Deformable Registration for Generating Dissection Image of a Tube from Annular Image Sequence

Suchit PONGNUMKUL†, Ryusuke SAGAWA†, Tomio ECHIGO†, and Yasushi YAGI†

† Department of Intelligent Media, The Institute of Scientific and Industrial Research, Osaka University 8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, JAPAN tel: +81-6-6879-8422 fax: +81-6-6877-4375

Abstract Examination inside a tubular object is difficult and time consuming, because the whole image of the tube 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 tube. 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 fitting two-dimensional-polynomial function. We evaluated our method by registering annular image sequences and generated dissection images, as presented in this paper.

Key words deformable registration, annular image, tube object, mosaicing

1. Introduction

The examination of a generalized cylinder (in this paper called a "tube object") using a camera is required for many applications. For example, a medical doctor examines the intestine of a patient using an endoscope, and some engineers inspect the insides of gas and water pipes with cameras. These examinations are difficult and time consuming because the field of view of a camera is restricted, and the whole image of the tube cannot be taken at one time. It would be desirable for an easy-to-understand view to be available for both medical doctors and engineers. This is especially important in the case where there is a lot of data, for example, a screening test by a medical doctor to select patients from a large study population.

In this paper, we propose a novel method to generate a dissection image of a tube from an image sequence; in this the dissection image is obtained by cutting the tube lengthwise and opening out its inside. We acquire an image sequence of a tube using an omnidirectional and 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. 2(a) shows an example of the image of an endoscope. To generate the dissection image of a tube, we utilize the annular area of the input image. By opening the annular image, we obtain Fig. 2(b).

The tube object that we generate the dissection image of is a generalized cylinder, such as an intestine; therefore, an accurate shape is unknown. Moreover, the trajectory of the camera takes a meandering path as shown in Fig. 3, since the shape of the tube is not a simple cylinder.

We generate the dissection image of the tube by registering and mosaicing the annular images. Because of the meandering camera motion and the shape of the tube, it is necessary for the image registration to deform images. In this paper, our method accomplishes the deformable registration of images by fitting a two-dimensional-polynomial function.

A number of methods have been explored and proposed in the video mosaicing literature. S. Peleg and J. Herman[7] developed a panoramic mosaicing method using manifold projection. Szeliski[10] developed an image-based video mosaicing method using 8-DOF projective image transformation parameters between pairs of input images. His method
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 characteristics 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 flow 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 [6] and the feature points are detected using [9]. 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 [11] or 8 point matches [5]. Next, we refine the solution by reprojecting the feature points using the estimated parameters [8]. Since our method accomplishes a robust estimation with a RANSAC approach [4], 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.

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 sev-
eral 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 [3]. Even if we use a different type of camera, the computation is similar.

Figure 5 shows 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 center of 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:

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

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 \((\theta, \phi)\) of the spherical coordinates.

A sample of a captured annular image is shown in Fig 6, and Fig 7 shows the image after projecting.

2.3 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 Mosaicing

After image registration, the remaining operation is to merge the registration results into the mosaic image. As Fig 8 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 frames. 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 mosaicking 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 $\theta$-axis and the $\phi$-axis, we can define the model of optical flows depending on $\theta$ and $\phi$ position of the optical flows; because the object and the motion of the camera are constrained to a smooth tube and continuous motion, respectively.

3.1.1 Motions along $\theta$-axis

Because the input image is an annular image sequence taken from an omnidirectional or wide-angle camera, the motion vectors have periodicity along the $\theta$-axis. The motion vectors have periodicity along the $\theta$-axis after projecting into a spherical coordinate. Because our target object is a smooth tube, the $\theta$-component of the motion vectors has 2 extrema as shown in Fig 9, which shows the $\theta$-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 $\theta$-axis and is extracted from an annular image sequence, we use 2 cycles of the data, as shown in Fig 10, 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 $\theta$-axis, the $\theta$-component of the motion vectors becomes a sine function as follows:

$$\Delta \theta = \sin(\theta),$$

where $\Delta \theta$ is the $\theta$-component of a motion vector and $\theta$ 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 completely a cylinder, the function between $\theta$ 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.

3.1.2 Motion along $\phi$-axis

The $\phi$-component of the motion vectors, depend on the motion along the $\phi$-axis. Fig 11 shows the situation where the camera moves along the $\phi$-axis.

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

$$C = \tan(\phi) - \tan(\phi')$$

where $\phi$ and $\phi'$ are the $\phi$-positions of a feature point before and after the motion, respectively. After algebraic manipulation, Eq. 3 becomes

$$\tan(\Delta \phi) = \frac{2C(\cos(2\phi) - 1)}{C \sin(2\phi) - 2}$$
\[ f(\phi) = \tan^{-1}\left(\frac{2C(\cos(2\phi) - 1)}{C \sin(2\phi) - 2}\right) \quad \text{when} \quad C = \frac{\pi}{180} \]

Fig 12 shows the function of Eq.4 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 13 shows the \(\phi\)-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 \(\phi\) is around or less than zero. Since the distribution of the \(\phi\)-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:

\[ \theta_{\text{new}} = \sum_{i=0}^{n} \sum_{j=0}^{m} C_{ij}^{[\theta]} \theta^i \phi^j \quad \text{(5)} \]

\[ \phi_{\text{new}} = \sum_{i=0}^{n} \sum_{j=0}^{m} C_{ij}^{[\phi]} \theta^i \phi^j \quad \text{(6)} \]

where \(\theta\) and \(\phi\) are the position of a feature point before the motion, \(\theta_{\text{new}}\) and \(\phi_{\text{new}}\) are those after the motion, and \(C_{ij}^{[\theta]}\) and \(C_{ij}^{[\phi]}\) are the coefficients of the polynomial function.

By substituting the position of a feature point for \(\theta\), \(\phi\), \(\theta_{\text{new}}\) and \(\phi_{\text{new}}\), we obtain two linear equations of \(C_{ij}^{[\theta]}\) and \(C_{ij}^{[\phi]}\). If the degree of \(\theta\) and \(\phi\) 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

For testing our method, we initially used the annular image sequence of a checker-pipe (shown in Fig 14) acquired by the hyper-omnidirectional camera. The size of the input image is \(640 \times 480\) pixels and the size of the panoramic image of spherical coordinates is \(720 \times 360\) 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 \(\theta\) and \(\phi\) 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 \(\theta\), and a 2-degree polynomial for \(\phi\). Our method successfully reduces the errors.

Fig 15 shows one part of the mosaicing result after registration. Fig 15(a) is same part of the result, only by translation,
### Table 1: Comparison of root of the mean square of the size of the optical flow after transformation

<table>
<thead>
<tr>
<th>Registration</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translation</td>
<td>1.265132995</td>
</tr>
<tr>
<td>Our approach</td>
<td>0.612071392</td>
</tr>
</tbody>
</table>

(a) Mosaic image by translation

(b) Mosaic image by our approach

**Figure 15: Comparison of the Results**

and Fig 15(b) is that for our deformable registration. These results are generated from 250 images. The distortion in the results of the deformable registration is reduced from that done by only translation.

Next, we experimented with an endoscope and a model of an intestine. An endoscope has a wide-angle camera, with an image angle of about 140 degrees (diagonal). Fig 2 shows the input image and the projection into spherical coordinates. Fig 16 shows the mosaicing result from 150 images. Though the Object is not completely a cylinder and the motion is not straight and at a constant speed, our method successfully generates a mosaicing image.

5. Conclusion

This paper describes a novel method to generate a dissection image of the inside of a tube. 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 a tube. Since a perspective or affine transformation is not suitable for registration of annular images, we propose 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 mosaics images. As a 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


