considering different points of the object. For example, Camera B in Fig. 1(b); its working distances for the crack points P′, B′, and Q′ are the length of BP′, BB′, and BQ′, respectively. In this case, the crack width cannot be precisely scaled using the pin-hole model. Therefore, when it is impossible to reach a

[1]Ph.D. Candidate, Key Laboratory of Civil Engineering Safety and Durability of China Education Ministry, Dept. of Civil Engineering, Tsinghua Univ., Beijing 100084, China. E-mail: liuyf06@gmail .com

[2]Research Assistant Professor, Ulsan National Institute of Science and Technology, Ulsan 689-798, Republic of Korea. E-mail: soojin.mail@ gmail.com

[3]Nathan M. and Anne M. Newmark Endowed Chair in Civil Engineering, Dept. of Civil and Environmental Engineering, Univ. of Illinois, Urbana, IL 61801. E-mail: bfs@illinois.edu

[4]Professor, Key Laboratory of Civil Engineering Safety and Durability of China Education Ministry, Dept. of Civil Engineering, Tsinghua Univ., Beijing 100084, China (corresponding author). E-mail: fanjsh@tsinghua .edu.cn

© ASCE 04014124-1 J. Comput. Civ. Eng.
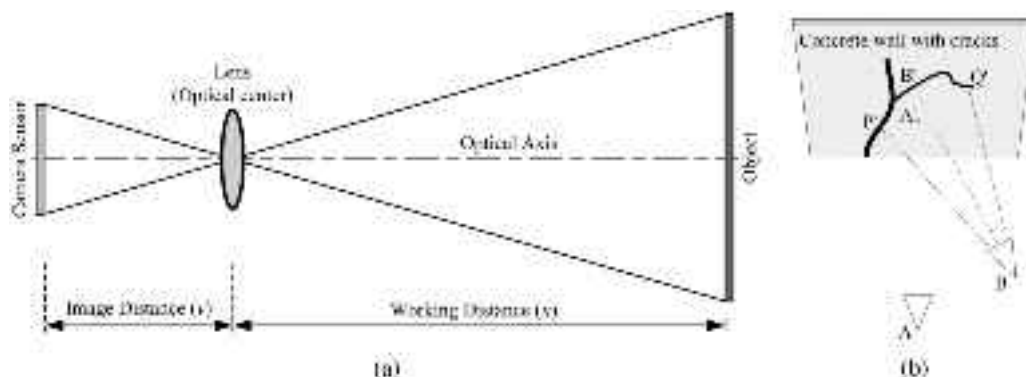
J. Comput. Civ. Eng., 2016, 30(1): 04014124

Fig. 1. Imaging model and camera locations: (a) pin-hole model; (b) different camera locations

position perpendicular to the concrete wall with the camera, the crack assessment would be difficult to carry out in practice. Another disadvantage for crack detection using 2D images is that if multiple crack images are taken during the assessment, the working distances have to be measured each time when taking images because the imaging model of each image is individual. With only the 2D DIP method, the previously mentioned problems are insolvable.

To overcome these challenges, three-dimensional (3D) reconstruction (i.e., 3D modeling) techniques employing the structure from motion (SFM) algorithm are introduced to build an integrated and composite imaging model, as well as the 3D scene. In this paper, the imaging model refers to the spatial relationship of the objects, the optical center, and the imaged objects. The 3D scene refers to the retrieved 3D model. Jahanshahi and Masri (2012) first utilized 3D reconstruction techniques in the field of image-based crack assessment. In their study, 3D reconstruction was used to retrieve the integrated imaging model so that the working distance could be scaled using absolute distances (e.g., the distance between two locations of the roving camera) in the 3D imaging model. However, this research only considers the case when the crack images were taken with the optical axis of the camera perpendicular to the concrete surface. In this way, all the points on the crack image should have one single working distance, such as Camera A in Fig. 1(b).

In this paper, an automatic crack assessment approach combining the 2D DIP technique and 3D scene reconstruction technique is proposed, which overcomes the disadvantages of existing algorithms. The assessment data of the crack width and distribution are obtained using the proposed approach. In detail, first the preprocessing of crack images is investigated, including the selection of key frames and image distortion correction; next is a detailed description of 3D scene reconstruction and imaging model retrieval using the SMF algorithm. Subsequently, DIP techniques are adopted for crack assessment using the clearest images obtained from the key frames extraction method and unscaled crack widths (and crack distribution) are obtained from the 2D images. Finally, combined with the retrieved imaging model and reconstructed 3D scene, crack edge and crack distributions are projected onto the fitted planes in the 3D scene to obtain the original crack shape, and accurate data of crack width. The final crack information is calculated from the projected 3D crack so that the crack images which are taken obliquely to the concrete surface can be used in the second step to process the 2D crack assessment. Because the accuracy of the retrieved 3D scene is low, the reconstructed 3D scene cannot be used in the crack assessment directly. In the proposed method, only the planes fitted from the 3D point cloud are adopted, which

have relatively higher accuracy. Several examples and field tests are given to illustrate the efficacy of the proposed approach.

Three-Dimensional Reconstruction Technique

During the past decades, image-based 3D scene reconstruction technique has been rapidly developed (Zisserman 2003) and widely researched for civil structure condition assessment (Uslu et al. 2011), topographic mapping (James and Robson 2012), civil structure modeling (Furukawa et al. 2010), and so on. The technique is to reconstruct the camera locations and the 3D scene cloud including the object shape and appearance, from various images taken around the target object. In this section, the 3D scene reconstruction technique adopted in the paper is presented, which begins with two preprocessing steps, as follows: (1) key frame extraction for the 3D reconstruction, and (2) image distortion correction. The key frame extraction method which is used to automatically select the proper images for 3D reconstruction is investigated in the subsection, "Acquisition of Proper Images for 3D Reconstruction and Crack Detection." The image distortion correction is discussed in the subsection, "Camera Calibration and Image Distortion Correction." Eventually, the 3D scene and the imaging model will be retrieved and ready to be used for the projection in crack assessment. The two preprocessing steps as well as the 3D reconstruction technique are subfields of their own and have been widely studied. Fig. 2 provides a flow chart for the proposed approach.
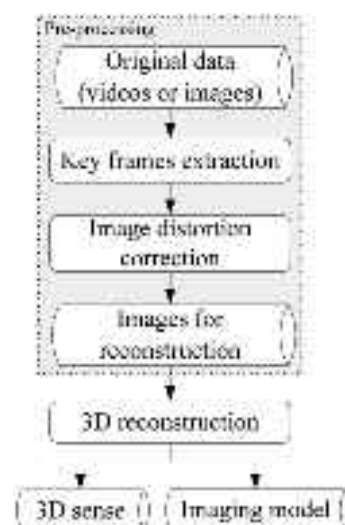

Fig. 2. Flow chart of 3D reconstruction including preprocessing

Acquisition of Proper Images for 3D Reconstruction and Crack Detection

The images used for 3D reconstruction and crack detection should be properly selected from the original collection of images, depending on the number of required images and the image quality. This step is called the so-called key frames extraction. Usually the images for 3D reconstruction are obtained as a collection of single-shot images, a burst of images, or a sequence decomposed from a video. The form of the obtained data depends on the sheer size and complexity of the civil structure to be represented. Generally, for small objects such as a piece of a concrete wall, less than 100 single-shot images are sufficient. For 3D reconstruction of complete structures or large size structure members, more images are need. In this case videos and bursts will be better because they have higher so-called sampling frequency and contain more information comparing with single-shot images. Usually more than 25 frames in a video or five images in a burst can be obtained in 1 s.

The quality of the images seriously impacts the accuracy of crack detection and 3D reconstruction. An important criteria used in this paper to quantize the degree of image blurring is the blurry criteria proposed by Crete et al. (2007). For the choosing of images used for crack detection, the images with the least blurry degree is employed, and usually one or two images are enough for 2D crack detection around a local area such as a piece of wall. For the images used for 3D reconstruction, the blurry criteria as well as some other indicator are adopted, which depend on the procedure of the way to do 3D reconstruction.

The first step in the 3D reconstruction will be so-called feature extraction and matching. In this step, different images are matched and compared, and the matched point pairs are produced as a result, as follows: (1) When single-shot images are employed, which means that only a few images are adopted, the matching process will be conducted on every two images, i.e., the crossing matching. In this situation, the preprocess to acquire the proper images (key frames) for 3D reconstruction is filtering out the blurry images. (2) If videos or bursts are employed, they can be decomposed into thousands of images. In this case, the crossing match will not be good choice in the step of feature extraction and matching considering the computational time. Instead, the sequential matching will be conducted, where each time two neighboring images are compared. In this situation, to pick out the key frames, image blur, image overlap, proper baseline length (i.e., the distance between two optical centers) between two neighboring images, as well as some other indices [proposed in Seo et al. (2008) and Rashidi et al. (2013)] are taken into consideration concurrently. In this paper, the tested images used for 3D reconstruction is less than 20 so that only the blurry criteria is considered.

Camera Calibration and Image Distortion Correction

Before conducting 3D reconstruction, image distortion should be corrected to improve the reconstruction precision. Distortion exists in all images, especially when employing the wide-angle lens such as the fish-eye lens. To conduct the image distortion correction, the image distortion parameters and the intrinsic parameters of the camera need to be calculated first, which can be obtained using the camera calibration process (Tsai 1987; Weng et al. 1992). In this paper, the camera calibration toolbox for *Matlab* developed by Bouguet (2013) is employed to both acquire the calibration parameters and remove images distortions. This toolbox is developed on the basis of the work of Zhang (1999, 2000) and uses checker board images for calibration. The calibration parameters obtained include the following: (1) camera intrinsic parameters such as the focal lengths of horizontal and vertical directions
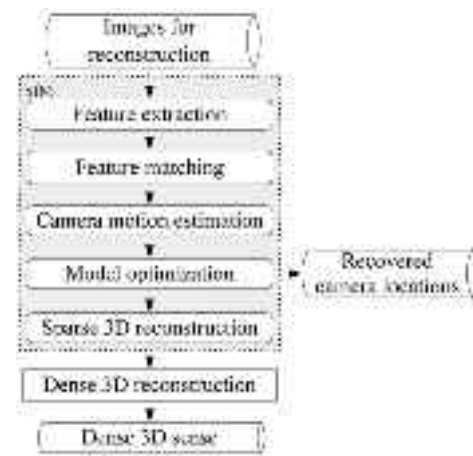

Fig. 3. Flow diagram of 3D reconstruction employing SFM technique and dense 3D reconstruction

(and skew coefficient), and (2) lens distortion parameters such as the radial distortion and the tangential distortion.

Three-Dimensional Reconstruction and Imaging Model Retrieval

The SFM technique derives the sparse 3D scene as well as the camera positions based on multiple images, drawing on efforts in numerous research fields (Huang and Netravali 1994; Zisserman 2003). In detail, four main parts are involved in the SFM technique, as follows: (1) image feature extraction (or key point detection), (2) feature matching (or key point matching), (3) camera motion estimation, and (4) model optimization. After the sparse 3D point cloud and the integrated imaging model are obtained using SFM, a dense 3D point cloud model can be further derived. The retrieved 3D scene, including both the sparse model and the dense model, is a point cloud representation rather than a continuous scene. The flow diagram of the 3D reconstruction employing the SFM technique and dense 3D reconstruction is shown in Fig. 3. In particular, the retrieved camera locations will help to build the imaging model when projecting the cracks from the 2D images onto the 3D scene. This part will be discussed detailed in the section, "Projection of Crack from Images onto 3D Scene."

Feature Extraction and Matching

The scale invariant feature transform (SIFT) proposed by Lowe (2004) is the most popular algorithms in the field of image feature detection. In this paper, SIFT is employed for its stability in feature detection under different image transformations such as rotations and scale changes comparing to some other image feature detection methods (Bay et al. 2006; Juan and Gwun 2009). The matched key points are usually called point correspondences of the two images. Fig. 4 shows an example of the feature extraction and matching using two crack images employing the SIFT method, where the short stubs in Fig. 4(a) indicate the locations of the detected feature points and their corresponding gradient orientation, the circle symbols, and the connecting lines marked the matched feature point pairs in Fig. 4(b).

Camera Motion Estimation

Camera motion estimation is aimed to get the extrinsic camera matrices which describe the relative positions and orientations of the camera optical centers. The multiple view geometry (i.e., the multiview stereo) is developed and applied to solve the
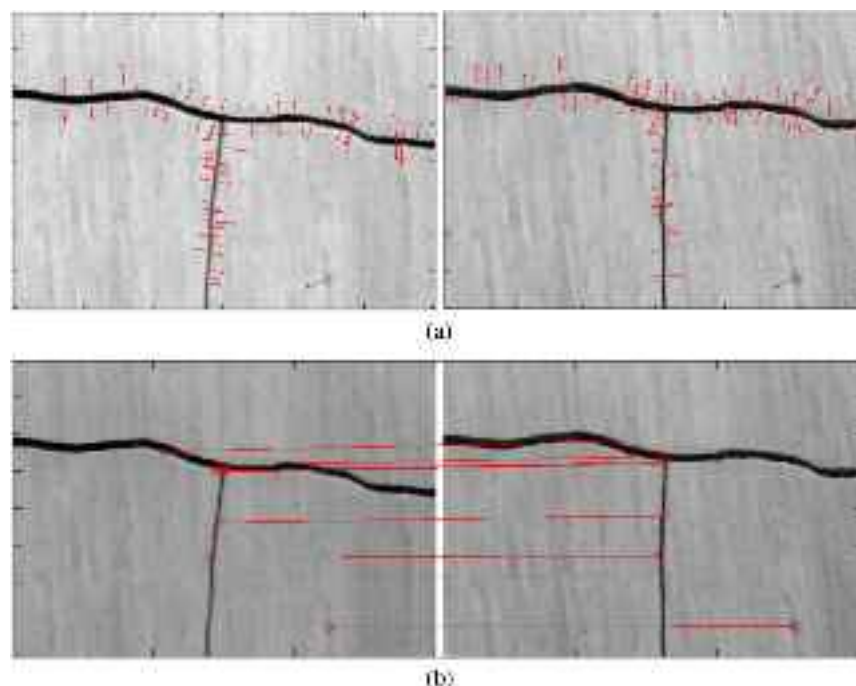
© ASCE   04014124-3   J. Comput. Civ. Eng.

J. Comput. Civ. Eng., 2016, 30(1): 04014124

Fig. 4. Example of feature extraction and matching using SIFT: (a) keypoints; (b) keypoint matching

problem of the camera motion estimation under the restriction of the two-view geometry, three-view geometry, and *n*-view reconstruction (Shashua 1997; Zisserman 2003), where *n* refers to the number of views. Usually *n*-view reconstruction could be decomposed into the combination of two-view and three-view problems.

Two important matrices are involved in the two-view geometry, as follows: (1) foundation matrix **F**, and (2) essential matrix **E**. These two matrices are estimated from the point correspondences of the two images and can be used to get the extrinsic camera matrices.

The foundation matrix is the mathematical description of the epipolar geometry relationship between the imaging models of the two images covering the same scene (Fig. 5). It has seven degrees of freedom (DOF) and can be estimated by at least seven point correspondences using

$$\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0 \tag{1}$$

where $\mathbf{x}'$ and $\mathbf{x}$ are the homogeneous pixel coordinate of a pair of feature points from the two images. Fig. 5 represents the relationship of the point correspondences O and O' are the optical centers of two cameras, where $\mathbf{x}$ and $\mathbf{x}'$ are the photographed points from point X onto the images of the two cameras.

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

$$\mathbf{E} = \mathbf{K}'^T \mathbf{F} \mathbf{K} \tag{2}$$

where **K** and $\mathbf{K}'$ are the intrinsic camera matrices of the two cameras, respectively. The essential matrix is used to calculate the extrinsic camera matrices which describe the camera motion. The estimated camera location is in relative positions, which can be scaled using any known absolute distance in the scene.

The two matrices, the (1) foundation matrix **F**, and (2) essential matrix **E**, are the bridge from the feature matching to the extrinsic

camera matrices, and they enshrine the principle of multiple view geometry. In the three-view geometry, the trifocal tensor replaces the foundation matrix in the two-view geometry.

Model Optimization and Sparse 3D Reconstruction
Model optimization seeks to minimize the projection errors for given images with respect to the noise, which is conducted mostly using the bundle adjustment algorithm (Triggs et al. 2000; Wu et al. 2011). After the model optimization, the sparse 3D model is reconstructed based on the triangulation of the key point correspondences. Fig. 5 presents a case of the triangulationl knowing the related key points $\mathbf{x}$ and $\mathbf{x}'$ as well as the two camera models, the position of the original 3D point X is recovered. An example of nine images of several books (shown in Fig. 6) is used for displaying the 3D reconstruction and the book cover of Lutes and Sarkani (2004) will be used to simulink the crack surface. The sparse 3D reconstruction from the nine images is presented in Fig. 7(a).

Dense 3D Reconstruction
With the imaging model known, more pairs of the matching points can be obtained. To improve the searching efficiency, image rectification (Ayache and Hansen 1988; Papadimitriou and Dennis 1996; Pollefeys et al. 1999; Zisserman 2003) is applied before
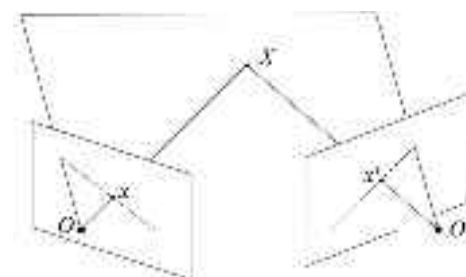

Fig. 5. Epipolar geometry restriction in two-view geometry

© ASCE 04014124-4 J. Comput. Civ. Eng.

J. Comput. Civ. Eng., 2016, 30(1): 04014124

Fig. 6. Images of two books used for 3D reconstruction [this cover was published in Loren, L., and Sarkani, S. (2003). *Random Vibrations: Analysis of Structural and Mechanical Systems*, Elsevier Butterworth-Heinemann, Burlington, Massachusetts, Copyright Elsevier (2004)]

the dense points matching. After the rectification, all the epipolar lines are parallel and the searching space is reduced to one dimension. Meanwhile, the epipolar equation is no longer needed to be calculated each time. Fig. 7(b) shows the 3D dense reconstruction result for the book presented in Fig. 6.

Dense 3D reconstruction is not precise and elaborate enough to be used directly in assessing the characteristics of subtle objects, especially like the width or length of cracks. The 2D DIP technique is still needed to obtain accurate crack information.

In this paper, the *FIT3D* toolbox (Esteban et al. 2010) and *VisualSFM* (Wu 2011; Wu et al. 2011) employing the SIFT feature extraction algorithm are used for SFM, and the clustering views for multiview stereo (CMVS) algorithm proposed by Furukawa and Ponce (2010) is used for the dense 3D reconstruction.

Adaptive Crack Detection Using the DIP Techniques Based on 2D Images

To assess the crack distribution along with accurate widths, the crack images need to be preprocessed. During the preprocessing step, the image noise can be eliminated and the crack objects can be enhanced. In this paper, the adaptive crack detection method proposed in Liu et al. (2014) is adopted. In details, two steps are involved for the preprocessing step, as follows: (1) image subtraction with a median filter, and (2) Niblack's local binarization method. When employing the median filter and local binarization method, the size of window to apply the filter and the local binarization is the most critical parameter to determine the crack width with minimal noise. If the window sizes are not tuned properly, the crack image may be distorted and the crack shape
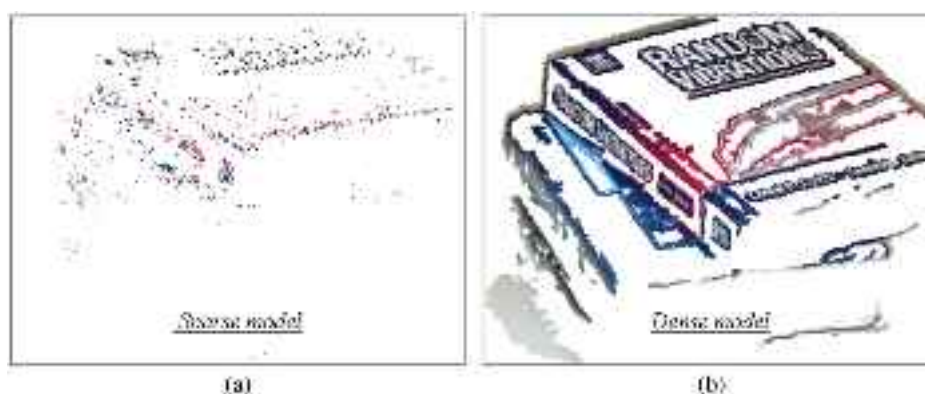

Fig. 7. Three-dimensional reconstruction using book images [this cover was published in Loren, L., and Sarkani, S. (2003). *Random Vibrations: Analysis of Structural and Mechanical Systems*, Elsevier Butterworth-Heinemann, Burlington, Massachusetts, Copyright Elsevier (2004)]
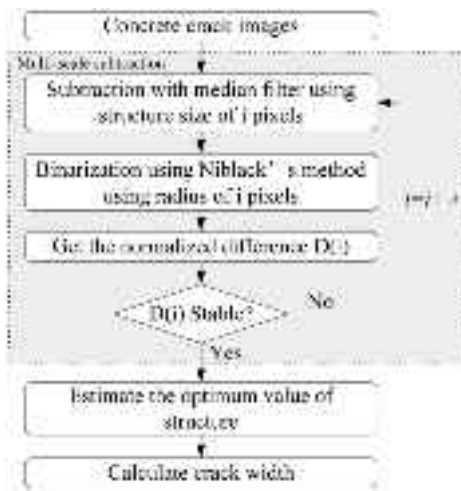
© ASCE   04014124-5   J. Comput. Civ. Eng.

J. Comput. Civ. Eng., 2016, 30(1): 04014124

Fig. 8. Flow diagram of adaptive crack assessment

may be changed, which leads to inaccurate assessment of the crack widths.

To address this problem, the optimal filter size index (OFSI) in Liu et al. (2014) is used in this study to realize the automated tuning of window sizes. The OFSI index works for both the median filter and the local binarization, and is defined as

$$D_N(i) = \frac{|N(i + \Delta i) - N(i)|}{w \times h} \tag{3}$$

where w and h = width and height of the image in pixels (in the image coordinate), respectively; $N(i)$ = total pixel number of detected objects (cracks and noises) after the subtraction and the binarization using window size of i (i.e., i-pixel window); and Δi is the increment of the window size. With incremental window sizes there are three considerations, as follows: (1) when two neighboring processed images vary a lot, $D_N(i)$ will get a relative high value; (2) when the two neighboring processed images appear to be similar, $D_N(i)$ will get a relative low value; and (3) $D_N(i)$ will become stable after the stable stage begins. If $D_N(i)$ is large and unstable, it infers that the window size has not reach the optimal value.

Fig. 8 presents the flow diagram of the adaptive crack assessment approach. After obtaining the binary image of the crack, the

crack decomposition and crack edge extraction are applied, and the image coordinates of the crack width points can be obtained. The obtained points on the crack edge used for width calculation are named so-called crack width points in the next part. The precision analysis of the 2D crack width detection method can refer to Liu et al. (2014).

Because the 2D crack detection is based on the square pixel, the width of most narrow detectable crack is related to the parameters of camera and the imaging model. Assumed that at least three pixels (two for crack edges and one for crack skeleton) are needed in the width direction of crack to be detected, the width of most narrow detectable crack is calculated and presented in Fig. 9, where a Canon EOS 5D II and a 500D are used as examples. In the calculation, the physical size (in millimeters) and image resolution (in pixels) of the camera sensor are needed to scale the crack width from pixel units to millimeters, Define P_{CMOS} as the number of pixels per millimeter of the camera sensor. For 5D II with maximum resolution (5,616 × 3,744) and 500D with maximum resolution 4,752 × 3,168, P_{CMOS} = 156 and 213 pixel/mm, respectively. By using the pin-hold model [Fig. 1(a)], the minimum detectable crack width is calculated as

$$W_{mm} = \frac{W_{pixel}}{P_{CMOS}} \cdot \frac{u}{v} \tag{4}$$

where W_{mm} and W_{pixel} = crack width (units of mm and pixels, respectively; $W_{pixel} = 3$); and u and v = working and image distances, respectively. Moreover, u and v are controlled by the thin lens model

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f} \tag{5}$$

where f = focal length.

Projection of Crack from Images onto 3D Scene

As discussed in the Introduction, the crack assessment based on 2D images is restricted to using the images taken perpendicularly to the concrete wall and the working distances of different crack images have to be measured each time. This fact is a significant limitation to the automated assessment of cracks using 2D DIP techniques. However, if combined with 3D reconstruction techniques, this limitation can be overcome. By employing the
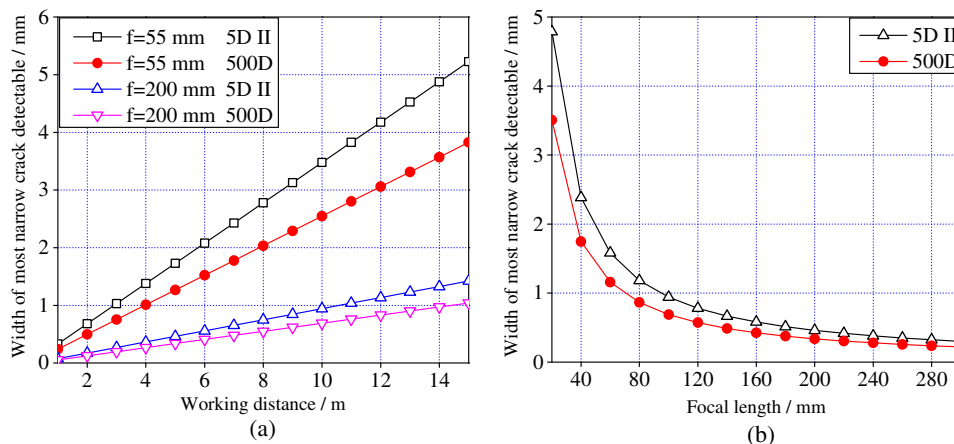

Fig. 9. Width of most narrow crack detectable: (a) when focal length is constant and working distance varies; (b) when working distance is a constant 5 m and focal length varies
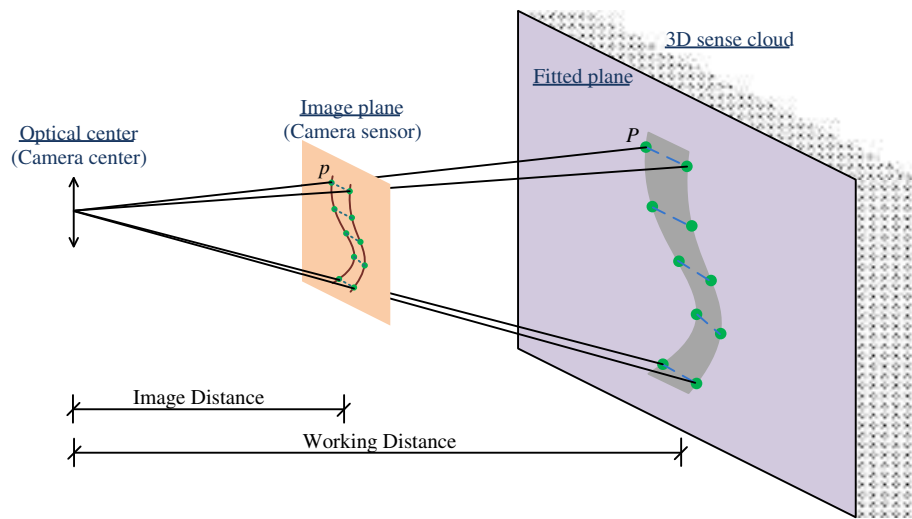
© ASCE 04014124-6 J. Comput. Civ. Eng.

J. Comput. Civ. Eng., 2016, 30(1): 04014124

Fig. 10. Projection of crack from 2D image onto 3D scene

3D reconstruction and getting the integrated imaging model, the whole imaging model and any object size can be scaled using only one known absolute length. Meanwhile, if the crack can be located in the 3D sense to calculate the crack information from the 3D model rather than the 2D image, those crack images taken obliquely to the concrete surface can also be used in crack detection.

To achieve this goal, a projection algorithm of the 2D crack images onto the reconstructed 3D scene on the basis of dense reconstruction scene is proposed in this section. The key step of the projection is to deduce the relationships between the 2D crack width points and their 3D coordinates in the 3D scene. The crack width points and the 3D scene point cloud are obtained by two independent algorithms, so the crack width points may not have correspondences in the 3D scene cloud. In this paper, the projection is executed by the following four steps: (1) plane fitting of the 3D point cloud, (2) finding the crack planes from the fitted planes using proximate key points, (3) locating the crack width points in the image plane under 3D coordinates, and (4) projecting the located crack width points onto the 3D point cloud. Fig. 10 shows a schematic diagram for the projection from the crack width point p on the 2D image plane onto the 3D point P on the crack surface. The detailed four procedures will be described in the subsequent part of this section.

Plane Fitting on the 3D Scene Cloud

Plane fitting is to fit the 2D planes based on the distribution of key points in the 3D scene cloud. This problem has been studied variously in research field of surface reconstruction using the 3D point cloud (Esteban et al. 2010; Gallo et al. 2011). In this paper, the plane fitting starts with randomly choosing points in a subregion of the 3D point cloud. Subsequently, the random sample consensus (RANSAC; Fischler and Bolles 1981) is applied to fit planes using these selected points. The RANSAC is an iterative method used for estimating the model parameters. Unlike to the least-squares estimation that fits a model from all data, the RANSAC estimates the model parameters only from the so-called inliers (i.e., data that contributes to the model) with discarding the so-called outliers (i.e., data that do not fit the model). By executing the RANSAC, the plane function and inliers 3D points can be obtained. Next, all points in the 3D point cloud will be checked to see whether they

belong to this fitted plane, the tolerance used in the investigation comes from the variance of the distances between inliers 3D points to the fitted plane. Finally, after all the points on the fitted plane are found out, they will be removed from the 3D point cloud to start another circulation of the plane fitting.

Fig. 11 shows the result of the plane fitting of the book images of Fig. 6 and five planes colored differently are found out reasonably.

Determining Crack Planes Using Proximal Key Points

In most case, the crack width points are not the key points; they only appear on the 2D images rather than in the 3D point cloud. So it may be hard to determine the fitted plane on which the crack width points lie, i.e., which fitted plane is the crack surface. To automatically select the crack surface, this subsection uses the key points nearest to the crack width point on the 2D image to do the evaluation. These nearest key points are named so-called proximal key points in this paper, which can be approximately considered to be in the same plane as the crack width point. Meanwhile, because these key points also exist in the 3D point cloud, their located plane can be easily found and this plane is
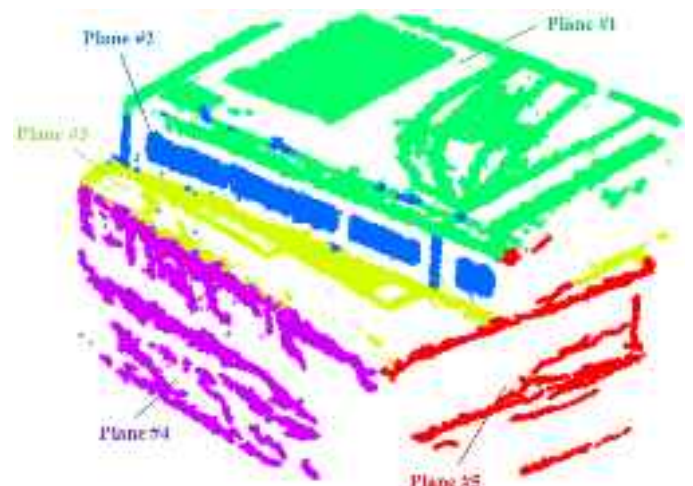

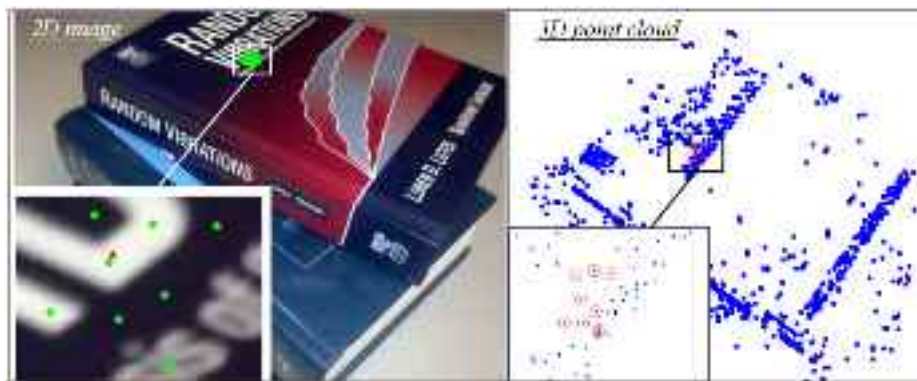**Fig. 11.** Plane fitting of the 3D point cloud of books

Fig. 12. Proximal key points used for determining the crack surface [this cover was published in Loren, L., and Sarkani, S. (2003). *Random Vibra-tions: Analysis of Structural and Mechanical Systems*, Elsevier Butterworth-Heinemann, Burlington, Massachusetts, Copyright Elsevier (2004)]

supposed to contain the crack width point, i.e., to be the crack sur-face. In this evaluation method, these proximal key points act as a bridge to help determine the crack surface in the 3D scene.

In this subsection, it is assumed that for crack width point p, the 3D plane contains the most proximal key points of p will also con-tains points p. See Fig. 12 as an illustration example; the triangle in the enlarged drawing on the left-hand side of Fig. 12 indicates the crack width point. This point is not a key point and only exists on the 2D image, not the 3D point cloud. To determine which 3D plane this crack width point lies on, nine proximal key points are pick out. The nine proximal key points are the nearest points to the crack width point on the 2D image, which are marked as a filled circle in the 2D image of Fig. 12 (two points are almost coincident). According to the assumption mentioned previously, these nine proximal key points will mostly be on the same fitted plane as the selected key points. Therefore, it can be determined that the crack width point lies on the Plane 1 (as shown in Fig. 11) because all nine proximal key points lie on Plane 1. The 3D point cloud of Fig. 12 gives the corresponding locations of the nine proximal key points, which are shown inside the circles in the right-hand side of Fig. 12. The crack width point does not exist in the 3D point cloud.

Tests show that usually five proximal key points are enough to determine the plane of the crack surface when using the dense 3D point cloud. If the crack width point is just on the edge of the wall, i.e., the intersection area of two planes, the determination of the plane may make mistakes. However, it will not bring too much error because the point will still almost be on the edge.

Locating Camera Sensor under the 3D Coordinate System

Before this step is taken, the coordinate systems in the imaging model must be established. As shown in Fig. 13, four coordinate systems are involved in the camera imaging model. The four co-ordinate systems are as follows: (1) pixel coordinate system (shown as $o_p - u - v$), (2) image coordinate system (shown as $o - x - y$), (3) camera coordinate system (shown as $O_c - X_c - Y_c - Z_c$), and (4) world coordinate system (shown as $O_w - X_w - Y_w - Z_w$). Point p is the crack width point on the image plane and point P is its corresponding 3D location. The units of these coordinate systems are either in pixels or millimeters, which can be converted into each other by using dx and dy given in the subsequent part.

In this subsection, the location of the camera sensor (the image plane) in the 3D scene will be calculated and the pixel coordinates of the points on the camera sensor will be transformed into the world coordinate system. As a result, the world coordinates of
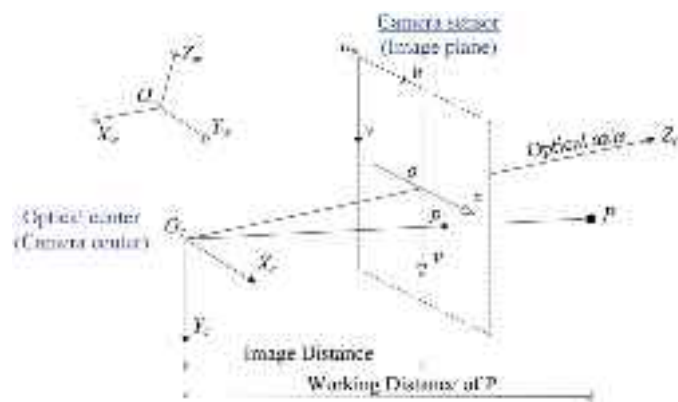

Fig. 13. Coordinate systems of imaging model

the crack width points (obtained from the image-based adaptive crack assessment) can be obtained, such as the world coordinate of point p in Fig. 13. The involved five parameters of the camera and imaging model during the transformation are as follows:
1. Physical size of the camera sensor (in millimeters), i.e., M (in u-direction) and N (in v-direction);
2. Pixel number of the camera sensor (in pixels), i.e., m (in u-direction) and n (in v-direction);
3. Focal length of the camera (in millimerters), f;
4. Calibrated camera intrinsic parameter matrix, **K**; and
5. Camera extrinsic parameter matrix of the camera from the camera coordinate system to the world coordinate system, i.e., $[\mathbf{R}|\mathbf{T}]$, where **R** and **T** denote the 3×3 rotation matrix and the 3×1 translation matrix, respectively (the camera ex-trinsic parameter matrix is obtained from the camera motion estimation in the 3D reconstruction).

Among the five parameters, the first two can be obtained from the manual book of the camera and the last one is determined in the imaging model during the camera motion estimation. The focal length f and the matrix **K** can be obtained during the camera cal-ibration process which will be more realistic, or directly from the manual of the camera if ignoring the distortion parameters.

Define dx (in u-direction) and dy (in v-direction) as the physical length of one pixel on the camera sensor, whose unit is millimeters per pixel. Thus, $dx = M/m$ and $dy = N/n$. The dx and dy terms could be used to convert the units between pixels and millimeters.

Assume that the coordinate of the crack width point on the im-age plane under the pixel coordinate system, the image coordinate

© ASCE

04014124-8

J. Comput. Civ. Eng.

system, the camera coordinate system, and the world coordinate system is (u, v), (x, y), (x_c, y_c, z_c), and (x_w, y_w, z_w), respectively. The unit of (u, v) is in pixels and the other coordinates as well as the translation matrix $\mathbf{T}$ are in millimeters. For linear imaging model without affine distortion (i.e., calibrated), the transform of the homogeneous coordinate between the pixel coordinate system and the image coordinate system is

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/dx & 0 & m/2 \\ 0 & 1/dy & n/2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (6)$$

Between the image coordinate and the camera coordinate

$$z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \quad (7)$$

Known the world coordinate is calculated using

$$\left(\mathbf{R} \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} \right) + \mathbf{T} = \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \quad (8)$$

or

$$\begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \quad (9)$$

Plug in, thus

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K}[\mathbf{R}|\mathbf{T}] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (10)$$

where

$$\mathbf{K} = \begin{bmatrix} f/dx & 0 & m/2 \\ 0 & f/dy & n/2 \\ 0 & 0 & 1 \end{bmatrix} \quad (11)$$

Because the crack width point is on the image plane, the restriction to its camera coordinate is $z_c = f$. Consequently

$$\begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = \mathbf{R}^{-1} \left(f \mathbf{K}^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - \mathbf{T} \right) \quad (12)$$

When (u, v) is the pixel coordinate of the crack width point obtained from the 2D crack assessment discussed in the section, "Adaptive Crack Detection Using the DIP Techniques Based on 2D Images," the world coordinate of the crack width point could be calculated using the Eq. (10).

Projection Using Pin-Hole Model

After obtaining the world coordinates of the crack width points, the crack width points on the camera sensor can be projected onto the 3D sence using the pin-hole model [Fig. 1(a)]. With the world coordinates of both the optical center of a roving camera and the crack width points on the camera sensor known, lines starting from the optical center can be drawn over the crack width points on

the camera sensor and then intersect the fitted crack surface plane in the 3D point cloud. Eight projected intersection points of the books model are shown in Fig. 14, where the two top images are the 2D images used to pick out the imaginary so-called crack width points (marked as dots on the book cover), and the bottom drawing is the projection of the imaginary crack width points onto the 3D scene. In the bottom projection drawing of Fig. 14, the camera optical centers are shown as a circle, while the crack width points on the image plane are shown as a plus sign.

Furthermore, because multiple planes will be fitted as already discussed at the beginning of this section, the projection is able to locate cracks piercing through the structure edge, i.e., appearing on two connected planes (see the section, "Test of Methodology for Cracks on Multiple Concrete Surfaces," as an example). Morever, crack width points on different images can be merged and projected onto the same 3D scene (Fig. 14).

After the projection the real crack width will be calculated using the scaled 3D coordinates of the projected crack width points. The crack nephogram is the final presitation for the crack assessment including the information of crack distribution and crack width.

A summary of the proposed projection approach in this section is presented in the flow diagram of Fig. 15.

Validation Test of Proposed Method

To validate the proposed crack assessment method, experiments were conducted on two concrete structures. Tests 1 and 2 were carried out on a concrete reaction wall with various cracks. The wall is in the Newmark Structural Engineering Laboratory at the University of Illinois at Urbana-Champaign and possesses lots of cracks, caused by loading and shrinkage. A Canon EOS 500D with EF-S 18-55 lens whose focal length is fixed at 55 mm was employed for Tests 1 and 2. The image resolution is $4{,}752 \times 3{,}168$ pixels. The crack images for Tests 1 and 2 were taken on a single concrete surface, and a corner of the concrete wall with three surfaces, respectively. Test 3 was carried out on a concrete flange near a concrete-steel joint. For Test 3, a canon EOS 5D Mark II with EF 24-105 lens was used, and the focal length was not fixed during Test 3. Crack images used for crack detection were down sampled from the original images. The final resolution of the crack images is $1{,}404 \times 936$ pixels.

For Tests 1–3, crack images were directly applied for both the 3D reconstruction and the crack assessment without any trimming. The camera and the lens were calibrated before tests using checker board images. The crack images with the lowest blurry degree discussed in the section, "Three-Dimensional Reconstruction Technique," are applied in the proposed crack assessment approach. The widths of the cracks were measured at several points using a crack ruler or a vernier caliper to be used as references.

Test 1, Single Concrete Surface

The cracks on a single surface of the wall were identified by the proposed approach. In Test 1, the cracks located about 1.9 m high from the floor, while the camera was set lower than 1.7 m high from the floor and around 0.25 m away from the concrete surface. Therefore, the crack images were taken upward sloping. The six crack images after the distortion correction are shown in Fig. 16.

Step 1, Crack Assessment Using 2D DIP Technique
Images 4–6 in Fig. 16 were used for the 2D crack assessment individually. By conducting the adaptive crack assessment using the
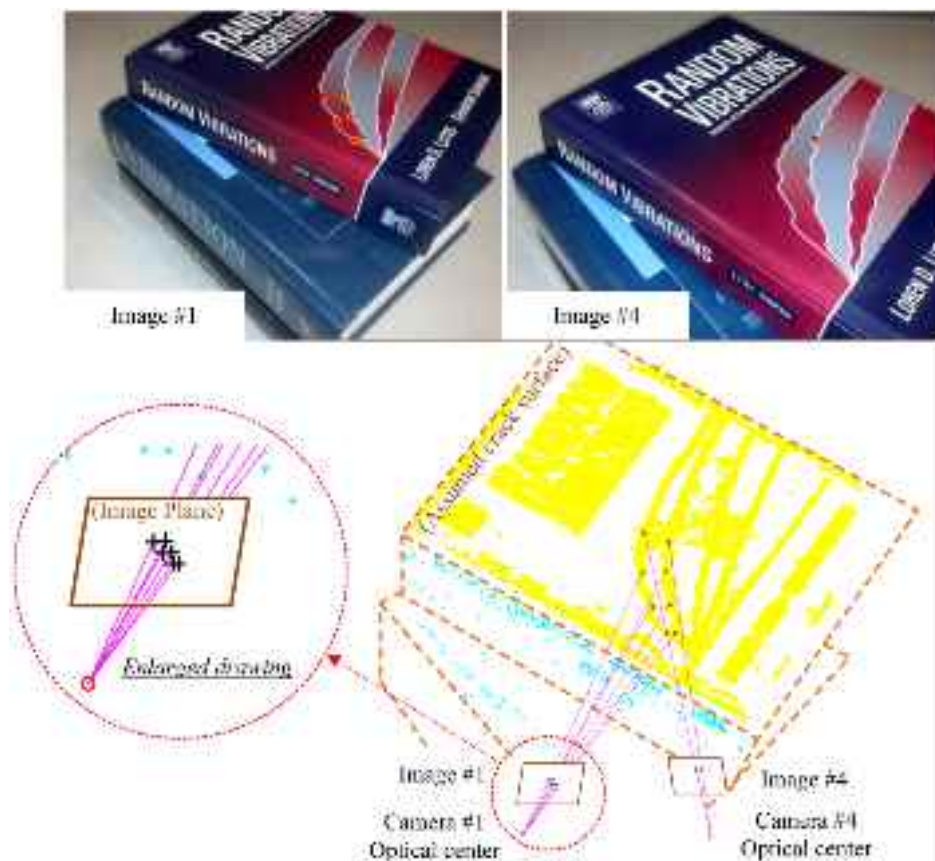
Fig. 14. Projection from images onto 3D scene using the pin-hole model; points used for projection are marked as red dots in Images 1 and 4 [this cover was published in Loren, L., and Sarkani, S. (2003). *Random Vibrations: Analysis of Structural and Mechanical Systems*, Elsevier Butterworth-Heinemann, Burlington, Massachusetts, Copyright Elsevier (2004)]

DIP technique, precise coordinates of the crack width points under the image coordinate system can be obtained. The OFSI values of the three crack images were chosen as 80 pixels. After the noise elimination and binarization, 2D crack width calculation was conducted and the locations of the crack width points under the pixel coordinate were obtained. Take Image 4 as an example; the process of noise elimination and binarization is presented in Fig. 17(a) while the 2D crack width calculation is presented in Fig. 17(b). The crack width points used in the projection were obtained for every 15 pixels on the skeleton line.

Step 2, Imaging Model Retrieval and 3D Scene Reconstruction of Concrete Surface

The SFM technique was used for retrieving the 3D imaging model and the sparse 3D scene from the six images shown in Fig. 16. The key points were obtained by the SIFT technique and they are matched by crossing matching. Fig. 18 shows the obtained sparse 3D scene point cloud with the roving camera locations in frontal and side views. In Fig. 18, the triangles indicate the location and orientation of cameras. The number near the camera sequentially corresponds to the images listed in Fig. 16, and C refers to camera.
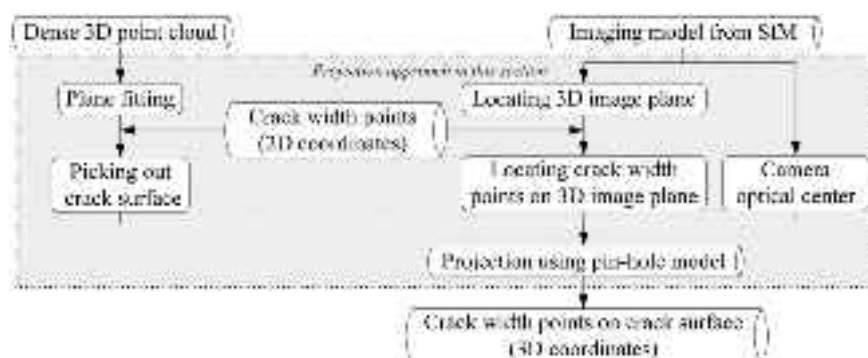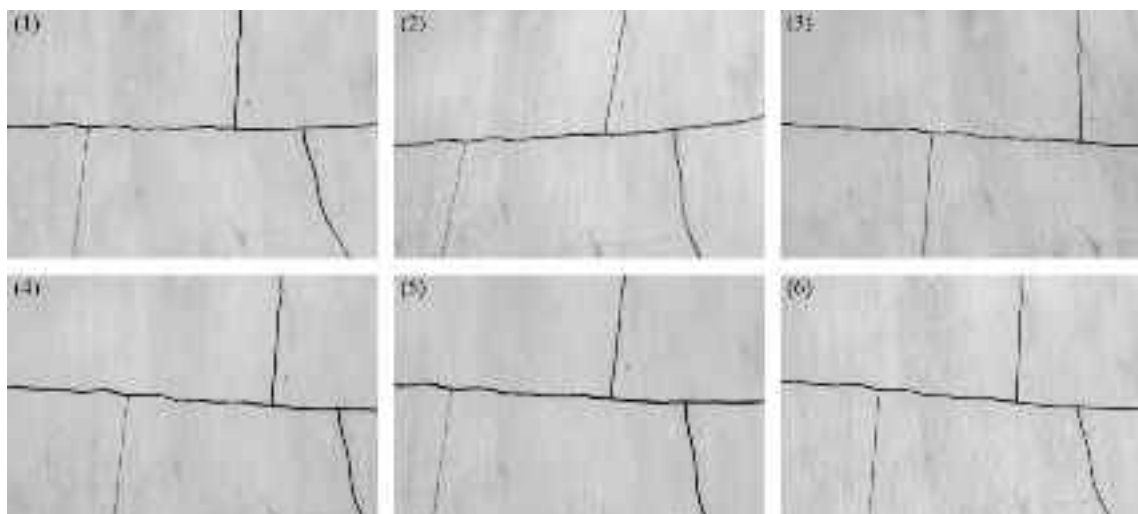

Fig. 15. Flow diagram of projection approach

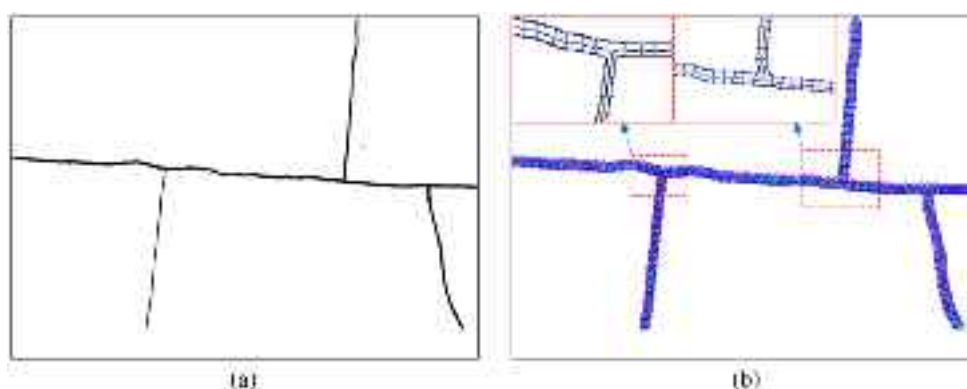**Fig. 16.** Undistorted crack images taken on a single surface

Fig. 17. Crack detection and width calculation of Crack Image 4: (a) crack image after DIP; (b) width calculation

The point cloud in Fig. 18 is formed of the matched key points right after retrieving the imaging model; then, some more points are obtained after the image rectification and added into the sparse scene to form the dense 3D scene (Fig. 19). The cloud in the side view also shows that the images are taken upward sloping.

The dense 3D scene is composed using CMVS. Unlike the sparse 3D point cloud which hardly shows the characteristics of the crack object, the dense 3D scene roughly presents the distribution of cracks. However, the accurate locations of the crack width points cannot be assessed from the 3D scene due to the large tolerances of the point locations in the dense 3D reconstruction.

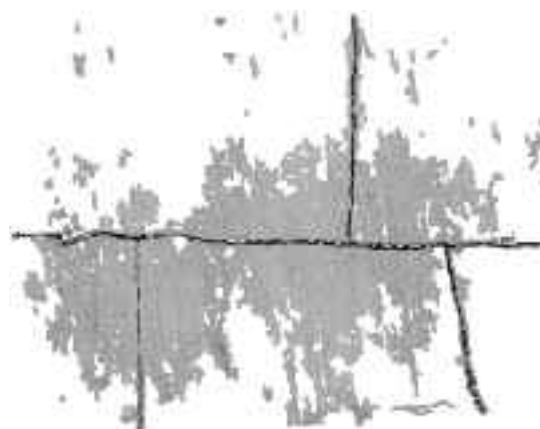**Fig. 18.** Sparse 3D scene point cloud with roving camera locations in frontal and side views

Fig. 19. Dense 3D scene of concrete surface

© ASCE 04014124-11 J. Comput. Civ. Eng.
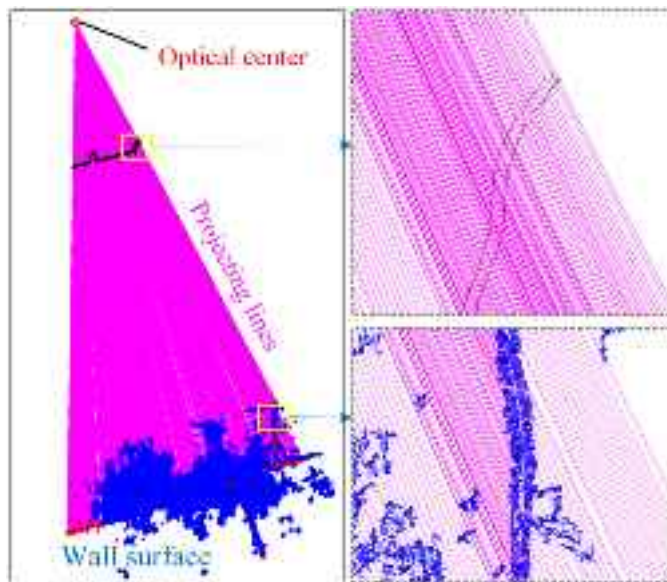
J. Comput. Civ. Eng., 2016, 30(1): 04014124

Fig. 20. Projection of crack width points from Image 4 onto dense 3D scene

Step 3, Crack Projection and Final Results of Crack Assessment

Only one plane (wall surface) was fitted from the dense 3D scene and was used for the projection of crack. The crack width points of the 3 images obtained in Step 1 were projected onto the plane individually. The distance from the imaging sensor of Crack Image 6 to the center of the imaging scene was measured as 24.1 cm. This distance will be used to scale the imaging model. The crack from Image 4 projected onto the fitted plane is shown in Fig. 20. The circle on the top indicates the camera location (optical center) of

Image 4, and the radials are the projection lines. The small dots located on the radius are the crack width points on the image plane (in the world coordinate system) obtained by Step 1. The large dots on the wall surface are the 3D points involved in the plane fitting.

Fig. 21 shows the crack nephograms (i.e., crack distribution map) obtained in the world coordinate system from the three crack images separately. With the plane equation of the concrete surface known from Step 2, the crack skeleton (the middle line between two edge lines) in the 3D scene can be transformed in the 2D nephogram. The shapes of the crack in the three nephograms are slightly different from any of the original crack images for they represent the real crack shape which can be captured perpendicularly against the crack wall.

Results and Errors

For the validation of the obtained crack nephogram, the real crack width was measured at seven points (Fig. 22), using the crack width ruler shown in Fig. 22(b). Table 1 shows the comparison of the assessed crack widths from the three crack images with the measured ones at the seven points. For the assessed result using Image 4, the errors are quite negligible (absolute errors less than 0.05 mm). For Test Points T1-2 and T1-7, the relative errors of Images 5 and 6 are a little large, which are possibly caused by local out of focus of the crack images. In Test 1, although the two points have high relative errors, their absolute errors are less than 0.15 mm, which is still acceptable considering the narrow crack widths of the two points [i.e., (1) 0.25, and (2) 0.50 mm].

Test 2, Multiple Concrete Surfaces

The cracks propagated along a corner of the concrete wall were identified in Test 2. The trimmed concrete corner has three surfaces to be inspected. Multiple images were taken around the corner using the same camera and lens as in Test 1, and the blurry criteria

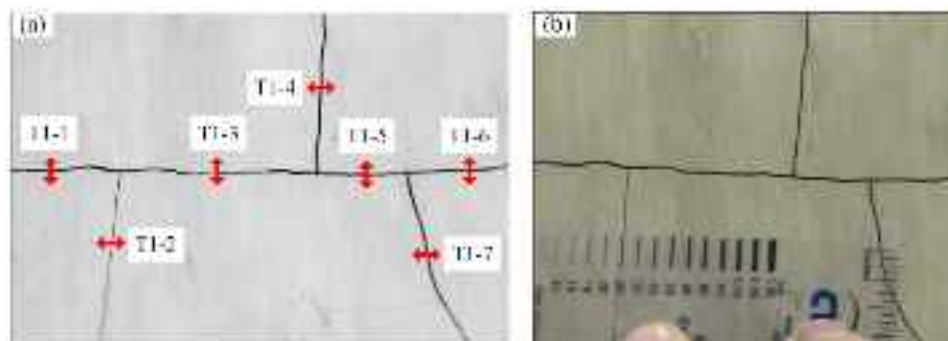**Fig. 21.** Crack nephogram; units are in millimeters

Fig. 22. (a) Measurement locations; (b) crack ruler used in measurement

© ASCE   04014124-12   J. Comput. Civ. Eng.

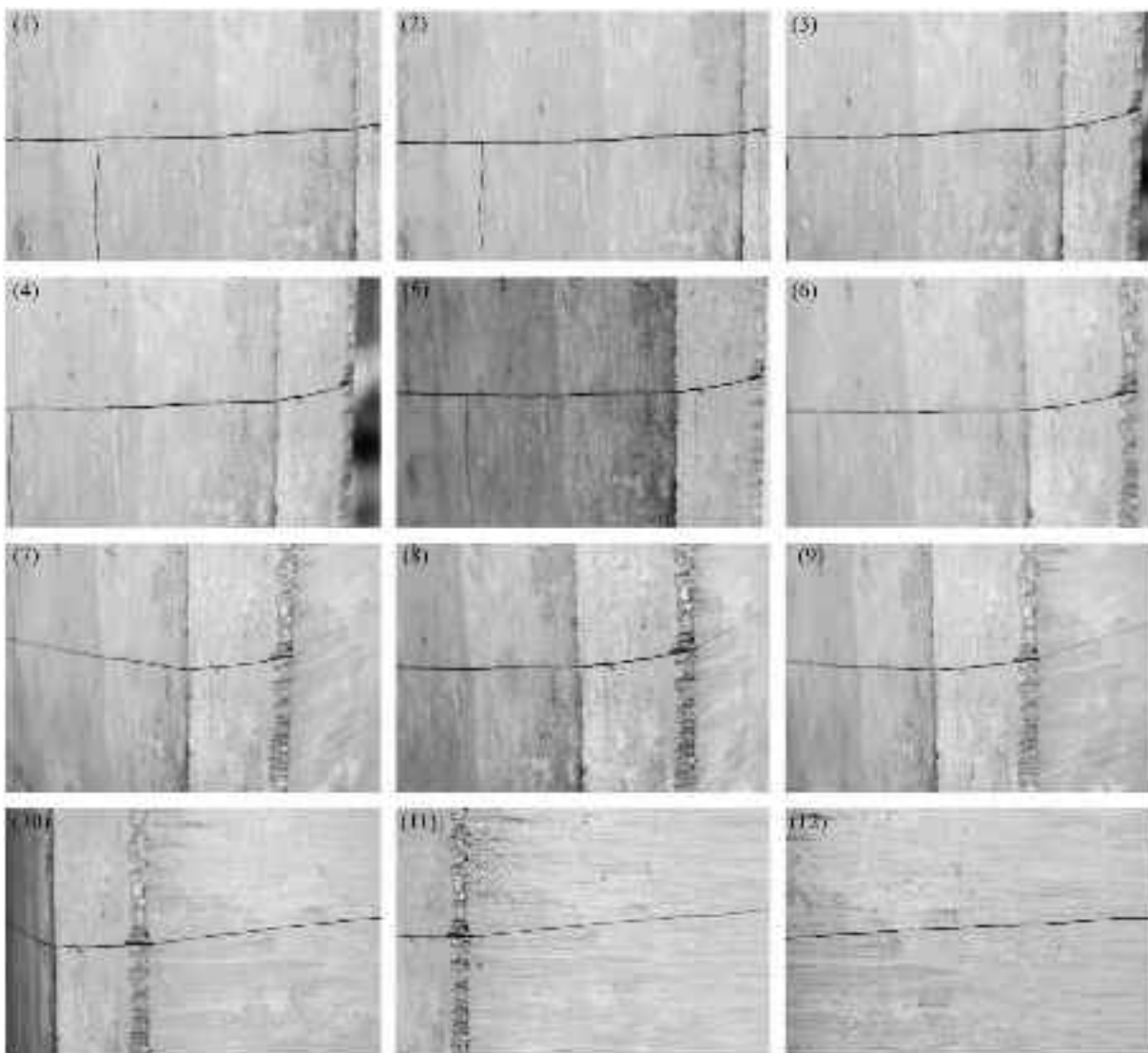J. Comput. Civ. Eng., 2016, 30(1): 04014124

Table 1. Results of the Assessed Crack Width and Comparison with Measured Data

Point number	Measured crack width (mm)	Assessed crack width (mm)			Absolute error (mm)			Relative error (%)			Average relative error (%)
		d	e	f	d	e	f	d	e	f	
T1-1	0.90	0.90	0.90	0.83	0.00	0.00	−0.07	0	0	−8	3
T1-2	0.25	0.30	0.34	0.38	0.05	0.09	0.13	20	36	52	46
T1-3	0.90	0.90	0.83	0.83	0.00	−0.07	−0.07	0	−8	−8	5
T1-4	0.50	0.50	0.53	0.53	0.00	0.03	0.03	0	6	6	4
T1-5	0.80	0.80	0.72	0.75	0.00	−0.08	−0.05	0	−10	−6	5
T1-6	0.80	—	0.75	0.75	—	−0.05	−0.05	—	−6	−6	−4
T1-7	0.50	0.50	0.53	0.35	0.00	0.03	−0.15	0	6	−30	−8

were applied to obtain the clear images. As a result of the influence of background bokeh due to the short working distance, images taken on one side of the wall will has the scene on the other sides blurred. To precisely capture all the cracks, among the obtained clear images, 12 images taken from different angles towards the corner of the wall were finally selected for the crack assessment. Fig. 23 shows the selected images; they were also undistorted by the preliminary camera calibration process.

Step 1, Crack Assessment Using the 2D DIP Technique

Two individual crack assessments were conducted using the crack images. To ensure that in each assessment case all the crack objects can be captured and detected, at least two images with different views are needed. The crack images used for the 2D crack assessment are (1) Images 2 and 11 for Assessment Case 1, and (2) Images 1, 9, and 12 for Assessment Case 2. The process of Case 1 is presented next.

Fig. 23. Undistorted crack images taken on corner of wall (from the left view to the right view)
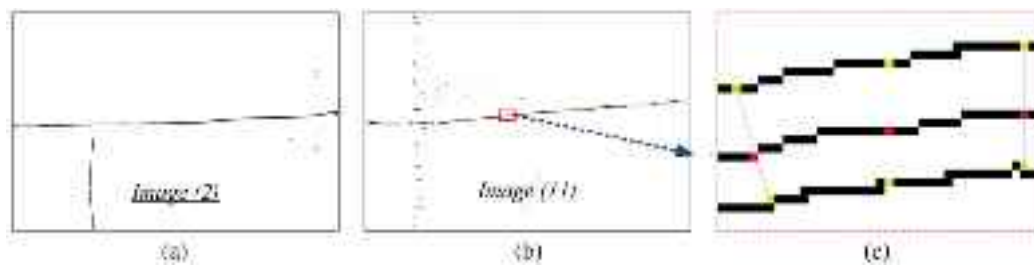
Fig. 24. Crack width points obtained by 2D crack detection approach (Case 1): (a) crack after noise elimination; (b) crack after noise elimination; (c) crack width calculation (enlarged drawing)

Similarly to Test 1, the crack detection was conducted and the locations of the crack width points under the pixel coordinate were obtained using the DIP crack assessment approach (Fig. 24). The OFSI values of the crack images were all chosen as 70 pixels. After the noise elimination using OFSI, few noise points still existed on the binarization crack image [Fig. 24(b)]. While conducting the crack width calculation, these short stubs are removed using a threshold of relative skeleton length. In this way, most of the left noise can be reduced.

Step 2, Imaging Model Retrieval and 3D Scene Reconstruction of Concrete Surface

The retrieved 3D sparse model together with the camera locations; the dense model is presented in Figs. 25 and 26, respectively. In Fig. 25 the triangles indicate the location and orientation of cameras. The number near the camera sequentially corresponds to the images listed in Fig. 23, and C refers to camera. The locations of Cameras 1 and 2 are very close so that they are displayed almost at the same place. The dense 3D scene roughly shows the distribution of cracks with some errors. From Fig. 26, the cracks on the left-side surface should be continuous rather than broken. This kind of error indicates that the accuracy of the 3D reconstruction is not enough to directly calculate the crack width.
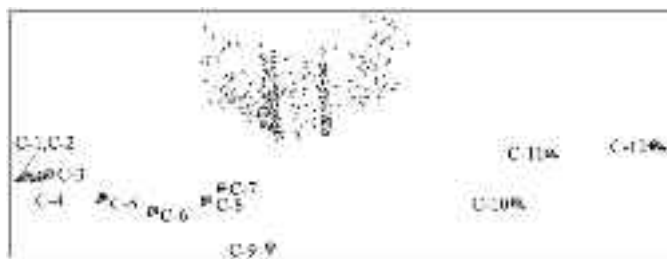
Step 3, Crack Projection and Final Results of Crack Assessment

Three planes (wall surfaces) were fitted from the dense 3D scene and were used for the projection of crack (Fig. 27). The final crack was stitched using the projection of 2D crack width points obtained from Step 1. The crack width nephograms for Cases 1 and 2, which is unfolded from the original three planes, are presented in Figs. 28(a and b), respectively.

Results and Errors

The four points shown in Fig. 29(a) were measured to validate the result from the proposed process. The real crack width was measured using the crack width ruler shown in Fig. 29(b). Table 2 presents the comparison of the assessed crack widths with the measured ones at four points. The relative errors of Test Points T2-1 and T2-4 are larger comparing to Points T2-2 and T2-3, but the absolute errors (0.06 and 0.05 mm, and 0.08 and 0.07 mm) are still acceptable.
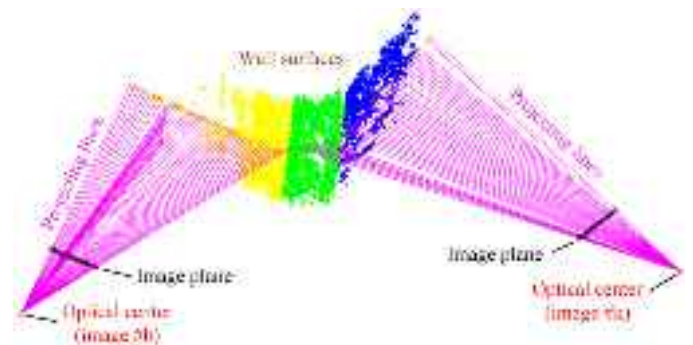

Fig. 27. Projection of crack width points from 2D image onto dense 3D scene (Case 1)

Fig. 25. Sparse 3D scene point cloud with roving camera locations

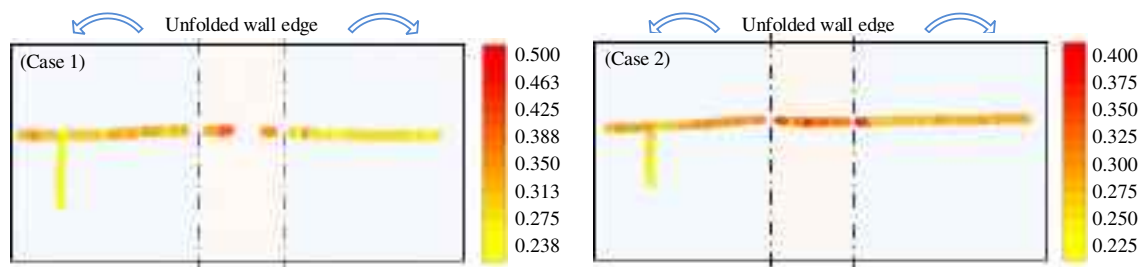**Fig. 26.** Dense 3D scene in two views

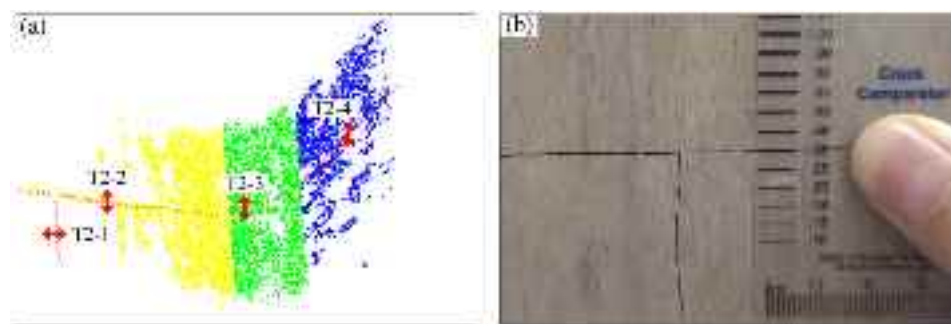**Fig. 28.** Crack nephogram; units are in millimeters

Fig. 29. (a) Measurement locations; (b) crack ruler used in measurement

Test 3, Complex Crack Shapes

In Test 3, crack of complex shapes distributing along a concrete top flange near a concrete-steel composite joint is processed. This joint was an experiment-used structure member under cyclic loading. During the loading test, more than one crack appeared on the top flange near the joint. Refer to Nie et al. (2008) for more details about the load test and composite joint. Similarly to Test 2, the crack exists on multiple surfaces. Nine crack images (Fig. 30) were final employed for the crack detection and 3D reconstruction. In Test 3, the shape of the crack is tortuous which may introduce errors in both the calculation and the measurement of crack width.

Step 1, Crack Assessment Using the 2D DIP Technique

Image 1 in Fig. 30 was used for 2D crack detection. In Test 3, all the crack images contain multiple objects including the flange and some other scene such as the steel column, which forms large area of shadings in the crack images. Under such circumstances, the proposed background subtraction plays an important role in removing the complex background of the crack image. The OFSI values of Crack Image 1 were chosen as 50 pixels. To better show the effect of background subtraction, the crack image binarized directly by the Otsu and Niblack (correction factor = −0.15, structure size = 15 pixels) methods are presented in Figs. 31(a and b) for

comparison, while Fig. 31(c) is the one processed using the proposed background subtraction and the binarization method. Background subtraction maximally removed the large area shadings. Crack width calculation was conducted employing Crack Image 1 in Fig. 30, parts of the calculation are shown in the enlarged drawing in Fig. 31(c). Considering the viewing angle and scope, only one crack image (i.e., Image 1) was used in 2D crack detection.

Step 2, Imaging Model Retrieval and 3D Scene Reconstruction of Concrete Surface

The retrieved 3D sparse model together with the camera location is presented in Fig. 32. The triangles indicate the location and orientation of cameras. The number near the camera sequentially corresponds to the images listed in Fig. 30 and C refers to camera.

Step 3, Crack Projection and Final Results of Crack Assessment

The projection of the crack width points obtained from Step 1 is shown in Fig. 33. Three surfaces are involved in the scene and two of them have a crack. From the side view, one can easily find that the image plane is slopy. As mentioned in Tests 1 and 2, crack distribution and width obtained on a slopy image does not mean the real case and projection to the 3D fitted plane can make the correction. After the projection the real crack distribution was obtained

Table 2. Results of the Assessed Crack Width and Comparison with Measured Data

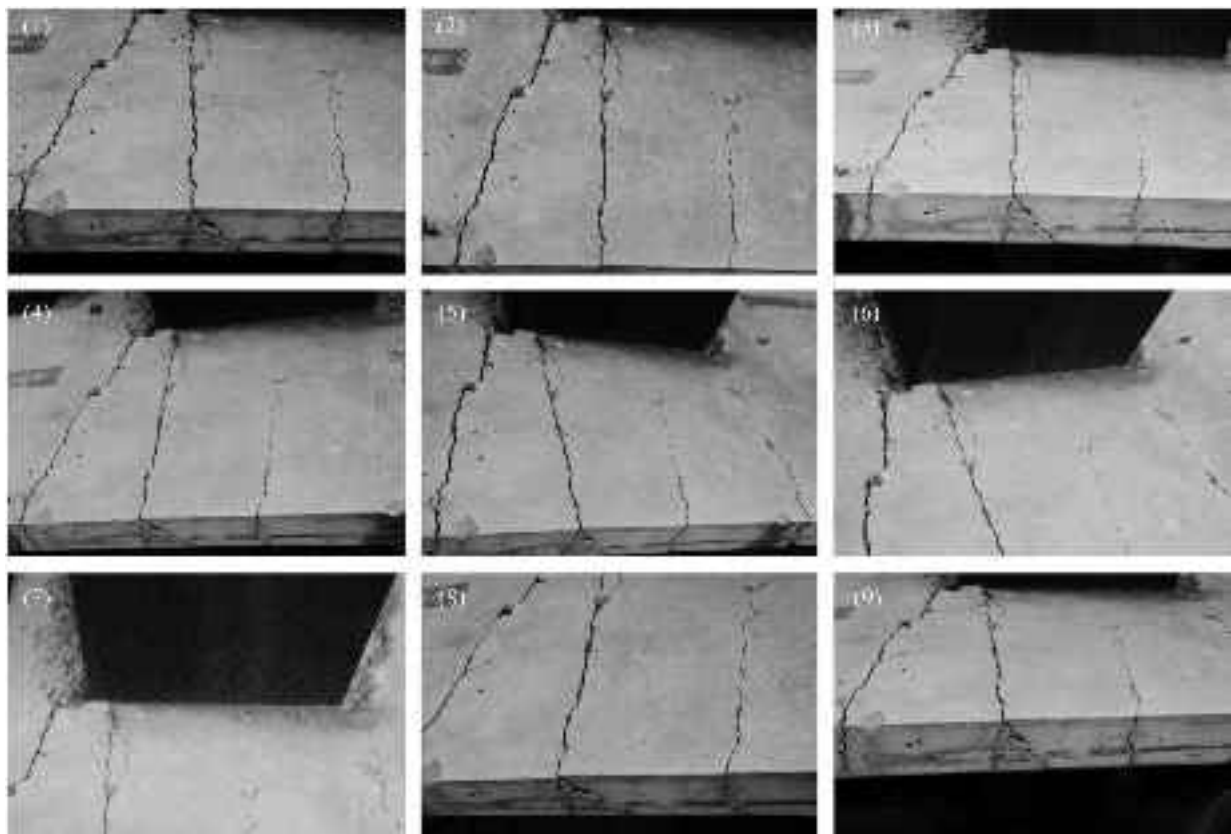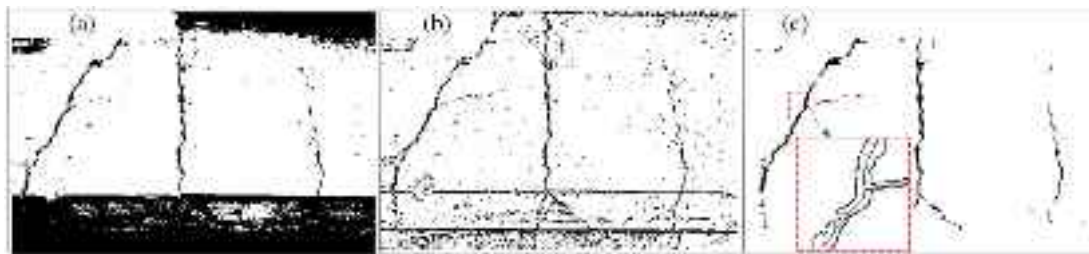| Point number	Measured crack width (mm)	Assessed crack width (mm)		Absolute error (mm)		Relative error (%)		Average relative error (%)
		Case 1	Case 2	Case 1	Case 2	Case 1	Case 2	
T2-1	0.18	0.24	0.23	0.06	0.05	33	28	31
T2-2	0.30	0.31	0.30	0.01	0.00	3	0	2
T2-3	0.40	0.46	0.38	0.06	−0.02	15	−5	5
T2-4	0.23	0.31	0.30	0.08	0.07	35	30	33

Fig. 30. Undistorted crack images

Fig. 31. Two-dimensional image processing and crack width calculation: (a) Otsu method; (b) Niblack method; (c) proposed DIP and width calculation

Fig. 32. Sparse 3D scene point cloud with roving camera locations

and the real crack width was calculated (Fig. 34). The width of the top surface of flange (distance from the column surface to the side face of flange), which is 26.4 cm, is measured to scale the 3D scene.

Results and Errors

Eight points were measured for the validation of the crack width assessment (Fig. 35). Because the crack width in this test can be as large as 2 mm and exceed the measuring range of a crack ruler, a vernier caliper was used to measure the real crack width. Table 3 presents the comparison of the assessed crack widths with the measured ones. For Test Points T3-1, T3-2, T3-4, T3-5 and T3-6, the absolute error (less than 0.08 mm) and relative error (less than 3%) are negligible. For Point T3-3, T3-7, and T3-8, the relative error is a lithe higher, which is possibly caused by the tortuous crack shape. To improve the assessment accuracy in this case, images with higher sharpness are recommended.
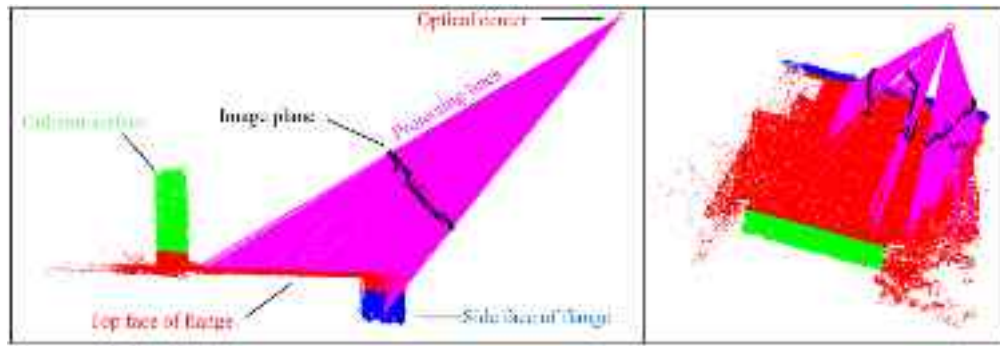
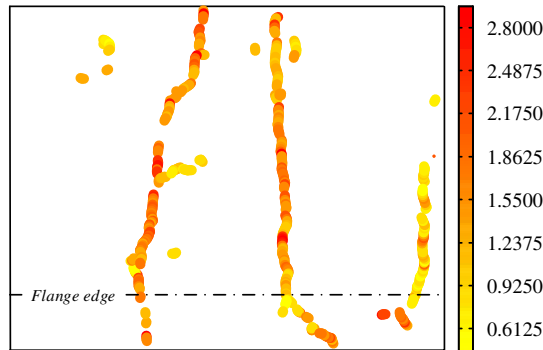**Fig. 33.** Projection of crack width points from 2D image onto dense 3D scene

Fig. 34. Crack nephogram; units are in millimeters

Table 3. Results of the Assessed Crack Width and Comparison with Measured Data

Point number	Measured crack width (mm)	Assessed crack width (mm)	Absolute error	Relative error (%)
T3-1	2.57	2.49	−0.08	−3
T3-2	2.49	2.49	0	0
T3-3	1.06	1.40	0.34	32
T3-4	2.45	2.49	0.04	2
T3-5	1.89	1.86	−0.03	−2
T3-6	0.96	0.93	−0.03	−3
T3-7	0.74	0.61	−0.13	−18
T3-8	0.77	0.61	−0.16	−21

Efficiency and Accuracy of Assessment

The computational time for different steps in Tests 1–3 is presented in Table 4. For two or more individual assessments contained in one test, the value in Table 4 refers to one assessment procedure. For the other individual assessments, usually only the 2D crack width detection and the projection needs to be done again. During Tests 1–3, a desktop with Intel an i7 quad-core 3.5 GHz CPU was used (no parallel computation except for the 3D dense reconstruction) for all the procedure. The median filtering operation medfilt2 of *Matlab* works quite slowly for large images and large filter structures. In this paper, the algorithm median filtering in constant time (MFCT) by Perreault and Hébert (2007) was employed for median filtering operation. The algorithmic complexity of the MFCT algorithm is O (1) when the size of filter window varies, and about 0.9 s is needed

for each median filtering processing using the crack image in Test 1.

The mean of relative errors for each tests are presented in Table 4 as well. The mean of relative error for all the crack points (36 points) is 4.74%, which is pretty small and acceptable, comparing to −5.45% (eight test points) by Jahanshahi et al. (2013) and 31.41% (18 test points) by Nishikawa (2012).

Summary and Conclusion

This paper proposes an automated concrete crack assessment method using digital image processing and 3D scene reconstruction. Two main parts are included in the approach, as follows: (1) 2D crack assessment technique where the pixel coordinates of the crack width points are calculated, and (2) projection of the crack width point using the 3D reconstruction technique. After
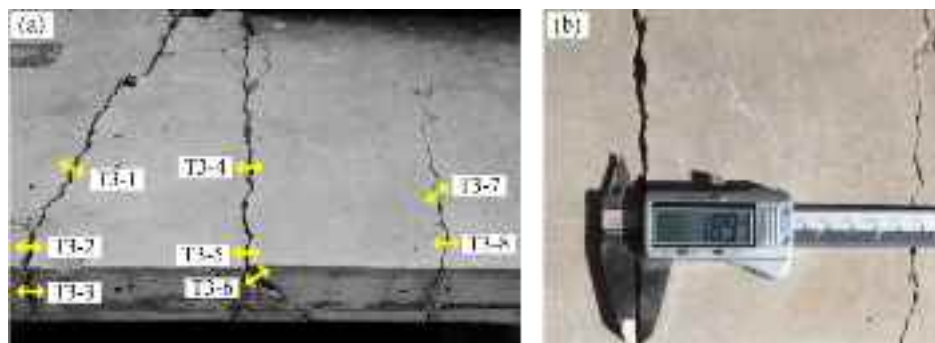

Fig. 35. (a) Measurement locations; (b) vernier caliper used in measurement

© ASCE   04014124-17   J. Comput. Civ. Eng.

J. Comput. Civ. Eng., 2016, 30(1): 04014124

Table 4. Computational Time and Mean of Relative Error for Tests 1–3

Test number	Computational time/min					Mean of relative error (%)
	Optimal filter size index calculation	Three-dimensional reconstruction	Three-dimensional plane fitting	Other procedure[a]	Total	
1	3	1.5	2	2	8	2.20
2	3	5	4	2	15	17.46
3	2	0.8	2	2	7	−1.63

[a]Includes 2D crack width detection, projection, and data processing.

projection, the precise locations of the crack width point under the 3D world coordinate are found, which will be used for the calculation of crack information and the presentation of crack distribution.

The proposed crack assessment approach significantly improves the applicability of the traditional image-based crack assessment approaches. The four main benefits from the proposed approach are as follows: (1) images used for 2D crack assessment do not need to be taken perpendicularly to the concrete crack wall (images taken on any angle can be used for the accurate crack assessment), (2) imaging model is ntegrated and composite for multiple crack images (the working distances for different images do not need to be measured each time and can be scaled using any absolute length in the scene), (3) cracks on the corner of the concrete structure member and piercing through multiple intersected surfaces are able to be detected, and (4) multiple images used in the 2D crack assessment can be stitched into a same 3D scene to deal with situations of large area cracks. Because of these benefits, the proposed methodology can be applied in the damage detection or regular health monitoring of concrete structure members, which are difficult for operators to get near or perpendicularly taking crack images.

The validation tests were carried out on a concrete reaction wall as well as a concrete flange and relatively precise detection results were obtained. The approach proposed in this paper supplies an automated, convenient, and reliable solution to crack assessment of concrete structures and offers significant potential for practical implementation.

Acknowledgments

References

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

Ayache, N., and Hansen, C. (1988). "Rectification of images for binocular and trinocular stereovision." *Proc., Int. Conf. on Pattern Recognition*, IEEE, New York.

Bay, H., Tuytelaars, T., and Van Gool, L. (2006). "Surf: Speeded up robust features." *Computer Vision—ECCV 2006*, Springer, Amsterdam, Netherlands, 404–417.

Bouguet, J.-Y. (2013). "Camera calibration toolbox for Matlab." ⟨http://www.vision.caltech.edu/bouguetj/calib_doc/⟩ (Oct. 31, 2014).

Crete, F., Dolmiere, T., Ladret, P., and Nicolas, M. (2007). "The blur effect: Perception and estimation with a new no-reference perceptual blur metric." *Human vision and electronic imaging*, 6492, 64920I.

Esteban, I., Dijk, J., and Groen, F. (2010). "FIT3D toolbox: Multiple view geometry and 3D reconstruction for matlab." *Int. Symp. on Security and Defence Europe*, Toulouse, France, 78350J.

Fischler, M. A., and Bolles, R. C. (1981). "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography." *Commun. ACM*, 24(6), 381–395.

Fujita, Y., and Hamamoto, Y. (2011). "A robust automatic crack detection method from noisy concrete surfaces." *Mach. Vision Appl.*, 22(2), 245–254.

Furukawa, Y., Curless, B., Seitz, S. M., and Szeliski, R. (2010). "Towards internet-scale multi-view stereo." *Proc., IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, New York.

Furukawa, Y., and Ponce, J. (2010). "Accurate, dense, and robust multiview stereopsis." *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(8), 1362–1376.

Gallo, O., Manduchi, R., and Rafii, A. (2011). "CC-RANSAC: Fitting planes in the presence of multiple surfaces in range data." *Pattern Recogn. Lett.*, 32(3), 403–410.

Huang, T. S., and Netravali, A. N. (1994). "Motion and structure from feature correspondences: A review." *Proc. IEEE*, 82(2), 252–268.

Jahanshahi, M. R., and Masri, S. F. (2012). "Adaptive vision-based crack detection using 3D scene reconstruction for condition assessment of structures." *Autom. Constr.*, 22(SI), 567–576.

Jahanshahi, M. R., and Masri, S. F. (2013). "Parametric performance evaluation of wavelet-based corrosion detection algorithms for condition assessment of civil infrastructure systems." *J. Comput. Civ. Eng.*, 10.1061/(ASCE)CP.1943-5487.0000225345–357.

Jahanshahi, M. R., Masri, S. F., Padgett, C. W., and Sukhatme, G. S. (2013). "An innovative methodology for detection and quantification of cracks through incorporation of depth perception." *Mach. Vision Appl.*, 24(2), 227–241.

James, M., and Robson, S. (2012). "Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application." *J. Geophys. Res. Earth Surf.*, 117(F3), in press.

Juan, L., and Gwun, O. (2009). "A comparison of sift, pca-sift and surf." *Int. J. Image Process.*, 3(4), 143–152.

Liu, Y., Cho, S., Spencer, J. B. F., and Fan, J. (2014). "Automated assessment of cracks on concrete surfaces using adaptive digital image processing." *Smart Struct. Syst.*, 14(4), 719–741.

Lowe, D. G. (2004). "Distinctive image features from scale-invariant keypoints." *Int. J. Comput. Vision*, 60(2), 91–110.

Lutes, L. D., and Sarkani, S. (2004). *Random vibrations: Analysis of structural and mechanical systems*, Elsevier Butterworth-Heinemann, Burlington, MA.

Matlab 8.2 [Computer software]. Natick, MA, MathWorks.

Nie, J., Qin, K. and Cai, C. S. (2008). "Seismic behavior of connections composed of CFSSTCs and steel-concrete composite beams—Experimental study." *J. Constr. Steel Res.*, 64(10), 1178–1191.

Nishikawa, T., Yoshida, J., Sugiyama, T., and Fujino, Y. (2012). "Concrete crack detection by multiple sequential image filtering." *Comput. Aided Civ. Infrastruct. Eng.*, 27(1), 29–47.

Papadimitriou, V., and Dennis, T. J. (1996). "Epipolar line estimation and rectification for stereo image pairs." *IEEE Trans. Image Process.*, 5(4), 672–676.

Perreault, S., and Hébert, P. (2007). "Median filtering in constant time." *IEEE Trans. Image Process.*, 16(9), 2389–2394.

© ASCE   04014124-18   J. Comput. Civ. Eng.

J. Comput. Civ. Eng., 2016, 30(1): 04014124

Pollefeys, M., Koch, R., and Van Gool, L. (1999). "A simple and efficient rectification method for general motion." *Proc., IEEE Int. Conf. on Computer Vision*, New York.

Rashidi, A., Dai, F., Brilakis, I., and Vela, P. (2013). "Optimized selection of key frames for monocular videogrammetric surveying of civil infrastructure." *Adv. Eng. Inform*, 27(2), 270–282.

Seo, Y.-H., Kim, S.-H., Doo, K.-S., and Choi, J.-S. (2008). "Optimal keyframe selection algorithm for three-dimensional reconstruction in uncalibrated multiple images." *Opt. Eng.*, 47(5), 053201.

Shashua, A. (1997). "Trilinear tensor: The fundamental construct of multiple-view geometry and its applications." *Algebraic frames for the perception-action cycle*, Springer, Amsterdam, Netherlands, 190–206.

Sinha, S. K., and Fieguth, P. W. (2006). "Segmentation of buried concrete pipe images." *Autom. Constr.*, 15(1), 47–57.

Sonka, M., Hlavac, V., and Boyle, R. (2008). "Image processing, analysis, and machine vision." Thomson Engineering, Toronto.

Tanaka, N., and Uematsu, K. (1998). "A crack detection method in road surface images using morphology." *Proc., IAPR Workshop on Machine Vision Applications*, IAPR MVA, Tokyo, 154–157.

Triggs, B., McLauchlan, P. F., Hartley, R. I., and Fitzgibbon, A. W. (2000). "Bundle adjustment—A modern synthesis." *Vision algorithms: Theory and practice*, Springer, Amsterdam, Netherlands, 298–372.

Tsai, R. Y. (1987). "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses." *IEEE J. Rob. Autom.*, 3(4), 323–344.

Uslu, B., Golparvar-Fard, M., and de la Garza, J. M. (2011). "Image-based 3D reconstruction and recognition for enhanced highway condition assessment." *Proc., ASCE Int. Workshop on Computing in Civil Engineering*, Reston, VA.

Weng, J., Cohen, P., and Herniou, M. (1992). "Camera calibration with distortion models and accuracy evaluation." *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(10), 965–980.

Wu, C. (2011). "VisualSFM: A visual structure from motion system." ⟨http://ccwu.me/vsfm/⟩ (Nov. 6, 2014).

Wu, C., Agarwal, S., Curless, B., and Seitz, S. M. (2011). "Multicore bundle adjustment." *Proc., IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, New York.

Zhang, Z. (1999). "Flexible camera calibration by viewing a plane from unknown orientations." *Proc., IEEE Int. Conf.on Computer Vision*, New York.

Zhang, Z. (2000). "A flexible new technique for camera calibration." *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(11), 1330–1334.

Zisserman, A. (2003). *Multiple view geometry in computer vision*, Cambridge University Press, Cambridge, U.K.

© ASCE   04014124-19   J. Comput. Civ. Eng.

J. Comput. Civ. Eng., 2016, 30(1): 04014124