

Research Article

Performance Improvement for Two-Lens Panoramic Endoscopic System during Minimally Invasive Surgery

Dinh Thai Kim ^{1,2}, Ching Hwa Cheng ³, Don Gey Liu,^{1,3} Kai-Che Jack Liu,⁴ Shih Wei Wayne Huang,⁴ and Song Toan Tran¹

¹Department of Electrical and Communications Engineering, Feng Chia University, 40724 Taichung, Taiwan

²University of Information and Communication Technology, Thai Nguyen, Vietnam

³Department of Electronic Engineering, Feng Chia University, 40724 Taichung, Taiwan

⁴IRCAD Taiwan/AITS, Lukang, Taiwan

Correspondence should be addressed to Dinh Thai Kim; kimdinhthaibk@gmail.com

Received 15 July 2019; Revised 18 September 2019; Accepted 16 October 2019; Published 19 November 2019

Academic Editor: Angkoon Phinyomark

Copyright © 2019 Dinh Thai Kim et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

One of the major challenges for Minimally Invasive Surgery (MIS) is the limited field of vision (FOV) of the endoscope. A previous study by the authors designed a MIS Panoramic Endoscope (MISPE) that gives the physician a broad field of view, but this approach is still limited, in terms of performance and quality because it encounters difficulty when there is smoke, specular reflections, or a change in viewpoint. This study proposes a novel algorithm that increases the MISPE's performance. The method calculates the disparity for the region that is overlapped by the two cameras to allow image stitching. An improved evaluation of the homography matrix uses a frame-by-frame calculation, so the stitched videos are more stable for MIS. The experimental results show that the revised MISPE has a FOV that is 55% greater, and the system operates stably in real time. The proposed system allows a frame rate of 26.7 fps on a single CPU computer. The proposed stitching method is 1.55 times faster than the previous method. The stitched image that is obtained using the proposed method is as similar as the ground truth as the SURF-based stitching method that was used in the previous study.

1. Introduction

Minimally Invasive Surgery (MIS) is becoming a gradually preferred option to traditional open surgery because it involves decreased blood loss, decreased postoperative pain, fast recovery time, and less scarring. However, one of the major obstacles for MIS is limited FOV. MIS is difficult to implement in operations because of its nonintuitive nature, so less experienced physicians encounter greater risks during MIS surgery.

To solve the problem of limited FOV in the current endoscope, we consider other studies that proposed the use of a panoramic endoscope that has a special design. One example of such endoscopes was the one developed by Yamauchi et al. [1]. The group was the first to design an endoscope that was able to produce a wider view of an image through the help of an image-shifting prism. In another

study, Roulet et al. [2] proposed the design of a 360° endoscopy that would employ panomorph lens to produce wide-view images. Recently, a panoramic endoscope that used convex parabolic mirrors was developed by Tseng and Yu [3]. These studies are important in increasing the viewing angle of the surgical images as well as the entire working area. This makes laparoscopic surgery safer. These studies have focused on designing optical systems through the application of multiple prisms, lens, and mirrors to generate a larger surgical view of images. However, this approach also encountered many issues from aberrations to blind zones, and image quality is often affected by noise or distortion.

To increase the image viewing angle, an image stitching technique (mosaicking) is used in computer vision [4]. Many studies use this technique in MIS. Behrens et al. [5, 6] demonstrated the mosaicking of a sequence of endoscopic bladder images from a video. A Global and Local Panoramic

View for Gastroscopy is proposed in [7]. In [8], a scene-adaptive feature-based approach was proposed for the mosaicking of placental vasculature images that are obtained during computer-assisted fetoscopic procedures. These studies perform the image stitching using the movement of a monocular endoscope. This yields only panoramic static images that do not reflect the changes that may occur in the shape of the organs or blood vessels outside the camera's FOV. Takada et al. [9, 10] proposed a Hybrid Tracking and Matching Algorithm for Mosaicking Multiple Surgical Views. They combined the feature-based image registration and the optical flow tracking algorithm to evaluate the homography matrices for frames acquired from different trocar-retractable cameras. This approach has improved speed and robustness for cases in which the overlap size is small compared to the feature-based method only. However, this approach can lead to errors in results if the tracking process is failed.

Previous studies by the authors proposed a MIS panoramic endoscope (MISPE) that gives physicians a broad area of view [11] using a feature-based image-stitching algorithm. Our MISPE is made up of two lenses which are mounted on its tip and have a tunable distance between them. The shortest distance is 5 mm. The endoscope is then connected to a PC using a USB to enable it to simultaneously capture the surgical images. The video-stitching module is applied at this point. It stitches the images to enable physicians to obtain a wide view of the image. The stitched image represents every transformation that occurs in the FOVs of the two lenses mounted on the endoscope. Figure 1 shows a schematic diagram of this MISPE system.

However, this feature-based image-stitching approach does not perform well for MIS, which is often affected by smoke, a change in viewpoint, and specular highlights. These problems are a function of the tissue characteristics, the proximity of the light source, and the proximity of the cameras. In this situation, the distribution of features in images is ambiguous, and the precision with which feature pairs are matched is decreased, so few features are detected or they are unevenly distributed. This results in a failure in image registration and less accurate stitching results. Figure 2 shows the stitching result for the Speeded-Up Robust Features- (SURF-) based stitching method that is used in the previous study. Figure 2(b) shows an erroneous patch because there are only a few common feature points.

In the previous study, video stitching is performed using a frame-by-frame evaluation of a homography matrix. This approach ensures correct stitching when the two cameras move toward or away from the surgical area. However, this approach changes the shape of the stitched images because there is a significant change in the homography matrix, even though the cameras are fixed. The matching points change frequently because of environmental factors such as brightness, smoke, and specular reflections. Figure 3 shows the changes in the dark area from Figures 3(a) and 3(b). This produces unstable stitched images in a video, which distract the physician's concentration while observing the video during surgery. The change in the homography matrices

between consecutive frames must be smooth when stitching videos.

This study proposes several improvements in the revised MISPE system that address the issues of the previous MISPE system. A new algorithm based on calculating the disparity map is used to stitch images. The speed, quality, and stability with which the video is stitched is increased.

The remainder of this paper is organized as follows. Section 2 presents the proposed image-stitching algorithm. Section 3 presents the proposed video-stitching algorithm. The experimental results are presented and discussed in Section 4. Finally, conclusions are drawn in Section 5.

2. The Proposed Image-Stitching Algorithm

The image-stitching algorithm comprises two stages: image registration and image compositing [11].

Image registration is the most important element of the image-stitching process because it directly affects the accuracy of the image-stitching results. This involves searching for pixels (e.g., feature points) or objects in the two different camera views. However, searching the entire image is a complex process that requires high-performance computing. Images that are captured directly from the camera must also be corrected to account for lens distortion. Therefore, this study uses an image rectification technique [12] to transform the endoscope system into an aligned-undistorted configuration. The search is then simplified to a one-dimensional problem with a horizontal line parallel to the baseline between the cameras. A video-stitching module then uses the aligned-undistorted configuration. The algorithm involves three steps, as shown in Figure 4. The details of the processes are described in the following subsections.

2.1. Rectifying Images. This step corrects any distortion in the image that occurs due to distortion of the lens and aligns the two cameras into one viewing plane, so that the pixel rows between the cameras are exactly aligned with each other. This study uses Bouguet's algorithm [12] from the OpenCV library. In order to ensure precise rectification, 20 concurrent images were captured of a 16×11 chessboard at a distance from 3 cm to 15 cm and at different angles. This step was performed offline. Figures 5(a) and 5(b) show the input images and the undistorted and rectified images with the corresponding pixels on the same horizontal line (epipolar line). When the rectification process is complete, two rectified images are used for image stitching.

2.2. Image Registration. Image registration matches points in two overlapping images to evaluate a homography matrix. For two rectified images, the image registration involves two steps:

Step 1. Compute the disparity map.

The disparity map shows all of the pixels that are different in the two rectified images. This study uses the block matching (BM) algorithm from OpenCV as a StereoBM module because this algorithm is fast and effective and similar to the one

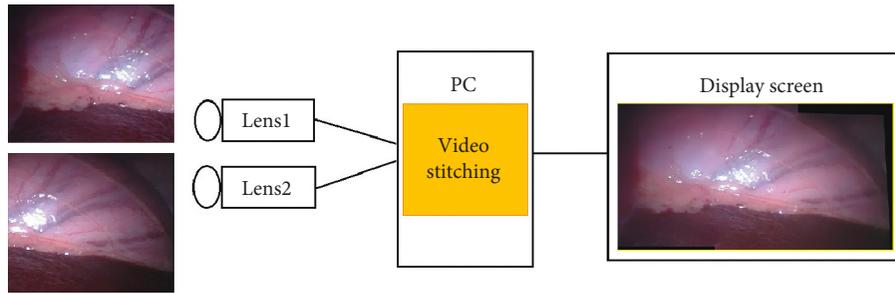


FIGURE 1: MISPE system.

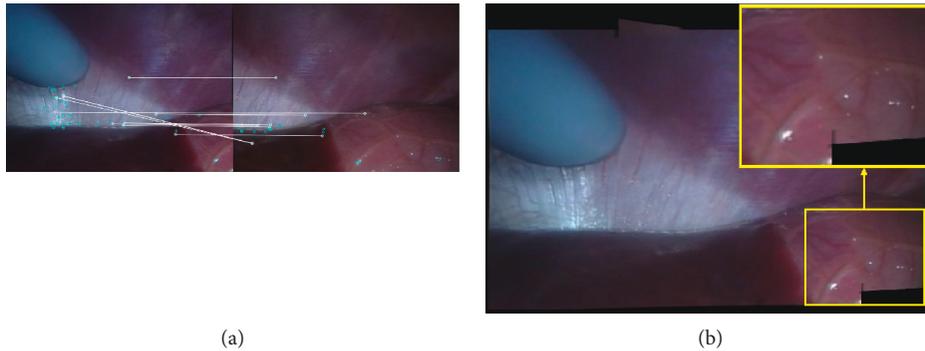


FIGURE 2: An illustrative example of the SURF-based stitching algorithm for MIS showing (a) only a few matching features are obtained and distributed unequally and (b) error stitching.

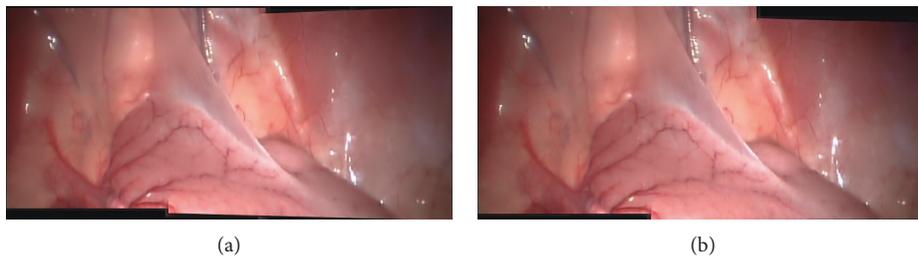


FIGURE 3: Stitched images of two consecutive frames. (a) Stitching image at frame 36. (b) Stitching image at frame 37.

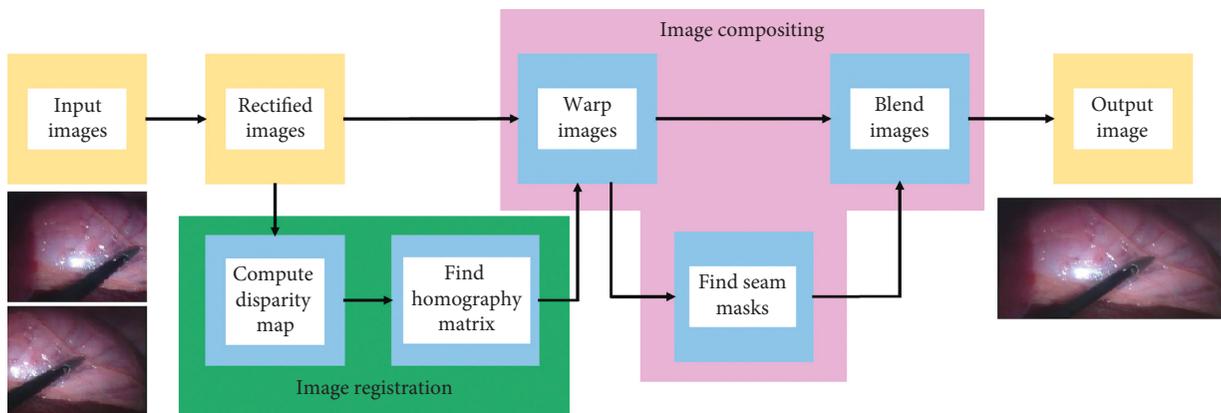


FIGURE 4: Proposed image-stitching algorithm.

that was developed by Konolige [13]. It uses a small “sum of absolute difference” (SAD) window to find points in the left and right images that match, and then the disparity is calculated as the actual horizontal pixel difference.

A disparity map that is computed using by StereoBM usually contains invalid values (holes), which are usually concentrated in uniform texture-less areas, half-occlusions, and regions near depth discontinuities. Therefore, Fast

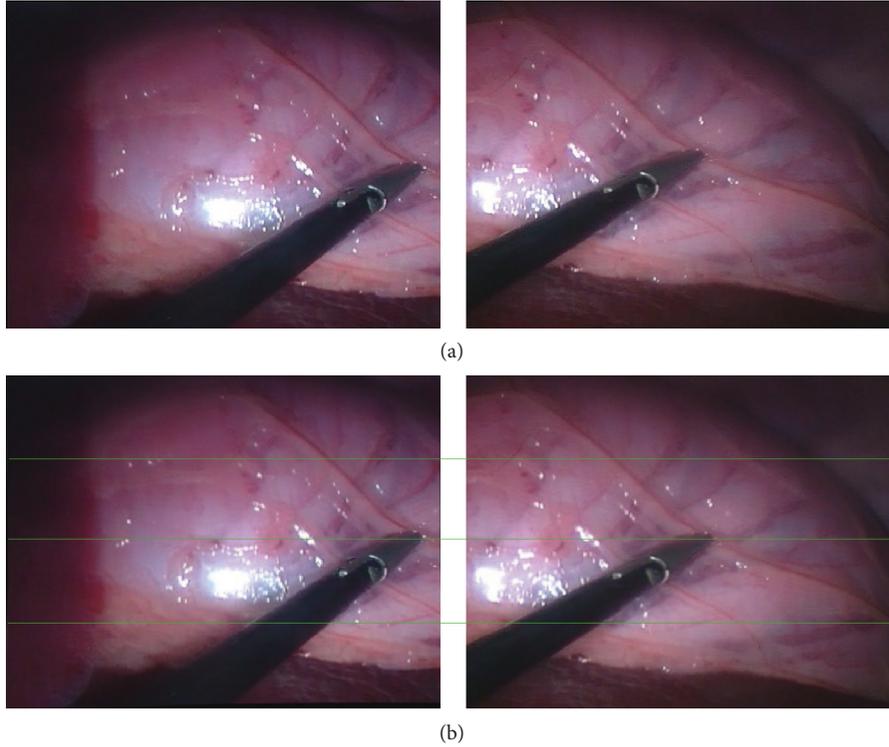


FIGURE 5: Image rectification. (a) Two input images. (b) Two output aligned images.

Global Smoothing (FGS) [14] is used as the postfiltering module in OpenCV to filter the disparity map. This module enables this type of postfiltering for real-time processing on the CPU. Figure 6 shows that two disparity maps that are computed using StereoBM: one that uses the left image as the reference image (left disparity map) and a second disparity map that uses the right image as the reference (right disparity map). The left-right-consistency-based confidence [15] is then used to refine the disparity map (refined disparity map) in half-occlusions and uniform areas.

Step 2. Determining the homography matrix.

It is not necessary to use all of the matching pixels in the overlap region to evaluate the homography matrix because the computational burden is large. Therefore, this study proposes an ROI-grid method to determine the corresponding point pairs for two rectified images to evaluate the homography matrix.

A region of interest (ROI) where the pixels are calculated for disparity is defined as a region in the overlapping part of the left resized-rectified image. This study assumes that the minimum width of the overlapping area is equal to 30% of the width of the rectified image, so the ROI established as shown in Figure 7 is a region at position A ($7w/10, 5$) with width $(3w/10 - 5)$ and height $(h-10)$, where w and h are the width and height of the left rectified image. The points at the edge of the image are removed because the disparity value for these points is often unstable.

This ROI is then divided into $m \times n$ grids. The peak intensity of the grid is used to extract its corresponding

point in the right rectified image by calculating its disparity value as follows:

$$(x_Q, y_Q) = (x_P - \text{disparity}(P), y_P). \quad (1)$$

This gives a set of $(m \times n)$ corresponding point pairs for two overlapped rectified images. Because the homography matrix is a (3×3) matrix with 8 degrees of freedom (DoF), at least four corresponding point pairs are required to determine the matrix, so the rank of the $(m \times n)$ matrix must ensure that the number of corresponding point pairs was not less than 4. This study uses a (9×24) matrix. Figure 7 shows that this approach ensures that a large number of corresponding pairs are evenly distributed in the ROI. This makes stitching more accurate and more stable.

Because there are still some mismatched pairs that have invalid disparity values, the RANdom SAMple Consensus (RANSAC) algorithm [16] is used to determine the well-matched pairs (inliers) by removing the mismatched pairs (outliers). The homography matrix is then evaluated using the set of the inliers and the least-squares method to give the least reprojection error [17].

2.3. Image Compositing. After image registration, the image-compositing stage yields wide-angle images. This step uses the same process that is described in a previous paper by the authors [11]. The graph-cut technique algorithm [18] is used to determine an optimal seam to eliminate the appearance of “artifacts” or “ghosting.” The multiband blending method [19] is then used to smooth the stitching results.

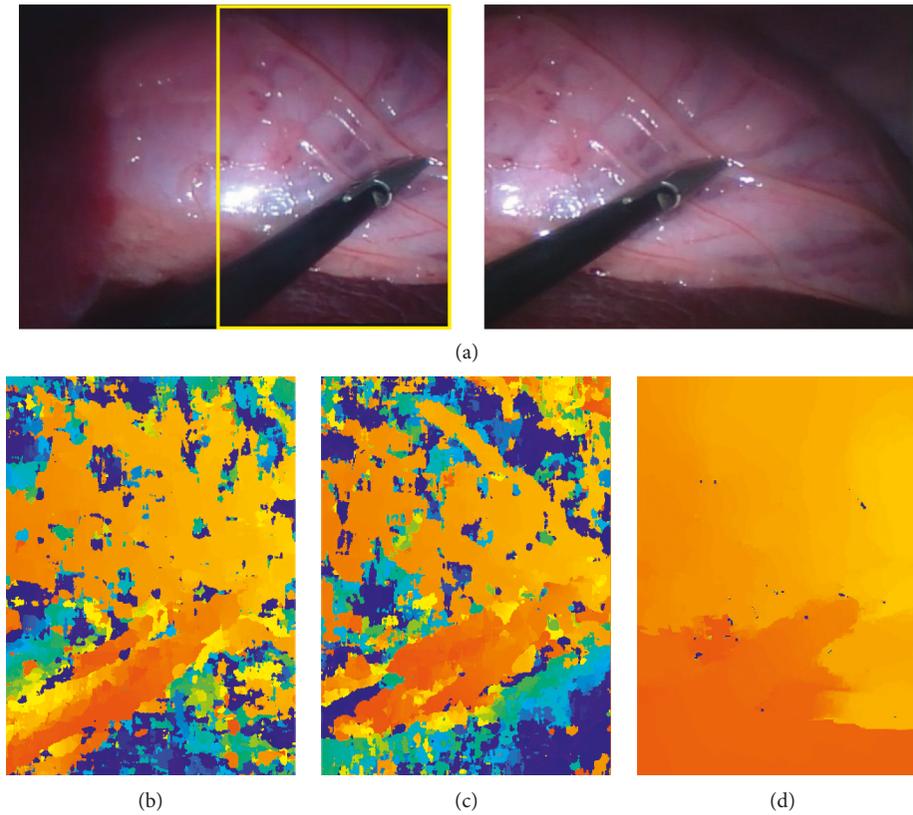


FIGURE 6: Compute disparity map using stereoBM and FGS filter. (a) Two rectified images with an overlap area (yellow color). (b) Left disparity map. (c) Right disparity map. (d) Refined disparity map.

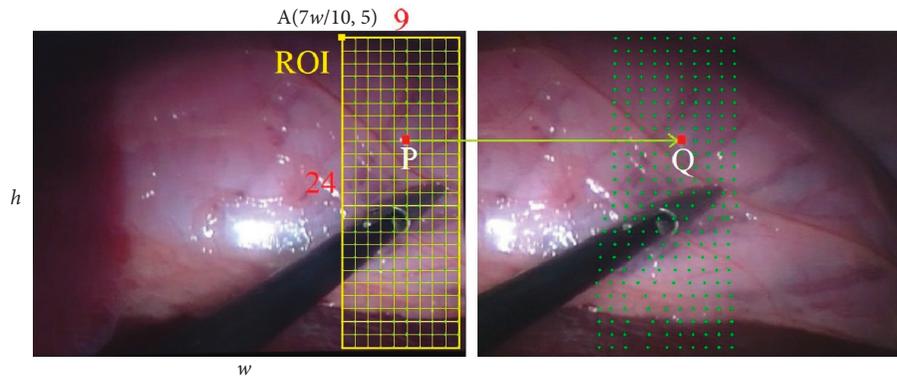


FIGURE 7: The ROI-grid method. The corresponding point pairs are determined based on the disparity value. The ROI (dark yellow) is the region that is used to calculate disparity. The ROI is divided into a 9×24 grid and each grid's peak (P) is used to determine the corresponding point (Q) in the right rectified image.

In summary, the proposed image-stitching algorithm for this study includes the following six steps as described in (Algorithm 1):

3. The Proposed Video-Stitching Algorithm

The proposed image-stitching algorithm stitches video by stitching images that are captured from videos or cameras frame-by-frame. For practical applications, there are two requirements for the proposed MISPE system: stability and fast processing time for stitching videos.

3.1. Stitching Video at Increased Speed. For the proposed image-stitching algorithm, the two most time-consuming steps involve computing the disparity map and determining the seam mask, especially for high-resolution images. Therefore, the proposed method accelerates the video-stitching process by reducing the time that is required to calculate the disparity map and determine the seam mask, as shown in Figure 8.

This study uses a downsizing technique to transform the processed images into low-resolution images using an image-resize function and the bilinear interpolation algorithm in

Input: Two input images

- (1) Rectify two input images to two rectified images
- (2) Compute the disparity map using the stereoBM and FGS filter in OpenCV
- (3) Establish a homography matrix using the proposed ROI-grid method and a RANSAC algorithm
- (4) Transform the right rectified image into the same left rectified image plane using the estimated homography matrix
- (5) Determine an optimal seam to prevent the possibility of “ghosting” in the stitched image using the graph-cut technique
- (6) Render panorama using the multiband blending method

Output: Panoramic image

ALGORITHM 1: Proposed image-stitching algorithm.

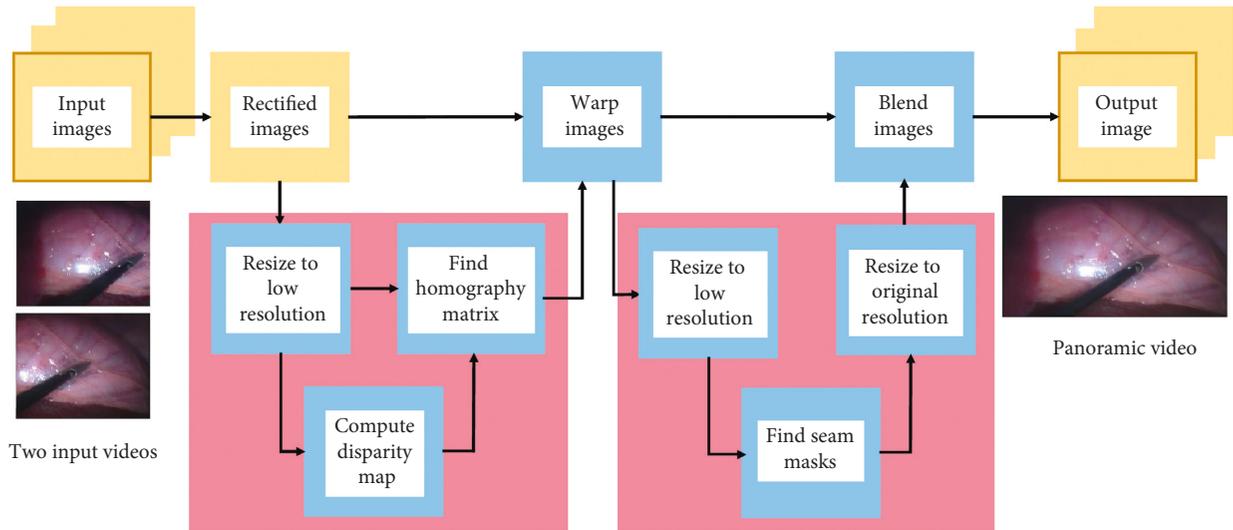


FIGURE 8: Decreasing the time that is required to stitch video using the downsizing technique (pink area).

OpenCV. The resized-scale value is input manually to accelerate the process and maintains the required image quality.

To decrease the time that is required to produce the disparity map, the resized-scale value (k_1) is selected such that the rectified images can be resized to a resolution of (320×240) in many of the experimental cases. The disparity map is then computed, and the homography matrix for these two resized-rectified images is calculated. It is assumed that the homography matrix for the two resized-rectified images is H_{resize} . Because the coordinates of a point in the resized image are proportional to the coordinates of that point in the original image with a resized-scaling factor k_1 , the homography matrix that transforms the two original rectified images on the same plane is defined as follows:

$$H = \begin{bmatrix} k_1 & 0 & 0 \\ 0 & k_1 & 0 \\ 0 & 0 & k_1 \end{bmatrix} H_{\text{resize}} = \begin{bmatrix} \frac{1}{k_1} & 0 & 0 \\ 0 & \frac{1}{k_1} & 0 \\ 0 & 0 & \frac{1}{k_1} \end{bmatrix}. \quad (2)$$

To decrease the time that is required to determine the seam mask, the resized-scale value (k_2) is used to transform

the two warped images to two low resolution (64×48) resized-warped images. The two seam masks are then calculated using the graph-cut algorithm. The two masks are then resized to the original resolution, and the warped images are blended. For this study, the resized-scale value (k_2) can be adjusted to ensure that the quality of the seam estimation is high.

3.2. Increasing the Stability of the Stitched Video. The proposed method rectifies two frames that are captured from two cameras and computes the disparity map. The matching point pairs in two rectified frames are then determined using the proposed ROI-grid method. A RANSAC algorithm is used to determine the well-matched pairs (inliers) and to remove the mismatched pairs (outliers) using a RANSAC threshold. This study uses a RANSAC threshold of 3.0. It is assumed that there are N inliers, $\{(p_1, q_1), (p_2, q_2), \dots, (p_N, q_N)\}$, in the two frames. The medium reprojection error of a 3×3 homography matrix (H) for the set of N inliers is defined as follows:

$$ME(H) = \sum_{i=1}^N \frac{d(p_i, Hq_i)}{N}, \quad (3)$$

where H is a 3×3 homography matrix, p_i and q_i are the corresponding homogeneous coordinates of the i -th pair in

Input: Inliers in two current frames, the homography matrix of the previous two frames (H_{prev}), the threshold value (μ), and the RANSAC threshold (α)

(1) Repeat for k iterations (RANSAC loop)

(a) Select four random matching point pairs

(b) Compute the exact homography matrix H for the four matching point pairs

(c) Determine the number of inliers such that $H: d(p_i, Hq_i) < \alpha$

(2) Choose the value of H with the largest number of inliers N

(3) Calculate the medium reprojection error for H_{prev} for the N inliers:

$$ME(H_{prev}) = \sum_{i=1}^N d(p_i, H_{prev}q_i) / N$$

(4) If $ME(H_{prev}) < \mu$

(4.1) Then the homography matrix is assigned as H_{prev}

(4.2) Else recompute the homography matrix using the least-squares method and all of the N inliers and a minimum cost function

$$\sum_{i=1}^N d(p_i, Hq_i) / N$$

Output: Homography matrix for the two current frames

ALGORITHM 2: Increasing the stability of the stitched video.

the N inliers, and $d(p_i, Hq_i)$ is the Euclidean distance between points p_i and Hq_i .

It is assumed that H_{prev} is the homography matrix for the previous two frames. The medium reprojection error for H_{prev} on the set of N inliers is as follows:

$$ME(H_{prev}) = \sum_{i=1}^N \frac{d(p_i, H_{prev}q_i)}{N}. \quad (4)$$

The error $ME(H_{prev})$ is large when the cameras are moved significantly, so the coordinates of the inliers change significantly. The error is small when the cameras are almost fixed. If this error is sufficiently small, it is possible to use H_{prev} to stitch the two current frames. Therefore, the proposed algorithm compares $ME(H_{prev})$ with a specified threshold value to reduce the variation in the homography matrix which increases the stability of the stitched videos.

For the two initial frames, the homography matrix is calculated as the minimum value of the cost function $ME(H)$ for the set of inliers in the two initial frames.

For the next frames, the error $ME(H_{prev})$ is calculated using the same set of inliers in the current frames. If the error is less than a specified threshold value, H_{prev} is used to stitch the two current frames. If the cameras move by a significant amount, the homography matrix is calculated again as the minimum value of the cost function $ME(H)$ for a set of inliers in the two current frames. This study uses a threshold value of 2.0.

Detailed proposed algorithm is described in (Algorithm 2):

4. Results and Discussion

The proposed method was tested using a PC with an Intel i5-4590 3.4 GHz CPU 16 GB of RAM memory on an Ubuntu 16.04 system. The program is implemented using C++ with OpenCV 3.4.0. The default parameters for the OpenCV function are used. These parameters can be adjusted on the control panel to achieve the best quality disparity.

4.1. Video-Stitching Results. To validate the method, experiments for in vivo animal trials were performed with the support of the IRCAD MIS research center, Show Chwan Memorial Hospital, Taiwan. The two endoscopic cameras that were used in the experiments were 2.0MP USB Digital Endoscopes from Oasis Scientific Inc.

Figure 9 shows the animal trial results: Left and middle show two input images that are captured using the two endoscopic cameras during the in vivo animal experiment. Right shows the image stitching results. The results confirm that the proposed method can increase the FOV by 155%.

We performed the video-stitching process on various kinds of video samples, with three of these shown in Figure 9. The video samples shown in the figure are the ones that occur under the conditions of presence of smoke, the appearance of specular reflections, the appearance of moving surgical tools, and the presence of a moving camera during MIS. The first sample represents a situation where the cameras are held still, but there is a movement of the heart and presence of smoke. The second sample describes the situation in which our endoscope moves toward or away from the surgery area. The third one is filmed under various camera movements, specular reflection appearance, and the appearance of a moving tool in MIS. These input videos and results can be found in [20–22].

4.2. Comparison with the Previous Method. To determine the effectiveness of the proposed algorithm, the proposed method is compared with the SURF feature-based stitching method that is described in a previous study by the authors [11]. In this study, we only made comparisons with [11] because the stitched images from two endoscopic cameras for MIS are similar.

First of all, we evaluate the percentage of stitchable frames for the three samples, as shown in Table 1. In sample 1, there are a lot of frames in the two input videos that are not stitched by SURF. This is due to the occurrence of environmental factors such as smoke or specular reflection appearances resulting in very few matching features detected correctly, while all the frames are stitched using the proposed method for the three samples.

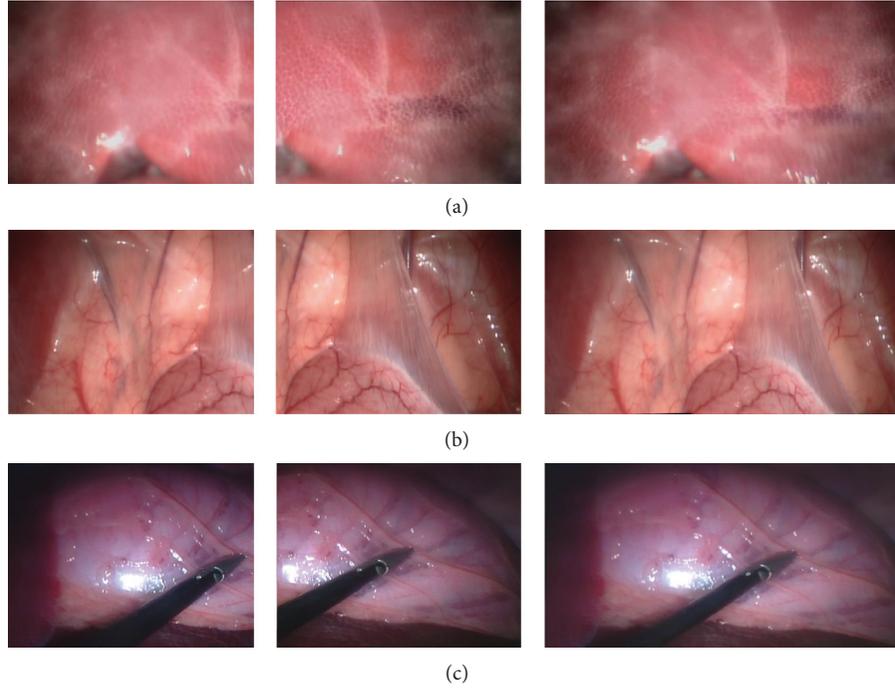


FIGURE 9: Video-stitching results from various samples for in vivo animal trials. Left and middle: two input images captured from two endoscopic cameras. Right: the stitching results. (a) Sample 1. (b) Sample 2. (c) Sample 3.

TABLE 1: The percentage of stitchable frame of SURF [11] and our method on the same dataset.

The percentage of stitchable frame	Sample 1 (%)	Sample 2 (%)	Sample 3 (%)
SURF [11]	20	95	87
Our method	100	100	100

The second, the quality of the two methods is evaluated in comparison with ground truth for the same dataset. Figure 10 shows the stitching results for both methods in the three samples: Left shows the SURF-based stitching results, Middle shows the stitching results for the proposed method, and Right shows the ground truths. It is seen that the stitching results for the proposed method are similar to the ground truth, and the SURF-based stitching result is distorted in shape because only a few matching feature pairs are detected and these are unevenly distributed.

To determine the alignment accuracy for both methods, the average error for every pixel in the overlapping area after the warping transformation is calculated using the evaluated homography matrix. (f) is the left image and (g) is the result for the right image after transforming on the same plane as the left image. A maximum rectangular area within the overlapping area is then defined to calculate the corresponding pixel difference between images (f) and (g). The alignment error is calculated as follows:

$$\text{alignment error} = \frac{\sum_{m \times n} |f(x, y) - g(x, y)|}{m \times n}, \quad (5)$$

where $f(x, y)$ and $g(x, y)$ are the grayscale pixel values for position (x, y) in the defined rectangular area of the two images (f) and (g) and $m \times n$ is the size of the selected rectangular area.

The alignment error of the stitching video is evaluated as the average alignment error of the stitchable frames. For the three datasets in Figure 9, the alignment error is shown in Table 2. It is clear that the alignment errors for the proposed method are all smaller than those for SURF.

The third, the computational time or the video-stitching rate for both methods is determined. Input videos that were used in the experiments were simultaneously acquired by the two endoscopic cameras with a resolution of 640×480 pixels at 30 fps.

Figure 11 shows the video-stitching rate after 1500 consecutive frames. These graphs show that the proposed method operates at 26.7 fps and the SURF-based method operates at 17.2 fps. The proposed method stitches video 1.55 times faster than the SURF-based method.

Finally, to compare the stability as well as quality of the stitched videos for both methods, three additional demonstration videos were posted on Youtube [20–22]. These results show that the shape of frames in the stitched videos by our method only changes when the cameras move significantly away or toward the surgical area. Meanwhile, the frames in the stitched video by the previous approach always change the shape frame-by-frame. It means that the video that is processed using the proposed method has a more stable picture than that processed using the previous method.

Therefore, these results demonstrate the feasibility of the proposed method in MIS to expand the narrow FOV of a conventional endoscope.

TABLE 2: Alignment error of SURF [11] and our method on the same dataset.

Alignment error	Sample 1	Sample 2	Sample 3
SURF [11]	7.25	3.23	4.45
Our method	5.02	2.74	3.31

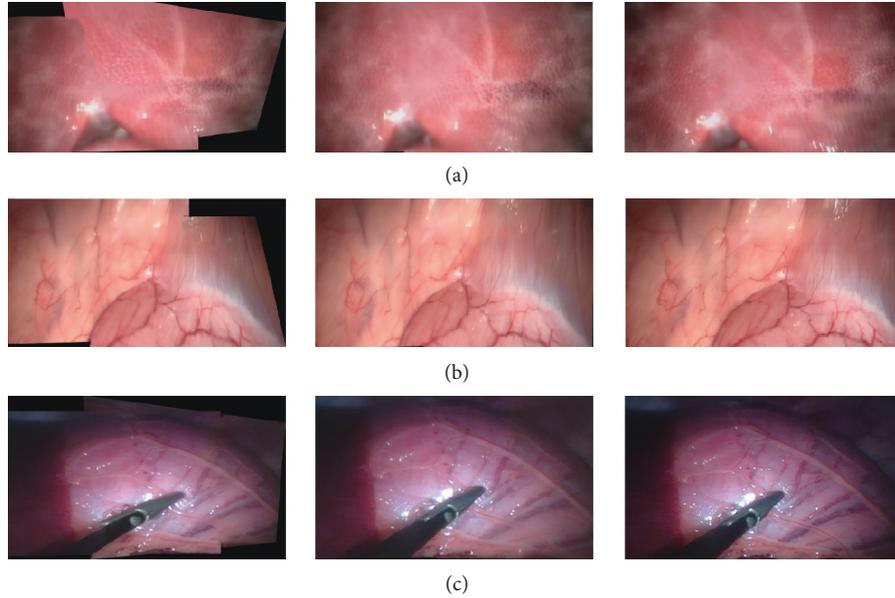


FIGURE 10: Comparison of the stitching result for the SURF-based method and the proposed method. Left: SURF-based stitching result. Middle: the proposed stitching result. Right: ground truth. (a) Sample 1. (b) Sample 2. (c) Sample 3.

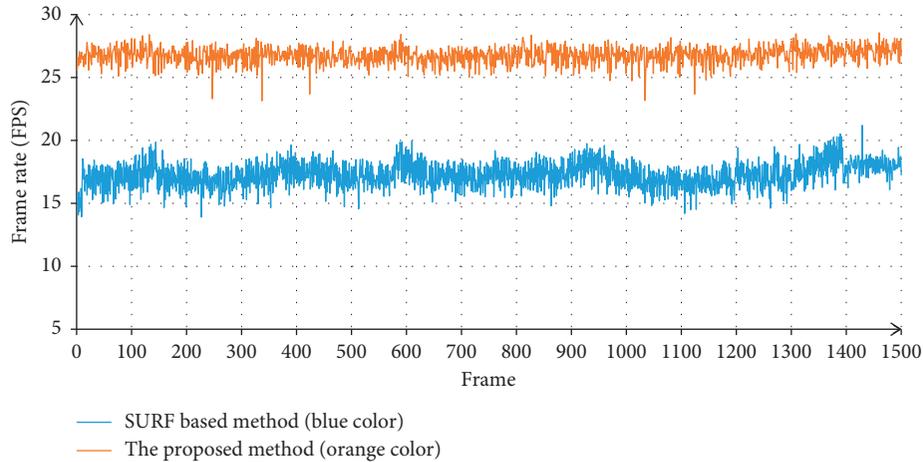


FIGURE 11: A comparison of the frame rate for stitching using the SURF-based method (blue color) and the proposed method (orange color) for 1500 consecutive frames performed using a computer with CPU only.

4.3. Discussion. The proposed method increases the FOV of the input image and supports the reconstruction of a 3D surface image of the overlapping area. Figure 12 (a) shows two input images with an overlapping area (yellow color), (b) shows the disparity map, (c) shows the stitching result, and (d) shows a 3D surface image as a direct result from the disparity map.

However, the proposed method still has some limitations. The method cannot stitch any two images unless the parameters for the rectification process are known. These parameters include the intrinsic and extrinsic

parameters for each camera. The feature-based stitching method does not have this requirement. The accuracy of the proposed method depends on the stereo matching algorithm that is used and on the accuracy of the rectification process, so if the rectification process is not sufficiently accurate, the two rectified images are not aligned and this affects the quality of the final stitched image. This study only uses the stereoBM algorithm and FGS filter in OpenCV to calculate the disparity map. A future study will involve a new proposal to improve the

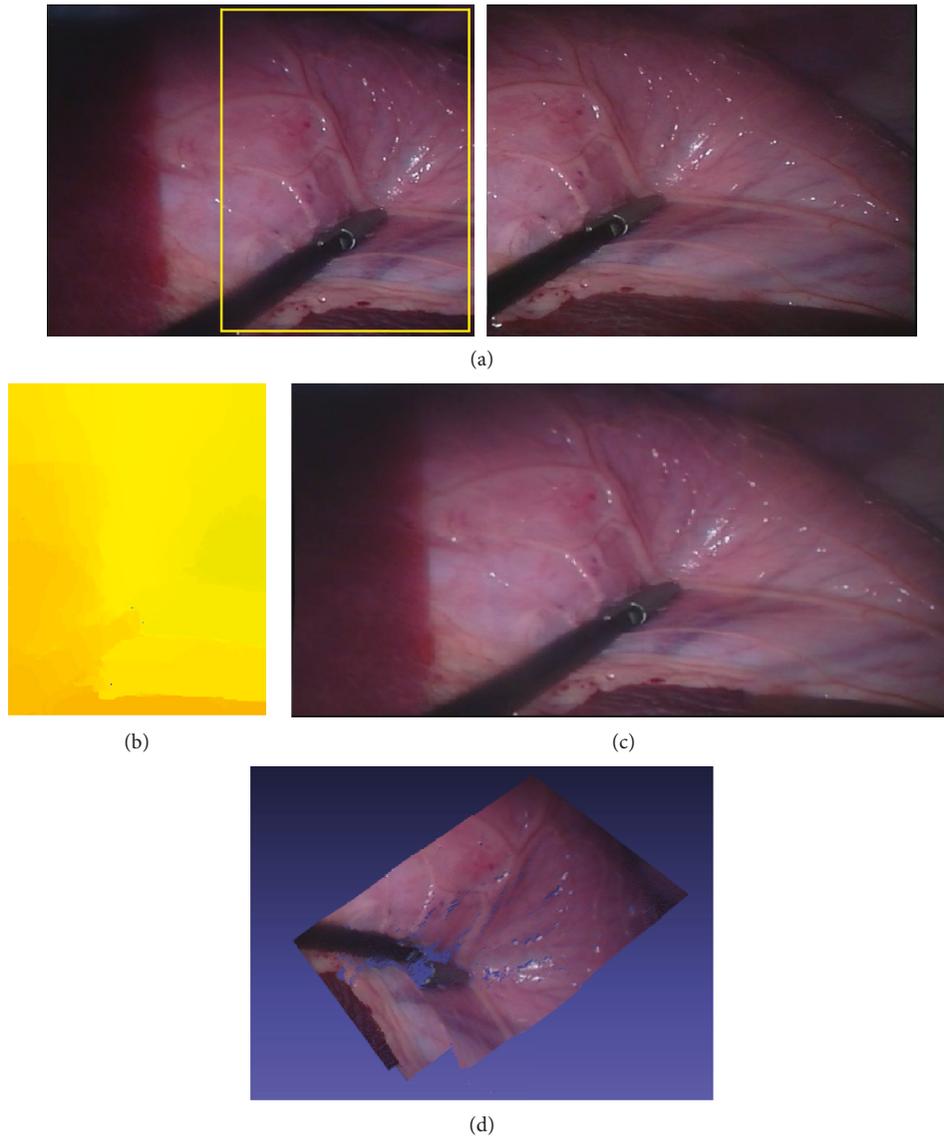


FIGURE 12: The proposed method increases the FOV of the input image by 35% and is used to reconstruct the dense surface image of the overlapping area. (a) Input images with an overlapping area (yellow color). (b) Disparity map. (c) Stitched image. (d) 3D reconstruction.

quality of the disparity map and the development of an additional 3D reconstruction module.

5. Conclusions

This study proposes a novel algorithm that increases the performance of a MISPE. The proposed stitching algorithm uses stereo-vision theory, so it also supports 3D reconstruction. The experimental results show that the revised MISPE system operates stably in real time and increases the endoscope's FOV by 155%. Using the proposed algorithm, the revised MISPE operates at a frame rate to 26.7 fps on a single CPU computer for two endoscopic cameras at a resolution of 640×480 . The proposed stitching method is 1.55 faster and produces results that are closer to ground-truth than the SURF-based method that was used in a previous study by the authors.

Data Availability

The image (video) data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to express their special thanks for research grant provided by the Ministry of Science and Technology (MOST) of Taiwan, Republic of China (ROC), under the contract number 106-2221-E-035-095.

References

- [1] Y. Yamauchi, J. Yamashita, Y. Fukui et al., "A dual-view endoscope with image shift," in *Proceedings of Computer Assisted Radiology and Surgery*, pp. 183–187, Paris, France, June 2002.
- [2] P. Roulet, P. Konen, M. Villegas, S. Thibault, and P. Garneau, "360° endoscopy using panomorph lens technology," in *Proceedings of the SPIE 7558, Endoscopic Microscopy V*, San Francisco, CA, USA, February 2010.
- [3] S.-M. Tseng and J.-C. Yu, "Panoramic endoscope based on convex parabolic mirrors," *Optical Engineering*, vol. 57, no. 3, p. 1, 2018.
- [4] R. Szeliski, "Image alignment and stitching: a tutorial," *Foundations and Trends in Computer Graphics and Vision*, vol. 2, no. 1, pp. 1–104, 2006.
- [5] A. Behrens, T. Stehle, S. Gross, and T. Aach, "Local and global panoramic imaging for fluorescence bladder endoscopy," in *Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6990–6993, Minneapolis, MN, USA, September 2009.
- [6] A. Behrens, M. Bommers, T. Stehle, S. Gross, S. Leonhardt, and T. Aach, "Real-time image composition of bladder mosaics in fluorescence endoscopy," *Computer Science-Research and Development*, vol. 26, no. 1-2, pp. 51–64, 2011.
- [7] J. Liu, B. Wang, W. Hu et al., "Global and local panoramic views for gastroscopy: an assisted method of gastroscopic lesion surveillance," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 9, pp. 2296–2307, 2015.
- [8] L. Yang, J. Wang, T. Ando et al., "Towards scene adaptive image correspondence for placental vasculature mosaic in computer assisted fetoscopic procedures," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 12, no. 3, pp. 375–386, 2016.
- [9] C. Takada, T. Suzuki, A. Afifi, and T. Nakaguchi, *Hybrid Tracking and Matching Algorithm for Mosaicking Multiple Surgical Views*, Springer, Berlin, Germany, 2017.
- [10] C. Takada, A. Afifi, T. Suzuki, and T. Nakaguchi, "An enhanced hybrid tracking-mosaicking approach for surgical view expansion," in *Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3692–3695, Seogwipo, South Korea, July 2017.
- [11] D. T. Kim, V. T. Nguyen, C.-W. Cheng, D.-G. Liu, K. C. J. Liu, and K. C. J. Huang, "Speed improvement in image stitching for panoramic dynamic images during minimally invasive surgery," *Journal of Healthcare Engineering*, vol. 2018, Article ID 3654210, 14 pages, 2018.
- [12] A. Kaehler and G. Bradski, *Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library*, O'Reilly Media, Inc., Sebastopol, CA, USA, 2016.
- [13] K. Konolige, *Small Vision Systems: Hardware and Implementation*, Springer, Berlin, Germany, 1998.
- [14] D. Min, S. Choi, J. Lu, B. Ham, K. Sohn, and M. N. Do, "Fast global image smoothing based on weighted least squares," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5638–5653, 2014.
- [15] X. Hu and P. Mordohai, "A quantitative evaluation of confidence measures for stereo vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2121–2133, 2012.
- [16] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [17] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, UK, 2003.
- [18] V. Kwatra, A. Schödl, I. Essa, G. Turk, and A. Bobick, "Graphcut textures: image and video synthesis using graph cuts," *ACM Transactions on Graphics*, vol. 22, no. 3, pp. 277–286, 2003.
- [19] M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *International Journal of Computer Vision*, vol. 74, no. 1, pp. 59–73, 2007.
- [20] D. T. Kim, Comparison-0 for Panoramic Video, <https://youtu.be/audXZXKdJco>.
- [21] D. T. Kim, Comparison-1 for Panoramic Video, https://youtu.be/XCEq_bceufs.
- [22] D. T. Kim, Comparison-2 for Panoramic Video, https://youtu.be/4_FFcJGUcIo.

