Deformable Registration for Generating Dissection Image of an Intestine from Annular Image Sequence

Suchit Pongnumkul, Ryusuke Sagawa,
Tomio Echigo, and Yasushi Yagi

The Institute of Scientific and Industrial Research, Osaka University,
8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, Japan
{p-suchit, sagawa, echigo, yagi}@am.sanken.osaka-u.ac.jp

Abstract. Examination inside an intestine by an endoscope is difficult and time consuming, because the whole image of the intestine cannot be taken at one time due to the limited field of view. Thus, it is necessary to generate a dissection image, which can be obtained by extending the image of an intestine. We acquire an annular image sequence with an omnidirectional or wide-angle camera, and then generate the dissection image by mosaicing the image sequence. Though usual mosaicing techniques transform an image by perspective or affine transformations, these are not suitable for our situation because the target object is a generalized cylinder and the camera motion is unknown a priori. Therefore, we propose a novel approach for image registration that deforms images by a two-dimensional-polynomial function which parameters are estimated from optical flow. We evaluated our method by registering annular image sequences and we successfully generated dissection images, as presented in this paper.

1 Introduction

Examining a patient's intestine using an endoscope is difficult and time consuming, because the field of view of a camera is restricted and the doctor can only look at a small part of the intestine at a time. It would be desirable for an easy-to-understand view to be available for medical doctors. This is especially important when a lot of data is involved, for example, when a doctor compares patients among those who have similar symptom.

The most easy-to-understand way for an examination is to operate and cut open the intestine for a view of the whole intestine, but it is impossible for many reasons. The generation of images, which look like cut-and-expanded intestine, is required. Thus by using this representation a doctor can examine the wide area of an intestine at a glance without performing an operation. Videos that are taken by an endoscope can be used to generate such images by utilizing an adapted video mosaicing technique.

A number of methods have been explored and proposed in the video mosaicing literature. S. Peleg and J. Herman [9] developed a panoramic mosaicing method using manifold projection. Szeliski [13] developed an image-based video mosaicing method using 8-DOF projective image transformation parameters between pairs of input images. His

Y. Liu, T. Jiang, and C. Zhang (Eds.): CVIBA 2005, LNCS 3765, pp. 271–280, 2005.

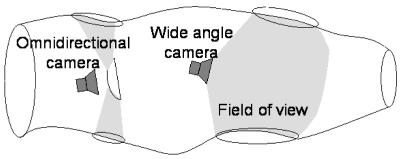


Fig. 1. Omnidirectional and wide-angle cameras inside an intestine

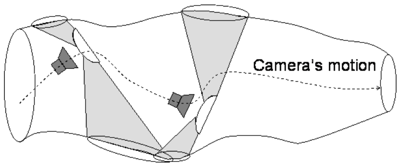
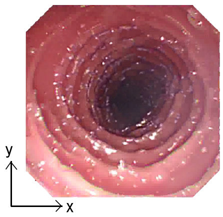
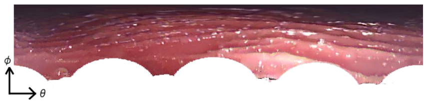


Fig. 2. The camera motion is unknown a priori when examining inside an intestine



(a) Original image



(b) Projected image

Fig. 3. An image of intestine's model taken by endoscope

method can be used when a target is a plane (planar image mosaicing) or the optical centers of images are approximately fixed throughout the video capture (panorama image mosaicing). Swaminathan and Nayar [12] developed a non-metric method for calibrating wide-angle lenses and polycameras, seeking the distortion parameters that would map the image curves to straight lines. But the method does not apply in our objective because of the lack of straight lines in an intestine.

On the other hand, a number of non-rigid body registration methods have been proposed in the field of medical image, for example, free-form deformation(FFD) [3], global transformation using affine transformation and local transformation controlled by FFD model based linear singular blending(LSB) B-spline [14], finite element method (FEM) [6]. These methods work well only if the motion between two images is small.

As mentioned earlier, existing methods make use of the knowledge of camera position or rigid-object shapes. However, the shape of intestine is non-rigid and an endoscopic image sequence contains large motions. Moreover, the trajectory of the camera takes a meandering path as shown in Fig. 2, since the shape of the tube is not a simple cylinder. Therefore, a new image registration approach is required. In this paper, we propose a novel method to generate a dissection image of an intestine from an image sequence, where the dissection image look like the image obtained by cutting the intestine lengthwise and opening out its inside. And we also propose a novel deformable registration method for non-rigid object using the multiple variable polynomial functions (two-dimensional-polynomial functions) which parameters are estimated from optical flow.

The examination of an intestine by an endoscope can be considered as capturing a sequence of images of the inside of a generalized cylinder (in this paper called a "tube object") using a camera. Each image in the sequence captures a part of the tube, and combining them will yield an image of the whole tube. We acquire an image sequence of a tube using an omnidirectional or a wide-angle camera. These cameras have a large field of view as shown in Fig. 1. Thus, they obtain the image of the inside wall of a tube object. Fig. 3(a) shows an image captured by an endoscope; from this image, we utilize the annular area and apply an appropriate formula to obtain a stripe of a dissection image, as shown in Fig. 3(b). After we obtain a sequence of parts of the dissection images, we need to register and mosaic them to form a full dissection image of the intestine. In this paper, our method accomplishes the deformable registration of images by fitting a two-dimensional-polynomial function, which restricted by shape of a generalized cylinder.

This paper is organized as follows. Section 2 gives an overview of our approach. Section 3 explains our deformable registration. Section 4 presents our experimental results. Finally section 5 summarizes our findings.

2 Generating a Dissection Image

In our approach, we take the annular images with a 360 degree view camera with no information about the camera's motion; however, we assume that the camera's motion is continuous, and the characteristic of the observed target is a smooth tube. Our approach to mosaic the annular image sequence is a feature-based approach with steps as shown below:

1. Detect optical flows
2. Project input image into panoramic image
3. Register the projected image
4. Mosaic the registered image

We detect the optical flows from our input image sequence, and then project the input image into panoramic image. After that, we use the detected optical flow in our image registration. Finally, we mosaic the transformed image into mosaicing image.

2.1 Computing Optical Flows and Estimating the Pose of the Camera

We compute optical flows by extracting and tracking feature points. Our method is based on the KLT feature tracker [8] and the feature points are detected using [11]. Fig. 4 shows an example of the result of computing optical flows. Each feature point is detected at the point where the texture changes steeply and the black lines depict the vectors of optical flows.

After computing optical flows, we estimate the pose of the camera using optical flows. The purpose is to reject the outliers of optical flows and cancel the rotation of the camera before projecting the input image into a panoramic image coordinate. Our method first estimates the epipolar geometry using 7 point matches [15] or 8 point matches [7]. Next, we refine the solution by reprojecting the feature points using the estimated parameters [10]. Since our method accomplishes a robust estimation with a RANSAC approach [5], we can reject false optical flows as outliers.

Because the deformation of an input image, which is caused by the rotation of the camera does not change according to the distance from the camera, an image without deformation can be created by applying the inverse rotation to the input image. On the other hand, the deformation by the translation of the camera cannot be easily restored because it depends on the distance from the camera to the object viewed. We explain the solution to this issue in section 3.

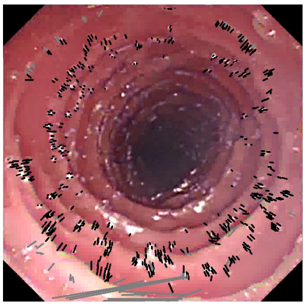


Fig. 4. Detected optical flow (black lines)

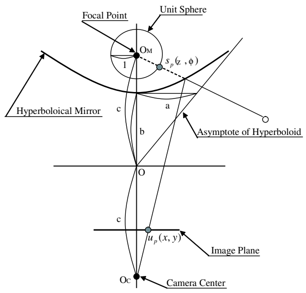


Fig. 5. Projection model of a hyper-omnidirectional camera

2.2 Projection into Spherical Coordinates

The next step is to project an annular image into spherical coordinates. If we use cylindrical coordinates, the mapping function can diverge when the direction of the camera changes. Thus, we use spherical coordinates. There are several ways to capture the annular image, using an omnidirectional or a wide-angle camera. In this section, we explain the method to generate a panoramic image in a spherical coordinate from the image of a hyper-omnidirectional camera [4]. Even if we use a different type of camera, we can generalize the image into spherical coordinate, which make the computation is similar.

Figure 5 shows the projection model of a hyper-omnidirectional camera system. Every ray point to the focal point of the hyperboloidal mirror through a point $S_p(\theta, \phi)$ on the spherical coordinates of the mirror. It reflects to the camera center through a point $u_p(x, y)$ in the annular image coordinates. Thus, we can project the annular image taken by a hyper-omnidirectional camera into a unit sphere by the following computation:

$$\begin{aligned}
\beta &= \arctan\left(\frac{-\left(\left(b^2+c^2\right) sin\left(\phi \cdot \pi/n_h - \pi/2\right) - 2bc\right)}{-\left(\left(b^2-c^2\right) cos\left(\phi \cdot \pi/n_h - \pi/2\right)\right)}\right) \\
r &= f \tan\left(\frac{\pi}{2} - \beta\right) \\
x &= c_x + r\cos\left(2\theta\frac{\pi}{n_w}\right) \\
y &= c_y + r\sin\left(2\theta\frac{\pi}{n_w}\right)
\end{aligned} \qquad (1)$$

where n_w and n_h are the width and height, respectively, of a projected image, a ,b and c are the parameters of the hyperboloidal mirror, and (c_x,c_y) is the center of the annular image. From these equations, we can project each pixel (x, y) of an annular image into a point (θ, ϕ) of the spherical coordinates. A sample of a captured annular image is shown in Fig 3(a), and Fig 3(b) shows the image after projecting.

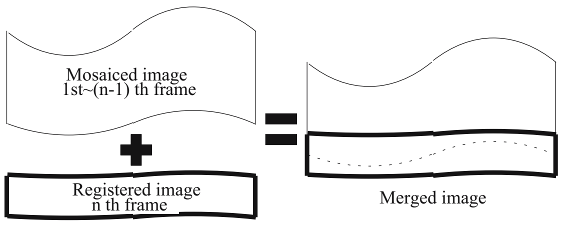


Fig. 6. Image mosaicing

2.3 Image Registration

The third step is image registration that determines the spatial transformation, which maps points from one image to corresponding points on the second image, and transforms the image from the determined map. Since our target object is a tube object, a perspective or affine transformation is not suitable. Therefore, we propose deformable registration by fitting a two dimensional polynomial function. The detail is described in section 3.

2.4 Image Mosaicing

After image registration, the remaining operation is to merge the registration results into the mosaic image. As Fig 6 shows the idea of merging a registered image from the n-th frame on the left side of the figure into a mosaic image, which is created by mosaicing from the first to the (n-1)-th images. We simply mosaic the registered images by overwriting the new registered image on the mosaic image to make a new mosaic image.

3 Deformable Image Registration by Polynomial Transformation

Most image mosaicing techniques transform images by perspective or affine transformations for registration. However, these transformations only work for images of stationary planar objects taken with a perspective camera. Since our images are annular images taken by omnidirectionally with a wide-angle camera, and the target is a tube object, these methods do not work. Therefore, we propose a novel approach for image registration that deforms images by fitting two-dimensional-polynomial functions.

3.1 Model of Optical Flow

If we separate the motion vector to the components along the θ-axis and the ϕ-axis, we can define the model of optical flows depending on θ and ϕ position of the optical flows; because the object and the motion of the camera are constrained to a smooth tube and continuous motion, respectively.

Motions Along θ-axis. Because the input image is an annular image sequence taken from an omnidirectional or wide-angle camera, the motion vectors have periodicity

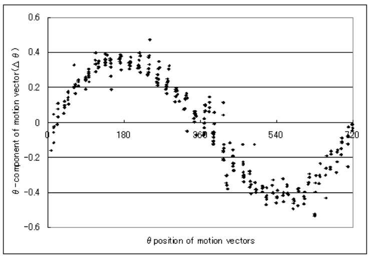


Fig. 7. Detected optical flow (black lines)

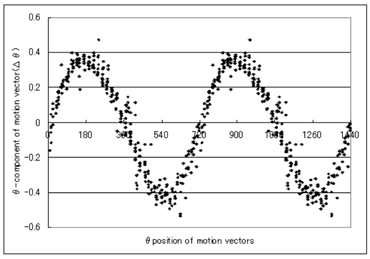


Fig. 8. Projection model of a hyper-omnidirectional camera

along the θ-axis after projecting into a spherical coordinate. Because our target object is a smooth tube, the θ-component of the motion vectors has 2 extrema as shown in Fig 7, which shows the θ-component of motion vectors that are detected from an input annular image sequence

When fitting this data into a polynomial function, since the data has periodicity along θ-axis and is extracted from an annular image sequence, we use 2 cycles of the data, as shown in Fig 8, to approximate the periodicity by a polynomial function. Thus, the number of extrema is 4. If we use a polynomial function for fitting, the degree must be at least 5.

If the object is a real cylinder and the camera moves minutely at the center of the object along the θ-axis, the θ-component of the motion vectors becomes a sine function $\Delta\theta = \sin(\theta)$, where $\Delta\theta$ is the θ-component of a motion vector and θ is the position of the optical flow. Even in other cases, for example, when the camera is not at the center of the tube or where the tube is not a perfect cylinder, the function between θ and $\Delta\theta$ is similar to a sine function. Therefore, when we fit a polynomial function to the data, the degree of the function is sufficient if we can fit the polynomial function to a sine function. We empirically find that a 10-degree polynomial function is sufficient for fitting the objects that we operate on in this paper with the function.

Motion Along ϕ-axis. The ϕ-component of the motion vectors, depend on the motion along the ϕ-axis. Fig 9 shows the situation where the camera moves along the ϕ-axis.

If the distance between the camera origin O and the wall of tube is r and the camera moves rC along the ϕ-axis, where C is an arbitrary constant, the positions of a feature point satisfies

$$C = \tan(\phi) - \tan(\phi') \tag{2}$$

where ϕ and ϕ' are the ϕ-positions of a feature point before and after the motion, respectively. After algebraic manipulation, Eq. 2 becomes

$$\tan(\Delta\phi) = \frac{2C(\cos(2\phi) - 1)}{C\sin(2\phi) - 2} \tag{3}$$

where $\Delta\phi = \phi' - \phi$. Fig 10 shows the function of Eq.3 where $-\pi/2 \leq \phi \leq \pi/2$ and $C = \pi/180$. Since it has only an extremum, we can estimate this function by fitting a second degree polynomial function.

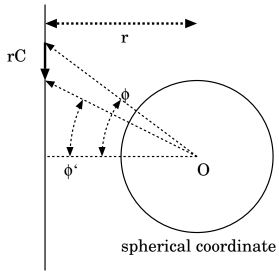


Fig. 9. Detected optical flow (black lines)

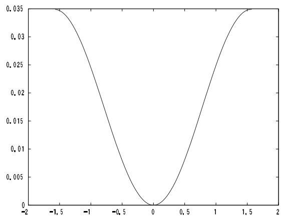


Fig. 10. Projection model of a hyper-omnidirectional camera

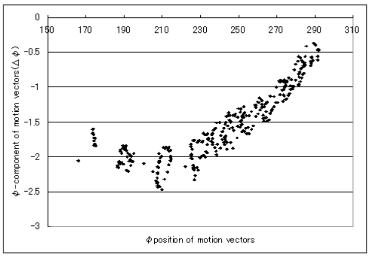


Fig. 11. Relation between ϕ position(ϕ) and ϕ-component($\Delta\phi$) of motion vectors

Fig 11 shows the ϕ-components of optical flows that are detected from an annular image sequence captured by an omnidirectional camera. As described above, the function has an extremum. Thus, it can be approximated by a second degree polynomial function. When we use a wide-angle camera, it cannot observe the lateral and backward directions, in which ϕ is around or less than zero. Since the distribution of the ϕ-components of optical flows are almost linear in such cases, we use a first degree polynomial function for fitting.

3.2 Polynomial Transformation

From the model of optical flows described above, we approximate the distribution of optical flows by a polynomial functions. By putting them together, we model the distribution of optical flows by two-dimensional-polynomial functions as follows:

$$\theta_{new} = \sum_{i=0}^{n} \sum_{j=0}^{m} C_{i,j}^{[\theta]} \theta^i \phi^j \tag{4}$$

$$\phi_{new} = \sum_{i=0}^{n} \sum_{j=0}^{m} C_{i,j}^{[\phi]} \theta^i \phi^j \tag{5}$$

where θ and ϕ are the position of a feature point before the motion, θ_{new} and ϕ_{new} are those after the motion, and $C_{ij}^{[\theta]}$ and $C_{ij}^{[\phi]}$ are the coefficients of the polynomial function.

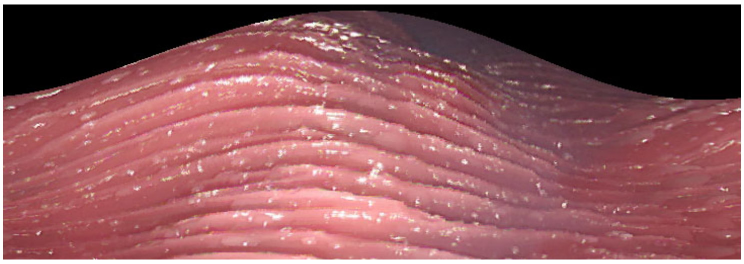


Fig. 12. Generated dissection image of an intestine model

By substituting the position of a feature point for θ, ϕ, θ_{new} and ϕ_{new}, we obtain two linear equations of $C_{ij}^{[\theta]}$ and $C_{ij}^{[\phi]}$. If the degree of θ and ϕ in Eq. (3) is N and M, the number of coefficients is $2NM$. Thus, if we have more than NM feature points, we can determine the coefficients. Since the system becomes linear equations that consist of $2NM$ variables, we estimate the coefficients by a least square method.

If we fit the data to polynomial functions with $2NM$ variables; that is, we model the shape of the object and the camera motion by the polynomial functions with $2NM$ variables. Though the degree of freedom of the shape is quite reduced, we can efficiently model a tube object to create the dissection image.

4 Experiments

To test our method, we set up an experiment using an endoscope and a model of an intestine. The endoscope has a wide-angle camera, with an image angle of approximately 140 degrees (diagonal). Fig 3 shows the input image and the projection into spherical coordinates. Fig 12 shows the mosaicing result from 150 images. Though the object is not a perfect cylinder and the motion is not straight and not at a constant speed, our method successfully generated a mosaicing image.

Next, to evaluate the error of our proposed method, we used the annular image sequence of a checker-pipe (shown in Fig 13) acquired by the hyper-omnidirectional camera. The camera moved along a meandering path. The size of the input image was 640×480 pixels and the size of the panoramic image of spherical coordinates was 720×360 pixels. We computed our deformable registration on a PC with a Pentium4 3.2GHz processor. The computational time for the two images was 680 msec for fitting 750 data and 0.11 msec to transform one pixel. We compared our method with a simple method that transform images only by translation along the θ and ϕ axes. Table 1 shows the roots of mean square (RMS) errors for both methods, which are computed from the differences of the positions of feature points after registration. We used a 5-degree polynomial for θ, and a 2-degree polynomial for ϕ, that is the minimum requirement as shown in 3.1. Our method successfully reduced the errors.

Fig 14 shows one part of the mosaicing results after registration. Fig 14(a) is registered by translation, and Fig 14(b) is the same part registered by our deformable registration. These results were generated from 250 images. Though the shape of the target

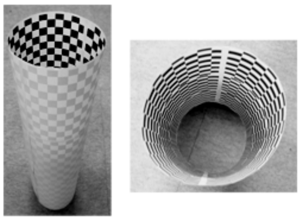
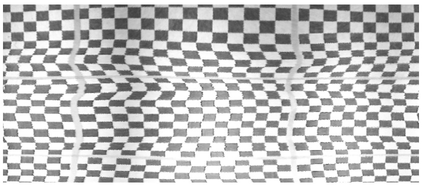


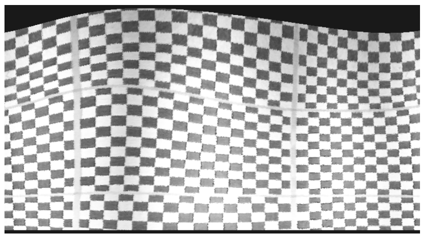
Fig. 13. Used target object

Table 1. Comparison of root of the mean square of the size of the optical flow after transformation

registration	RMS
translation	1.265132995
our approach	0.612701392



(a) Mosaic image by translation



(b) Mosaic image by our approach

Fig. 14. Comparison of the Results

object is a simple cylinder, the mosaicing result was distorted because of the meandering camera motion. On the other hand, the distortion in the results of the deformable registration was reduced from that done by only translation.

5 Conclusion

This paper described a novel method to generate a dissection image of the inside of an intestine. While examinations inside a tube object are difficult and time consuming due to the limited field of view, our method provides an easy-to-understand view for inspecting an intestine. Since a perspective or affine transformation is not suitable for registration of annular images, we proposed a deformable registration by fitting a two dimensional polynomial function. The proposed method successfully creates a dissection image of a generalized cylinder with arbitrary camera motion. Even if the object moves and deforms during observation, it is possible for our method to mosaic images.

As future work, since residual error exists after polynomial fitting, we will apply other deformable registration techniques that have more degrees of freedom, for example, a FEM-based method [1,2] to the imaging problem.

References

1. L. Cohen and I. Cohen. Finite-element methods for active contour models and balloons for 2-d and 3-d images. *PAMI*, 15(11):1131–1147, November 1993.
2. T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *Lecture Notes in Computer Science*, 1407:484–498, 1998.
3. R. D, S. LI, H. C, H. DL, L. MO, and H. DJ. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.
4. K. Daniilidis and C. Geyer. Omnidirectional vision: Theory and algorithms. In *ICPR*, pages 1089–1096, 2000.
5. M. Fischler and R. Bolles. Random sampling consensus: a paradigm for model fitting with application to image analysis and automated cartography. *Commun. Assoc. Comp. Mach.*, 24(6):381–395, 1981.
6. J. Gee and D. Haynor. Numerical methods for high dimensional warps, 1998.
7. R. Hartley. In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(6):580–593, 1997.
8. B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. International Joint Conference on Artificial Intelligence*, pages 674–679, 1981.
9. S. Peleg and J. Herman. Panoramic mosaics by manifold projection, 1997.
10. M. Pollefeys. *Self-calibration and metric 3D reconstruction from uncalibrated image sequences*. PhD thesis, ESAT-PSI, K.U.Leuven, 1999.
11. J. Shi and C. Tomasi. Good features to track. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 593–600, 1994.
12. R. Swaminathan and S. K. Nayar. Nonmetric calibration of wide-angle lenses and polycameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1172–1178, 2000.
13. R. Szeliski. Image mosaicing for tele-reality applications. In *WACV94*, pages 44–53, 1994.
14. S. Tang and T. Jiang. Nonrigid registration of medical image by linear singular blending techniques. *Pattern Recogn. Lett.*, 25(4):399–405, 2004.
15. Z. Zhang. Determining the epipolar geometry and its uncertainty: A review. *International Journal of Computer Vision*, 27(2):161 – 195, 1998.