Spherical Panoramic Mosaics with the Image Division and Warping Methods

Weihua An and Bingfeng Zhou
Institute of Computer Science and Technology, Peking University, Beijing, China

Abstract

Some limitations exist in previous approaches for spherical panoramic mosaics. They usually employ optimizing methods for image registration, and need particular imaging equipments. In this paper we present a new approach to automatically create seamless spherical panoramas from image sequences. These images are acquired from an approximately fixed point by using off-the-shelf cameras and tripods. This approach employs the FFT-based phase-correlation algorithm to approximately compute the transformation relationships between images. To avoid the accumulation of registration errors, we mosaic the images acquired along each vertical line into a vertical sub-image, and then generate a spherical panorama from all the vertical sub-images by employing the image division and warping methods. As demonstrated by the experiment results, the new approach can generate high-quality mosaics while requiring simpler manual operations and less time consumption.

Introduction

Panoramic image mosaicing has been extensively researched, and there have been many mosaic methods [1]. But most of them focus on the cylindrical mosaics instead of spherical ones. Moreover, the existing methods for creating spherical panoramas have some limitations.

For example, Brown [3] selects SIFT features for image matching, which is insensitive to the orientation, scale and illumination of the images. Thus, it brings the advantage of automatically recognizing images' sequences, and stitching them. However, when the amount of images increases, the time consumption increases exponentially.

Szelisk's method [2] for spherical panoramic mosaics uses the Levenberg-Marquardt algorithm to compute the transform relationships between images. This algorithm iteratively estimates the camera parameters by minimizing an error function based on the intensity difference in the overlapped area. The advantage of the algorithm is that it can register the images accurately, and adapts to various environment mapping models. However, it is quite difficult to choose a proper initial value for the iterative algorithm.

In order to obtain an initial value close enough to the exact one, Coorg [4] has designed an equipment to capture images. It can not only ensure that all the captured images have a fixed view point, but also approximately record the camera parameters, which can be used as the initial value for the iterative algorithm.

In this paper, we present a new approach to automatically create a spherical panorama from image sequences. These images are acquired from an approximately fixed view point. This approach requires the equipments including common cameras and tripods, and involves manual operation of just rotating the camera and taking images. Our approach adopts the FFT-based phase-correlation algorithm to approximately compute the transform relations between images [5]. To avoid the accumulation of errors, we firstly create a set of vertical sub-images from the input images, and then stitch them into a spherical panorama by the horizontally division and warping method.

Overview of Our Approach

Capturing Images

In our approach, we use the spherical panorama shown in Figure 1 as our representation of choice. This representation is convenient as it is very easy to construct the mapping relations between the input images and the panorama.

An overview of our hardware setup is shown in Figure 2, where a camera is mounted on a tripod. To collaborate with the panoramic representation, we take images as follows. Firstly, we vertically rotate the camera for different α angles, and take images. Secondly, we horizontally rotate the camera for different β angles, and take another column of images. The two steps are repeated until all the direction at the view point is covered by the images. While taking images, we must ensure that the adjacent images are overlapped partly. The captured images can be intuitively considered as a 2D image matrix as shown in Figure 3.

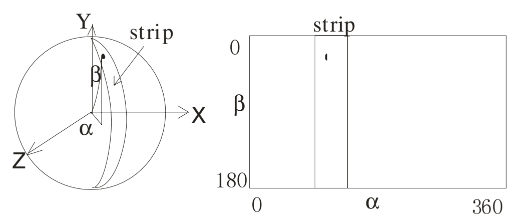


***Figure 1**. the representation of a spherical panorama. (the vertical block on the panorama is corresponding to the strip of the sphere)*

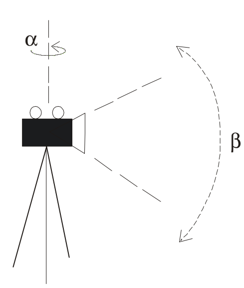


***Figure 2**. the equipment and the way to capture images.*

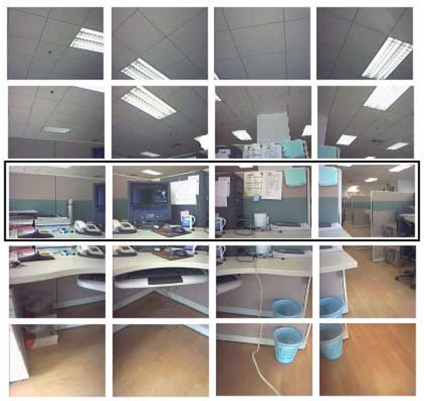


Figure 3*. the partial captured images. (The images at the same column belong to the same group, and the images in the black box are horizontal images)*

To describe conveniently, we divide all the images into different groups according to the following rule: the images located at the same column belong to the same group. In each group, we choose the image most parallel to the horizon as the horizontal image.

Projection Transformation

Given that the origin of camera coordinates equals to that of the world coordinates, the relationship between a 3D point $\vec{p}$ and its homogeneous coordinates $\vec{q}$ in the captured image can be described as [2]:

$$\vec{q} = TVR\vec{p} \tag{1}$$

where

$$T = \begin{bmatrix} 1 & 0 & c_x \\ 0 & 1 & c_y \\ 0 & 0 & 1 \end{bmatrix},\ V = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix},\ R = \begin{bmatrix} r_{00} & r_{01} & r_{02} \\ r_{10} & r_{11} & r_{12} \\ r_{20} & r_{21} & r_{22} \end{bmatrix}$$

The variants T, V, R are respectively the translation, focal length scaling, and 3D rotation matrices. We assume that the origin of the image coordinates is located at its center, and then the matrix T can be ignored.

Reversely transforming the equation (1), we can obtain the 3D direction corresponding to an image pixel.

$$\vec{p} = R^{-1}V^{-1}\vec{q} \tag{2}$$

We can denote $\vec{p}$ as spherical coordinates, i.e.

$$\alpha = \tan^{-1}(X / Z) \tag{3}$$

$$\beta = \tan^{-1}(Y / \sqrt{X^2 + Z^2}) \tag{4}$$

Its corresponding coordinates on the spherical panorama can be described as

$$x_p = \alpha r \tag{5}$$

$$y_p = \beta r \tag{6}$$

where r is the spherical radius.

According to above derivation, we can see that if we have obtained the focal lengths and the rotation matrices of input images, the panorama can be created by projective geometry. So, our task is to compute the parameters and to project the input images onto the panorama.

In this paper, we suppose that the focal lengths of all the images are equal, and that they have a common view point. So, we can define r=f. According to the manner of capturing images, the rotation matrix can be expressed as

$$R_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi_i & -\sin\varphi_i \\ 0 & \sin\varphi_i & \cos\varphi_i \end{bmatrix} \begin{bmatrix} \cos\theta_i & 0 & \sin\theta_i \\ 0 & 1 & 0 \\ -\sin\theta_i & 0 & \cos\theta_i \end{bmatrix} \tag{7}$$

where θ_i, φ_i is the horizontal and vertical rotation angles.

Therefore, our task can be further simplified to compute the focal length f and the images' rotation angle θ_i, φ_i.

Estimating Parameters

Estimating the Focal Length

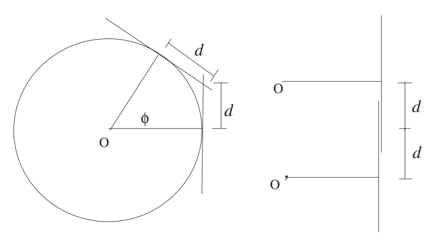
The focal length can be deduced according to the relationships between the horizontal images, which is shown in Figure 4 (for convenience, we substitute the 2D format for 3D).

There is the rotation relationship between the adjacent horizontal images, and the rotation radius is the focal length f (Figure 4(a)). If the rotation angle is small enough, the relationship can be considered as translation [2] (Figure 4(b)). We adopt the FFT-based phase-correlation algorithm to obtain the translation distance d' between adjacent images, and substitute d' for d in Figure 4(a).

After obtained all the translation distances between each two adjacent horizontal images, we can estimate the focal length, using the following equation.

$$2\pi - 2\sum \tan^{-1}(d_i / f) = 0 \tag{8}$$

where d_i denotes the translation distance, and f denotes the focal length.



(a) (b)

Figure 4*. the relationship between adjacent input images. ((a) denotes the real geometry relation, and (b) denotes the approximate translation)*

Estimating Transformation Matrices

We firstly define the rotation angles of horizontal images, and then deduce the rotation angles of other images. For convenience, we arbitrarily select a horizontal image, and assume that its camera coordinate system coincide with the world coordinate system. Therefore, its horizontal and vertical rotation angles are both equal to zero.

As mentioned in last section, there is the approximate translation relationship between adjacent horizontal images, and their translation distances can be computed by the FFT-based phase-correlation algorithm. Therefore, the recursive equations on the rotation angles of the horizontal images can be described as

$$\gamma_{i+1} = \gamma_i + 2\tan^{-1}(\frac{\Delta w_{i+1}}{2f}) \quad (9)$$

$$\eta_{i+1} = \eta_i + 2\tan^{-1}(\frac{\Delta h_{i+1}}{2f}) \quad (10)$$

where γ_i and η_i denote the horizontal and vertical rotation angles of the *i*th horizontal image respectively. Δw and Δh denote the translational distances of adjacent horizontal images.

Similarly, the images in the same group also have the approximate translation as shown in Figure 4. Therefore, the recursive equations of images in the same group can be described as

$$\theta_{i+1} = \theta_i + 2\tan^{-1}(\frac{\Delta w_{i+1}}{2f}) \quad (11)$$

$$\eta_{i+1} = \eta_i + 2\tan^{-1}(\frac{\Delta h_{i+1}}{2f}) \quad (12)$$

where θ_i and φ_i denote the horizontal and vertical rotation angles of the *i*th image in a group.

Creating Vertical Sub-images

Due to the assumption, the recursive equations in above section are not accurate, and have some errors. To avoid the error accumulation, we stitch each group of images into a vertical sub-image, instead of stitching all the images into a whole spherical panorama.

According to the projection relationship between the input images and the vertical sub-image, we can create the vertical sub-image by projection transform. The procedure is described as follows.

1. Process every pixel in the vertical sub-image as follows.
2. Compute the world coordinates (*x*, *y*, *z*) of the pixel, and deduce its spherical coordinates.
3. Compare the spherical coordinates with every rotation angle of the input images, and select the closest one.
4. Project the coordinates (*x*, *y*, *z*) onto the selected image, and use the corresponding pixel value to shade the vertical sub-image.

During the procedure of creating vertical sub-images, it is necessary to take the image fusion process. In our approach, we adopt the multi-resolution pyramid algorithm to perform image fusion [6]. Figure 5 shows some vertical sub-images created by our approach.

Creating Spherical Panoramas

In theory, there is translation relationship between adjacent vertical sub-images. However, the translation is not accurate due to the assumption and computation errors. Therefore, the results obtained by directly computing the translation distances between two adjacent vertical sub-images are not satisfactory (the left image in Figure 6).

We propose a stitching method with division and warping. Firstly, we divide one vertical sub-image into *n* horizontal sub-images. Secondly, we use the FFT-based phase-correlation algorithm to register them. Finally, we stitch them by warping method.

After registration of all the horizontal sub-images, we can map them onto the spherical panorama. During mapping, the horizontal sub-images belonging to the same vertical sub-image are not connected to each other, as shown in Figure 7(a). We connect them by the warping method which includes two steps, i.e. vertically and horizontally warping.

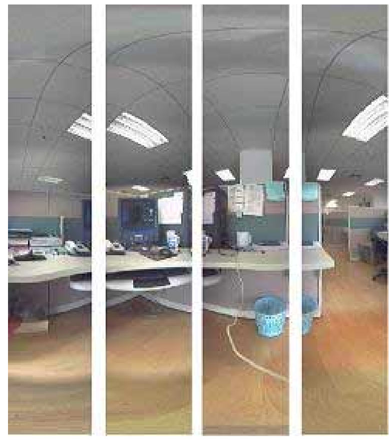


Figure 5. *the vertical sub-images created by our approach.*

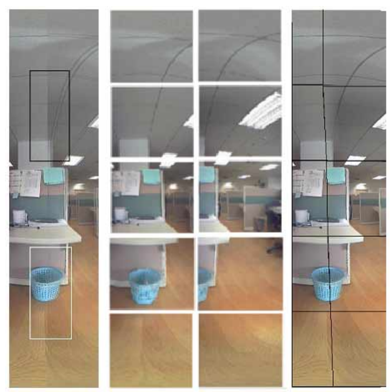
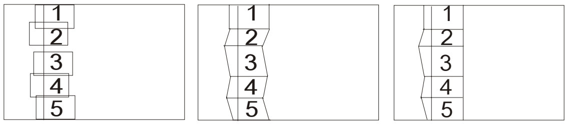


Figure 6. *Left: the result of directly computing the translation distances of two vertical sub-images. Middle: the results of dividing them. Right: the stitched result by division and warping method.*

The object of vertical warping is to seamlessly connect horizontal sub-images belonging to the same vertical sub-image. The procedure is shown in figure 7(b). Firstly, fixing the top edge of the *i*th horizontal sub-image, we adjust its bottom edge to coincide with the top edge of the *i*-1th horizontal sub-image. Secondly, the pixels' locations in the *i*th horizontal sub-image are adjusted accordingly [7].

The object of horizontally warping is to make the horizontal sub-images belonging to the same vertical sub-image locate in the same column. As shown in the Figure 7(c), the procedure is to warp every row of the horizontal sub-images until their right edges are coherent.

The right image in Figure 6 shows the result of stitching two adjacent vertical sub-images. To obtain a seamless spherical panorama, we also should cut out the overlapped areas located in its left and right sides. Figure 8 shows the spherical panorama processed by this algorithm.



(a) (b) (c)

Figure 7*. the procedure of stitching the horizontal sub-images.*

Figure 8*. the spherical panorama created by our method (the lines denote the boundaries of horizontal sub-images)*

Figure 9*. two spherical panoramas created by our approach.*

Experimental results

Figure 9 has shown another two panoramas created by our method. We have used Zhang's method [8] to adjust the brightness of the input mages.

Our approach is implemented on a computer with Pentium4 2.0GHz CPU and 512MB RAM, and each input image has the resolution of 320×240 pixels. As shown in Table 1, the time consumption is no more than 2 minutes.

In order to improve the registration of adjacent vertical sub-images, we horizontally divide them. For a vertical sub-image, the horizontal sub-image amount n should be appropriate. In our approach, the default value of n is 5.

Table 1: the experimental data

Example	Resolution	Images	Time
Fig. 9	2720×1360	90	111.9s
Fig. 10(up)	2376×1188	81	99.5s
Fig. 10(down)	2588×1294	87	104.5s

Conclusion

We present a new approach to automatically create a spherical panorama from image sequences. Our approach does not require particular equipments to capture images, and the involved manual operation is very simple.

To avoid the accumulation of errors, we stitch each group of images into a vertical sub-image. They are finally stitched into a spherical panorama by the horizontal division and warping method. This strategy not only improves the registration accuracy, but also avoids the global optimization.

References

[1] S. E. Chen, QuickTime VR —an image-based approach to virtual environment navigation, Proceedings of SIGGRAPH'95, p. 29. (1995).

[2] R. Szeliski and H -Y Shum, Creating full view panoramic image mosaics and texture-mapped models, Proceedings of SIGGRAPH'97, p. 251. (1997).

[3] M. Brown, and D. G. Lowe, Recognising Panoramas, Proceedings of ICCV'03, p. 1218. (2003).

[4] S. Coorg, and S. Teller, Spherical mosaics with quaternions and dense correlation, International Journal of Computer Vision, 37, 3 (2000).

[5] C. D. Kuglin, and D. C. Hines, The phase correlation image alignment method, IEEE Conference on Cybernetics and Society, p. 163. (1975).

[6] P. J. Burt and E. H. Adelson, A multiresolution spline with application to image mosaics, ACM transaction on Graphics, 2, 4 (1983).

[7] George Wolberg, Digital image warping (IEEE computer society press, 1990) pg. 187.

[8] W. Zhang and B. Zhou, An HDR Brightness Correction Algorithm for Panorama Images, Proceedings of VIIP2004, p. 608. (2004).

Author Biography

Weihua An received his B.S. in computer science in 2001 from Beijing Jiaotong University, Beijing, China. He is now a PhD student in Peking University, Beijing, China. His research is focused on virtual reality and image based rendering.

Bingfeng Zhou received his PhD degree in Peking University. He is now a researcher and doctoral supervisor in the Institute of Computer Science and Technology, Peking University,China.his research is focused on graphics.