Automatic multi-image stitching for concrete bridge inspection by combining point and line features

Renping Xie, Jian Yao[*], Kang Liu, Xiaohu Lu, Yahui Liu, Menghan Xia, Qifei Zeng

Computer Vision and Remote Sensing (CVRS) Lab, School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, Hubei, PR China

ABSTRACT

Most of the current techniques for concrete bridge inspection are based on human visual interpretation, which often is dangerous and time-consuming. To address this problem, we introduce in this paper a newly developed vehicle-based robot inspection system that can automatically capture thousands of bottom surface images with a group of high-resolution industrial cameras, which are then stitched into a single composite image. However, traditional image stitching methods generally fail with large drift due to the great number (more than 2000) and sparse texture of linearly distributed images in sequence. Therefore, a novel image stitching method was developed for our robot inspection system, which combines both the 2D image point features and the 3D line features to reduce the drift. First, the bottom surface images are arranged into different strips based on their acquisition order and rough poses, and images in a single strip are divided into several groups. Then, the proposed image stitching method is performed in a bottom-up way, as follows: 1) the images within a single group initially are aligned via their point and line features; 2) the groups within a single strip are then stitched together via a homographic refinement procedure; 3) the strips are aligned into a single composite image that completely covers the bottom surface of the bridge; and 4) after all the stitching procedure are complete, a multi-band blending algorithm is applied to generate the mosaicked panorama as seamlessly as possible. The experimental results on a set of representative images acquired from the bottom surfaces of a real bridge demonstrate the capabilities and the limitations of the proposed approach.

1. Introduction

Concrete and cement are the most widely used man-made materials for constructed systems, especially bridges. Due to excessive usage, overloading, and aging, as well as insufficient maintenance and inspection deficiencies [1], many defects (e.g., cracks) appear on the surfaces of these materials. According to statistics of the National Bridge Inventory (NBI) of the United States Federal Highway Administration [2,3], more than 30,000 bridges on which people rely daily are vulnerable to failure. There is an urgent need to develop more effective approaches for the inspection and evaluation of these bridges. In addition, periodic inspections and maintenance of bridges are necessary to prolong their service lives [4–6]. Although there are many sophisticated nondestructive evaluation methods [7–12] in engineering practice, visual inspection by the human eye [13] is the predominant method still used for the inspection of bridges. However, manual inspection is a labor-intensive task and is also dangerous, qualitative, subjective, and time-consuming [14], especially for the complicated bottom structures of some bridges.

To make the bridge inspection process more effective and safer, various robot systems via remote sensing techniques have been developed in the past decade, which can detect defects (e.g., cracks) more accurately and assess the health status of bridges remotely and easily. Oh et al. [15] designed a robotic system for inspecting the safety status of bridges, which consists of three parts: a specially designed car, a robot mechanism and control system for mobility, and a machine vision system for automatic crack detection. Lim et al. [16] developed a system that used a mobile robot to conduct the inspection, where the robot collected bridge deck images with a high resolution camera. De-Vault [17] developed an automated robotic system to enable safe and cost-effective underwater inspections of bridge substructures. La et al. [18] presented a mechatronic system design for an autonomous robotic system for highly efficient bridge deck inspection and evaluation. Yang et al. [19] proposed an innovative approach for conducting bridge inspections by optimizing an unmanned aerial system (UAS) that comprised a rotorcraft prototype and a camera gimbal mechanism. In addition, other robot systems [20,21] have been developed for bridge inspection.

Cracks are the main defects that appear on the surfaces of concrete materials. Not only in roads [22–24] and buildings [25,26], but also in bridges [1,27,28], crack inspection is an essential task to monitor the health status of these structures. The existing crack detection methods can be classified into two categories: local-feature-based [22,29] and global-feature-based algorithms [30,31]. However, in order to capture high-resolution images of the bottom surface of a bridge, industrial cameras must be placed at close range with a mobile platform which only covers small regions each time. Thus, when using the robotic system to clearly inspect the defects (e.g., cracks) appearing on the bottom surface of a bridge, thousands of images usually are collected. Detecting cracks or other defects on all the images one by one would take a lot of time and cannot represent images as a whole. So, stitching all the images captured from the bridge bottom surface into a single wide-view panorama is necessary in order to comprehensively evaluate the health status of a bridge.

Image stitching technology is widely used in the fields of image processing [32,33], photogrammetry [34,35], and computer vision [36,37], and have been featured in the literature and commercial applications [38,39]. Recently, Brown and Lowe [36] developed a novel method that can recognize multiple panoramas in an unordered image dataset using invariant features for panoramic image stitching. Zaragoza et al. [40] developed an as-projective-as-possible warping method based on a novel moving direct linear transformation (Moving DLT) technique to seamlessly stitch image regions that are inconsistent with the projective model. Jahanshahi and Masri [41] presented novel integrated inspection software based on the use of inexpensive digital cameras, which are appropriately mounted on a structure and can zoom or rotate in three directions (similar to traffic cameras). Rankov et al. [42] proposed an optimized, automated, fast, and reliable method for both image joining and blending that overcame intensity discrepancies and geometric misalignments between the stitched images. Jia and Tang [43] achieved seamless image stitching without producing visual artifacts based on structure deformation and propagation. Although the above studies have solved some key problems in image stitching, which is not suitable for creating a complete single-view image for the bottom surface of a bridge from a large number of high-resolution images.

This paper introduces a specially designed bridge inspection system to automatically detect the bridge defects (e.g., cracks) from a complete single-view image of the bridge bottom surface, which is obtained via stitching thousands of high-resolution images. The remainder of this paper is organized as follows. The new bridge inspection system is briefly introduced in Section 2. Section 3 describes the proposed automatic multi-image alignment and stitching approach in detail, which consists of image alignment within a single strip, image alignment between neighboring strips, and image warping and blending. Experimental results on a large set of images acquired from the bottom surface of a real bridge are discussed and evaluated in Section 4. Conclusions and future work are provided in Section 5.

2. System overview

The designed bridge inspection system can inspect the bottom structure of a bridge remotely by controlling a specially designed truck with a huge and flexible mechanical arm and an intelligent inspection robot system with various sensors as shown in Fig. 1. The shape of the intelligent inspection robot mounted on the terminal of the multi-linkage arm looks like a cube box with holes in the surface. In total, there are two high-resolution industrial CCD cameras, two 3D cameras, one near infrared line laser and one ultrasonic range finder mounted on the top side of the "cube box", as shown in Fig. 2.

In this paper, a novel approach also is introduced to stitch the images of the concrete bottom surfaces of a bridge using the designed robot system. In the practical application, at least two CCD cameras are used to capture the images of the bottom surfaces of a bridge. With the
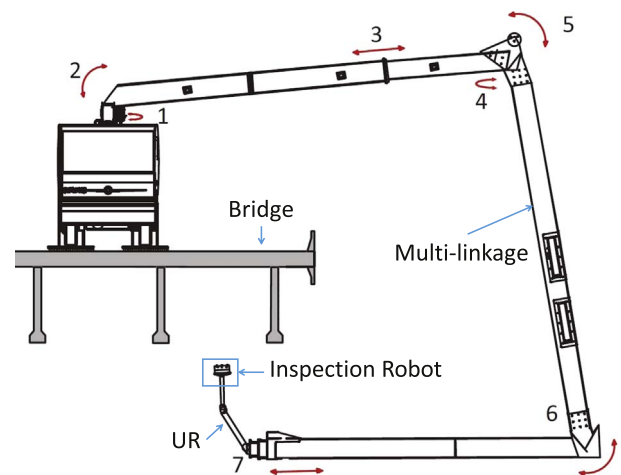

Fig. 1. An overview of the whole mechanism for the designed bridge inspection system. The arcs with double arrows refer to the rotation operation and the lines with double arrows refer to the elongation and shortening of each joint of the multi-linkage arm.

slow moving trunk, each CCD camera collects a large set of images at the low frequency of 6 Hz as shown in Section 4, and each 3D camera acquires 3D cloud points with each profile of 2048 points at a high frequency of 339 Hz. The operation and data acquisition procedures of the bridge inspection system can be summarized as follows. The operator first pulls the trunk to the side of a bridge and then the multi-linkage arm system carries the intelligent inspection robot to the bottom of the bridge. Next, the laser range finder (LRF) mounted on the inspection robot starts to scan the bridge via the rotation device. The LRF scanning data are then used to estimate the geometric information (e.g., the width, height, and length of the various object elements) of the bridge bottom structure via point cloud segmentation [44]. This bridge information then is utilized in the acquisition task planning for the truck and the intelligent inspection robot. Next, the images together with metadata (i.e., the poses of the cameras estimated by the LRF and high-precision incremental encoder (HPIE)) of the bridge bottom surfaces are collected by the sensors mounted on the intelligent inspection robot close to the surfaces with the slow moving of the trunk at a low speed of around 1 m per second. Fig. 3 shows the processing flow of the proposed image stitching approach, which is as follows. First, based on the rough pose information of each image obtained from both the LRF and the HPIE, the acquired images are divided into several strips. Second, images in a single strip are divided into multiple groups. The poses of each image in the same group are first optimized using point features extracted from 2D images and line features extracted from 3D point clouds, which then are used to project the image to the 3D bridge model, which is represented by a set of 3D planes extracted from the LRF scanning data. Next, a homographic refinement procedure is applied to increase the global consistency between different groups of images in a single strip. Third, the point features are used to align the multiple strips for generating a composite wide-view image covering the whole bottom surface of a bridge. Finally, a multi-band blending method [45] is applied with an open-source software called Enblend [1] to make the single-view composite image as seamlessly as possible, which can greatly eliminate both the luminance differences and the color deviations between images and further conceal image parallax.

3. Multi-image alignment and stitching

Image alignment and stitching, which mosaics a number of geometrically aligned images into a single-view composite image, is becoming increasingly popular in computer vision. It is widely used in

---

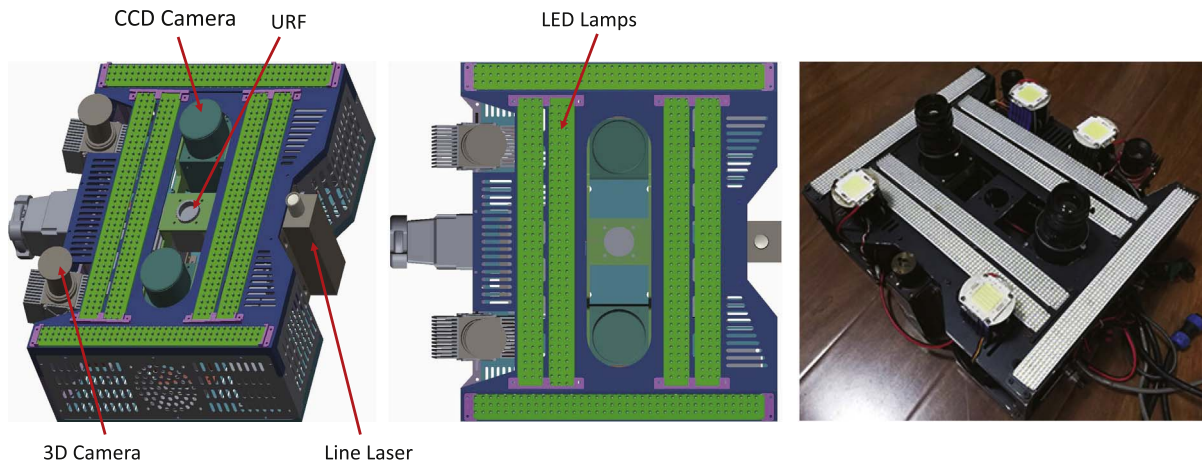[1] Available at http://enblend.sourceforge.net/.

Fig. 2. The structure diagram and real photo of the intelligent inspection robot mounted on the terminal of the multi-linkage arm.

daily life and can display wide-view or even 360° scenes vividly. For example, it can be used in virtual travel on the Internet, building virtual environments in games, and processing personal photos.

The current image alignment and stitching algorithms can be divided into two categories: direct stitching and feature-based stitching [32,36]. Direct stitching benefits from using all of the available image data and can achieve a highly accurate alignment, but it requires close initialization or accurate geometrical calibration. Feature-based stitching is faster and more robust, and has the ability to automatically detect overlapping relationships among a set of images. However, the alignment and stitching result will not be ideal or even will be failed when some of images are poorly textured. In this paper, a reliable alignment method for image stitching is proposed, which not only takes advantage of the above methods but also exploits the structure line information of the bridge. The experimental results in this paper demonstrate that the proposed method is suitable for the bridge inspection application.

With the slow moving of the trunk and the intelligent inspection robot, the CCD cameras acquire high-resolution images and the 3D cameras capture the dense 3D point clouds. The LRF mounted on a rotation platform is first used to obtain the 3D model of the bridge's bottom surface, as shown in Section 3.1.1, and is then combined with a HPIE to estimate the rough poses of the intelligent inspection robot at a high frequency of 40 Hz. With the acquired 3D bridge model, the
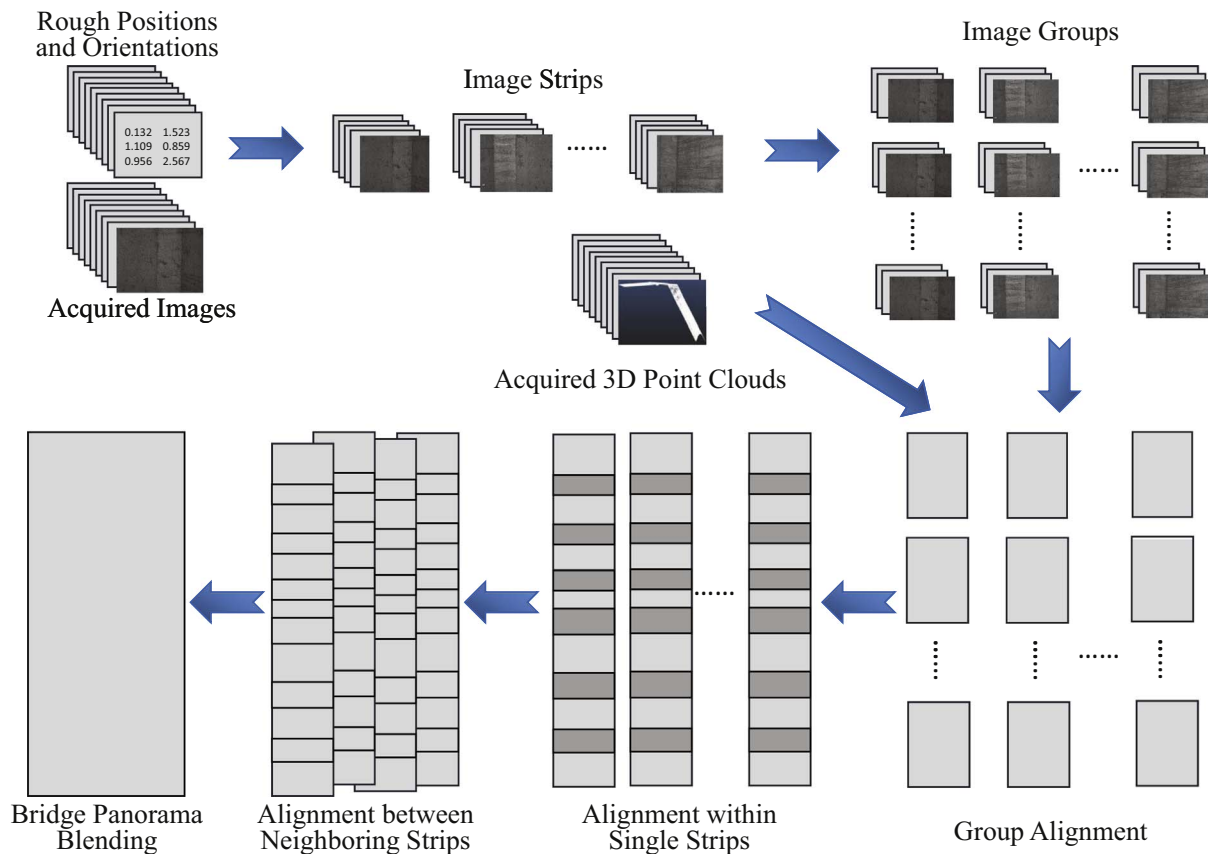

Fig. 3. A schematic overview of the proposed multi-image stitching algorithm for the designed bridge inspection system.
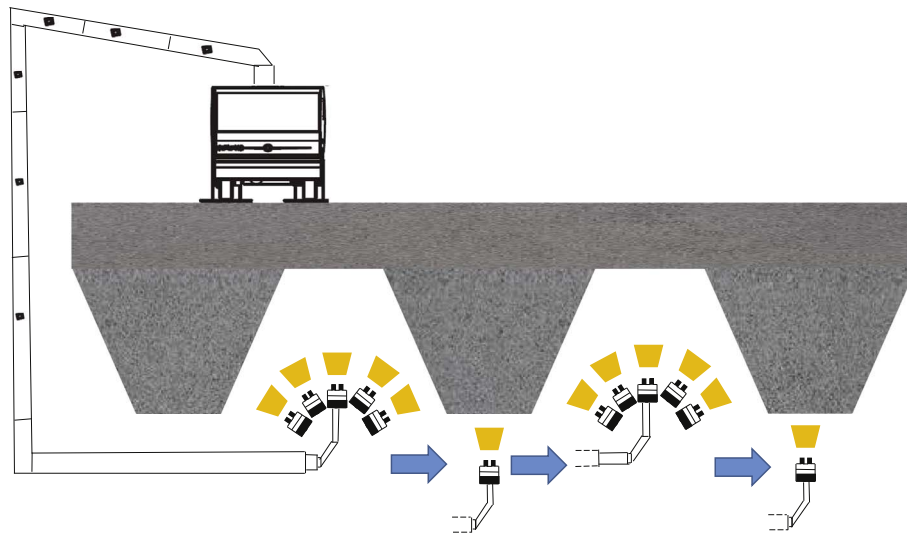
Fig. 4. An illustration of the data acquisition task planning of the designed bridge inspection system for a small box girder bridge.

acquisition task planning can be done, which consists of procedures including the path of the trunk, the elongation and shortening planning of the multi-linkage arm, and the orientation planning for the intelligent inspection robot. The acquisition process then begins as follows. First, the trunk moves to the starting point of the moving path and the multi-linkage arm dispatches the intelligent inspection robot to a planned position beneath the bridge. Second, the trunk slowly moves to the end point of the moving path, which is accompanied by the acquisition of 2D images, the 3D point clouds and the corresponding pose data. Third, the robot along with the multi-linkage arm prepare the intelligent inspection robot for the pose for the second strip, and then the trunk moves back to the starting point with the data of the second strip. The system repeats the above two steps with the back-and-forth movement of the trunk to acquire data which cover the whole bottom surface of a bridge to be inspected. In this way, the system collects all the needed data for inspecting a bridge in several minutes. Fig. 4 illustrates the elongation and shortening planning of the multi-linkage arm and the orientation planning of the intelligent inspection robot for a small box girder bridge.

In order to generate a single mosaicked image for the whole bottom surface of a bridge from a series of images acquired by high-resolution industrial CCD cameras, the images must be organized effectively before beginning. Based on the data acquisition task planning and the estimated rough poses, the images are divided into several strips. Next, the proposed image stitching algorithm for this specific application is conducted in three stages. First, a novel method is proposed to implement the geometrical alignment of the images in a single strip, which is described in Section 3.1. Second, multiple strips with geometrical dislocations are further aligned, which is discussed in Section 3.2. Finally, in Section 3.3, a multi-band blending algorithm [45] is applied to generate a mosaicked panorama as seamless as possible by blending all the warped images in multiple strips with Enblend, which can greatly eliminate both the luminance differences and color deviations between images and further conceal the image parallax.

3.1. Image alignment within a single strip

At this stage, how to align all the images within a single strip is the key problem to investigate. The images in a single strip first are sequentially divided into several groups with each group consisting of a moderate number of images. Then, the point features and the bridge structure lines are extracted from the 2D images and dense 3D point clouds, respectively. Next, the matching relationships between images
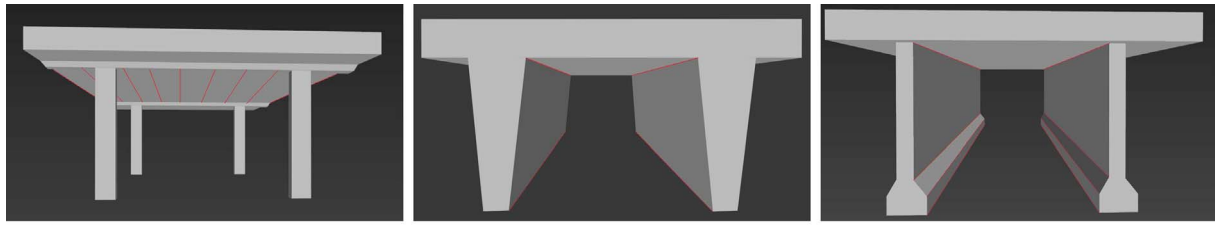
are automatically discovered by using the point features, after which the bundle adjustment is utilized to optimize the poses of images in a single group by using the corresponding point and line features. Then, using the optimized parameters, images in the same group are projected to the 3D bridge model and merged as a group panorama. Finally, the homography computation refinement among sequential groups is applied to increase the global consistency of whole strip of images.

3.1.1. Feature detection

To efficiently stitch images into a strip panorama and stitch multiple neighboring strip panoramas into a single composite image completely covering the bottom surface of a bridge, the scale invariant point features and the image corners extracted from 2D images and the bridge structure lines extracted from 3D point clouds are utilized for image alignment.

Both SIFT and SURF [46] are the most popular choices for scale invariant point feature detection. The SIFT features are identified by finding local extrema in a scale-space representation of an image while the SURF detector uses a Hessian matrix-based measure for detecting features. While the SIFT descriptor is not as efficient and not suited for illumination changes, it is invariant to rotation, scale changes, and affine transformations. The SURF descriptor takes less time than SIFT and provides good performance similar to SIFT, but it is not stable enough to rotation and illumination changes. Considering their advantages and disadvantages, in the proposed system in this paper, SIFT is mainly used to find the rough homographies between images within a single strip while SURF is mainly used to quickly find corresponding images between neighboring strips. However, considering the sparse texture on the bottom surfaces of concrete bridges, using only these feature points is far from enough. Thus, given the rough homographies between images, we also employed the Shi-Tomasi corner detector [47] to obtain more dense matches.

Within a single strip, there are thousands of images captured by each camera. If only point feature matching is applied for each pair of images sequentially, drifting errors will eventually appear due to the fact that matching errors between two images accumulate gradually. In order to solve this problem, more image features are required. Although the Shi-Tomasi corner detector [47] can easily detect adequate corners, a very non-even distribution may occur due to poor texture, which makes the optimization unstable. Fortunately, an interesting phenomenon was found when observing the bottom surface of concrete bridges carefully. No matter what type of bridge, there are usually some bevel structures on the bridge bottom, and the profile appears to be straight

(a) The sketch maps of three different types of typical bridges

(b) The real images of three different types of typical bridges

Fig. 5. The structures of the three main types of typical bridges: hollow-core slabs (Left), small box girder (Middle) and T-beam girder (Right). The red lines in (a) stand for the bridge structure lines used for geometrical alignment.

along the intersection line between the bevel structures. These intersection lines are regarded as the bridge structure lines, which are obviously seen on the bridges as shown in Fig. 5. This prior knowledge can be effectively utilized to enhance the optimization by matching the structure lines in 2D images to those of the 3D bridge model.

The bridge structure lines, which are extracted from the dense 3D cloud points acquired by 3D cameras, along with the matched points between images, are utilized to perform an alignment optimization in each image group. In practice, there are many texture differences between the images of adjacent bottom surface patches; thus, the 3D structure lines extracted from the 3D bridge model are projected on the image first based on its rough pose to get the initial 2D structure lines, and then a refinement procedure based on the image gradient information is conducted as shown in Fig. 6. Due to the stains, spots, and cracks commonly seen on the bridge bottom surface, however, the 2D structure lines sometimes can't be extracted robustly and accurately relying only on image gradients. Fortunately, the 3D structure lines can be robustly extracted from the dense 3D cloud points acquired by 3D cameras, so the 3D structure lines from the 3D point clouds were extracted first and then projected them onto the corresponding 2D images based on their relative calibration results to obtain the initial 2D structure lines, which are shown in Fig. 7.

3.1.2. Two-image matching

At this stage, for each image in one group, the point features extracted from the current view image were matched with the point features extracted from others. Because the images were captured with the slow moving trunk, the images in the same group were well organized. Thus, in order to reduce the computational cost, each image in the group was only matched with several neighboring images in the sequence. In this way, the feature points can be matched more efficiently and robustly from these tightly coupled images. Within a single strip, at least two industrial CCD cameras acquired images sequentially with the slow moving trunk. Thus, there were overlapped regions between these two CCD cameras due to a rigid fixture and a roughly fixed imaging distance. The acquired images were already well organized and were matched sequentially based on the acquisition sequence in each group of images.

To make the practical engineering application more feasible, a fast and efficient two-image matching method is proposed in this paper, as shown in Fig. 8. First, the three-level pyramid image for each image was created. The first level was the original image, the second level was the down-sampled image with a scale factor of 1/4, and the third level was the down-sampled image with a scale factor of 1/16. Next, the third level image was divided into several blocks for detecting SIFT features and matching corresponding images based on their descriptor vectors.

Fig. 6. An example of the obvious structure line in an image: (Left) the initial structure line generated by projection with the rough pose of the image; (Middle) the image gradient map; (Right) the refined structure line optimized by image gradients. The red lines represent the bridge structure lines.
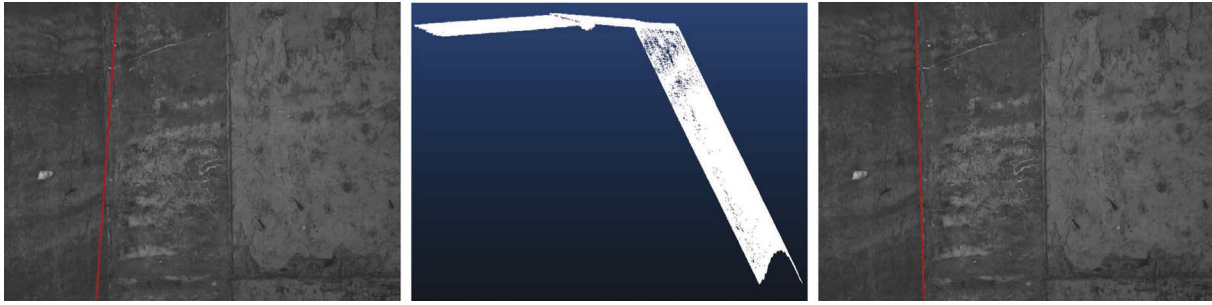
Fig. 7. An example of the obscure structure line in an image, but obvious enough in the corresponding 3D point clouds collected by 3D cameras: (Left) the initial structure line generated by projection with the rough pose of the image; (Middle) the 3D point clouds; (Right) the structure line projected from the extracted 3D structure line from 3D point clouds based on the relative calibration. The red lines stand for the bridge structure lines.
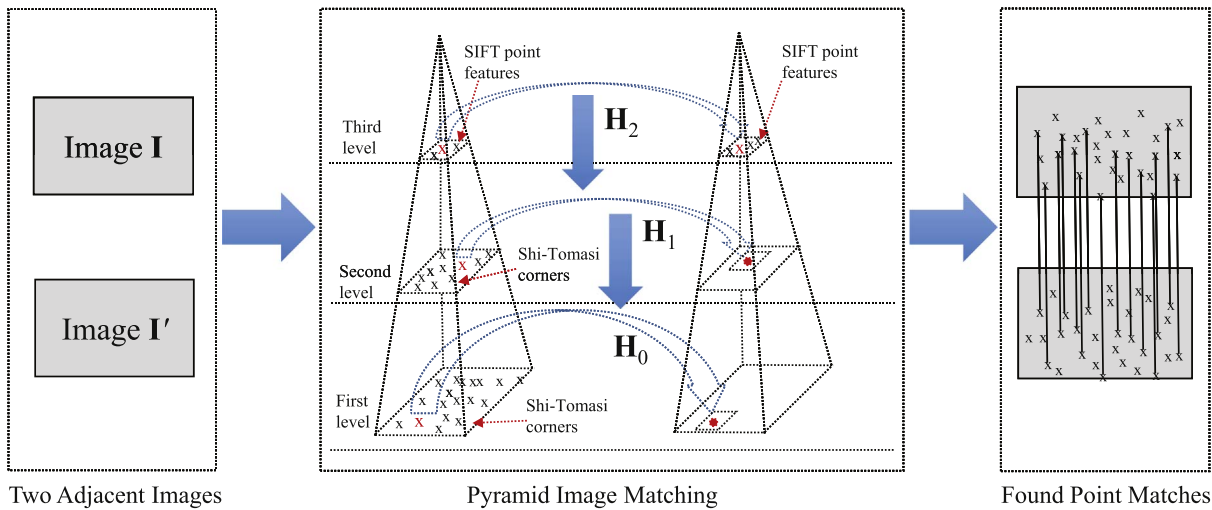

Fig. 8. The proposed three-level pyramid image matching strategy for efficiently matching two adjacent images.

A point feature $\mathbf{p}'$ in $\mathbf{I}'$ was regarded as the candidate correspondence of the point feature $\mathbf{p}$ in $\mathbf{I}$ if all the following three rules were satisfied [48]: (1) the Euclidean distance $d(\mathbf{p},\mathbf{p}')$ between the descriptor vectors of $\mathbf{p}$ and $\mathbf{p}'$ is less than some predefined threshold; (2) $\mathbf{p}'$ is the best one in $\mathbf{I}'$ matched to $\mathbf{p}$ in $\mathbf{I}$ (i.e., $d(\mathbf{p}, \mathbf{p}') = \min_{\mathbf{q}'_k \in \mathbf{I}'} d(\mathbf{p}, \mathbf{q}'_k)$); (3) the distance ratio between the first closest distance $d(\mathbf{p},\mathbf{p}')$ and the second closest is less than some predefined threshold (0.8 was used in this paper). With a set of candidate corresponding image points, the homography matrix $\mathbf{H}_2$ between two third level images can be easily recovered by using RANSAC [49]. On the second level of the pyramid, the Shi-Tomasi corner detector [47] is applied to extract sufficient quantity of corners from an image. Then, the homography matrix $\mathbf{H}_2$ is used to approximately locate the candidate regions in another image of a Shi-Tomasi corner $\mathbf{p}$ and then the candidate corresponding point can be found based on the matching evaluation criteria formulated as:

$$\rho(\mathbf{p}, \mathbf{p}') = \frac{\left| \sum_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} \sum_{\mathbf{q}'\in\mathcal{N}(\mathbf{p}')} (\mathbf{I}(\mathbf{q}) - \bar{\mathbf{I}}(\mathbf{p}))(\mathbf{I}'(\mathbf{q}') - \bar{\mathbf{I}}'(\mathbf{p}')) \right|}{\sqrt{\sum_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} (\mathbf{I}(\mathbf{q}) - \bar{\mathbf{I}}(\mathbf{q}))^2 \cdot \sum_{\mathbf{q}'\in\mathcal{N}(\mathbf{p}')} (\mathbf{I}'(\mathbf{q}') - \bar{\mathbf{I}}'(\mathbf{p}'))^2}}, \quad (1)$$

where $\bar{\mathbf{I}}(\mathbf{p})$ and $\bar{\mathbf{I}}'(\mathbf{p}')$ denote the average intensities of neighboring patches $\mathcal{N}(\mathbf{p})$ and $\mathcal{N}(\mathbf{p}')$ with the size of $m \times n$, i.e., $\bar{\mathbf{I}}(\mathbf{p}) = \frac{1}{m\times n} \sum_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} \mathbf{I}(\mathbf{q})$, $\bar{\mathbf{I}}'(\mathbf{p}') = \frac{1}{m\times n} \sum_{\mathbf{q}'\in\mathcal{N}(\mathbf{p}')} \mathbf{I}'(\mathbf{q}')$, respectively. When the value of $\rho(\mathbf{p},\mathbf{p}')$ is close to 1, the two feature points $\mathbf{p}$ and $\mathbf{p}'$ are most likely to be corresponding image points. A fixed threshold τ is set for ρ to avoid being trapped into a non-optimal "local maximum" (e.g., $\tau = 0.8$ in this paper). Then, with the newly found candidate matches in the second level, a better homography matrix $\mathbf{H}_1$ can be

recovered by using RANSAC. Finally, the above procedure in the first level of the pyramid is applied again to obtain the image point matches quickly and precisely.

3.1.3. Group alignment bundle adjustment

Bundle adjustment (BA) generally involves adjusting the extrinsic and intrinsic parameters of a set of cameras by minimizing the reprojection error. Given a set of 2D image point matches and 3D-to-2D bridge structure line matches, the extrinsic parameters of all the involved images can be optimally and jointly adjusted. In the designed bridge inspection system, the intrinsic parameters of the industrial CCD cameras used were precisely calibrated in advance and do not need to be optimally adjusted in bundle adjustment. Two industrial CCD cameras are used in the designed inspection robot system. In each image strip, each camera can acquire thousands of images. As mentioned, a single image strip is divided into several groups. For each group of images, the feature and corner detection and two-image pyramid matching are performed sequentially. Finally, the bridge 3D model, the projected structure lines and the found point matches are utilized to jointly optimize the extrinsic parameters (i.e., poses) of all the images in each group with the rough poses as their initialization.

Given a 2D image point $\mathbf{x} = [x,y,1]^\top$ and its corresponding 3D point $\mathbf{X} = [X,Y,Z,1]^\top$ on the object surface, their relationship can be expressed by a projective transformations as follows:

$$s\mathbf{x} = \mathbf{K}[\mathbf{R} \quad \mathbf{t}]\mathbf{X}, \quad (2)$$

where s is an arbitrary scale factor, $\mathbf{K}$ is a 3×3 intrinsic matrix, $(\mathbf{R},\mathbf{t})$ is the extrinsic parameters that comprised the 3×3 rotation matrix $\mathbf{R}$ and
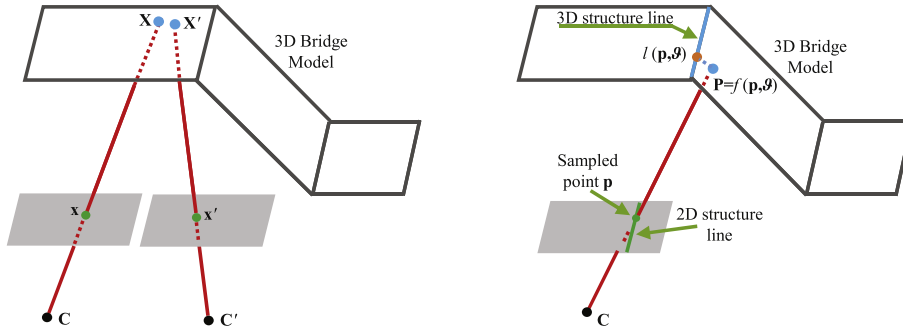
Fig. 9. Back-projection diagrams of image feature points (Left) and structure line (Right). The first optimization goal is to minimize the differences between the corresponding points **X** and **X′** for the former, and the second goal is to minimize the distance between the blue 3D structure line and the corresponding point **P** of the sampled point **p** located on the green 2D structure line in the image for the later.

the 3 × 1 translation vector **t**, which relate the world coordinate system (i.e., bridge 3D model) to the camera system. The intrinsic matrix **K** can be easily calibrated by Zhang's flexible camera calibration method [50]. The extrinsic parameters (i.e., (**R**,**t**)) are derived from the rough pose estimated by combining the LRF and HPIE sensor data.

With the intrinsic and extrinsic parameters, a 2D image point **x** can be back-projected to the 3D bridge bottom surface model $\mathscr{B}$ at its corresponding 3D point **X** based on the principle that the camera's optical center **C**, **x** and **X** are collinear in 3D space as shown in Fig. 9. As mentioned above, the 3D model $\mathscr{B}$ of the local bridge bottom was obtained by the LRF sensor in advance. All of the camera poses and the 3D bridge surface model are unified into the same world coordinate system, and then the intersection between the 3D back-projection line and the bridge bottom surface can be obtained along the vector derived from (**x**,**C**), which is represented as follows:

$$X = f(x, \mathscr{B}, K, \vartheta), \qquad x \in I, \tag{3}$$

where $f(\cdot)$ denotes the back-projection function with a 2D image point **x** in an image **I**, the bridge bottom surface model $\mathscr{B}$, and the intrinsic matrix **K** and the rough extrinsic parameters represented

$\vartheta = \{t_x, t_y, t_z, \theta_x, \theta_y, \theta_z\}$ consisting of three translations and three rotation angles. Due to that $\mathscr{B}$ and **K** are the same for all the images, the function $f(x, \mathscr{B}, \mathscr{K}, \vartheta)$ will be expressed by $f(x, \vartheta)$ for simplicity here. If ϑ is accurate, the perfect panorama can be easily generated by the back-projection method described in Eq. (3) with some image blending algorithm. We propose a perception criterion function to estimate the accuracy of the extrinsic camera parameters of K images (i.e., $\Theta = \{\vartheta_i\}_{i=1}^K$) as follows:

$$E_p(\Theta) = \sum_{i=1}^K \sum_{x \in I_i} \sum_{j=1}^K \sum_{x' \in I_j} \|f(x, \vartheta_i) - f(x', \vartheta_j)\|^2, \quad x \leftrightarrow x', \tag{4}$$

where $x \Leftrightarrow x'$ denotes a point match from two overlapping images as shown in Fig. 9 and described in Section 3.1.2.

The 2D structure lines projected from 3D structure lines are integrated into the whole optimization framework as follows. First, the equidistant points are extracted from each structure line in an image, which is projected onto the 3D bridge model. Each sampled point **p** in the structure lines $\mathscr{L}_i$ extracted from the image I_i has a corresponding 3D point $P = f(p, \vartheta_i)$ in the bridge surface model. If an error exists in the extrinsic parameters ϑ_i, the back-projected point on the bridge model $\mathscr{B}$

Fig. 10. The design drawing and real photos of the bridge inspection robot system: the design drawing (Left), a real photo in the entire trunk (Middle) and a real photo in the cab control space (Right) in the top row, and six photos captured during working for data acquisition of a test bridge in the middle and bottom rows.
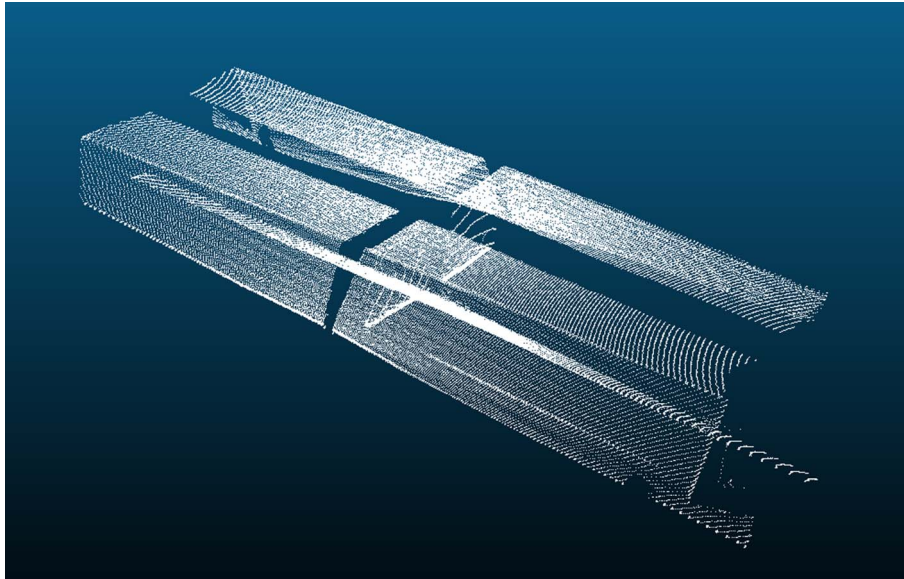
Fig. 11. The 3D point clouds of the test bridge acquired by the LRF sensor via a rotation device.

does not fall on the 3D structure lines of $\mathscr{B}$ accurately. So these deviations are utilized to evaluate the accuracy of the extrinsic parameters of K images as follows:

$$E_l(\Theta) = \sum_{i=1}^{n} \sum_{\mathbf{p} \in \mathscr{L}_i} \|f(\mathbf{p}, \vartheta_i) - l(\mathbf{p}, \vartheta_i)\|^2, \tag{5}$$

where $\mathbf{p}$ denotes the equidistantly sampled point on one of the extracted structure lines $\mathscr{L}_i$ from $\mathbf{I}$, and $l(\mathbf{p}, \vartheta_i)$ denotes the 3D corresponding point of the sampled $\mathbf{p}$, which is obtained by perpendicularly projecting $\mathbf{P}$ to the closest 3D structure line in the 3D model $\mathscr{B}$ as shown in Fig. 9.

The whole optimal solution:

$$\widehat{\Theta} = \arg \min_{\Theta}(E_p(\Theta) + E_l(\Theta)), \tag{6}$$

can be achieved by jointly minimizing the error functions in both Eq. (4) and Eq. (5) using the Levenberg-Marquardt (L-M) algorithm [51].

Within a single strip, each CCD camera can acquire a large set of

bridge bottom surface images with the slow moving trunk, as shown in Fig. 13. Because two CCD cameras are rigidly fixed and their relative pose can be precisely calibrated in advance, two image sets acquired from two CCD cameras in the same group have the same relationship. Thus, the poses of these two image sets can be jointly optimized under the constraint of their fixed relative pose.

After jointly optimizing the poses of all the images in a group, these images are directly projected onto the bridge plane model using the newly optimized poses. Thus, several group panoramas within a single strip can be achieved independently. Fig. 14 shows the simple mosaicked images with the rough and optimized poses by a simple superposition, from which the optimization greatly reduces the image parallaxes.

3.1.4. Strip alignment

To make the strip alignment algorithm more efficient, the images within a single strip were sequentially split into multiple groups and aligned in each group with the techniques described in Section 3.1.3. In
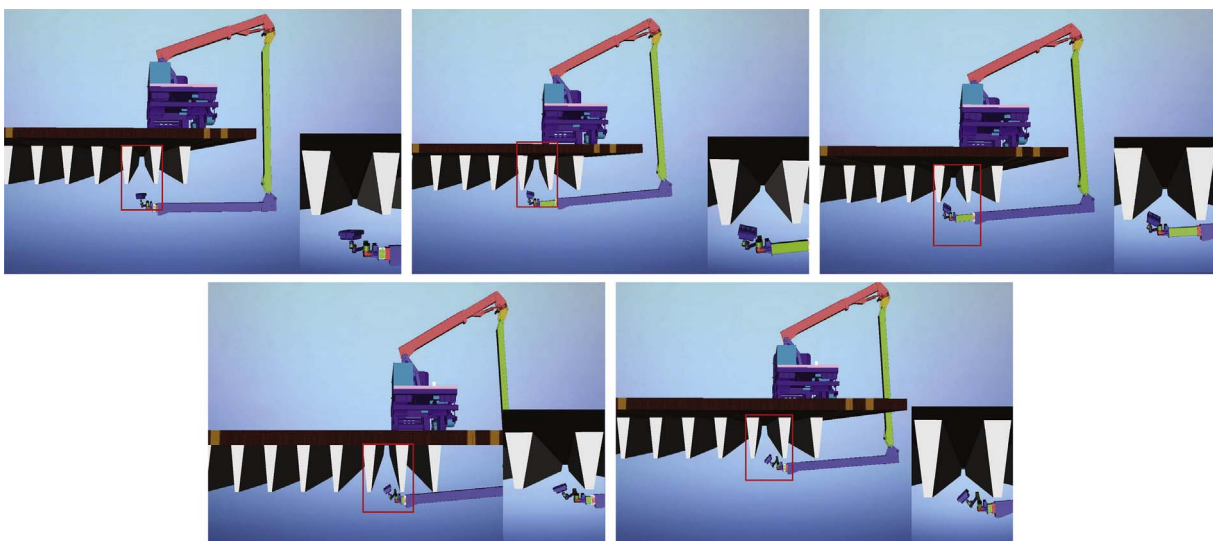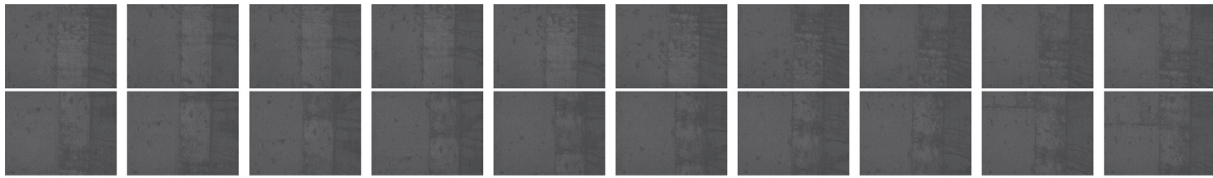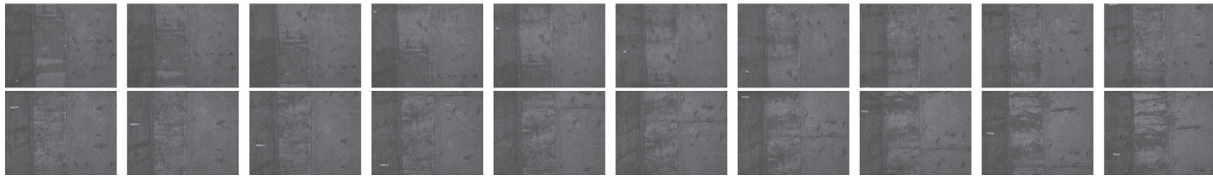

Fig. 12. An illustration of pose planning of the intelligent inspection robot to acquire different strips of data at each lane for the test small box girder bridge.

(a) Partial images captured by the first CCD camera in some strip.

(b) Partial images captured by the second CCD camera in some strip.

Fig. 13. Partial images captured by two CCD cameras in some strip consisting of 490 images.

this section, a simple method is utilized to align the images between the different groups. To easily implement the strip alignment, the group splitting strategy ensures that the first image of the current group is the same as the last one of the previous group. After projecting the last-and-first pair of images onto the bridge bottom surface planes, the homography matrix between them is estimated based on the same set of image point features. Finally, the homography matrices between neighboring groups can be utilized to align all the images within the strip. Fig. 15 (a) shows an example of the simple mosaicked strip panorama by a simple superposition.

3.2. Image alignment between neighboring strips

After independently aligning all the strips, if all these strips of images are directly merged together based on their own optimized image poses, it would result in the geometrical dislocations between neighboring strips along the moving direction of the trunk shown in Fig. 17 (a). There are two key causes for these dislocations: (1) the HPIE cannot provide accurate motion distance due to the synchronization error, and the wheel-spin and the out-of-straight moving of the trunk; (2) the accumulated drifting errors of different strips are not consistent. In this section, an efficient and simple strategy is proposed to remove these dislocations by aligning all the strips in a unified coordinate system. First, the strip whose corresponding plane parallel to the bridge road surface is selected as the reference. Then, the other strips are sequentially aligned to the reference strip based on the neighborhood relationship.

Given a strip $\mathscr{S}_A$ as the reference, the separated groups of its neighboring strip $\mathscr{S}_B$ close to $\mathscr{S}_A$ are sequentially aligned as follows. Let $\mathscr{G}$ be an image group in the strip $\mathscr{S}_B$ used in the single strip alignment. The first, the middle, and the last images of $\mathscr{G}$ are first selected. Second, the SURF detector is utilized to extract the image point features from each selected image $\mathbf{I}_B^s$ followed by searching its mostly overlapped image $\mathbf{I}_A^o$ in some corresponding image neighborhood in $\mathscr{S}_A$ based on the number of found point matches using the matching strategy described in Section 3.1.2. Then, the proposed three-level pyramid two-matching strategy is applied to find the point matches between $\mathbf{I}_B^s$ and $\mathbf{I}_A^o$ based on the largest overlapped bridge bottom plane determined by the back-projection to the 3D bridge model using their previously optimized image poses, which are used to recover the homography matrix between $\mathbf{I}_B^s$ and $\mathbf{I}_A^o$. In this way, three homography matrices for three selected images in each group of $\mathscr{S}_A$ can be recovered, after which the remaining images are locally grouped to one of the three selected ones. Therefore, each group of images in $\mathscr{S}_B$ can be easily aligned to $\mathscr{S}_A$ on the bridge bottom surface. In this way, the neighboring strips can be iteratively aligned to the selected reference strip. An example of aligning two strips is shown in Fig. 17 (b), which demonstrates that the geometrical dislocations between strips was eliminated.

3.3. Image warping and blending

After geometrically aligning all the strips of images, an efficient stitching scheme is needed to seamlessly generate a single composite panorama of the whole bridge bottom surface for visual inspection and/
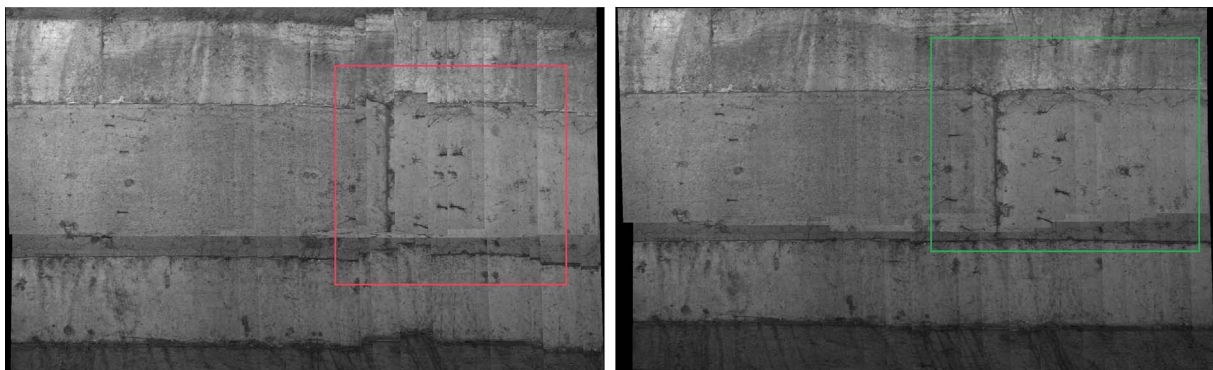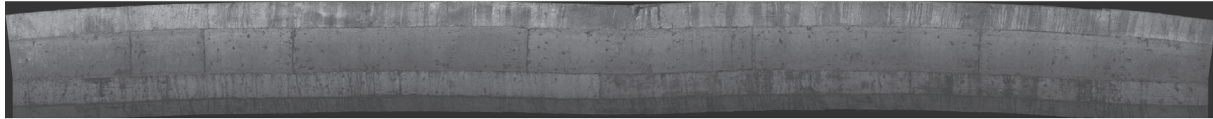

Fig. 14. The comparison between the mosaicked panoramas by a simple superposition from a group of images in some strip based on the rough poses (Left) and optimized ones (Right). The obvious improvement (i.e., the removal of geometrical dislocation) is obviously seen in the rectangular region in the right corresponding to the rectangular one in the left.

(a) The simply mosaicked strip panorama by a simple superposition with the optimized image

poses Θ (i.e., poses).

(b) The mosaicked strip panorama with PTGui.

(c) The mosaicked strip panorama with the open source software Enblend based on the

optimized image poses Θ (i.e., poses).

Fig. 15. The mosaicked strip panoramas under different conditions.

or automatic bridge defect detection (e.g., crack detection) on the whole bridge to comprehensively evaluate the health status of the bridge. Due to the different lighting conditions during the image acquisition and small geometrical dislocations in alignment, the directly mosaicked image by a simple superposition from the projected images on the 3D bridge plane based on their optimization parameters, as shown in Fig. 15 (a) and Fig. 16, will cause luminance differences and color deviations on the mosaicked image and also cannot conceal the small image parallax due to geometrical misalignment.

In this paper, a multi-band blending method [45] with an open-source software "Enblend" is applied to overcome these problems and generate a mosaicked image as seamlessly as possible. Enblend is a popular tool for compositing images in a way that the seams between images can be efficiently concealed by applying the multi-band blending on all the aligned images based on the graph-cut-based optimized stitching lines found from neighboring images. Enblend has been widely used to generate good stitching lines for image mosaicking in different areas, such as indoor image mosaicking [52], endomicroscopy video mosaicking [53], and UAV image mosaicking [54], among others.

In Enblend, a mask indicating the position of the corresponding image in the overall panorama is required, which is stored as an image channel. Therefore, the required memory is wasted if an overall mask is provided for each image. A novel method thus is presented here to solve this problem. For a panorama stitched from N input images, the images are first divided into several groups with the sequence information and each group contains $\lfloor \sqrt{N} \rfloor$ images. In each group, a mask with the size of this region is generated for every image. Then, Enblend is applied to generate seamless panoramas in each group. The final overall seamless panorama can be acquired by repeating the above procedures. Fig. 15 (c) illustrates the seamless mosaicking of a strip of images using Enblend which demonstrates that the luminance difference, color deviations, and geometrical dislocations between images were greatly reduced and effectively concealed.

4. Experimental results

In this section, the proposed method was evaluated on real bridge bottom surface images acquired by the proposed bridge inspection system. Fig. 10 shows the design drawing and a set of actual photos of the designed system. The chosen test bridge is the structure named

"small box girder" as shown in Fig. 5. In order to make the acquisition task planning easier and determine the poses of the platform in real time, the 3D planar model of a bridge was first recovered from the acquired 3D point clouds obtained by a LRF called SICK LMS151 via rotation on a platform, as shown in Fig. 11. After extracting 3D planar segments, the bridge bottom surface model was parameterized. Then, the pose (i.e., pose) planning for the intelligent inspection robot mounted on the terminal of the multi-linkage arm was made at five different poses for the test small box girder bridge to acquire five different strips of data in a lane as illustrated in the virtual environment shown in Fig. 12. With the slow back-and-forth movement of the trunk, the system acquired five different strips of data consisting of 2D images from CCD cameras, 3D point clouds from 3D cameras, and the LRF and HPIE sensor data for estimating the rough poses. In the designed system, the position accuracy in the moving direction (e.g., the X-axis) of the trunk achieved was around 30cm, and the LRF sensor acquiring the 3D surface profile cross the moving direction (e.g., the $Y\ OZ$-plane) of the trunk was used to locate the accuracy around 5cm. The orientation accuracy of the platform achieved by the Inertial Measurement Units (IMU) was around 0.5°, which possibly caused the orientation accuracy of the 2D and 3D cameras at the accuracy of $1 - 2$° due to the calibration error and the vibration of the platform itself. The 2D images were acquired by two JAI SP-20000C-CXP2 cameras, and each image was 5120 × 3840 pixels and a fixed frequency of 6 Hz. In total, a set of around 2500 images were acquired for the bridge bottom surface in a lane of the selected small box girder bridge. Partial images within some single strip are shown in Fig. 13. The 3D cloud points of the corresponding area were acquired by two AT C2-2040(HS)-GigE cameras with each profile of 2048 points at a fixed frequency of 339 Hz.

Fig. 14 shows a comparison between the mosaicked panoramas by a simple superposition from a group of images in some strip before and after group alignment bundle adjustment. The proposed bundle adjustment strategy for group alignment described in Section 3.1.3 greatly eliminated the geometrical dislocations. The mosaicked panoramas shown in Fig. 14 were generated from the projected images based on the rough image poses and the optimized image poses on some 3D bridge plane, respectively. Using the proposed strip alignment strategy to aligning multiple sequential groups in a single strip, the whole strip alignment was completed, as shown in Fig. 15 (a), in which the mosaicked strip panorama by a simple superposition was generated from
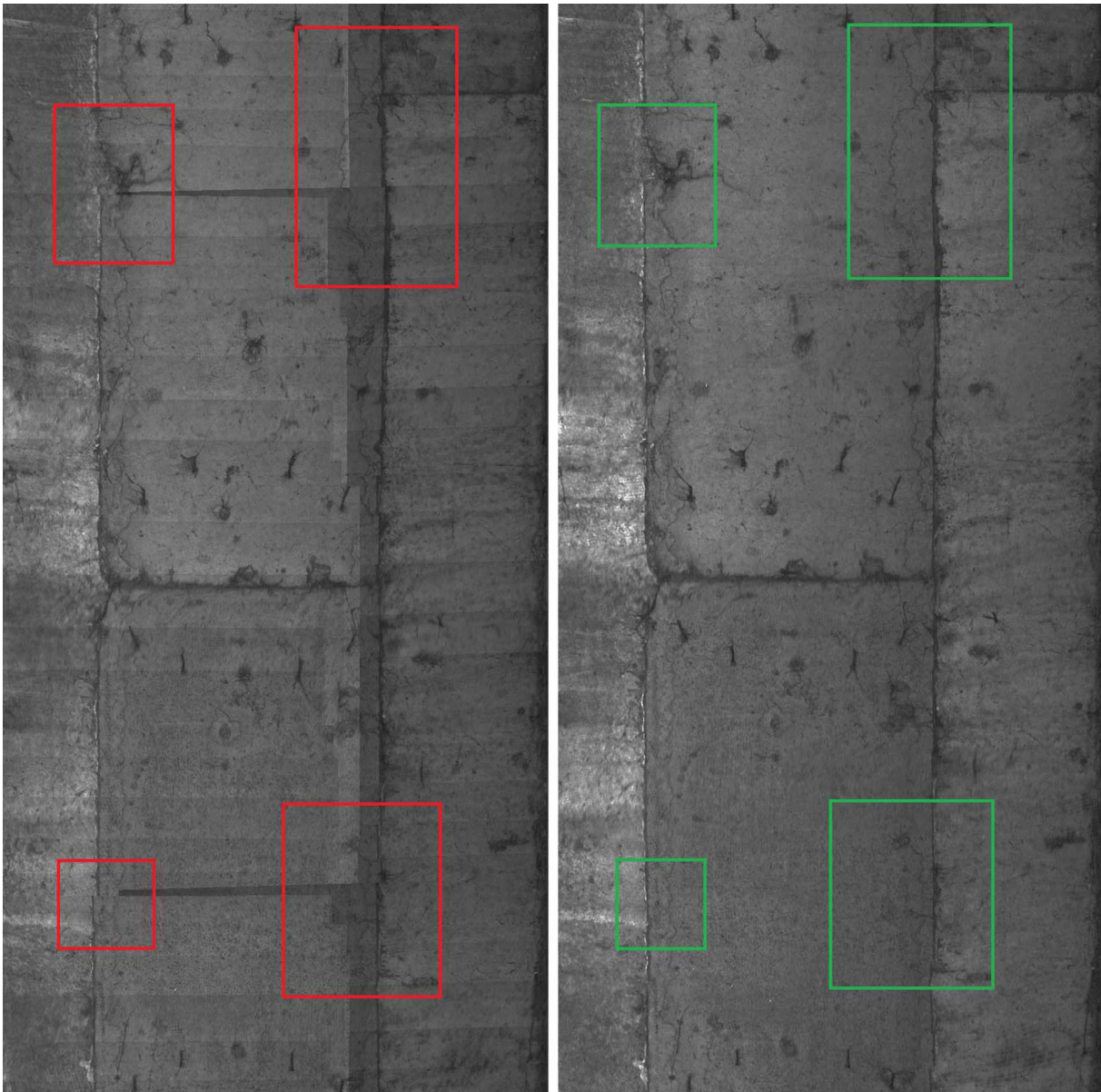
Fig. 16. A comparison of the detailed local views of two mosaicked strip panoramas shown in Fig. 15 (a) in the left and Fig. 15 (c) in the right. The image illumination differences and color deviations were greatly eliminated and the image parallax was further concealed by the utilized blending strategy.

the 490 images shown in Fig. 13 based on the optimized poses Θ by combining the image point features and structure lines in bundle adjustment.

In order to demonstrate the good performance of the proposed panorama stitching method, the commercial software "PTGui" [2] was used for comparison, which stitches the images only via image point features and creates perfect panoramas in different scenes [55,56]. Fig. 15 (b)–(c) shows the overall feature-points-based and structure-lines-aided mosaicked strip panoramas, which were generated by PTGui and Enblend with the global alignment obtained by the proposed method, respectively. Fig. 15 (c) demonstrates the luminance difference and color deviations, and also further concealed the image parallax compared to the simply mosaicked strip panorama by a simple superposition shown in Fig. 15 (a), which was obviously observed in the detailed local views as shown in Fig. 16. Fig. 15 (b) illustrates that

obvious distortions existed in the final mosaicked image obtained by PTGui, which used only the image point features alone for alignment. This distortion was almost eliminated by the proposed strip alignment combining both the image point features and the structure lines, as shown in Fig. 15 (c).

Fig. 17 (a) shows an example of directly mosaicking two neighboring strips by projecting them to a single bridge plane based on their own independent strip alignment results, in which there were obvious geometrical dislocations between strips mainly due to the accumulated drifting errors in both HPIE and matching. After applying the proposed image alignment strategy between strips described in Section 3.2, these geometrical dislocations were almost eliminated, as shown in Fig. 17 (b). This improvement can be more easily observed from the detailed local views as shown in Fig. 17 (c)–(d).

As mentioned, the bottom surface images of the test bridge were captured in five strips along the slow back-and-forth movement of the trunk with five different platform poses as shown in Fig. 12. With the proposed image alignment within a single strip, these images were

---

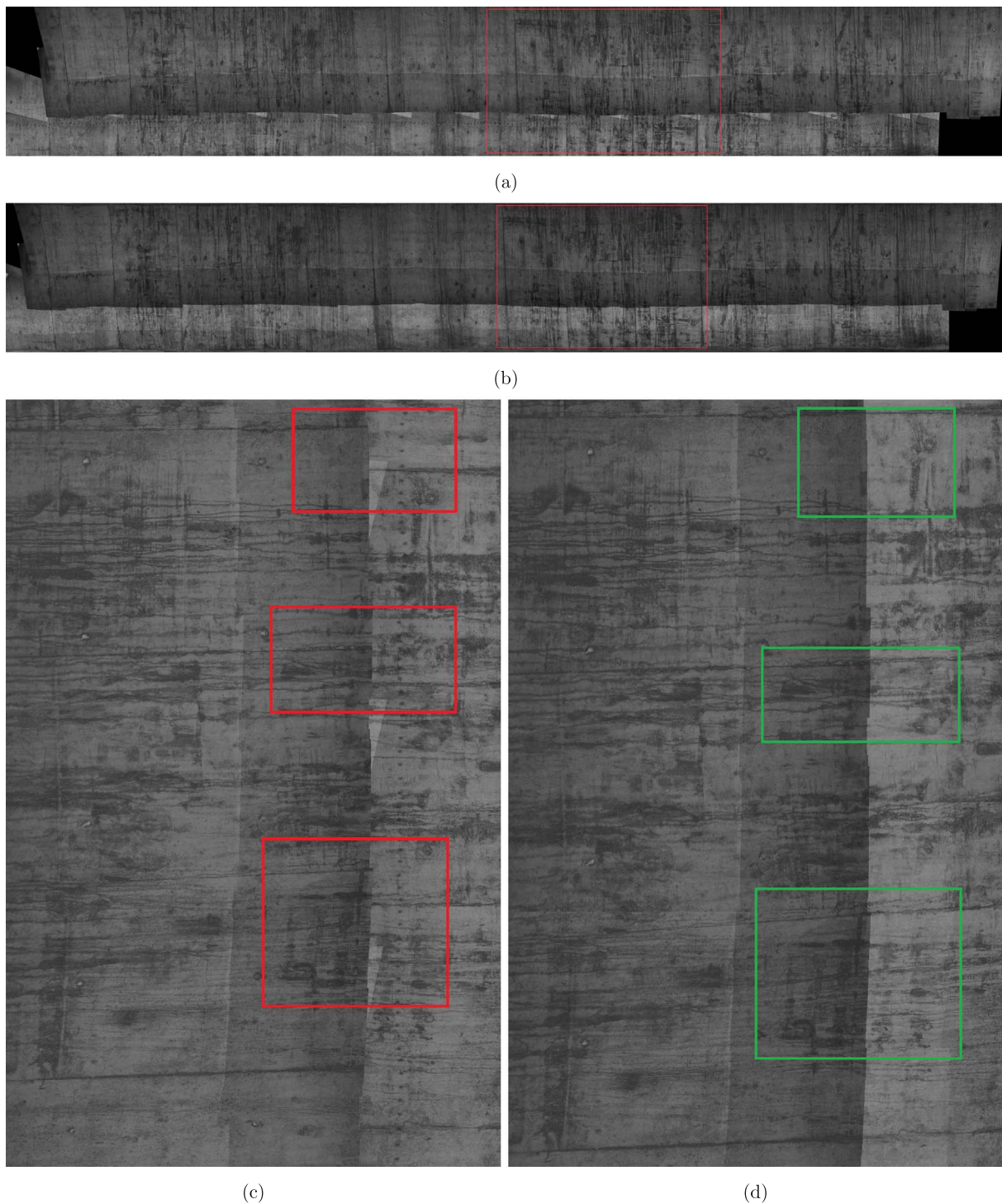[2] PTGui is available at http://www.ptgui.com/

Fig. 17. The visual comparison between the bridge plane panoramas simply mosaicked from two strips by a simple superposition: (a) direct mosaicking before aligning two neighboring strips; (b) mosaicking after aligning two neighboring strips; (c)–(d) the detailed local views corresponding to the red rectangular regions as shown in (a)–(b).

independently aligned in five strips, whose mosaicked strip panoramas are shown in Fig. 18, where the overexposures marked in the red rectangular regions were caused by the bar LED lamps (see Fig. 2) due to the fact that the CCD cameras were too close to the baffles in the small box girder bridge. These five strips covered the three main bottom surface planes of a lane of the test small box girder bridge, that comprised the top, left, and right surface planes as shown in Fig. 5. Next, the neighboring strips were stitched to generate three single panoramas based on their homography matrices, corresponding to the top, left, and

right surface planes as shown in Fig. 19 (a), (b), (c), respectively. Finally, these three surface planes were further stitched into a whole panorama as shown in Fig. 19 (d).

Given the mosaicked wide-view panoramas for the bottom surfaces of the bridge, the health status of the bridge was evaluated by visual inspection and/or automatic defect detection and analysis techniques. Cracks are one of the most important bridge defects for evaluating the health status of a bridge. The crack detection algorithm was applied on the whole mosaicked panorama. Fig. 20 shows the locally cropped

original image and its corresponding crack detection result. Fig. 21 shows the texture mapped results with the mosaicked panorama onto the 3D model of the test bridge. Based on the 3D models and the geographical coordinates of the bridges, all the detected bridge defects were digitally managed even at different periods. The detailed health status database of all the inspected bridges was easily built, which greatly improved the completeness and efficiency of both the health status evaluation and maintenance of bridges.

5. Conclusions and future works

Human visual inspection is the predominant approach for bridge inspection, but it is dangerous, qualitative, subjective, and time-consuming. In addition, the complicated bottom structures of some bridges may introduce difficulties for the human vision procedure. Therefore, using high-resolution industrial CCD cameras to acquire the images by a flexible and mobile platform for inspecting civil infrastructures like bridges can be a convenient, effective, and safer alternative. The proposed robotic bridge inspection system provides the inspector with a tool to comprehensively evaluate the health status of a bridge based on the detected defects (e.g., cracks) on the bridge bottom images acquired by an intelligent inspection robot mounted on the terminal of a multi-linkage arm.

The main goal of the study in this paper was to create a wide-view panorama of a concrete bridge bottom surface as seamlessly as possible from thousands of high-resolution images for bridge inspection purpose, especially crack detection and 3D visualization. The proposed system and approach is based on the assumption that the 3D model of a bridge can be obtained by an intelligent inspection robot and the rough pose information for each image have been obtained by combining the LRF and HPIE sensor data. In order to stitch all the images captured

from different views in different strips with the back-and-forth movement of the trunk. First, all the acquired images must be organized effectively based on the task planning and rough poses by dividing those images into multiple strips. Second, the images in each single strip are separated into multiple sequential groups and those images in each group are geometrically aligned. All the groups in each single strip are further aligned to depict the whole bridge system. Relying on the point features between neighboring strips in the same bridge plane, the whole bridge plane panorama can be achieved. Finally, a multi-band blending algorithm with Enblend is used to eliminate both the image luminance difference and color deviations and further conceal the image parallax.

The experimental results on a set of about 2500 images captured from a real bridge bottom surface demonstrated that the proposed method effectively dealt with the image stitching problem and provided good results for bridge inspection and 3D visualization applications. The major benefit of optimization by combining point features and structure line features is the local and global alignment consistency for image alignment. When compared to PTGui, which stitches the images only via image point features, the proposed method can obtain better stitching results via the combination of point and line features. The proposed method can not only overcome the large drifting error but also improve the optimized alignment results based on the image point features alone. Also, the mosaicked bridge bottom surface images are mapped on the reconstructed 3D bridge model, which can provide a valuable tool to inspector for comprehensive evaluation of the health status of a bridge.

In the future, several components of the proposed method can be improved. First, the combination of image point features and bridge structure line features is very important during the optimization of the poses of images. However, balancing weights between the two types of
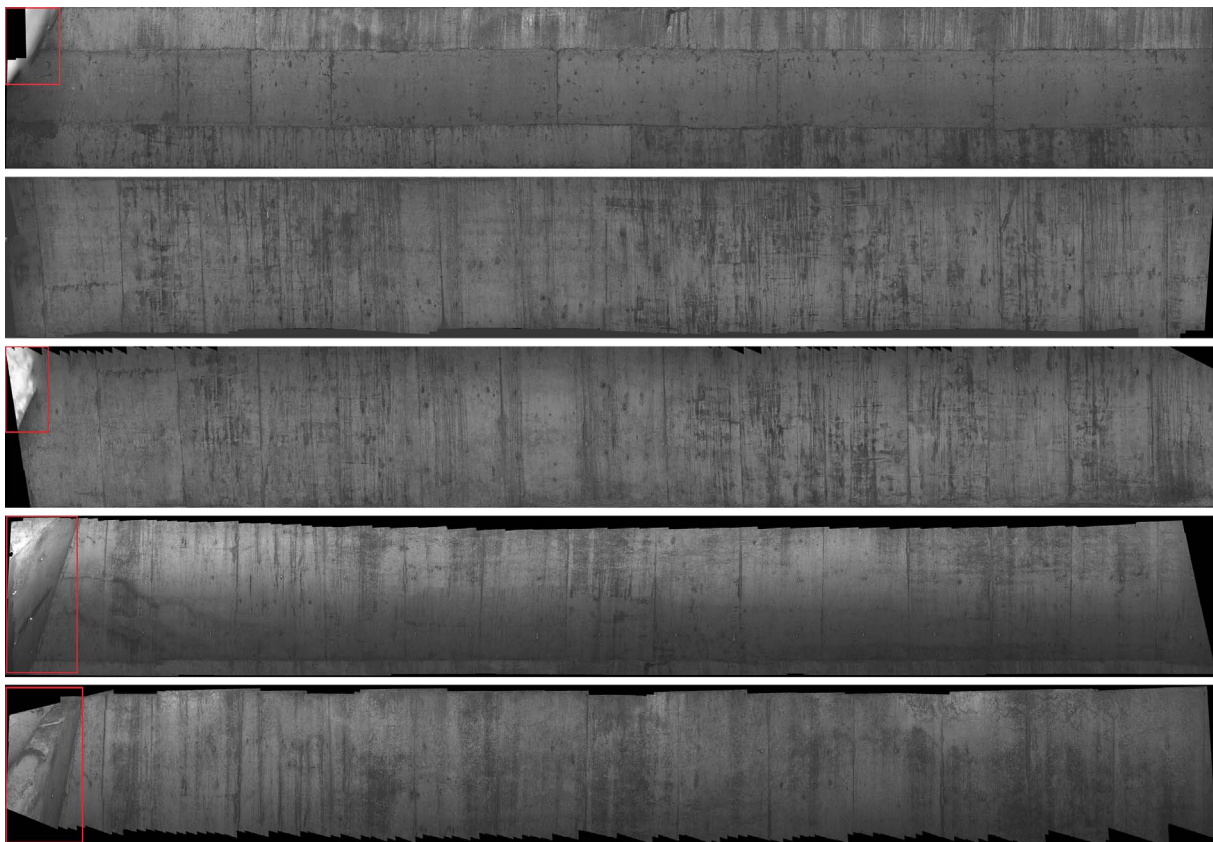

Fig. 18. The mosaicked strip panoramas in five different strips. The overexposure shown in the red rectangular regions were caused by the bar LED lamps (see Fig. 2) due to the fact that the CCD cameras were too close to the baffles existed in the small box girder bridge.
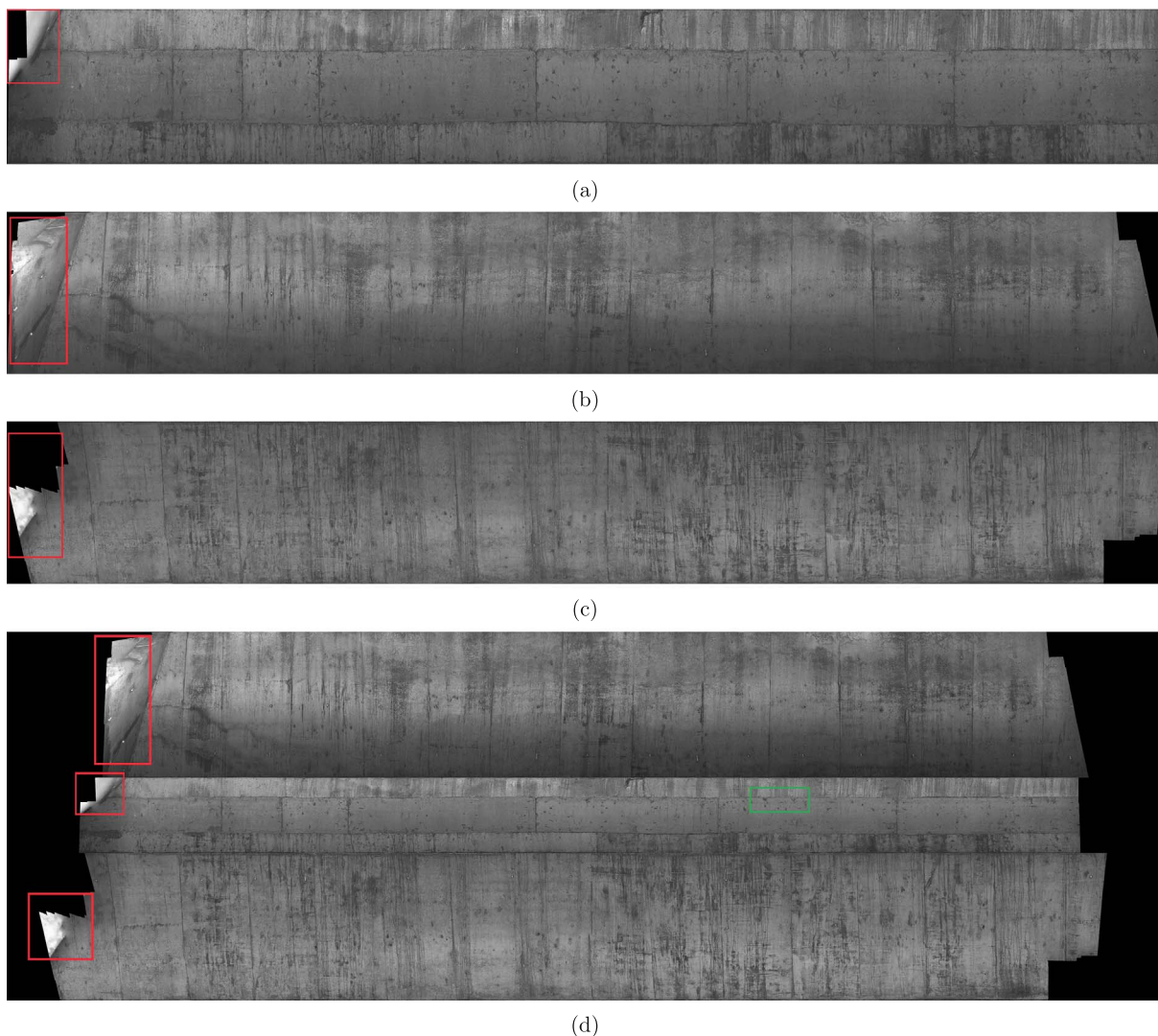
(a)

(b)

(c)

(d)

Fig. 19. The mosaicked panoramas of three bridge bottom surface planes (a)–(c) and the whole bottom surface of a lane of the test bridge. The overexposure shown in the red rectangular regions were caused by the bar LED lamps (see Fig. 2) due to the fact that the CCD cameras were too close to the baffles in the small box girder bridge.
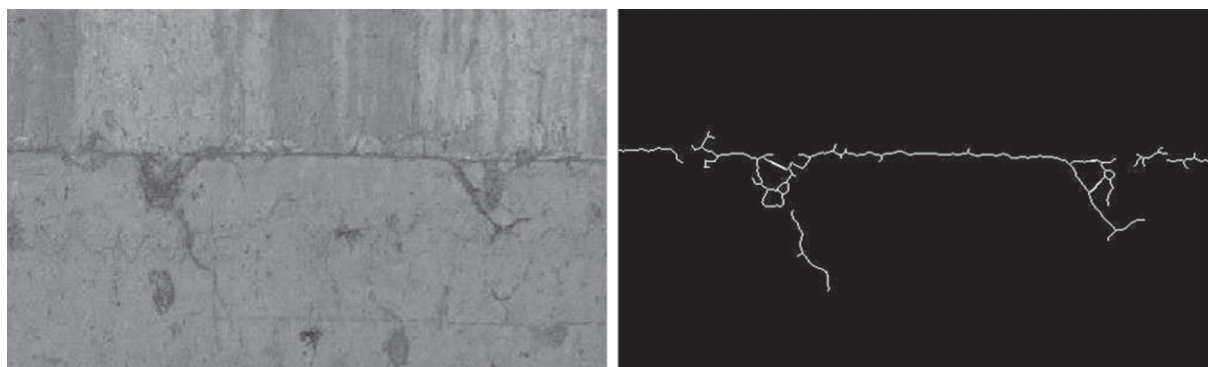

Fig. 20. An example of bridge crack detection: (Left) a locally cropped image region from the whole mosaicked panorama shown in Fig. 19 (d) and (Right) its corresponding crack detection result.

features is also important, which needs to be further studied by using a more effective and robust strategy to minimize the energy cost. Second, the optimization between groups in a single strip, which makes use of the homographic model only is not good enough. Some other transformation, such as the affine model, and the bridge structure line features also may be combined to achieve much better results. Finally, a more effective method to eliminate the luminance difference and color deviations is needed to carry out into just the overlapped regions between panoramas, not the overall images, which not only saves memory requirements and computational time, but also controls the stitching lines more easily.
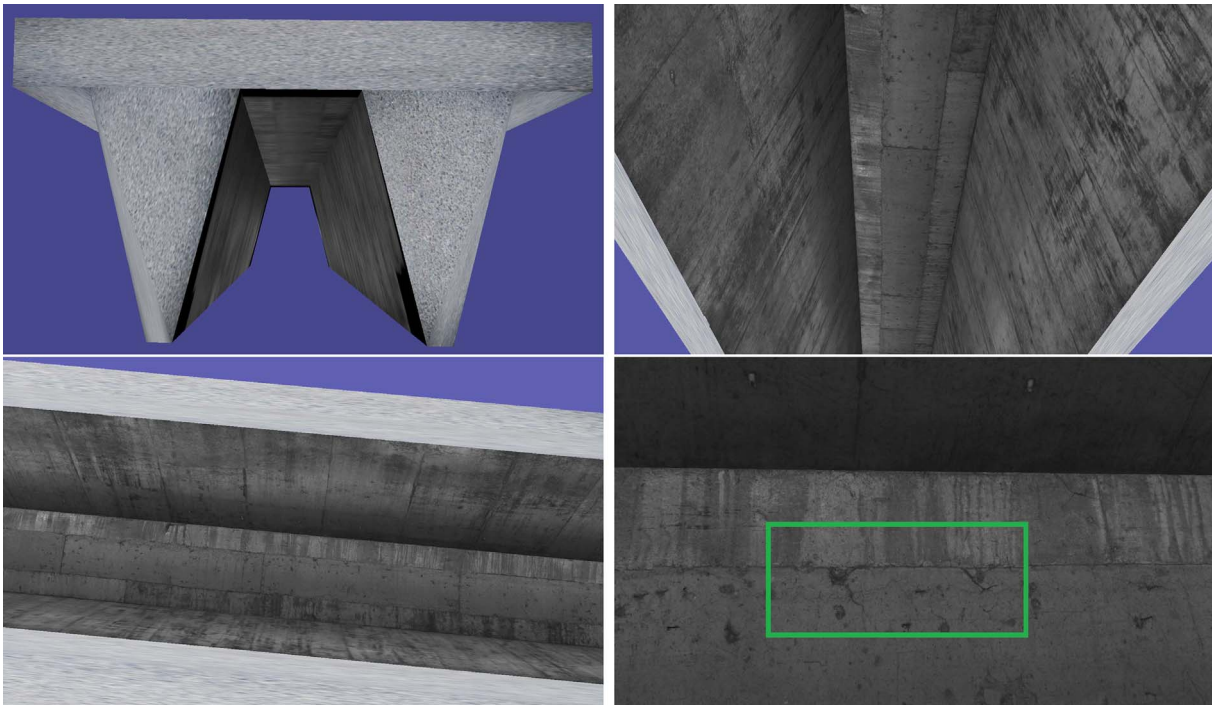
Fig. 21. The texture-mapped 3D models of the test small box girder bridge with the whole mosaicked panorama shown in Fig. 19 (d) at different viewpoints. The green rectangular region corresponds to the crack detection region as shown in Fig. 20.

Acknowledgments

References

[1] M.R. Jahanshahi, S.F. Masri, Adaptive vision-based crack detection using 3D scene reconstruction for condition assessment of structures, Automation in Construction 22 (2012) 567–576, http://dx.doi.org/10.1016/j.autcon.2011.11.018.

[2] W. Dekelbab, A. Al-Wazeer, B. Harris, History lessons from the national bridge inventory, Public Roads 0033-3735, 71 (6) (2008) 30–36 http://www.fhwa.dot.gov/publications/publicroads/08may/05.cfm.

[3] N.-C. Wu, S. Chase, An Exploratory Data Analysis of National Bridge Inventory, University of Virginia, 2010, https://trid.trb.org/view.aspx?id=1276845.

[4] G.A. Balaguer, C.A. Jardon, Climbing robots mobility for inspection and maintenance of 3D complex environments, Autonomous Robots 18 (2) (2005) 157–169, http://dx.doi.org/10.1007/s10514-005-0723-0.

[5] Y.N. Yang, H.J. Pam, M.M. Kumaraswamy, Framework development of performance prediction models for concrete bridges, Journal of Transportation Engineering 135 (8) (2009) 545–554, http://dx.doi.org/10.1061/(ASCE)TE.1943-5436.0000018.

[6] C. Pellegrino, A. Pipinato, C. Modena, A simplified management procedure for bridge network maintenance, Structure and Infrastructure Engineering 7 (5) (2011) 341–351, http://dx.doi.org/10.1080/15732470802659084.

[7] N. Gucunski, S.D. Boone, R. Zobel, H. Ghasemi, H. Parvardeh, S.-H. Kee, Nondestructive evaluation inspection of the Arlington Memorial Bridge using a robotic assisted bridge inspection tool (RABIT), SPIE Smart Structures and Materials + Nondestructive Evaluation and Health Monitoring, vol. 9063, 2014, http://dx.doi.org/10.1117/12.2063963.

[8] A.A. Hesse, R.A. Atadero, M.E. Ozbek, Uncertainty in common NDE techniques for use in risk-based bridge inspection planning: existing data, Journal of Bridge Engineering 20 (11) (2015), http://dx.doi.org/10.1061/(ASCE)BE.1943-5592.0000733.

[9] D. Rolander, B. Phares, B. Graybeal, M. Moore, G. Washer, Highway bridge inspection: state-of-the-practice survey, Transportation Research Record: Journal of the Transportation Research Board 1749 (1) (2001) 73–81, http://dx.doi.org/10.3141/1749-12.

[10] J.M. Ko, Y.Q. Ni, Technology developments in structural health monitoring of large-scale bridges, Engineering Structures 27 (12) (2005) 1715–1725, http://dx.doi.org/10.1016/j.engstruct.2005.02.021.

[11] B. Riveiro, D.V. Jauregui, P. Arias, J. Armesto, R. Jiang, An innovative method for remote measurement of minimum vertical underclearance in routine bridge inspection, Automation in Construction 25 (2012) 34–40, http://dx.doi.org/10.1016/j.autcon.2012.04.008.

[12] B. Yanev, G.A.C. Richards, Designing bridge maintenance on the network and project levels, Structure and Infrastructure Engineering 9 (4) (2013) 349–363, http://dx.doi.org/10.1080/15732479.2012.657651.

[13] D. Agdas, J.A. Rice, J.R. Martinez, I.R. Lasa, Comparison of visual inspection and structural-health monitoring as bridge condition assessment methods, Journal of Performance of Constructed Facilities 30 (3) (2015) 1–21, http://dx.doi.org/10.1061/(ASCE)CF.1943-5509.0000802.

[14] Z. Hu, Z. Hu, A fuzzy neural network technique for crack assessment of RC bridges, International Conference of Chinese Transportation Professionals, 2009, pp. 1–9, , http://dx.doi.org/10.1061/41064(358)349.

[15] J.-K. Oh, G. Jang, S. Oh, J.H. Lee, B-J. Yi, Y.S. Moon, J.S. Lee, Y. Choi, Bridge inspection robot system with machine vision, Automation in Construction 18 (7) (2009) 929–941, http://dx.doi.org/10.1016/j.autcon.2009.04.003.

[16] R.S. Lim, H.M. La, Z. Shan, W. Sheng, Developing a crack inspection robot for bridge maintenance, IEEE International Conference on Robotics and Automation, 2011, pp. 6288–6293, , http://dx.doi.org/10.1109/ICRA.2011.5980131.

[17] J.E. DeVault, Robotic system for underwater inspection of bridge piers, IEEE Instrumentation & Measurement Magazine 3 (3) (2000) 32–37, http://dx.doi.org/10.1109/5289.863909.

[18] H.M. La, R.S. Lim, B.B. Basily, N. Gucunski, J. Yi, A. Maher, F. Romero, H. Parvardeh, et al., Mechatronic systems design for an autonomous robotic system for high-efficiency bridge deck inspection and evaluation, IEEE/ASME Transactions on Mechatronics 18 (6) (2013) 1655–1664, http://dx.doi.org/10.1109/TMECH.2013.2279751.

[19] C.-H. Yang, M.-C. Wen, Y.-C. Chen, S.-C. Kang, An optimized unmanned aerial system for bridge inspection, ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction, vol. 32, 2015, pp. 1–6 https://search.proquest.com/openview/793a722cac341923687ecd643ddd891e/1?pq-origsite=gscholar&cbl=1646340.

[20] H.M. La, N. Gucunski, S.-H. Kee, L.V. Nguyen, Data analysis and visualization for the bridge deck inspection and evaluation robotic system, Visualization in Engineering 3 (1) (2015) 1–16, http://dx.doi.org/10.1186/s40327-015-0017-3.

[21] J. Shen, Y. Liu, Design and analysis of an obstacle-crossing wall-climbing robot mechanism, IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems, 2015, pp. 2067–2072, , http://dx.doi.org/10.1109/CYBER.2015.7288267.

[22] H. Oliveira, P.L. Correia, Automatic road crack detection and characterization, IEEE Transactions on Intelligent Transportation Systems 14 (1) (2013) 155–168, http://dx.doi.org/10.1109/TITS.2012.2208630.

[23] M. Gavilán, D. Balcones, O. Marcos, D.F. Llorca, M.A. Sotelo, I. Parra, M. Ocaña, P. Aliseda, P. Yarza, A. Amírola, Adaptive road crack detection system by pavement classification, Sensors 11 (10) (2011) 9628–9657, http://dx.doi.org/10.3390/s111009628.

[24] E. Zalama, J. Gómez-García-Bermejo, R. Medina, J. Llamas, Road crack detection using visual features extracted by Gabor filters, Computer-Aided Civil and Infrastructure Engineering 29 (5) (2014) 342–358, http://dx.doi.org/10.1111/mice.12042.

[25] E. Khanmirza, N. Khaji, E. Khanmirza, Identification of linear and non-linear physical parameters of multistory shear buildings using artificial neural network, Inverse Problems in Science and Engineering 23 (4) (2015) 670–687, http://dx.doi.org/10.1080/17415977.2014.933829.

[26] M.M. Torok, M. Golparvar-Fard, K.B. Kochersberger, Image-based automated 3D crack detection for post-disaster building assessment, Journal of Computing in Civil Engineering 28 (5) (2013) 1–13, http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.0000334.

[27] S. Chanda, G. Bu, H. Guan, J. Jo, U. Pal, Y.-C. Loo, M. Blumenstein, Automatic bridge crack detection - a texture analysis-based approach, Artificial Neural Networks in Pattern Recognition, vol. 8774, Springer, 2014, pp. 193–203, , http://dx.doi.org/10.1007/978-3-319-11656-3_18.

[28] R.S. Adhikari, O. Moselhi, A. Bagchi, Image-based retrieval of concrete crack properties for bridge inspection, Automation in Construction 39 (2014) 180–194, http://dx.doi.org/10.1016/j.autcon.2013.06.011.

[29] Q. Zou, Y. Cao, Q. Li, Q. Mao, S. Wang, CrackTree: automatic crack detection from pavement images, Pattern Recognition Letters 33 (3) (2012) 227–238, http://dx.doi.org/10.1016/j.patrec.2011.11.004.

[30] Q. Li, Q. Zou, D. Zhang, Q. Mao, FoSA: F* Seed-growing approach for crack-line detection from pavement images, Image and Vision Computing 29 (12) (2011) 861–872, http://dx.doi.org/10.1016/j.imavis.2011.10.003.

[31] J. Wang, P. Qiao, Improved damage detection for beam-type structures using a uniform load surface, Structural Health Monitoring 6 (2) (2007) 99–110, http://dx.doi.org/10.1177/1475921706072062.

[32] R. Szeliski, Image alignment and stitching: a tutorial, Foundations and Trends® in Computer Graphics and Vision 2 (1) (2006) 1–104, http://dx.doi.org/10.1561/0600000009.

[33] A. Levin, A. Zomet, S. Peleg, Y. Weiss, Seamless image stitching in the gradient domain, Computer Vision-ECCV 2004, vol. 2, Springer, 2004, pp. 377–389, , http://dx.doi.org/10.1007/978-3-540-24673-2_31.

[34] Y. Li, L. Ma, A fast and robust image stitching algorithm, The Sixth World Congress on Intelligent Control and Automation, vol. 2, 2006, pp. 9604–9608, , http://dx.doi.org/10.1109/WCICA.2006.1713865.

[35] J. De Villiers, Real-time photogrammetric stitching of high resolution video on COTS hardware, International Symposium on Optomechatronic Technologies, 2009, pp. 46–51, , http://dx.doi.org/10.1109/ISOT.2009.5326097.

[36] M. Brown, D.G. Lowe, Automatic panoramic image stitching using invariant features, International Journal of Computer Vision 74 (1) (2007) 59–73, http://dx.doi.org/10.1007/s11263-006-0002-3.

[37] A. Mills, G. Dudek, Image stitching with dynamic elements, Image and Vision Computing 27 (10) (2009) 1593–1602, http://dx.doi.org/10.1016/j.imavis.2009.03.004.

[38] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, M. Cohen, Interactive digital photomontage, ACM Transactions on Graphics 23 (3) (2004) 294–302, http://dx.doi.org/10.1145/1015706.1015718.

[39] W. Uricchio, The algorithmic turn: photosynth, augmented reality and the changing implications of the image, Visual Studies 26 (1) (2011) 25–35, http://dx.doi.org/10.1080/1472586X.2011.548486.

[40] J. Zaragoza, T.-J. Chin, Q.-H. Tran, M.S. Brown, D. Suter, As-projective-as-possible image stitching with moving DLT, IEEE Transactions on Pattern Analysis and Machine Intelligence 36 (7) (2014) 1285–1298, http://dx.doi.org/10.1109/TPAMI.2013.247.

[41] M.R. Jahanshahi, S.F. Masri, Multi-image stitching and scene reconstruction for evaluating change evolution in structures, Computing in Civil Engineering, vol. 10, ASCE Publications, 2011, pp. 643–657, , http://dx.doi.org/10.1061/41182(416)46.

[42] V. Rankov, R.J. Locke, R.J. Edens, P.R. Barber, B. Vojnovic, An algorithm for image stitching and blending, Biomedical Optics, vol. 5701, 2005, pp. 190–199, , http://dx.doi.org/10.1117/12.590536.

[43] J. Jia, C.-K. Tang, Image stitching using structure deformation, IEEE Transactions on Pattern Analysis and Machine Intelligence 30 (4) (2008) 617–631, http://dx.doi.org/10.1109/TPAMI.2007.70729.

[44] Z. Jakovljevic, R. Puzovic, M. Pajic, Recognition of planar segments in point cloud based on wavelet transform, IEEE Transactions on Industrial Informatics 11 (2) (2015) 342–352, http://dx.doi.org/10.1109/TII.2015.2389195.

[45] P.J. Burt, E.H. Adelson, A Multiresolution spline with application to image mosaics, ACM Transactions on Graphics 2 (4) (1983) 217–236, http://dx.doi.org/10.1145/245.247.

[46] L. Juan, O. Gwun, A comparison of SIFT, PCA-SIFT and SURF, International Journal of Image Processing 3 (4) (2009) 143–152 http://www.cscjournals.org/library/manuscriptinfo.php?mc=IJIP-51.

[47] J. Shi, C. Tomasi, Good features to track, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1994, pp. 593–600, , http://dx.doi.org/10.1109/CVPR.1994.323794.

[48] D.G. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision 60 (2) (2004) 91–110, http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94.

[49] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Communications of the ACM 24 (6) (1981) 381–395, http://dx.doi.org/10.1016/B978-0-08-051581-6.50070-2.

[50] Z. Zhang, A flexible new technique for camera calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (11) (2000) 1330–1334, http://dx.doi.org/10.1109/34.888718.

[51] J.J. Moré, The Levenberg-Marquardt algorithm: implementation and theory, Numerical analysis, 1978, pp. 105–116, , http://dx.doi.org/10.1007/BFb0067700.

[52] K. Nur, M. Morenza-Cinos, A. Carreras, R. Pous, Projection of RFID-obtained product information on a retail stores indoor panoramas, IEEE Intelligent Systems 30 (6) (2015) 30–37, http://dx.doi.org/10.1109/MIS.2015.90.

[53] D. Chang, N. Linard, M. Jessie, T. Vercauteren, J. Dauguet, Robust SIFT-based hierarchical video mosaicing for endomicroscopy, ISBI-International Symposium on Biomedical Imaging, 2014 inserm-01017954 http://www.hal.inserm.fr/inserm-01017954/.

[54] X. Ren, M. Sun, C. Jiang, L. Liu, H. Zheng, X. Li, A method of fast mosaic for massive UAV images, SPIE Asia Pacific Remote Sensing, vol. 9260, 2014, pp. 1–9, , http://dx.doi.org/10.1117/12.2069201.

[55] U. Krispel, H.L. Evers, M. Tamke, R. Viehauser, D.W. Fellner, Automatic texture and orthophoto generation from registered panoramic views, The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 40 (5) (2015) 131–137, http://dx.doi.org/10.5194/isprsarchives-XL-5-W4-131-2015.

[56] A. Fineschi, A. Pozzebon, A 3D virtual tour of the Santa Maria della Scala Museum Complex in Siena, Italy, based on the use of Oculus Rift HMD, 2015 International Conference on 3D Imaging, 2015, pp. 1–5, , http://dx.doi.org/10.1109/IC3D.2015.7391825.