Three Dimensional Panorama Visualization for Endoscopic Video

Atul Kumar[1,2], Yen-Yu Wang[1,2,4], Kai-Che Liu[1,2], Wan-Chi Hung[1,2], Shih-Wei Huang[2],
Wen-Nung Lie[3], and Ching-Chun Huang[3]

[1] IRCAD-Taiwan
[2] Chang Bing Show Chwan Memorial Hospital, Taiwan
[3] National Chung Cheng University, Taiwan
[4] National Changhua University of Education-Taiwan

Abstract—**Introduction: Endoscopy is widely used in the surgical world. Reduced surgical injury during endoscopic surgery makes the patient's life easy. However, the endoscopic surgery becomes a challenge for the novice surgeons because of the narrow field of view and the lack of 3D perception in the 2D endoscope image. Such limitations of the endoscope may be addressed with a 3D panorama created with the endoscopic 2D images. Few studies have been reported to create a 3D panorama from endoscopic images. However, to our knowledge, no study has reported a 3D panorama system which uses only a single camera image from the conventional endoscope to make a 3D shape of the organs, and further stitch those 3D shapes to create a 3D panorama of the surgical scene. Such system would enable surgeons to have an extended field of view of the surgical scene with a 3D perception.**

Method: Images from the endoscopic surgery video were used in this study. A shape from shading (SfS) algorithm was applied to create a 3D shape of the organs in the images. Characteristic feature points were identified on the images using a feature detection algorithm (SURF). The matching feature points in the consecutive images were found using a feature matching algorithm (BRIEF). An iterative closest point (ICP) algorithm was applied to stitch 3D shapes from the consecutive images to synthesize the 3D panorama.

Results: The method was applied on 100 consecutive video frames from an endoscopic video of a patient. The root mean square error for the registration of the consecutive image feature points was <4 mm. The 3D panorama from the method helps in visualizing a larger area of the surgical anatomy. An improved version of the method may be applied to a real time video from the endoscope.

Keywords— **Panorama · Mosaicing · 3D surface reconstruction.**

I. INTRODUCTION

Endoscopic surgery has revolutionized the surgical world. However, it has several limitations creating challenges to the surgeons.

A conventional (single camera) endoscope has limited field of view which can be increased using techniques such as mounting a fish eye lens, and mosaicing techniques. Mosaicing is a technique for stitching the images together to create a larger scene image. The images are acquired from different views with a partial overlap in the images. A panorama image is created by stitching the overlapping images together and making a larger image.

A conventional (single camera) endoscopic image (2D image) does not provide the depth perception of a scene in the image. However, there are several ways of calculating or retrieving the depth map of a scene using its 2D image, which include methods of shape from shading and depth from linear perspective [20]. A 3D structure can be reconstructed by using the depth map with the 2D color image. The 3D structures (surfaces) from overlapping images can be stitched together to form a 3D panorama of the scene.

A panorama using the 3D structures (surfaces) would not only increase the field of view, would also be more informative and suggest the relationship among the surgical anatomies. Moreover, such panorama can be useful in retrieving the complete 3D structure of an organ. Several studies have reported for 2D panorama using conventional endoscope [3],[4],[12],[15],[19] and few studies have reported 3D panorama (reconstruction) using stereo or conventional endoscope [9],[11],[13],[16].

The current study describes a technique which creates a 3D panorama of the surgical anatomy with the 2D endoscopic images and its depth information calculated from the 2D image. To our knowledge, no such study has been reported.

II. METHOD

The method used in the study includes, acquisition of video image using endoscope, applying shape from shading algorithm to each of the video frames to make a 3D structure of the surgical anatomy, features matching between consecutive video frames and aligning 3D structure using the matched feature points. A brief description of the algorithms used in the method is included in the respective section.

A. 3D surface reconstruction

Shape from shading: A shape from shading algorithm (*SfS*) [18] was applied to each of the image frames of the

J. Goh and C.T. Lim (eds.), *7th World Congress on Bioengineering 2015*,
IFMBE Proceedings 52, DOI: 10.1007/978-3-319-19452-3_11

video sequence for calculating the depth map of the scene in the image. The depth map was further used in reconstructing a 3D structure of the surgical anatomy in the image frame. The *SfS* algorithm finds the depth u(x) for each pixel coordinate $\mathbf{x} = (x, y)$ by solving the Hamiltonian equation (Eq. 1)

$$H(\mathbf{x}, \nabla v) = I(\mathbf{x}) \frac{1}{\rho} \sqrt{\left(v_x^2 + v_y^2\right) + J(\mathbf{x}, \nabla v)^2} \cdot Q(\mathbf{x})^{\frac{3}{2}} \quad (1)$$

where $H(\mathbf{x}, \nabla v)$ is the Hamiltonian, $J(\mathbf{x}, \nabla v) = \frac{v_x(x+\alpha) + v_y(y+\beta)+1}{(f+\gamma)}$ and $Q(\mathbf{x}) = (x+\alpha)^2 + (y+\beta)^2 + (f+\gamma)^2$, $v = \ln \mathrm{u}$, $I(\mathbf{x})$ is the intensity at $\mathbf{x}$, α, β and γ are the camera position with respect to the light (in our method camera and light were considered at the same point), v_x, v_y are the partial derivatives of v with respect to the x and y and ρ is the surface albedo (1 in our method). Depth map for one of the image frame is show in Figure 1.

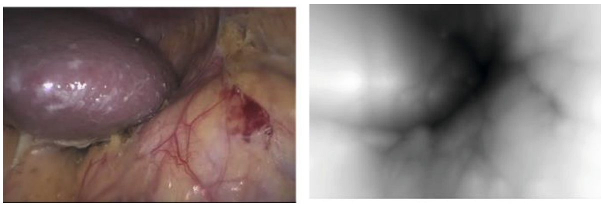

Fig. 1: RGB image (left) and its depth map calculated with SfS algorithm (right)

B. Triangulation and color mapping

The pixel coordinates in a 2D image and its corresponding depth calculated with *SfS* provided the x, y and z point coordinates of each pixel. A Delaunay triangulation technique [8] was used to make a 3D surface from the point coordinates. The points on the surface were mapped with the color of the respective pixel in 2D image. The triangle surface color was interpolated using the bilinear interpolation method [1]. The example surface without color mapping are shown in Figure 2.

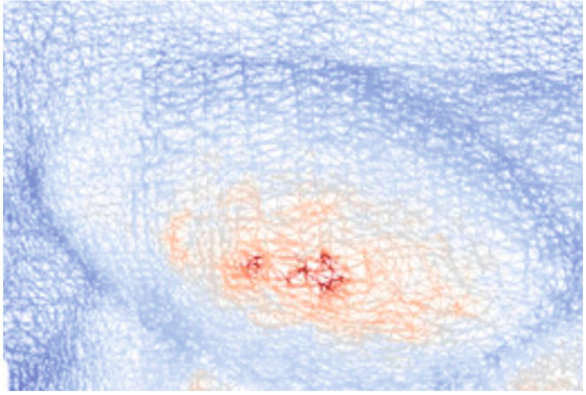

Fig. 2:3D surface mesh generated after Delaunay triangulation

C. Feature detection and matching

Unique feature points on the image frames were detected using the speeded-up robust feature (SURF) detection algorithm [2]. The SURF algorithm uses a Hessian detector where the Hessian is defined as the $\mathcal{H}(\mathbf{x}, \sigma)$ in x at different scale factor σ

$$\mathcal{H}(\mathbf{x}, \sigma) = \begin{bmatrix} L_{xx}(\mathbf{x}, \sigma) & L_{xy}(\mathbf{x}, \sigma) \\ L_{xy}(\mathbf{x}, \sigma) & L_{yy}(\mathbf{x}, \sigma) \end{bmatrix} \quad (2)$$

where $L_{xx}(\mathbf{x}, \sigma)$ is the convolution of the second order Gaussian derivative $\frac{\partial^2}{\partial x^2} g(\sigma)$ in the image I at point $\mathbf{x}$, and similar relationship is to $L_{xy}(\mathbf{x}, \sigma)$ and $L_{yy}(\mathbf{x}, \sigma)$ [28]. The detected feature points were assigned a descriptor known as Binary Robust Independent Elementary Features (BRIEF) which is based upon the intensity differences [7] among the pixels near the feature points. An image patch around the feature point is tested with τ (Eq. 3) to create a descriptor of that patch. In a patch p, τ is defined as

$$\tau(\mathrm{p}; x, y) := \begin{cases} 1 \text{ if } \mathrm{p}(x) < p(y) \\ 0 \quad \text{otherwise} \end{cases} \quad (3)$$

where $\mathrm{p}(\mathbf{x})$ is the pixel intensity at $\mathbf{x}$. The descriptor for the patch is defined as

$$f_{n_d}(p) := \sum_{1 \le i \le n_d} 2^{i-1} \tau(p; x, y) \quad (4)$$

where n_d is the number of binary tests and in our method n_d=128. The feature points (Figure 3) with matching descriptors were then identified with a nearest neighbor search using Hamming distance measurements in the descriptor space [29].

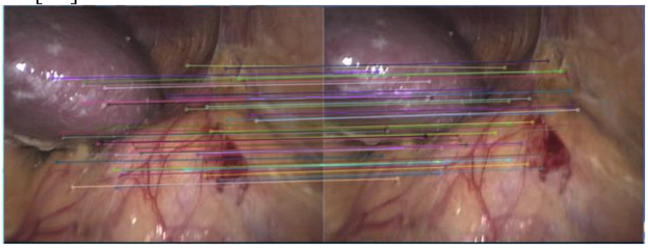

Fig. 3: Matched features in two consecutive image frames

D. Stitching the 3D surface

The pixel position of the points of correspondence (between two consecutive images calculated by the feature matching) provided their 2D coordinates (X,Y). Their Z coordinates were retrieved from their respective 3D surfaces. A rigid body transformation matrix between these matched points was calculated using ICP algorithm [5]. The matrix was used to stitch the 3D surfaces.

III. RESULTS

A recorded video of the laparoscopic surgery, acquired with Karl Storz® laparoscope system (single camera system) (frame size: 720 x 480) was used in this study. The laparoscope camera intrinsic parameters were calculated before the surgery using a camera calibration method proposed by Zhang et.al. [21].

A customized software using libraries of VTK [14] and OpenCV [6] in C++ was prepared for applying the algorithms and to display the 3D panorama. The current method was applied to 100 consecutive frames of the video (with Intel CoreTM i7 960 @ 3.20 GHz, 6.00RAM 64 bit Windows 7, NVidia TESLA C2075) to visualize the 3D panorama (Figure 4). Root mean square error for the alignment between two consecutive frames after ICP was between 1 to 4mm.

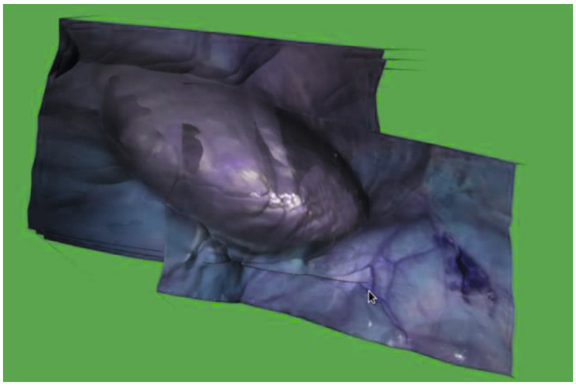

Fig. 4: 3D surface panorama

IV. DISCUSSION

The current work describes a method to create a 3D panorama using images from a single camera endoscope system. The method was applied on 100 consecutive frames of laparoscope video. An improved version of the system will be useful for most of the conventional endoscope systems where surgeons need to visualize a 3D structure of large part of the surgical area. The algorithms with good computational performance [17] were selected for the current system.

Previous studies which describe 3D surface reconstruction from endoscope images have used techniques such as structure from motion (*SfM*) while the current study used shape from shading (*SfS*) method. Compared to the *SfS*, the *SfM* provides a sparse set of 3D points making it difficult to reconstruct a smooth surface of the object. There are several algorithms currently available for the shape-from-shading [10, 20], however, we used the one presented by Visentini-Scarzanella et.al. which is among the recent algorithms for the *SfS* problem and the depth map calculated with the algorithm is more accurate than those from other algorithms [18].

For the feature detection and the feature points matching SURF and BRIEF were used, respectively. The SURF is faster than the many other available algorithms for feature detection [2] and with BRIEF it is promising for the near real-time applications [7].

The presented system would be useful when the surgeons need to visualize 3D shape of the surgical anatomy and an extended view of the surgical field to know the spatial relationship of important structures.

The current method has several limitations such as: (i) it assumes that the organs have moved linearly; (ii) it needs detectable features in every frame of the video; (iii) the specular areas of the endoscope may not have correct depth map; (iv) the cumulative error in the registration (stitching) after some duration of the video was not corrected in this system. In our future work, we would be applying non-linear registration method, the algorithms at GPU level and may include the bundle adjustment in the method.

V. CONCLUSION

A method to create a 3D panorama of the surgical anatomy using single camera endoscope video is reported. The method can be useful for surgeons in visualizing the extended view and the 3D structures of surgical anatomy. It can be improved further by including non-linear registration and bundle adjustment techniques.

ACKNOWLEDGMENT

The study was supported by the National Science Council (Grant application number:103WHK5400012) and Chang Bing Show Chwan Memorial Hospital (Grant application Number:RA 14003)..

CONFLICT OF INTEREST

Conflict of interest: Atul Kumar, Yen-Yu Wang, Kai-Che Liu, Ching-Chun Huang,Wen-Nung Lie, Wan-Chi Hung andShih-Wei Huang declare that they have no conflict of interest.

REFERENCES

1. Allebach, J., Wong, P.W.: Edge-directed interpolation. In: Image Processing, 1996. Proceedings., International Conference on, vol. 3, pp. 707–710. IEEE (1996)
2. Bay, H., Tuytelaars, T., Van Gool, L.: Surf: Speeded up robust features. In: Computer Vision–ECCV 2006, pp. 404–417. Springer (2006)
3. Behrens, A., Bommes, M., Stehle, T., Gross, S., Leonhardt, S., Aach, T.: A multi-threaded mosaicking algorithm for fast image composition of fluorescence bladder images. In: SPIE Medical Imaging, p. 76252S. International Society for Optics and Photonics (2010)
4. Bergen, T., Wittenberg, T., Munzenmayer, C., Chen, C.C.G., Hager, G.D.: A graph-based approach for¨local and global panorama imaging in cystoscopy. In: SPIE Medical Imaging, p. 86711K. International Society for Optics and Photonics (2013)
5. Besl, P.J., McKay, N.D.: Method for registration of 3-d shapes. In: Robotics-DL tentative, pp. 586–606. International Society for Optics and Photonics (1992)
6. Bradski, G.: Dr. Dobb's Journal of Software Tools
7. Calonder, M., Lepetit, V., Strecha, C., Fua, P.: Brief: Binary robust independent elementary features. In: Computer Vision–ECCV 2010, pp. 778–792. Springer (2010)
8. De Berg, M., Van Kreveld, M., Overmars, M., Schwarzkopf, O.C.: Computational geometry. Springer (2000)
9. Duan, M., Xu, R., Ohya, J.: 3d surface reconstruction based on image stitching from gastric endoscopic video sequence. In: SPIE Optical Engineering+ Applications, p. 88561I. International Society for Optics and Photonics (2013)
10. Durou, J.D., Falcone, M., Sagona, M.: Numerical methods for shape-from-shading: A new survey with benchmarks. Computer Vision and Image Understanding 109(1), 22–43 (2008)
11. Ishii, T., Zenbutsu, S., Nakaguchi, T., Sekine, M., Naya, Y., Igarashi, T.: Novel points of view for endoscopy: Panoramized intraluminal opened image and 3d shape reconstruction. Journal of Medical Imaging and Health Informatics 1(1), 13–20 (2011)
12. Konen, W., Breiderhoff, B., Scholz, M.: Real-time image mosaic for endoscopic video sequences. In: Bildverarbeitung fur die Medizin 2007, pp. 298–302. Springer (2007)¨
13. Mountney, P., Yang, G.Z.: Dynamic view expansion for minimally invasive surgery using simultaneous localization and mapping. In: Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE, pp. 1184–1187. IEEE (2009)
14. Schroeder, W., Martin, K., Lorensen, B.: An Object-Oriented Approach To 3D Graphics. Prentice hall (2013)
15. Soper, T.D., Chandler, J.E., Porter, M.P., Seibel, E.J.: Constructing spherical panoramas of a bladder phantom from endoscopic video using bundle adjustment. In: SPIE Medical Imaging, p. 796417. International Society for Optics and Photonics (2011)
16. Soper, T.D., Porter, M.P., Seibel, E.J.: Surface mosaics of the bladder reconstructed from endoscopic video for automated surveillance. Biomedical Engineering, IEEE Transactions on 59(6), 1670–1680 (2012)
17. Vemuri, A.S., Liu, K.C., Ho, Y., Wu, H.S., Ku, M.C.: Endoscopic video mosaicing: application to surgery and diagnostics. In: Living Imaging Workshop, December, pp. 1–2 (2011)
18. Visentini-Scarzanella, M., Stoyanov, D., Yang, G.Z.: Metric depth recovery from monocular images using shape-from-shading and specularities. In: Image Processing (ICIP), 2012 19th IEEE International Conference on, pp. 25–28. IEEE (2012)
19. Yamauchi, Y.: Non-optical expansion of field-of-view of the rigid endoscope. In: World Congress on Medical Physics and Biomedical Engineering 2006, pp. 4184–4186. Springer (2007)
20. Zhang, R., Tsai, P.S., Cryer, J.E., Shah, M.: Shape-from-shading: a survey. Pattern Analysis and Machine Intelligence, IEEE Transactions on 21(8), 690–706 (1999)
21. Zhang, Z.: A flexible new technique for camera calibration. Pattern Analysis and Machine Intelligence, IEEE Transactions on 22(11), 1330–1334 (2000)

The address of the corresponding author:

Author: Atul Kumar
Institute: IRCAD-Taiwan
Chang Bing Show Chwan Memorial Hospital, Taiwan
Street: No 6-1 Lugong Road, Lukang Township
City: Changhua
Country: Taiwan
Email: sharmaatul11@gmail.com