EPI Analysis of Omni-Camera Image

Hiroshi KAWASAKI,   Katsushi IKEUCHI,   Masao SAKAUCHI
Institute of Industrial Science, University of Tokyo
7–22–1 Roppongi, Minato-ku, Tokyo, 106-8558, Japan
h-kawa@sak.iis.u-tokyo.ac.jp

Abstract

This paper describes an efficient method to obtain 3D information from omni images using EPI analysis. Two types of omni cameras are employed to make a spatio-temporal volume, which is a sequence of omni images stacked in the spatio-temporal space. For the EPI analysis of omni image, we examine different types of cross sections in such spatio-temporal volumes. To conduct the EPI analysis realistic, we must find out the cross section on which the vertical straight lines in the real world are preserved as straight lines. We define such cross section as omni EPI.

To acquire 3D information using the characteristics of the omni EPI, we propose model based epipolar-plane image (EPI) analysis. To demonstrate the effectiveness of this method, we present experimental result using an omni video of outdoor environments.

1 Introduction

Progress of the computational efficiency and expansion of the storage capacity has made it possible to analyze a huge amount of video data efficiently. As a result of these improvements, acquisition of 3D information from video data seems to be a realistic goal, a goal that many researchers have attempted to achieve. However, several difficulties still remain in applying these techniques to the real-world environment as opposed to using them in the experimental indoor laboratory environment.

This paper describes how we use an omni-directional camera to take video data and proposes a more efficient and robust method for acquiring 3D information from video data than previous methods. Section 2 describes two different kinds of cross sections in a spatio-temporal volume obtained by stacking acquired omni-images in the spatio-temporal space. Section 3 presents an algorithm for acquiring 3D information from omni EPI which is one of the cross sections stated in section 2. Section 4 describes several experiments we performed to demonstrate the effectiveness of our method in acquiring 3D information. Section 5 concludes the paper.

1.1 Overview of Previous Work

In the past, various researchers have investigated the possibility of the acquisition of three dimensional(3D) information from video data. These methods are classified into two types:the EPI analysis [2][1][8] and the factorization method [7].

The EPI method recovers depth information from known motion such as speed and route of the camera and, in the actual analysis, it is usually assumed that the camera speed is constant and the route is straight.

In contrast, the factorization method has no restrictions on camera motion. However, in the real-world environment, stable extraction of the feature points which is necessary for this method are usually difficult due to noise, many obstacles, complicated structures and complex textures.

Based on the facts: that extracting feature points in the real-world is usually very difficult and that the target video satisfies the constraints of EPI (constant speed and a straight route), we adopted the EPI analysis for our research rather than the factorization method.

So far, there have been several attempts to extend or improve the EPI analysis. Zhigang *et al.* [10] use the Gaussian-Fourier orientation detector to assume depth layer from the EPI. This method, unlike other EPI analyses, does not need to detect lines or edges from EPI. We also propose a method which doesn't have to detect lines as accurately as is usually required for EPI analysis.

1.2 Omni directional camera

The omni directional camera is a camera designed to take an omni-directional view of the environment all at once by taking the reflected view on a symmetrical mirror. Many kinds of omni cameras which are characterized by the shape of mirror was propsed [5] [9] [6] and the analysis of the calibration was also employed [3].

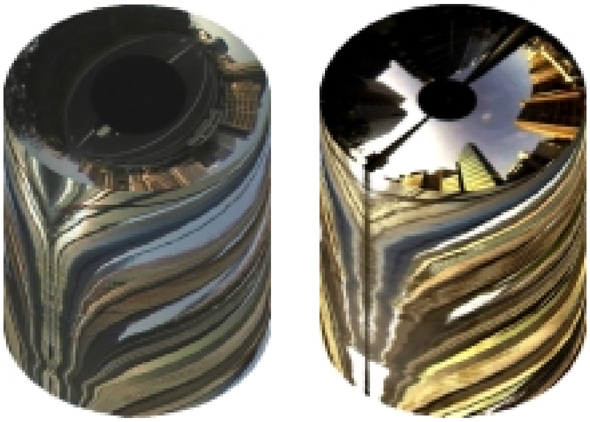
In this paper, we select the hyperboloidal mirror and paraboloidal mirror for our analysis.

The omni camera with a hyperboloidal mirror (Fig.2) has two foci and easy to reconstruct the perspective view. The hyperboloidal omni camera, which can see a certain degree of downward direction, can take a panoramic view of its surroundings.

The omni camera with a paraboloidal mirror (Fig.3) has only one focus and can make an orthogonal projection image from the mirror and can easily reconstruct the perspective view. The paraboloidal omni camera can take one hemisphere of the environmental view at a time.

2 Cross Section of Spatio-Temporal Volume

By accumulating omni images along the time axis, we can obtain a spatio-temporal volume of omni images (STVO) as shown in Fig.1(a) and (b). For the EPI analysis of this 3D volume, we make omni EPI which is a cross sections of this STVO. In the following section, we will describe two types of omni EPI; how to make it and its characteristics. We define a moving direction of the camera as the vertical direction of the volume, and image axes as the horizontal directions.



(a) (b)

Figure 1. (a)Hyperboloidal omni spatio-temporal volume (b) Paraboloidal omni spatio-temporal volume

2.1 Stream curves

In this subsection, we derive equations of stream curves, a 2D projection of a 3D curve along which an image feature flows in the spatio-temporal space. This stream curve corresponds to an epipolar line in the traditional EPI analysis.

2.1.1 Hyperboloidal Omni

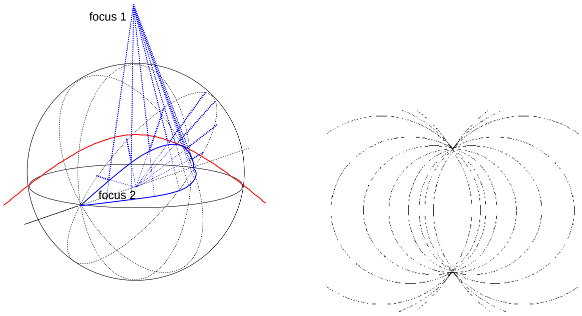
The equation of a solid of revolution, formed by rotating a hyperbolic curve, is represented as:

$$\frac{x^2+y^2}{a^2} - \frac{z^2}{b^2} = -1 \qquad (1)$$

A hyperboloidal omni consists of a mirror, the shape of which shape is a hyperboloidal solid, and a TV camera. This hyperboloidal solid has two foci and the camera is placed at one of these foci, facing the mirror. In Fig.2(a), the higher point indicates the camera position and the upper hyperbolic curve depicts the mirror. Images taken by this camera, in which rays are depicted as dotted lines, are the same as the images taken by the another camera, located at another focus, with a center projection, of which rays are depicted as broken lines.

By considering this reflection mechanism, we can obtain the equation of a stream curve from the equation of a hyperboloidal omni system (1). As shown in Fig.2(b), the shape of a stream curve is an ellipse (zenith angle $|\theta| < arctan(\frac{b}{a})$) or a hyperbolic (zenith angle $|\theta| \geq arctan(\frac{b}{a})$).

A stream curve starts from a focus of expansion (FOE) and ends at the vanishing point (VP). The center of this ellipse is located on the bisector of the diameter connecting the FOE with the VP. This center moves away from the diameter according to the zenith angle of image features. Here, the north pole is defined as the direction of the optical axis of the camera.



(a) (b)

Figure 2. (a)model of hyperboloidal mirror (b)locus of horizontal line on omni-camera

2.1.2 Paraboloidal Omni

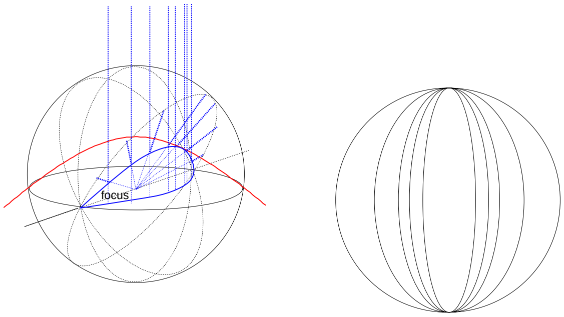
The paraboloidal omni camera, the use of which is described in this paper, has a paraboloidal mirror, a solid of revolution generated from a parabolic curve. The equation of a paraboloidal mirror is represented as:

$$\frac{x^2+y^2}{h^2} + \frac{2 \cdot z}{h} = 1 \qquad (2)$$

Here, we use the same parameterization as [5]. As shown in Fig. 3(a), under this parameterization, all the incoming rays toward the focus of the paraboloidal mirror become parallel and orthogonal against the image plane of the camera after being reflected by the mirror. And the shape of a stream curve is also an elliptic curve as shown in Fig.3(b).

2.1.3 Examples

By cutting STVO, given by hyperboloidal and paraboloidal omni, along a stream curve, we can obtain cross sections,



(a) (b)

Figure 3. (a)model of paraboloidal mirror (b)locus of horizontal line on omni-camera

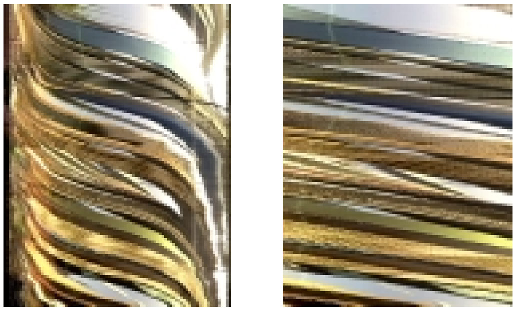
as shown in Fig.4(a) and 5(a), respectively. Here we denote those cross sections as hyperbolic and parabolic EPIs, respectively.

In the STVO, a far away feature moves slowly, while a nearby feature moves rapidly. Due to the difference in the distance, a trace of an image feature, as shown in Fig.4(a) and 5(a), has a S curve on a cross section.

Of course, by applying a perspective transformation, we can successfully transform this S curve into a straight line and can have usual EPIs(Fig.4(b) and 5(b)) which are equivalent to those defined in [2]. We define these transformed cross sections with straight lines as omni EPIs.

(a) (b)

Figure 4. Hyperbolic EPI



(a) (b)

Figure 5. Parabolic EPI

2.2 Straight line

In this section, we try to find out simple straight line to make cross section of the STVO which preserves vertical straight lines.

We will examine the relationship between an arbitrary straight line and radius lines which represent vertical lines in the 3D world. The vertical straight line located at (X_0, Y_0) in the 3D world draws a radius line on an omni image as:

$$y = \frac{Y_0}{X_0} \cdot x \qquad (3)$$

We can express an arbitrary straight line as:

$$ax + by + c = 0 \qquad (4)$$

We define the $u - t$ coordinate system, where the U axis is defined along the arbitrary straight line, and the T axis is the time axis. And if the motion speed of the camera is constant (v) and the moving path is straight, we can substitute Y_0 as $Y_{0original} - vt$, and we obtain the equation:

$$u = \frac{-c/b}{\frac{Y_{0original}-vt}{X_0}\cos(\theta) - \sin(\theta)} \quad (\text{if } \ b \neq 0) \qquad (5)$$

$$u = -\frac{c}{a}\frac{Y_{0original} - vt}{X_0}, \quad (\text{if } \ b = 0) \qquad (6)$$

where $\theta = \arctan(-\frac{a}{b})$. This equations are effective in any omni camera of which the mirror has a solid of revolution. We can apply these equations to both the hyperboloidal and the paraboloidal omni camera.

2.2.1 Vertical straight line

The equation (6) shows that, if $b = 0$ then, the vertical edge in the real world describes a straight line on the $u - t$ plane. We cut the STVO at vertical straight line and obtain cross section as Fig.6 and 7. In these figures, we can see that all edges and boundaries of buildings are preserved in a straight line on the cross section. Here, we also define these cross sections as omni EPIs.

3 Retrieval of 3D information

In this section,we propose a model based EPI analysis to obtain 3D information from omni EPI. Efficient method to employ 2D image matching is also proposed.

3.1 Model-based EPI analysis

Under the circumstances that the camera speed is constant and the camera is moving along a straight route, the traces of objects describe straight lines on EPI. The coefficients of these straight lines represent the depth of the objects; we can determine the depth of the objects by assuming the incline of these lines.

Figure 6. Hyperboloidal omni EPI: vertical cross section

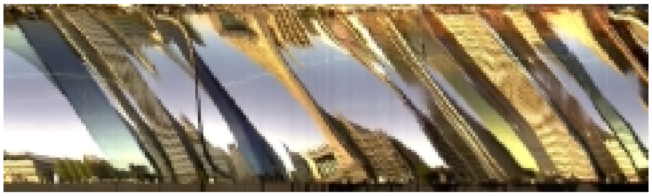


Figure 7. Paraboloidal omni EPI: vertical cross section

On the other hand, the EPI analysis is sometimes very difficult to apply to real world data, because keeping the camera speed almost constant and moving along the route in a straight direction is difficult in reality. Also, objects in the real world have complicated shapes and textures.

Further, we can obtain models for most objects in the real world, especially models of buildings. And in the actual situation, when we use the video data taken from the vehicle as it travels along the street, we can use map data as a model.

Considering these situations, we propose a model-based analysis to avoid the difficulties of image processing caused by real world complexity. To put this analysis concretely, we perform matching between video data and models, and then retrieve 3D information from video data using the matching results.

3.2 EPI - EPI matching

Since models and video data are different representations, it is impossible to perform direct matching between them. We have to define some common basis for both sets of data. As stated in sec.2.2.1, we can easily make EPI (exactly cross section) with straight lines from omni video data. We can also make EPI from models by simulating the camera motion and parameters. Thus, we will employ this EPI as the common basis for matching.

There are also some advantages in using EPI for matching compared with the usual EPI analysis to extract 3D features from EPIs. Namely, the route does not have to be exactly straight and the camera motion does not have to strictly keep the velocity constant, because this EPI-based matching compares regions generated from both real and simulated EPIs. This method does not need precise line detection, because we are not going to extract parameters of lines; rather we will compare regions bounded by these lines for matching.

Thus, 3D information retrieval problem results in a 2D matching problem. By assuming that the order of objects does not change, we can use the DP matching. Next we will show actual method for DP matching adapted for the 2D pattern.

3.3 2D DP matching

We must carry out matching between the EPI made from the omni-video camera and the EPI made from a model. EPI is a 2D image, so we have to carry out matching between 2D images. In this paper, we propose a 2D image matching method using the DP matching; the actual process of this method is as follows.

1. Make a rectangular parallelepiped from two EPIs(Fig.8)
2. Cut this 3D polygon by a horizontal plane and make a 1D search plane
3. Carry out usual 1 dimensional DP matching for each 1D search plane(Fig.9)
4. Iterate **1** to **3** using the constraints so that the matching path makes a continuous plane

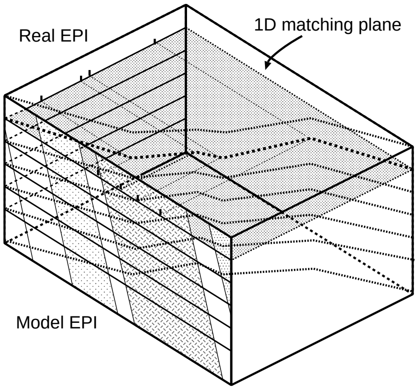


Figure 8. Concept of 2D DP matching

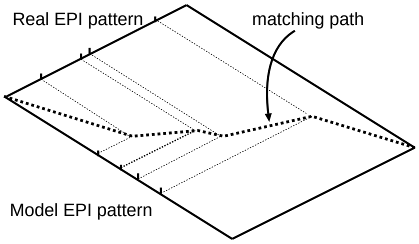


Figure 9. 1D search plane of 1D DP matching

4 Experiments

We have performed some experiments to demonstrate the effectiveness of our method. Target omni video data we use in this experiment is taken from a vehicle which runs along a city street.

4.1 Outline of actual STVO analysis

1. Make the omni EPI from the video data

   Simply cut the STVO at a vertical line and we can have one easily as stated in Sec. **2.2.1**

2. Make a model EPI from 2D map data

   To make the model EPI, we make the virtual 3D city map in the computer by using digital 2D map data and simulating the motion and the camera parameters of the virtual camera.

3. Perform matching between two EPIs

   Carrying out the DP matching, requires that we give the corresponding point. In this experiment we give the corresponding point manually.

4. Restore 3D information to video the data

4.2 Result of 2D DP matching

To achieve high accuracy and robustness of the matching, we use three different patterns for matching, such as the edge of the building, the boundary of the building and the sky pattern. Fig.10 shows the result of DP matching and by using this matching result we made a virtual 3D city map successfully(Fig.11).

In Fig.10, vertical axis represents the coordinate of EPI from the video data and horizontal axis represents the coordinate of the map data. If the model data is precisely corresponding to the real world, this matching path describes a straight line from the origin of the coordinate axes to the end point. But there are some errors in the map data, and the EPI made from video also contains some distortion errors. Thus, the actual path does not describe a straight line; rather, it describes the line as a solid line in the figures. A broken line is the matching result obtained by our matching method.

We can see that the matching result still has small errors, but nevertheless achieves sufficient accuracy for our proposed system.

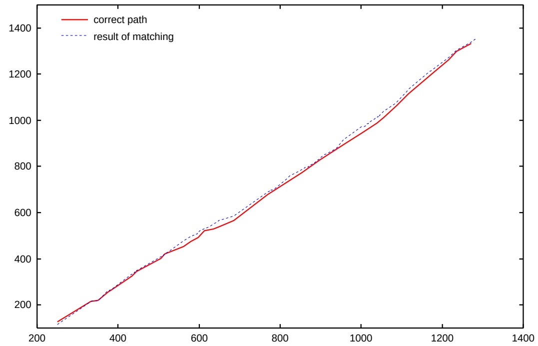


Figure 10. Result of the DP matching.

Figure 11. Virtual 3D map.

5 Conclusion

We made two types of cross sections from the spatiotemporal volume of omni-images (STVO) and made analyses. We stated some characteristics for these two planes and define these cross sections as omni EPIs. On these omni EPIs, the vertical straight lines in the real world are preserved as straight lines; the degree of this slope represents the depth of the object.

We can estimate the depth with analysis of these omni EPIs. We proposed an efficient and robust method based on EPI analysis to obtain 3D information; we also used models to make the analysis realistic. To improve the result, we also proposed an efficient matching algorithm between EPIs, that is, 2D DP matching.

To demonstrate the effectiveness of our proposed method, we conducted several experiments using real world omni video data and a digital map as a model. The result of the experiments show the effectiveness of our proposed method and matching algorithm.

References

[1] H. Baker and R. Bolles. Generalizing epipolar plane image analysis on the spatiotemporal surface. *Int.J.of Computer Vision*, 3:33–49, 1989.

[2] R. Bolles, H. Baker, and D. Marimont. Epipolar plane image analysis:an approach to determining structure from motion. *Int.J.of Computer Vision*, 1:7–55, 1987.

[3] C. Geyer and K. Daniilidis. Catadioptric camera calibration. In *International Conference on Computer Vision*, volume 1, pages 398–404, Sept. 1999.

[4] H. Kawasaki, K. Ikeuchi, and M. Sakauchi. Spatio-Temporal analysis of omni image. In *Computer Vision and Pattern Recognition*, June 2000.

[5] S. Nayar. Catadioptric omnidirectional video camera. In *Computer Vision and Pattern Recognition*, pages 482–488, June 1997.

[6] Y. Onoue, K. Yamasawa, H. Takemura, and N. Yokoya. Telepresence by real-time view-dependent image generation from omnidirectional video streams. *Computer Vision and Image Understanding*, 71(2):154–165, Aug. 1998.

[7] C. Thomasi and T. Kanade. Shape and motion from image stream under orthography: A factorization method. *Int.J.of Computer Vision*, 9:137–189, 1992.

[8] M. Yamamoto. Determining three-dimensional structure from image sequences given by horizontal and vertical moving camera. *(D)*, J-69(11):1631–1638, Nov. 1986.

[9] J. Y. Zheng and S. Tsuji. Panoramic representation of scenes for route understanding. In *International Conference on Pattern Recognition*, pages 161–167, June 1990.

[10] Z. Zhu, G. Xu, and X. Lin. Panoramic EPI generation and analysis of video from a moving platform with vibration. In *Computer Vision and Pattern Recognition99*, pages 531–537, 1999.