International Journal of Innovative Computing, Information and Control
Volume **16**, Number **3**, June **2020**
ICIC International ©2020 ISSN 1349-4198
pp. 799–812

THREE-DIMENSIONAL PANORAMA IMAGE OF TUBULAR STRUCTURE USING STEREO ENDOSCOPY

Kayo Yoshimoto[1], Kenji Watabe[2], Mizuki Tani[2], Tetsuji Fujinaga[3]
Hideki Iijima[2], Masahiko Tsujii[3], Hideya Takahashi[1]
Tetsuo Takehara[2] and Kenji Yamada[4]

[1]Department of Electrical and Information Engineering
Graduate School of Engineering
Osaka City University
3-3-138 Sugimoto, Sumiyoshi-ku, Osaka 558-8585, Japan
yoshimoto@eng.osaka-cu.ac.jp

[2]Department of Gastroenterology and Hepatology
Graduate School of Medicine
Osaka University
2-2 Yamadaoka, Suita, Osaka 565-0871, Japan

[3]Department of Gastroenterology and Hepatology
Higashiosaka City Medical Center
3-4-5 Nishiiwata, Higashiosaka, Osaka 578-8588, Japan

[4]Faculty of Engineering
Hiroshima Institute of Technology
2-1-1 Miyake, Saeki-ku, Hiroshima 731-5193, Japan

Received October 2019; revised February 2020

Abstract. *The narrow viewing angle of endoscopes impedes obtaining complete images of the digestive tract. In this article, we propose a method to acquire 3D panorama images of the digestive tract using stereo endoscopy. The method consists of acquiring sequential images of the digestive tract by moving the endoscope, reconstructing 3D surfaces for each image by stereo techniques, estimating the endoscope position using scene flow, and reconstructing a 3D image of the digestive tract by merging the surfaces. We verified the method using cylindrical pipes and a pig esophagus covered with printed squares with sides of 10 mm. The estimation error of the endoscope position was within 3% after moving 80 mm. In addition, we mapped the panorama image into a planar image to evaluate the method outcomes. The panorama procedure reduced the number of missing points owing to stereo matching failure or deficiency of resolution and improved the quality of the image. The estimation error of the sides of the reconstructed square was within 2%. The proposed method might be suitable to obtain an expanded view of the digestive tract from stereo endoscopy. This method has the potential to map and quantify the pathologic lesions of the digestive tract.*

Keywords: 3D reconstruction, Stereo endoscope, Panorama image, Digestive tract, Scene flow

1. **Introduction.** Endoscopy is a well-established procedure based on precise images for diagnosing various diseases in the digestive tract. However, endoscopists might find it difficult to map and quantify pathologic lesions when using conventional endoscopes owing to the limited field of view and inability to retrieve the orientation from the images. To overcome these problems, several endoscopic images can be combined to create a broader

DOI: 10.24507/ijicic.16.03.799

field of view [1]. In this study, we adopted a similar combination method to generate 3D panorama images.

In principle, a panorama image is generated in two stages, namely, reconstructing the underlying surface geometry and estimating the camera motion. Previous efforts by other researchers aimed to generate panorama images of tubular-shaped organs such as the esophagus using 2D images [2,3]. Given that the direction of view in a tube usually coincides with that of the motion, these approaches employ a cylindrical model (e.g., pipe projection) to approximate the organ shape. In this study, they mapped the images into a cylindrical surface by unwrapping them around an estimated projection center and estimated the camera motion between sequential images using an affine optical-flow technique.

We previously developed a prototype stereo endoscopy using a compound eye system called the Thin Observation Module by Bound Optics (TOMBO) [4,5]. Based on triangulation, TOMBO enables mapping of the depth of each point in the image and performs 3D measurements. Following that development, in this article, we propose a method to generate 3D panorama images of the digestive tract using stereo endoscopy. A preliminary version of this work has been reported [6]. In this paper, we described mathematical basis of our method in detail. We validated our method using a phantom not only by size estimation of the texture as in a preliminary version but also by estimation of moving distance of the endoscope in this study. We demonstrated the advantage of our proposed method using 3D images over the previous one using 2D images. We further validated our method using biomaterial: pig esophagus.

2. **Methods and Procedures.** Figure 1 illustrates the proposed method to generate 3D panorama images of the digestive tract. The stages are described as follows. First,

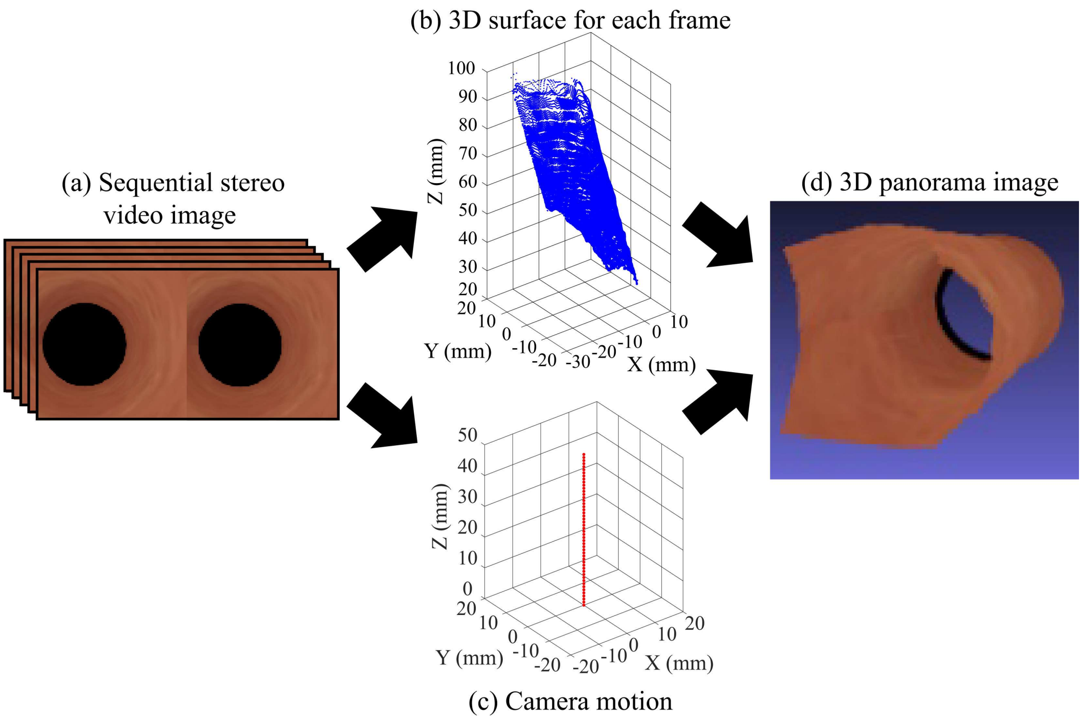


FIGURE 1. Proposed method to generate 3D panorama image of digestive tract: (a) capture sequential video frames using endoscope, (b) reconstruct 3D surfaces for each frame, (c) estimate motion of endoscope camera, and (d) reconstruct 3D panorama image by overlaying surfaces

sequential endoscopic video frames of the digestive tract are captured by moving the endoscope (Figure 1(a)). Then, 3D points of the surface are reconstructed for each frame using triangulation and stereo imaging (Figure 1(b)). Next, the camera motion is estimated using the scene flow of consecutive frames (Figure 1(c)). Finally, the 3D panorama image is retrieved by overlaying the surfaces obtained for each frame (Figure 1(d)). In the sequel, we detail each stage of the proposed method.

2.1. **Reconstruction of 3D surfaces.** For 3D measurements, the proposed system employs a compound eye system that is called TOMBO. TOMBO is comprised of a microlens array, partition, and CMOS image sensor. For stereo matching, we apply an area-based method with the Sum of the Squared Difference (SSD) to obtaining the corresponding points in two consecutive images. In addition, we apply a subpixel estimation with a parabolic function and cross-checking by switching the left and right images to improve the measurement accuracy. By determining correspondent points among all pixels in an image, we estimate the 3D points $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$ corresponding to the object surface.

2.2. **Camera motion estimation.** To determine the endoscopic camera 3D motion, we calculate the 3D scene flow from sequential stereo images. As in the reconstruction, we apply the area-based method with SSD to identifying a point in the previous and current left and right images, as illustrated in Figure 2(a). Likewise, we improve the measurement accuracy by applying subpixel estimation with a parabolic function and cross-checking by switching the left and right images. Then, we estimate the 3D scene flow by comparing the correspondent points between the previous and current images, obtaining the motion direction shown in Figure 2(b). Considering a digestive tract neither deformed by peristalsis nor decompressed by surrounding organs, we can assume that the scene flow is mostly composed of camera motion expressed as

$$\boldsymbol{p}' = \boldsymbol{p}\boldsymbol{R} + \boldsymbol{1}\boldsymbol{T}, \tag{1}$$

where $\boldsymbol{p} = [\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]$ is an $n \times 3$ matrix of previous 3D points estimated by triangulating stereo images; $\boldsymbol{p}' = [\boldsymbol{X}', \boldsymbol{Y}', \boldsymbol{Z}']$ is an $n \times 3$ matrix of correspondent current points; $\boldsymbol{R}$ and $\boldsymbol{T}$ are a 3×3 camera rotation matrix and camera translation vector with three elements, respectively; and $\boldsymbol{1}$ is the $n \times 3$ matrix of ones. Rotation matrix $\boldsymbol{R}$ and translation vector $\boldsymbol{T}$ are obtained from the following equation [7]:

$$\min \left\| \boldsymbol{p}' - (\boldsymbol{p}\boldsymbol{R} + \boldsymbol{1}\boldsymbol{T}) \right\|_{\mathrm{F}}^2 . \tag{2}$$

The relation between $\boldsymbol{T}$ and $\boldsymbol{R}$ is expressed as

$$\boldsymbol{T} = \bar{p'} - \bar{p}\boldsymbol{R}, \tag{3}$$

where $\bar{p'}$ and $\bar{p}$ are the averages of $\boldsymbol{p}'$ and $\boldsymbol{p}$, respectively. Finally, $\boldsymbol{R}$ is estimated using singular value decomposition. Furthermore, considering the influence of outliers caused by matching errors, we use the Random Sample Consensus (RANSAC) algorithm for the final camera motion estimation [8]. The rotation matrix $\boldsymbol{R}$ and the translation matrix $\boldsymbol{T}$ are estimated using three randomly sampled points, and the parameter with the smallest sum of squares of the error for all points is adopted. Thus, RANSAC reduces the influence of outliners when robustly estimating camera motion parameters.

2.3. **Merging 3D surfaces.** After estimating $\boldsymbol{R}$ and $\boldsymbol{T}$ from sequential images using scene flow, the measured 3D points can be arranged with respect to a starting position, that is, the points are transformed from the camera to a world coordinate system. We define the lens center at the initial position as the origin of the world coordinate system,

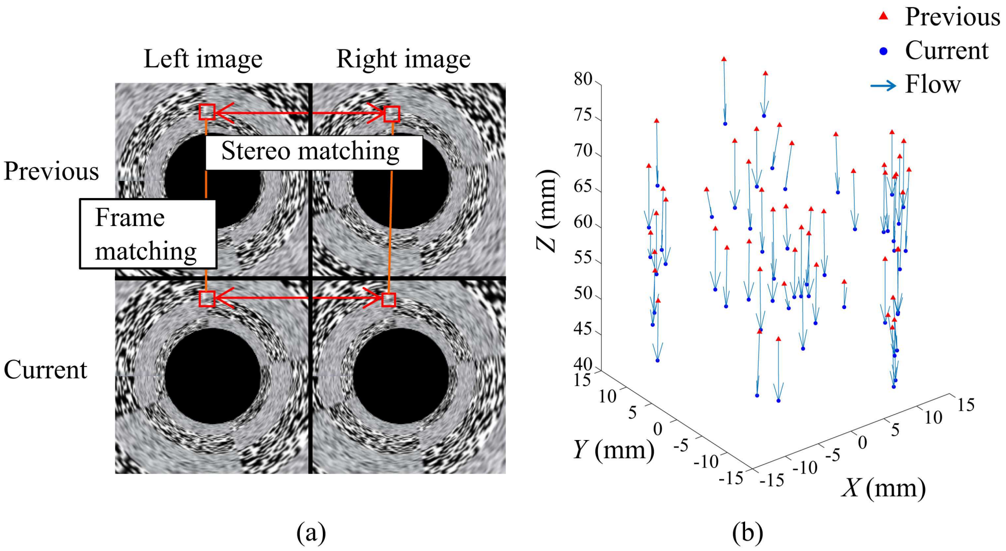


FIGURE 2. Method to obtain 3D flow of object surface: (a) correspondent points in sequential stereo images and (b) estimation of 3D scene flow

and express the transformation of 3D point $\boldsymbol{p}_t$ in the camera coordinate system into point $\boldsymbol{pw}_t$ in the world coordinate system as

$$\boldsymbol{pw}_t = \boldsymbol{A}_t^{\mathrm{T}} \boldsymbol{p}_t, \tag{4}$$

where $\boldsymbol{A}_t$ is the transformation matrix between the world and camera coordinate systems, and t represents the frame number ($t = 0, 1, 2, \ldots$). Transformation matrix $\boldsymbol{A}_t$ can be calculated from the estimated camera motion as

$$\boldsymbol{A}_t = \begin{bmatrix} & \boldsymbol{R}_t & & \boldsymbol{T}_t \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} \boldsymbol{A}_{t-1} \ (t = 1, 2, 3, \ldots), \tag{5}$$

$$\boldsymbol{A}_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{6}$$

where $\boldsymbol{R}_t$ and $\boldsymbol{T}_t$ are, respectively, the camera rotation matrix and translation vector between the t-th and $(t-1)$-th frame. Therefore, a panoramic 3D surface can be obtained by expressing all of the points with respect to the starting position.

2.4. **Mapping 3D panorama image to planar image.** To evaluate the outcome of the tubular 3D panorama image, 3D panorama images are mapped to planar images. Here, the tubular 3D panoramic image was mapped using cylindrical coordinates by approximating the tubular 3D panoramic image to a cylinder.

To map the 3D panorama images, the following cylinder parameters were estimated: three parameters that determine cylinder direction axis $\boldsymbol{a}$, three that determine point $\boldsymbol{b} = (b_x, b_y, b_z)$ on the cylinder axis, and radius r, as shown in Figure 3. We approximated the cutting plane of the cylinder as a circle and estimated the center position of the circles on the merged 3D surface at each distance D from the origin in the world coordinate system using the least square method. The center position of the circle is expressed in space as (c_x, c_y, D). By determining the line that passes through the center of all the circles, we estimated direction axis $\boldsymbol{a}$. Likewise, we estimated radius r of the cylinder using the distance between merged tubular 3D surface and axis $\boldsymbol{a}$. Point $\boldsymbol{b}$ was estimated as the intersection of the z-axis and the cylinder axis.

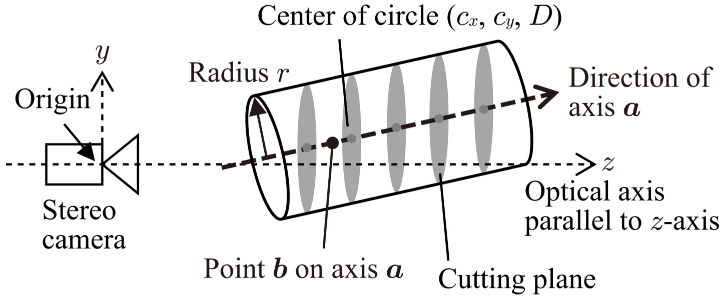


FIGURE 3. Parameters to determine a cylinder

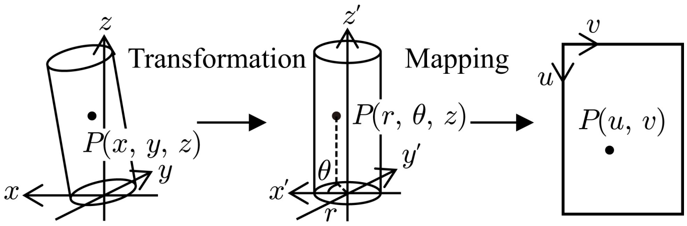


FIGURE 4. Methods to map from 3D tubular panorama to planar image

Finally, to unroll the obtained 3D points, we transformed the points from rectangular into cylindrical coordinates, as depicted in Figure 4. The merged 3D points given by (x, y, z) on the world coordinate system were transformed to 3D points given by (x', y', z') on the coordinate system that z' axis is parallel to the cylinder axis using the following equations:

$$[x', y', z', 1] = [x, y, z, 1]\boldsymbol{T}_b\boldsymbol{R}_y\boldsymbol{R}_x, \tag{7}$$

$$\boldsymbol{T}_b = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -b_x & -b_y & -b_z & 1 \end{bmatrix}, \tag{8}$$

$$\boldsymbol{R}_y = \begin{bmatrix} \cos\theta_y & 0 & \sin\theta_y & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{9}$$

$$\boldsymbol{R}_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_x & \sin\theta_x & 0 \\ 0 & -\sin\theta_x & \cos\theta_x & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{10}$$

θ_x and θ_y are the angle between the z-axis and a vector obtained by orthographic projection of vector $\boldsymbol{a}$ onto a y-z plane and a x-z plane, respectively. $\boldsymbol{R}_x$, $\boldsymbol{R}_y$, and $\boldsymbol{T}_b$ are rotation matrix and translation matrix for aligning the z-axis with z'-axis.

Then, the rectangular coordinate system (x', y', z') is converted into the cylindrical coordinate system (r, θ, z') by

$$r = \sqrt{(x')^2 + (y')^2}, \tag{11}$$

$$\theta = \tan^{-1}\left(\frac{y'}{x'}\right), \tag{12}$$

$$z' = z'. \tag{13}$$

In the mapped planar image, the horizontal axis represents circumferential direction θ, and the vertical axis represents axial direction z':

$$u = sz', \tag{14}$$

$$v = sr(\theta + \pi). \tag{15}$$

Coordinates (u, v) determine the planar image, where s is a scaling factor given in pixel/mm.

2.5. **Mapping 2D endoscopic image to planar image.** In order to compare our proposed method with conventional method, a planar image is generated from 2D endoscopic images [2]. By assuming that the image is obtained from a point on the center axis of the cylinder, the "unrolled" image is created from the 2D endoscopic image by the following equation:

$$\Theta = \tan^{-1}\left(\frac{v'}{u'}\right), \tag{16}$$

$$R = \sqrt{(u')^2 + (v')^2}, \tag{17}$$

where (u', v') is the coordinates of the 2D endoscope image, and (Θ, R) is the coordinates of the unrolled planar image. (u', v') has the origin at the center of the circle. The horizontal information Θ is the rotation angle from the u'-axis, and the vertical information R is the distance from the center of the circle. After the images are captured and unrolled, unrolled images are manually stitched to make planar image.

2.6. **Statistical analysis.** Results are expressed as mean $\pm$ Standard Error of the Mean (SEM). Comparisons between more than three groups were performed by Tukey test. A P value of less than 0.05 was considered statistically significant. All data analysis was performed with the JMP Pro 13.2.1 statistical package (Statistical Analysis Systems, Cary, NC).

3. Experiment.

3.1. **Experimental setup.** We confirmed the validity of the proposed method by acquiring the image of the tubular structure from the stereo endoscope system. To examine the basic performance, 26-mm, 31-mm, and 36-mm cylindrical pipes were used (Figure 5(a)). For evaluation, a square pattern of 10 mm on a side (Figure 5(b)) was placed inside the pipe. In addition, to verify the adaptability to mucous tissue, a similar experiment using a pig esophagus was performed. Because the pig esophagus was too small for our endoscope, the esophagus was placed on a semicylindrical structure created using a 3D printer, as shown in Figure 6. Square patterns were placed on the esophagus as a substitute for pathological lesions. The ethics committee for animal experimentation of Osaka University Graduate School of Medicine approved the animal experiments that were reported in this study (27-062-002).

Table 1 lists the specifications of the endoscopic system. We defined the coordinate system for the endoscope as follows: z-axis parallel to the optical axis, and x- and y-axes parallel to the image sensor. The experiment was conducted after the optics were calibrated. The 3D endoscope was fixed on a stage moving along the z-axis and obtained images while moving by 2 mm from the start position (0 mm) to the end position (80 mm).

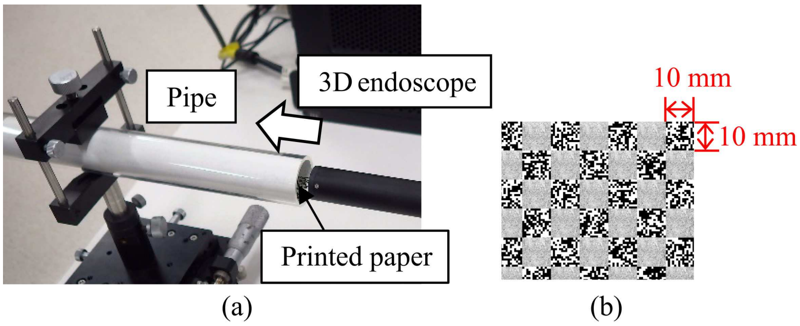


FIGURE 5. Experimental setup: (a) 3D endoscope fixed for 1D manual translation within pipe, and (b) printed square patterns inside pipe

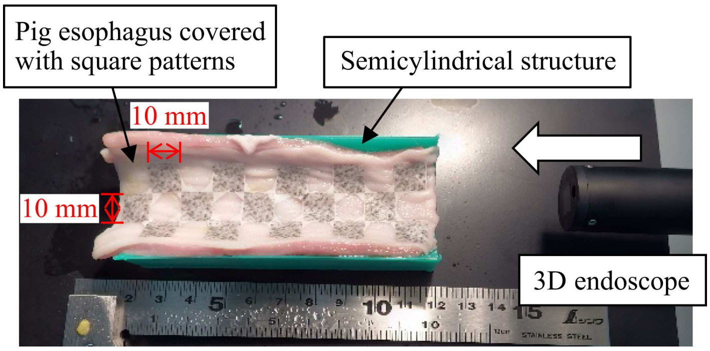


FIGURE 6. Experimental setup using pig esophagus. Pig esophagus covered with square patterns was placed on semicylindrical structure.

TABLE 1. Specifications of stereo endoscopic system

Parameter	Value
Focal length	1.5 mm
F-number	7.5
Angle of field	46°
Depth of field	20-50 mm
Pixel size	2.2 μm
Pixel count (per unit)	415 × 415
Tip diameter	20 mm

In general, the resolution of stereo measurement degrades with distance between the target and camera. In addition, the viewing angle of the endoscope limits the depth range of the target used for the 3D panorama image depending on the cylinder radius. We experimentally set the depth range such that the rate of area where the cylinder wall appears on the image sensor is 60% when the cylinder and optical axes coincide. The depth range used in the experiment was from 20 mm to 60 mm for 26-mm-diameter pipe, from 36 mm to 71 mm for 31-mm-diameter, and from 42 mm to 83 mm for 36-mm-diameter, respectively.

3.2. **Experimental results for pipes.** Experiments using pipes were performed three times for each size. Figure 7 shows the results of using a 26-mm-diameter pipe. Figure 7(a) shows images captured by the two eyes of the endoscopic system. Figure 7(b) shows a 3D panorama image merged by the proposed method from the 41 video frames obtained by moving the endoscope along the z-axis in increments of 2 mm from 0 mm to 80 mm.

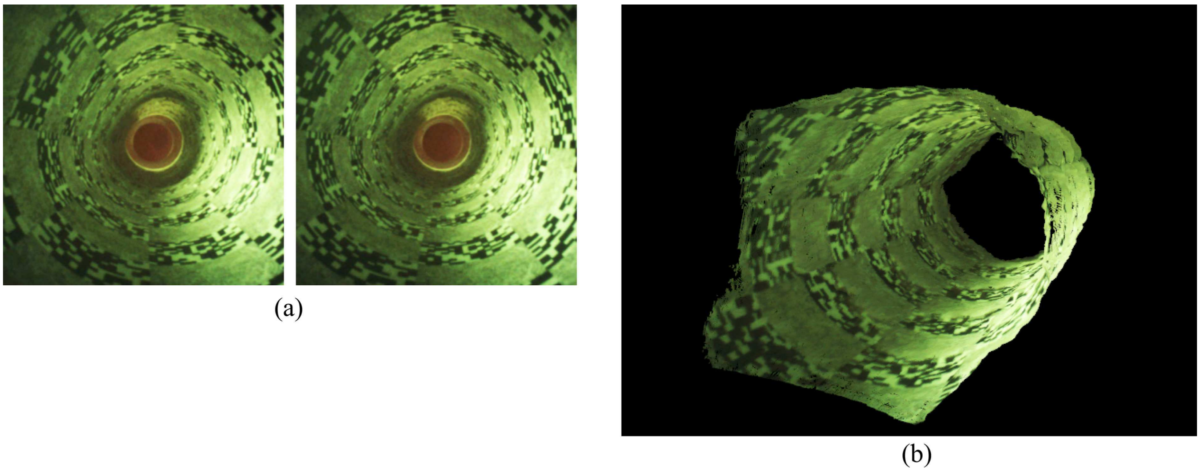


Figure 7. Three-dimensional panorama image of pipe with diameter of 26 mm and printed squares with sides of 10 mm. (a) Images acquired from two eyes of endoscopic system. Endoscope was moved inside cylinder to capture sequential video frames. (b) Three-dimensional panorama image merged from sequential video frames (41 frames). Images shown after adjusting brightness for better identification of patterns.

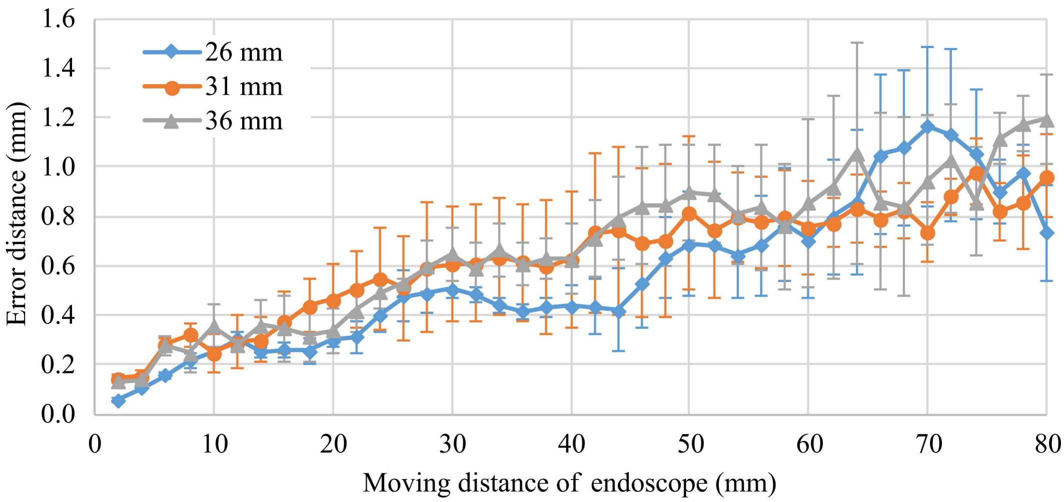


Figure 8. Error of estimated endoscope position using pipes. Error bars represent SEM.

To verify the accuracy of the camera position estimation, we calculated the error distance. We defined the error distance as that between the true position and the estimated position of the endoscope. Figure 8 shows the error distance when using 26-mm, 31-mm, and 36-mm pipes. The error distance was within 2 mm (error rate 2.5%). There was no significant difference in the error distance at 80 mm between pipes of different diameters.

Then, to verify the reconstruction accuracy of the 3D panorama image, we generated a planar image from the 3D panorama image. Figure 9 shows 2D planar image of a 26-mm-diameter pipe. Figure 9(a) shows the 2D planar image mapped from 3D image using a single video frame with a resolution of 1 mm per 5 pixels. Black areas indicate missing points owing to stereo matching failure or deficiency of resolution. In the single-frame image shown in Figure 9(a), 39.5% of the points are missing at distances

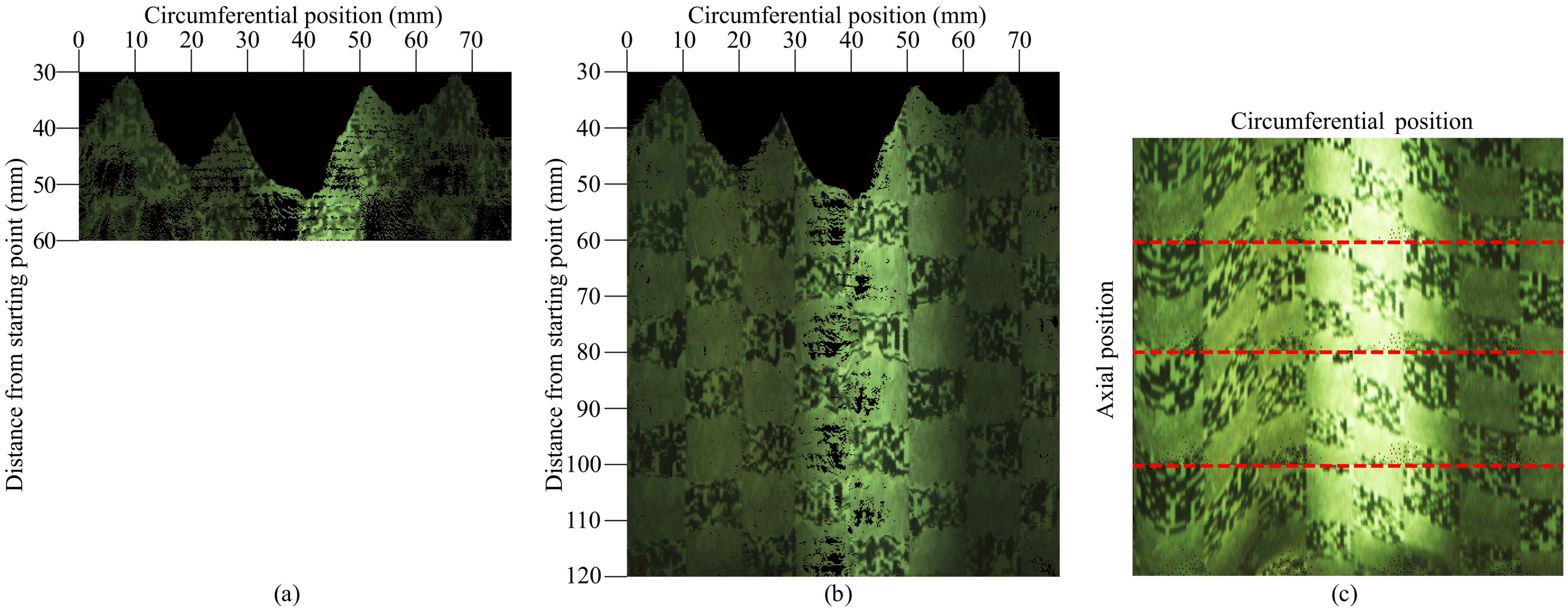


FIGURE 9. Planar image of 26-mm-diameter pipe. (a) Mapped image from 3D image of single frame. (b) Mapped image from 3D panorama image with 41 frames merged. (c) Unrolled image from 2D endoscopic image using the conventional method. The red dotted line represents the position where unrolled images are stitched manually.

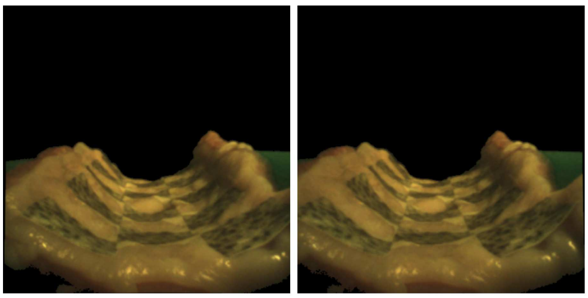
from 50 mm to 60 mm, thus severely compromising the visual recognition of the square pattern. Figure 9(b) shows the 2D planar image mapped from the 3D panorama image using sequential frames. In this case, the rate of missing points decreases to 5.6%, and the square pattern is easily recognized. The missing rate decreases from 47.4% to 1.5% at a diameter of 31 mm (distance 61-71 mm), and from 59.8% to 8.4% at a diameter of 36 mm (distance 73 mm-83 mm). Then, we determined the square sides by dividing the number of pixels between their vertices on the mapped image. Table 2 lists the square estimations. The mean error rate of the measurement size of 10-mm squares was less than 1%. Figure 9(c) shows the 2D planar image generated from 2D endoscopic images using the conventional method described in Section 2.5. The 2D image was captured and reconstructed at intervals of 20 mm. The red dotted line represents the position where unrolled images are stitched manually. The ruler cannot be placed on this image because the original 2D images do not hold no length information. While the actual squares are equally sized and located on a line, the squares in Figure 9(c) are different in size and are located on a curve. In addition, the square pattern is discontinuous at the junction of stitching.

3.3. **Experimental results for pig esophagus.** Figure 10 shows the results of using a pig esophagus. Figure 10(a) shows images captured by the two eyes of the endoscopic

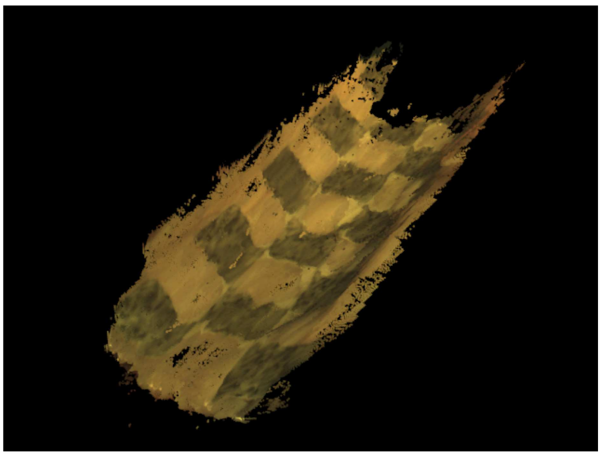
TABLE 2. Measured size of 10-mm squares in mapped image from 3D panorama image using pipes

Pipe diameter	Circumferential			Axial		
	Mean ± SEM (mm)	Error rate (%)	n	Mean ± SEM (mm)	Error rate (%)	n
26 mm	9.95 ± 0.035	0.51	120	10.00 ± 0.049	0.01	114
31 mm	10.06 ± 0.025	0.63	189	10.09 ± 0.044	0.94	162
36 mm	10.05 ± 0.032	0.47	159	9.97 ± 0.051	0.34	159

n: Number of measured sides



(a)



(b)

FIGURE 10. Three-dimensional panorama image of pig esophagus covered with printed squares with sides of 10 mm. (a) Images acquired from two eyes of endoscopic system (first frame). Endoscope was moved along z-axis to capture sequential video frames. (b) Three-dimensional panorama image merged from sequential video frames (41 frames).

system. Figure 10(b) shows a 3D panorama image merged by the proposed method from 41 video frames obtained by moving the endoscope along the z-axis in increments of 2 mm from 0 mm to 80 mm. Figure 11 shows the error distance with respect to the displacement of the endoscope. The error distance was also within 2 mm from the start position to the end position (80 mm). Figure 12 shows a 2D planar image mapped from

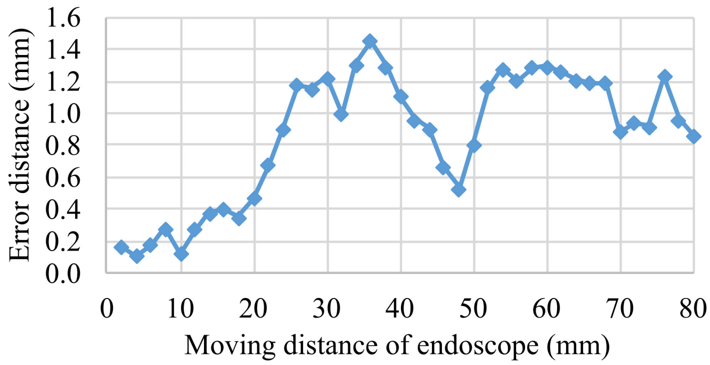


FIGURE 11. Error of estimated endoscope position using a pig esophagus

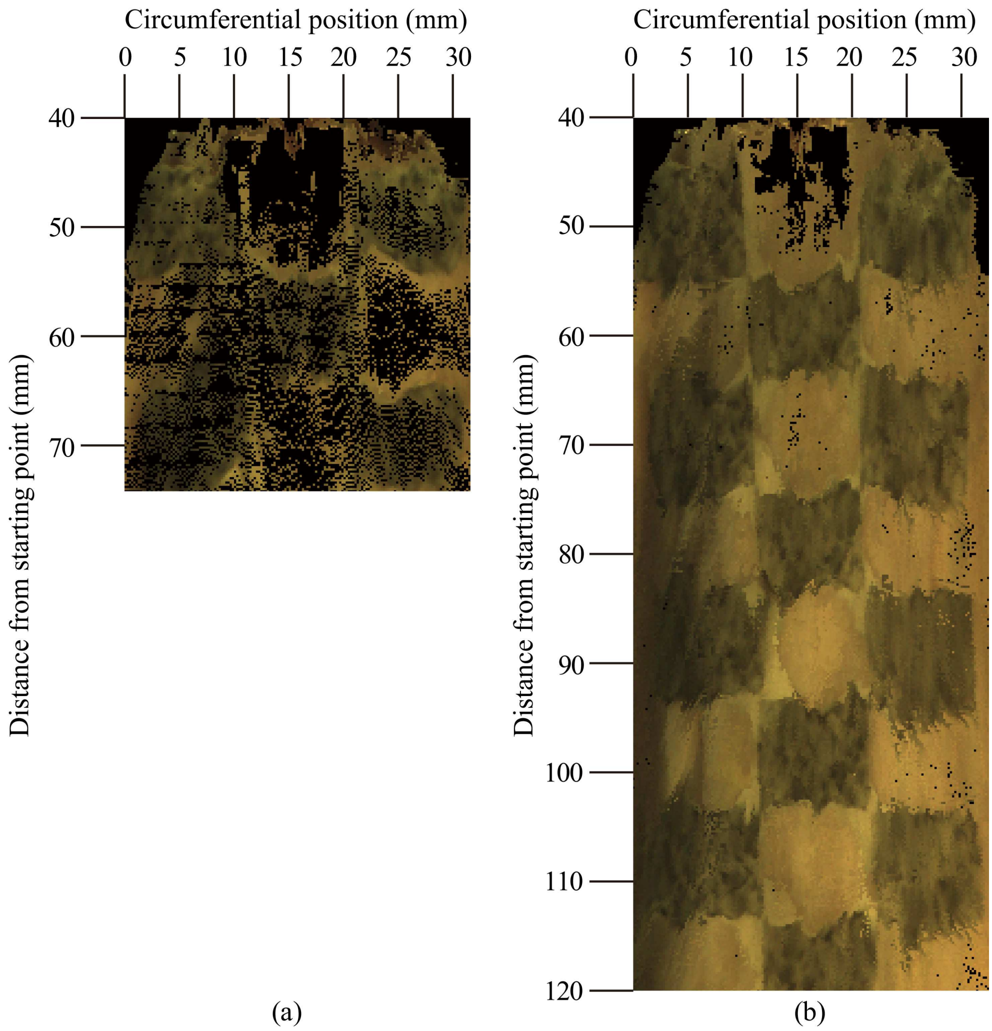


FIGURE 12. Planar image mapping from 3D images of pig esophagus: (a) mapped image from 3D image of single frame, and (b) mapped image from 3D panorama image with 41 frames merged

the 3D panorama image of the semicylindrical pig esophagus. The accuracy of the camera position estimation is shown in Figure 11. The error distance after moving 80 mm was within 1.6 mm (error rate 2%). Figure 12(a) shows the mapping 3D panorama image from a single video frame, and Figure 12(b) shows the image mapped from the 3D panorama image using sequential frames. In this case, the missing rate at distances from 50 mm to 60 mm decreases from 54% to 0.4%. Even in this case, it is possible to easily recognize a square pattern by merging a plurality of frames. Table 3 lists the square estimations for the pig esophagus. These results confirm that this system can measure the size of 10-mm squares within an error rate of 2% even for mucosal tissue.

Table 3. Measured size of 10-mm squares in mapped image from 3D panorama image using pig esophagus

	Mean $\pm$ SEM (mm)	Error rate (%)	n
Circumferential	9.93 ± 0.09	0.66	7
Axial	9.82 ± 0.16	1.78	12

n: Number of measured sides

4. **Discussion.** The proposed method to generate panorama images provides complete views of tubular structures that are usually hindered by the limitations on the field of view and orientation of conventional endoscopy. In this study, we demonstrated the generation of 3D panorama images of tubular structures using a stereo endoscope and estimated the camera position using scene flow to then reconstruct a complete 3D surface. The 3D panorama image is obtained by overlaying the 3D surfaces corresponding to sequential frames of endoscopic images. To test the proposed method, we generated a 3D panorama image of a tubular structure 12 cm long and thus expanded the endoscopic view. Furthermore, the validation of the proposed method revealed a seamless and accurate reconstruction of 3D tubular structures.

The unrolled 3D panorama image generated from a single video frame resulted in several missing points, thus preventing visual recognition of the printed square pattern inside the cylinder. This suggests that further improvements can be implemented in stereo matching for the proposed method. The missing points notably decreased in the unrolled 3D panorama image generated from 41 sequential video frames, thus allowing for clear pattern identification. This improved reconstruction could be related to the increased amount of data because the reconstructed 3D surfaces overlapped within the depth of field for the panorama image generation. In fact, a point in the cylinder is measured from different viewpoints when located within the depth of field, and this information was combined for reconstruction and to generate the unrolled image.

Previous works on panorama images of tubular structures using endoscopy provided methods to obtain 2D panorama images based on a pipe projection model and a conventional 2D endoscope [2]. In this study, we compared our proposed method with previous works. In the unrolled image generated by the conventional method, the image was distorted and discontinuous. Conventional method employs a cylindrical model where the direction of the camera view in a tube coincides with that of the camera motion. However, in the real setting of video capturing, it is hard to keep the camera view on the axis of the tube. The result demonstrates that deviation of the camera view from the axis significantly impairs the quality of panorama image in the conventional method. On the other hand, in our proposed method, the cylindrical axis is automatically corrected to match the endoscope axis using 3D information. Furthermore, because 2D images created by the conventional method do not have length information, it is impossible to measure the length of the pattern in the unrolled image. Stitching procedure was performed manually and resulted in discontinuation of the planar image. Therefore, our proposed method shows advantage over the conventional method of automatic axis correction, automatic merging of consecutive frames, and length measurement.

The clinical application of the 3D panorama images remains to be evaluated. In this article, we demonstrated that the panorama images faithfully reproduce a square pattern in size and shape. These results suggest that 3D panorama images may allow for the accurate quantification of areas of pathologic lesions such as Barrett's esophagus. This can lead to improved diagnoses as the prevalence of adenocarcinoma in patients with long segment Barrett's esophagus is significantly higher than in those with short segment [9]. However, conventional 2D endoscopic observations result in reported interobserver differences of the Barrett's esophagus length of up to 1.4 cm [10], and 2D panorama images have been shown to reduce this difference [11]. Given that Barrett's esophagus is located in the lower part of the esophagus, its shape is constricted by the lower esophageal sphincter, and hence generating 2D panorama images based on a pipe projection model cannot be suitably applied in such constricted conditions. In principle, 3D panorama images could be used to measure the area of Barrett's esophagus more accurately than their 2D counterparts.

In this study, we demonstrated that 3D panorama images reproduced small squares in both size and distribution throughout the tubular structure. Therefore, the proposed method could be used to completely record Lugol-voiding lesions in the esophagus. Multiple Lugol-voiding lesions represent a risk factor in the development of squamous cell carcinoma in the esophagus, head, and neck [12]. However, a scoring system for the severity of these lesions has not been established, and their clinical evolution has not been investigated. As 3D panorama images can potentially register a complete tubular inner surface, they may be used to track the evolution of Lugol-voiding lesions and establish a scoring system to determine the risk of developing a related cancer.

Overall, this study was a proof-of-concept on the generation of 3D panorama images of tubular structures. Future developments will aim to demonstrate the generation of such images to measure complex tubular structures, such as those deriving from decompression by surrounding organs, depressed lesions such as ulcers, and protruded lesions such as polyps.

Acknowledgment. This work was supported in part by JSPS KAKENHI grant numbers JP15K09057, JP15K16339, and JP18K08003.

REFERENCES

[1] T. Bergen and T. Wittenberg, Stitching and surface reconstruction from endoscopic image sequences: A review of applications and methods, *IEEE J. Biomed. Heal. Informatics*, vol.20, no.1, pp.304-321, 2016.

[2] R. Kim, B. B. Baggott, S. Rose, A. O. Shar, D. L. Mallory, S. S. Lasky, M. Kressloff, L. Y. Faccenda and J. C. Reynolds, Quantitative endoscopy: Precise computerized measurement of metaplastic epithelial surface area in Barrett's esophagus, *Gastroenterology*, vol.108, no.2, pp.360-366, 1995.

[3] E. J. Seibel, R. E. Carroll, J. A. Dominitz, R. S. Johnston, C. D. Melville, C. M. Lee, S. M. Seitz and M. B. Kimmey, Tethered capsule endoscopy, a low-cost and high-performance alternative technology for the screening of esophageal cancer and Barrett's esophagus, *IEEE Trans. Biomed. Eng.*, vol.55, no.3, pp.1032-1042, 2008.

[4] Y. Kitamura, R. Shogenji, K. Yamada, S. Miyatake, M. Miyamoto, T. Morimoto, Y. Masaki, N. Kondou, D. Miyazaki, J. Tanida and Y. Ichioka, Reconstruction of a high-resolution image on a compound-eye image-capturing system, *Appl. Opt.*, vol.43, no.8, pp.1719-1727, 2004.

[5] K. Yoshimoto, K. Yamada, K. Watabe, M. Takeda, T. Nishimura, M. Kido, T. Nagakura, H. Takahashi, T. Nishida, H. Iijima, M. Tsujii, T. Takehara and Y. Ohno, Gastric contraction imaging system using a 3-D endoscope, *IEEE J. Transl. Eng. Heal. Med.*, vol.2, p.1800208, 2014.

[6] K. Yoshimoto, K. Watabe, T. Fujinaga, H. Iijima, M. Tsujii, H. Takahashi, T. Takehara and K. Yamada, Evaluation of endoscopic entire 3D image acquisition of the digestive tract using a stereo endoscope, *Proc. of SPIE*, vol.10054, p.1005412, 2017.

[7] K. S. Arun, T. S. Huang and S. D. Blostein, Least-squares fitting of two 3-D point sets, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.PAMI-9, no.5, pp.698-700, 1987.

[8] M. A. Fischler and R. C. Bolles, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Commun. ACM*, vol.24, no.6, pp.381-395, 1981.

[9] A. P. Weston, P. T. Krmpotich, R. Cherian, A. Dixon and M. Topalosvki, Prospective long-term endoscopic and histological follow-up of short segment Barrett's esophagus: Comparison with traditional long segment Barrett' s esophagus, *Am. J. Gastroenterol.*, vol.92, no.3, pp.407-413, 1997.

[10] R. Dekel, D. E. Wakelin, C. Wendel, C. Green, R. E. Sampliner, H. S. Garewal, P. Martinez and R. Fass, Progression or regression of Barrett's esophagus – Is it all in the eye of the beholder?, *Am. J. Gastroenterol.*, vol.98, no.12, pp.2612-2615, 2003.

[11] R. Kim, S. Rose, A. O. Shar, M. Weiner and J. C. Reynolds, Extent of Barrett's metaplasia: A prospective study of the serial change in area of Barrett's measured by quantitative endoscopic imaging, *Gastrointest. Endosc.*, vol.45, no.6, pp.456-462, 1997.

[12] C. Katada, T. Yokoyama, T. Yano, K. Kaneko, I. Oda, Y. Shimizu, H. Doyama, T. Koike, K. Takizawa, M. Hirao, H. Okada, T. Yoshii, K. Konishi, T. Yamanouchi, T. Tsuda, T. Omori, N. Kobayashi, T. Shimoda, A. Ochiai, Y. Amanuma, S. Ohashi, T. Matsuda, H. Ishikawa, A. Yokoyama and M. Muto, Alcohol consumption and multiple dysplastic lesions increase risk of squamous cell carcinoma in the esophagus, head, and neck, *Gastroenterology*, vol.151, no.5, pp.860-869.e7, 2016.