MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11-13, 2002, Nara-ken New Public Hall, Nara, Japan

13—34

Cylindrical Panorama Using the Tilt of a Camera

Takashi Iida* and Naoki Chiba
Digital Systems Development Center
SANYO Electric Co., Ltd.

Abstract

We propose a method for extending the vertical field of view when constructing a cylindrical panorama. Cylindrical panorama construction is a simple technique with which we can construct a wide-view and high-resolution panoramic image from a collection of still images taken with a panning camera. This method, however, has a problem in that it lacks a field of view in the vertical direction. To solve this problem, we have to compute the tilting angle that is fixed during the panning. We therefore propose a method of computing the angle from images. Since our method is based on feature correspondence and a linear function, it is fast to compute. We show the effectiveness by experiments both on synthetic data and real images.

1 Introduction

Image mosaicing has become an active research area, because it is able to construct a wide-view and high-resolution panoramic image from a collection of still images. Applications include the construction of aerial and satellite photographs and the creation of virtual environments. The simplest method for creating virtual environments is cylindrical panorama construction[1], with which we can construct a panoramic image from natural images by projecting them onto a cylindrical surface.

This method, however, has a problem in that it lacks the vertical field of view. This is because the method limits the camera motion to horizontal panning with a tripod. To solve this problem, Szeliski et al. have proposed a method which can construct a full-view panoramic image by computing three-parameter rotational transformations between images taken with a hand-held camera [2]. This method, however, has the following two problems. First, the way of capturing images to cover the entire field of view in all directions is difficult for users with a hand-held camera. Second, the accuracy of estimating the transformation tends to be low when the images do not contain sufficient texture. This comes from the fact that the number of degrees of freedom in the geometric transformation is large (three).

On the other hand, the cylindrical panorama method has the following two advantages over Szeliski's method. First, it makes it easy to capture images, because it uses a tripod. Second, the accuracy of estimating the geometric transformation is better than Szeliski's method, because it only has one degree of freedom, which is fewer than Szeliski's (three).

*Address: 1-18-13, Hashiridani, Hirakata-City, Osaka 573-8534 Japan. E-mail: iida@hr.hm.rd.sanyo.co.jp

We can easily extend the cylindrical panorama method to cover the vertical field of view. If we tilt the camera and maintain the angle while panning, we can extend the vertical field of view. Since commercially available tripods have a function to tilt the camera, we can easily tilt it without any special devices.

Thre is another problem, however, in extending the vertical field of view. We need to compute the tilting angles from the captured images alone, because ordinary tripods do not have any scale for measuring the tilting angle. To solve this problem, one possible method is to estimate the angle by trial-and-error. This operation is tedious for the user and not very accurate. Until now, no methods have been proposed to compute a tilting angle that is fixed while panning a camera.

We propose a method for computing the tilting angle from images alone. Once we obtain the tilting angle, we can extend the vertical field of view of a cylindrical panorama by projecting the images onto a cylinder with the consideration of the tilt.

2 Cylindrical Panoramas

2.1 Principle of cylindrical panoramas

The process of constructing a cylindrical panorama consists of the following two steps. The first step is to project images, captured by a panning camera on a tripod, onto a cylindrical surface. If the camera focal length or the field of view is known, each perspective image can be warped into cylindrical coordinates. To project the image, we map world coordinates $\mathbf{p} = (X, Y, Z)$ to 2D cylindrical screen coordinates $\mathbf{u} = (\theta, \nu)$ by using the following equations.

$$\theta = \tan^{-1}\frac{X}{Z} \tag{1}$$

$$\nu = \frac{Y}{\sqrt{X^2 + Z^2}} \tag{2}$$

where θ is the panning angle and v is the scanline coordinate.

The second step is to align the projected images on the cylindrical surface for image stitching. The alignment is simplified to be horizontal translation, because the camera motion is limited to only horizontal rotation around the optical center with a tripod. The details can be found in [2].

2.2 Problem of cylindrical panoramas

Cylindrical panorama construction with panning alone has a problem when capturing the images of tall buildings from the ground or looking down from a sightseeing place at the top of a tower. The vertical field of view is too narrow to cover the scenes. We can solve this problem by tilting the camera.

If the tilting angle is unknown, the accuracy of constructing a panorama with tilt is poor. Figure 1 shows an example of a cylindrical panorama without considering the tilt. The content of the panorama in the overlapped region has ghost-like misregistration.

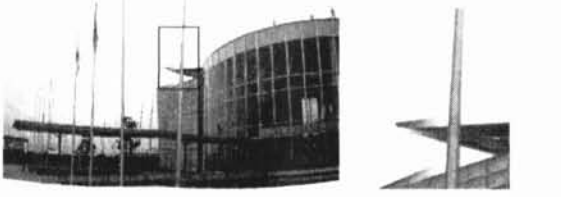


Figure 1: A panorama constructed without estimating the tilt

If we know the tilting angle ψ and the focal length, we can project the images onto the cylindrical surface by using the following equations.

$$\theta = \tan^{-1} \frac{X}{Z\cos\psi - Y\sin\psi} \quad (3)$$

$$\nu = \frac{Z\sin\psi + Y\cos\psi}{\sqrt{X^2 + (Z\cos\psi - Y\sin\psi)^2}} \quad (4)$$

This allows us to handle cylindrical panoramas with tilt.

No methods have been proposed to estimate the tilting angle ψ that is maintained while panning the camera. Several related methods have been proposed in [2, 3]. These methods compute not only tilting angles, but also all of the rotational angles as well as the intrinsic parameters of the camera. What they can compute, however, is the relative rotations between two images, not the titling angle that is maintained while panning the camera. Therefore we propose a novel method of computing the tilting angle from the images.

3 Computing the tilting angle for a cylindrical panorama

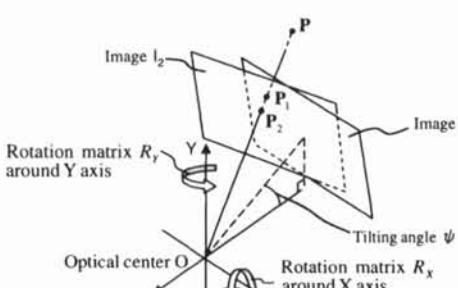


Figure 2: Relationship between an image with tilt and the image after panning while maintaining the tilt

Here we show how to estimate a tilting angle that is fixed while panning the camera. Figure 2 illustrates the geometric relationship between two images captured by a panning camera around the optical center O with the fixed tilt ψ. When a 3D point $\mathbf{P}$ is projected onto the image planes I_1 and I_2, the projected points $\mathbf{P_1}$ and $\mathbf{P_2}$ are equal in the projective sense. This is because these two points are on the same viewing ray from the optical center O through the 3D point $\mathbf{P}$.

Figure 3 illustrates this in detail. The left side of Figure 3 illustrates the first image I_1 captured with the camera tilt ψ around the X axis. Here we see the geometric relationship between the world coordinate system $O - XYZ$ and the camera coordinate system $O - U_1V_1W_1$. Since we rotate the camera around the optical center O, the relationship is simply rotation. If the camera focal length f is known, the point p_1 on the image screen can be represented as follows:

$$p_1 = (u_1, v_1, f). \quad (5)$$

Since the 3D point $\mathbf{P_1}$ after the tilt ψ on the image plane I_1 is the rotated point of $\mathbf{p_1}$ around the axis X, it can be written as follows:

$$\mathbf{P_1} = \mathbf{R}_X(\psi)\mathbf{p_1} \quad (6)$$

where $\mathbf{R}_X$ is a 3×3 rotation matrix around axis X.

The right side of Figure 3 illustrates that the second image I_2 captured by the camera pan ϕ while maintaining the tilt ψ. Here we see the relationship between the world coordinate system $O - XYZ$ and the camera coordinate system $O - U_2V_2Z_2$. Since we rotated the camera around the Y axis after rotating it around the X axis, this relationship can be written as follows:

$$\mathbf{P_2} = \mathbf{R}_Y(\phi)\mathbf{R}_X(\psi)\mathbf{p_2} \quad (7)$$

where $\mathbf{R}_Y$ is a 3×3 rotation matrix around axis Y.

Because the points $\mathbf{P_1}$ and $\mathbf{P_2}$ are equal projectively, we have the following equation.

$$\mathbf{P_2} = k\mathbf{P_1} \ \ or \quad \begin{pmatrix} X_2 \\ Y_2 \\ Z_2 \end{pmatrix} = k \begin{pmatrix} X_1 \\ Y_1 \\ Z_1 \end{pmatrix} \quad (8)$$

where k is a non-zero coefficient.

Since Equation (8) shows that the ratio of each element between $\mathbf{P_1}$ and $\mathbf{P_2}$ is equal, we can rewrite the equation by deleting the coefficient k as follows:

$$\begin{cases} \frac{X_2}{Z_2} = \frac{X_1}{Z_1} \\ \frac{Y_2}{Z_2} = \frac{Y_1}{Z_1}. \end{cases} \quad (9)$$

We then rewrite Equation (9) for the tilting angle ψ. By substituting Equation (6) and Equation (7) for Equation (9), we have the following equations.

$$\begin{cases} a_{11}\sin^2\psi + a_{12}\sin\psi\cos\psi + a_{13}\sin\psi + a_{14}\cos\psi = b_1 \\ a_{21}\sin^2\psi + a_{22}\sin\psi\cos\psi + a_{23}\sin\psi + a_{24}\cos\psi = b_2 \end{cases} \quad (10)$$

where the coefficients $a_{11}, a_{12}, \cdots, a_{24}, b_1$ and b_2 are fixed factors. These factors can be computed if we know the parameters of the focal length f, the corresponding image coordinates (u_1, v_1), (u_2, v_2) and the panning angle ϕ.

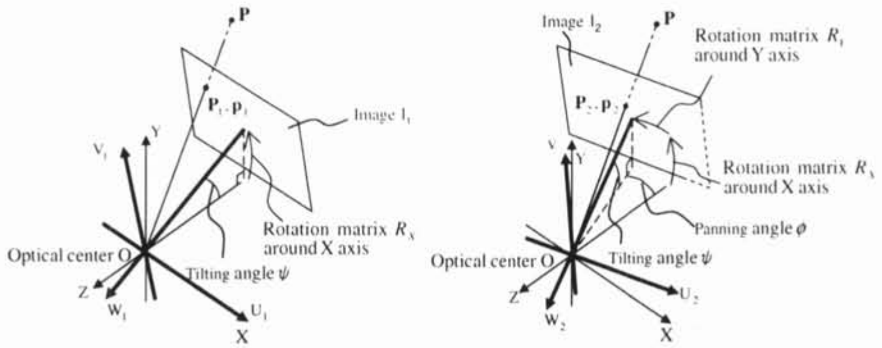


Figure 3: Relationship between the world coordinate system and the camera coordinate system (Left: First camera coordinate system, Right: Second camera coordinate system)

Since Equation (10) is good for all of the point correspondences, we have the following linear system.

$$\mathbf{AX} = \mathbf{B} \quad or$$
$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \vdots & \vdots & \vdots & \vdots \\ a_{2n1} & a_{2n2} & a_{2n3} & a_{2n4} \end{bmatrix} \begin{bmatrix} \sin^2\psi \\ \sin\psi\cos\psi \\ \sin\psi \\ \cos\psi \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \tag{11}$$

where n is the number of correspondences. We can compute the row vector X by the least squares method, if n is two or more. Then, we compute the tilting angle ψ from the vector X. Here we have four choices to compute, $\sin^2\psi$, $\sin\psi\cos\psi$, $\sin\psi$ and $\cos\psi$. We use $\sin\psi$, because this trigonometric function is the most linear among the four up to 45 degrees, which is sufficient for tilting the camera in real situations.

In this process, we assumed that the panning angle is known. In fact, we need to compute this angle as well. We can obtain an equation about the panning angle in exactly the same manner as Equation (11). The angles of panning and tilting are dependent on each other. Once we obtain the tilting angle, we can compute the panning angle from it. We perform this iterative estimation until the angles converge or the iterations reach the pre-defined number. We start this process by setting the tilting angle to 0 degree.

4 Cylindrical panorama using tilt

The processing steps for constructing a cylindrical panorama by estimating the tilting angles are given bellow.

1. Feature correspondence
   We automatically establish the feature correspondence by using optical flow estimation[4].
2. Computation of the tilting angle
   We compute the tilting angle by the proposed method above.
3. Projection to the cylindrical surface
   We project captured images onto the cylindrical surface considering the computed tilt by using Equation (3) and Equation (4).
4. Alignment
   We align the projected images onto the cylindrical surface to construct a panorama.

5 Experiments

Next, we show the effectiveness of our method with experiments by computer simulation and real images.

5.1 Computer simulation

By computer simulation, we evaluated the accuracy of computing the tilting angle with regard to the following two types of noise.

- Error in feature correspondence
- Radial lens distortion

The experimental setup is as follows. The image size of the simulated camera is 640 by 480 pixels. The focal length is 640 pixels. The first image was captured with the camera tilted by 30 degrees. The second image was captured with the camera panned by 30 degrees while maintaining the tilt. We placed points in 3D at an equal distance from the optical center. The points were aligned 50 pixels apart on the first image plane. We obtained 58 pairs of corresponding points between the first and second images.

In the first experiment, we evaluated the performance with regard to the error in feature correspondence. Figure 4 shows the result. We added gaussian noise with 0 mean and σ standard deviation to the point coordinates on the second image. We varied the noise level from 0 to 10 pixels at intervals of 0.5 pixel. For each noise level, we performed 100 independent trials, and the results are shown on the average. The errors were less than 7 % until the noise level reached 8. This shows that the effect of the error in feature correspondence is very small.

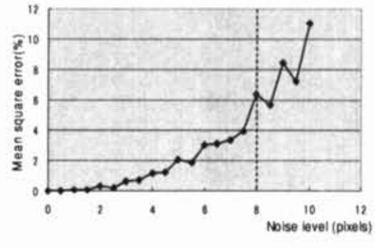


Figure 4: Error plot by varying noise levels of feature correspondences

In the second experiment, we evaluated the performance with regard to lens distortion. Figure 5 shows the result. We added lens distortion with the coefficient of the radial distortion k to the point coordinates on the first and second images. We varied the coefficient k from 0 to 6.0×10^{-7} at intervals of 0.5×10^{-7}. Until the coefficient of 1.5×10^{-7}, the error was less than 0.65 %, which is very small. The coefficient is the actual amount of the zooming lens for a digital still camera (SANYO DSC-MZ1). We can see that the effect of lens distortion is negligible, unless we use a wide-angle lens whose lens distortion might be large.

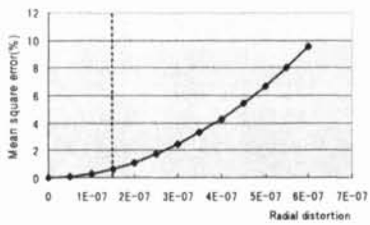


Figure 5: Error curve by varying lens distortion

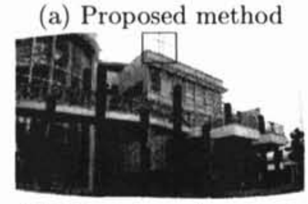
5.2 Real data

We then conducted experiments using real images. The images were captured by a digital still camera (SANYO DSC-SX550) mounted on a tripod. The image size is 640 × 480 pixels. The focal length is 640 pixels. The first image was captured with the camera tilted by 13 degrees. The second image was captured with the camera panned by 20 degrees while maintaining the tilt. These angles were measured by the scale printed on the tripod.

We computed the tilting angle by using the images. The result was 12.89 degrees, which should be 13 degrees. The error is very small. The processing time on a PC (Pentium 4, 2GHz) was 15 mseconds, which is very fast. We compared the constructed panoramas with and without estimating the tilting angle. Figure 6 shows the result. The panorama by our method is very good, while that by a conventional method disregarding the tilting angle is poor. The edges in the overlapped region are doubled.

(a) Proposed method

(b) Enlarged view of (a)



(c) Convetional method

(d) Enlarge view of (c)

Figure 6: Comparison between the proposed method and a conventional method (image: building)

We quantitatively evaluated the accuracy of our method. We measured the average error E of the pixel value differences in the RGB channels in the overlapped region ω between the images I_1 and I_2 by using the following equation.

$$E = \frac{\sum_{\mathrm{RGB}} \sum_{\omega} \sqrt{(I_1 - I_2)^2}}{3N} \tag{12}$$

Table 1 shows the result. The error of our method is smaller than that of a previous method not computing the tilting angle.

Table 1: Error comparison by pixel value differences

Images	Proposed	Previous
Building	14.22	25.71
Bridge	16.75	31.43

6 Discussion

We have proposed a method for extending the vertical field of view when constructing a cylindrical panorama. It estimates the tilting angle from the images. We confirmed that the accuracy of our method is very good even in the presence of error in feature correspondences and radial lens distortion.

Our method has the following two advantages over Szeliski's method. First, the way of capturing images is easier for users than that of Szeliski's. This is because we use a tripod to restrict the camera motion. The flexibility of Szeliski's method might be a problem in capturing all directions without any gaps. Second, we can expect that the accuracy for constructing a panorama is better than that of Szeliski's, because the number of degrees of freedom for the transformation is fewer (two) than that of Szeliski's (three). Experimental evaluation is a topic for the future.

We can say that our method is an intermediate method between the conventional cylindrical method and Szeliski's full-view creation method.

7 Conclusion

In this paper, we have illustrated a method for extending the vertical field of view when constructing a cylindrical panorama, by computing the tilting angle. The proposed method is fast to compute because it has the following two features. First, it is based on feature correspondence. Second, it is based on linear computation.

In future work, we plan to incorporate slight twisting angles which might occur during panning.

References

[1] S. E. Chen, "QuickTime VR - An Image-Based Approach to Virtual Environment Navigation," proc. of SIGGRAPH'95, pp. 29-38, 1995

[2] R. Szeliski and H. Y. Shum, "Creating Full View Panoramic Image Mosaics and Environment Maps," proc. of SIGGRAPH'97, pp. 251-258, 1997

[3] Y. Seo, "A Multiple View Approach for Auto calibration of a Rotating and Zooming Camera," IECIE Trans. Inf. & Syst., Vol. E83-D, No. 7, pp. 1375-1385, 2000

[4] N. Chiba, H. Kano, M. Minoh et. al, "Feature-Based Image Mosaicing," proc. of IAPR MVA'98, pp. 5-10, 1998