3D Scene Reconstruction from Reflection Images in a Spherical Mirror

Masayuki Kanbara, Norimichi Ukita, Masatsugu Kidode and Naokazu Yokoya
Graduate School of Information Science, Nara Institute of Science and Technology
8916-5 Takayama-cho, Ikoma-shi, Nara 630-0192, JAPAN
{kanbara, ukita, kidode, yokoya } @is.naist.jp

Abstract

This paper proposes a method for reconstructing a 3D scene structure by using the images reflected in a spherical mirror. In our method, the mirror is moved freely within the field of view of a camera in order to observe a surrounding scene virtually from multiple viewpoints. The observation scheme, therefore, allows us to obtain the wide-angle multi-viewpoint images of a wide area. In addition, the following characteristics of this observation enable multi-view stereo with simple calibration of the geometric configuration between the mirror and the camera; (1) the distance and direction from the camera to the mirror can be estimated directly from the position and size of the mirror in the captured image and (2) the directions of detected points from each position of the moving mirror can be also estimated based on reflection on a spherical surface. Some experimental results show the effectiveness of our 3D reconstruction method.

1. Introduction

This paper proposes a method for 3D reconstruction of a wide area from an image sequence capturing a spherical mirror by a camera whose projection center is fixed. In the field of computer vision, 3D scene reconstruction using images taken from different view points have attracted much attention [1, 2]. Especially, as computers and cameras have made remarkable progress in recent years, a large number of methods for reconstructing a 3D scene from multiple images have been proposed[3].

One of the major approaches to 3D reconstruction from multiple images is to use a static stereo vision [4, 5]. To reconstruct the whole 3D structure of a scene, a large number of cameras must be employed for wide observation. However, conventional methods cannot employ a large number of images because it is difficult to calibrate a large number of cameras accurately. These methods, therefore, are not appropriate for reconstructing the whole 3D structure of a scene.

One of other approaches is to use an image sequence taken by a moving camera, which is called shape-from-motion [6, 7]. The method can recover extrinsic camera parameters and 3-D positions of natural features simultaneously by tracking the 2D positions of the natural features in multiple images of the sequence. Therefore, the method allows us to move a camera freely and widely in order to observe multidirectional images of a scene. A factorization algorithm [8] is one of the well known shape from motion methods that can estimate a rough 3D scene model stably and efficiently by assuming an affine camera model. However, when the 3-D scene is not suitable for the affine camera model, the estimated results (i.e., both of extrinsic camera parameters and 3-D positions of feature points) are not reliable. Therefore, this method is not suitable for reconstructing a dense 3D structure in general. Although, some other methods of shape-from-motion are based on a projective reconstruction method [9], most of these methods reconstruct only a limited scene from a small number of images and are not designed to obtain a dense structure. We can summarize the above discussion as follows: The methods of 3D reconstruction from many images need an accurate calibration of a large number of cameras or undesired assumptions regarding camera/scene models, and thus these methods become usually complex and/or unstable.

On the other hand, 3D reconstruction methods using reflection images acquired by capturing a mirror surface by a camera have been also investigated [10, 11, 12]. Some of these methods reconstruct 3D information by using the relationship between real features, that are directly captured by a camera, and virtual features, that are points observed through a mirror surface [13, 14]. In these methods, both of the real and virtual features of a 3D point must be captured simultaneously by the camera in order to reconstruct its 3D position, and thus the reconstruction environment is limited to a local scene.

This paper proposes a 3D reconstruction method using only a combination of a spherical mirror and a high res-

0-7695-2521-0/06/$20.00 (c) 2006 IEEE

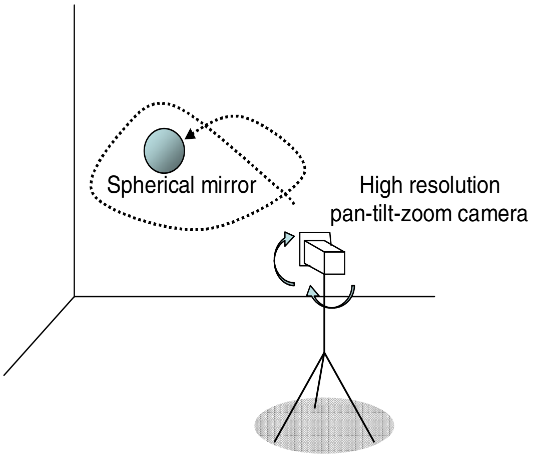

Figure 1. Approach of proposed reconstruction method. A pan-tilt-zoom camera tracks a spherical mirror which is moves freely.

olution video camera. The method obtains scene images observed from multiple viewpoints by capturing reflection images on a moving spherical mirror. Figure 1 shows a desired practical example of our method. In this case, a pan-tilt-zoom camera is employed to continuously observe the high-resolution images of the spherical mirror. As a first step, we used a pan-tilt camera to put our focus upon how to reconstruct 3D information from the multiple reflection images. Figure 2 shows an example image capturing the mirror sphere with the high resolution camera (1600×1200 pixels) at 7.5 frames per second. The spherical mirror can move on an arbitrary path, and the video camera captures it continuously. Because only one camera whose projection center is fixed is used to capture a real scene in our proposed method, calibration of multiple cameras is not required. Moreover, a perspective camera model is used and there is not any limitation regarding the structure of an observed scene in our method. Therefore, it is guaranteed that the reconstruction result is reliable if the geometric configuration (i.e., position) between a camera and a moving mirror can be estimated accurately at each capturing moment. In a practical way, the relative position of the spherical mirror to the camera can be estimated accurately from the position and size of the mirror in an observed image. Here, it should be noted that the accurate intrinsic camera parameters must be given for obtaining the precise position and size of the mirror in the image, and the parameters are estimated in advance. of a pan-tilt-zoom camera. It means that the complex calibration of the extrinsic camera parameters is not necessary if the intrinsic camera parameters of one fixed camera can be calibrated. In addition, our method is endowed with the following advantages:

Figure 2. Example image captured by high resolution video camera (1600×1200pixels, 7.5 fps).

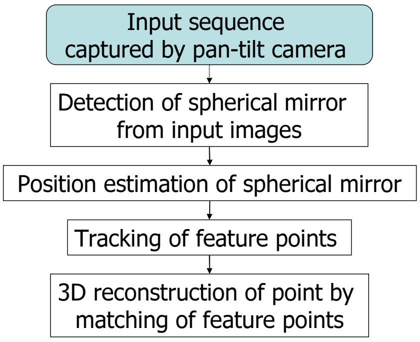

Figure 3. Flow diagram of proposed method.

- Since the images from many different views acquired from reflection regions in the images are used, the occluded regions can be reduced.
- Since the camera focused on the surface of the spherical mirror, all images are observed clearly without defocusing. Therefore, the method can reconstruct the 3D structure of a scene with a large depth.

The rest of this paper is structured as follows. Section 2 describes the 3D reconstruction method using reflection images of a spherical mirror in the images captured by a fixed camera. In Section 3, some experimental results with the proposed method are shown. Finally, Section 4 summarizes the present work.

2. 3D reconstruction from reflections of spherical mirror

In this section, 3D reconstruction from an image sequence that captures a spherical mirror moving along an ar-

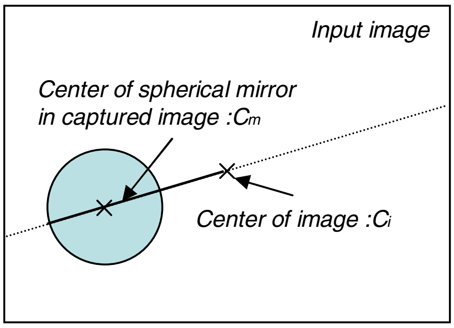

Figure 4. Reflection region in the image.

bitrary path is described. We assume that a projection center of a pan tilt camera is fixed and an observed scene is also static. This calibration can be achieved by using [15]. Figure 3 shows a flow diagram of the proposed method. In the proposed method, first, the region of the spherical mirror is extracted from an input image sequence captured by a high resolution camera at each frame in order to estimate the position of the mirror sphere in the camera coordinate system. Next, feature points in the extracted regions of the spherical mirror are detected and tracked in subsequent frames in order to acquire the point correspondences between the frames. Finally, the 3D positions of the feature points are estimated by obtaining the directions of the feature points in all the images and calculating the intersections of these directions.

2.1. Experimental environments

We use the high resolution digital camera to increase the resolution of the reflection region of spherical mirror. The high resolution camera captures a spherical mirror moving along an arbitrary trajectory as shown in Figure 1. As mentioned before, the proposed method does not need to calibrate the extrinsic parameters (i.e., the position and posture in a world coordinate system) of the camera since only one camera is used and 3D reconstruction is implemented based only on the relative position of the spherical mirror to the camera. What we have to do is only calibrating the internal parameters of the camera for acquiring the exact projection of the spherical mirror under the perspective projection. In our experiments, the parameters are practically estimated with the existing method[16] in advance.

2.2. Detection of spherical mirror

In order to acquire the relationship between the camera and the mirror surface, the center and diameter of the sphere in the input image are estimated by extracting the region of the spherical mirror from it. If the spherical mirror is observed by the camera, the camera itself is captured at the center of the extracted region. By using this property, we attempt to determine the center of the mirror sphere in the input image. To make the problem simple, a color marker is attached around the lens of the camera. The marker is then tracked in the input images by using the color information of the marker. First, in each frame, the region of the colored marker is searched for around the position of the marker in the previous frame. The center of the detected region is then assumed to be the center of spherical mirror. Next, a circle corresponding to the boundary contour of the mirror region is determined by analyzing edge lines in the image. For circle detection based on edge information, a large number of methods have been proposed and they are also effective for detecting the spherical mirror in our method[17].

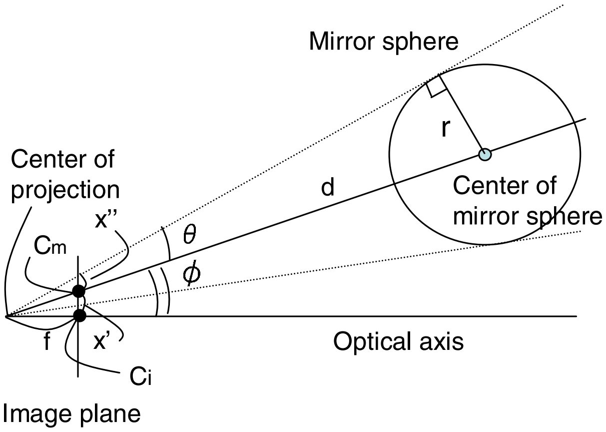

Figure 5. Relationship between camera and spherical mirror.

2.3. Position estimation of spherical mirror

The 3D position of the spherical mirror in the camera coordinate system is calculated from the 2D position of the center and radius of the spherical mirror in the image, which are estimated in the previous section. Figure 4 illustrates an input image including the mirror sphere. Let C_i and C_m denote the centers of the image and the region of the spherical mirror, respectively, as shown in Figure 4. Figure 5 shows a cross-section diagram of a scene along the plane which is determined by C_i, C_m and the center of the projection of the camera. Let θ and φ denote the following angles:

θ :The angle between the 3D lines from the projection center to the center of the sphere and the boundary of the sphere.

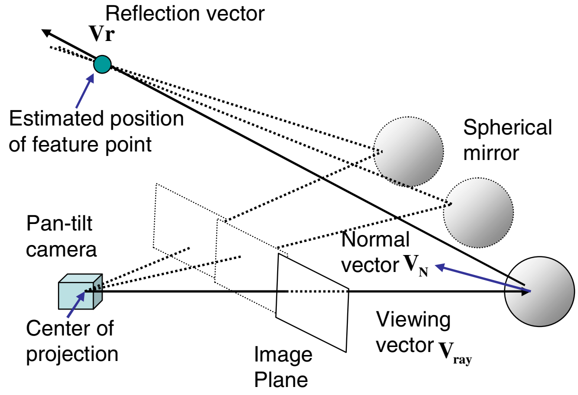

Figure 6. Relationship between a viewpoint and a feature point.

Figure 7. Mirror region is triangulated basedon tracked feature points with Delaunay triangulation method.

- φ :The angle between the 3D lines from the projection center to C_i and C_m.

Under the above condition, the following equations stand:

$$\begin{aligned} \tan\varphi &= \frac{x'}{f} && (1) \\ \tan(\varphi+\theta) &= \frac{x'+x''}{f}, && (2) \end{aligned}$$

where x', x'' and f denote the distance from C_i to C_m in the image, the apparent radius of the spherical mirror, and the focal length of the camera, respectively. Since x', x'', and f are known, θ and φ can be calculated. Next, let r and d denote the radius of the spherical mirror in a real world and the distance between the projection center and the center of the spherical mirror, respectively, as shown in Figure 5. Now, the following equation stands:

$$d = \frac{r}{\sin\theta}. \quad (3)$$

Therefore, if the radius of the sphere is known[1], the distance d can be estimated. The 3D position of the center of the spherical mirror is then able to be acquired from d and the direction of the center of the spherical mirror, which can be also calculated from the image coordinates of the center of the mirror sphere and f in the camera coordinate system.

2.4. Tracking of feature points

For 3D reconstruction, corresponding feature points between different images is required. Within the detected region of the spherical mirror, feature points detected in the first frame are tracked through subsequent frames in order to obtain the point correspondences in the whole sequence. In a frame, each feature point is searched for in an area in which that point is in the previous frame. In our implementation, tentative feature points are extracted by interest operator [18] in the first frame, and these points are then tracked by using a template matching approach.

2.5. 3D reconstruction of feature points

Next, the 3D directions of the detected feature points are estimated for 3D reconstruction. The directions can be estimated based on (1) the relationship between the camera and the spherical mirror and (2) a normal vector of the mirror surface in which the feature point in the image is headed from the projection center to the spherical mirror as shown in Figure 6. The direction $\mathbf{V_r}$ from the center of the spherical mirror to the 3D feature point can be calculated by using the viewing vector $\mathbf{V_{ray}}$ and the normal vector $\mathbf{V_N}$ at incident point on the mirror surface as follows:

$$\mathbf{V_r} = -\mathbf{2}(\mathbf{V_N} \cdot \mathbf{V_{ray}})\mathbf{V_N} + \mathbf{V_{ray}}. \quad (4)$$

With the above equation, the 3D line passing through each feature point can be determined in all the frames. The 3D position, which has the smallest summation of the distances to the 3D lines passing through corresponding points in the images, is considered to be the position of that feature point in space. Finally, in order to reconstruct the 3D scene, the image is triangulated by the feature points with Delaunay triangulation method as shown in Figure 7.

3. Experiments

We have carried out the 3D reconstruction experiments in an indoor scene illustrated in Figure 8. We have used an

---

[1] In our experiments, the radius of the spherical mirror is precisely measured in advance.

Figure 9. Examples of sampled input images.

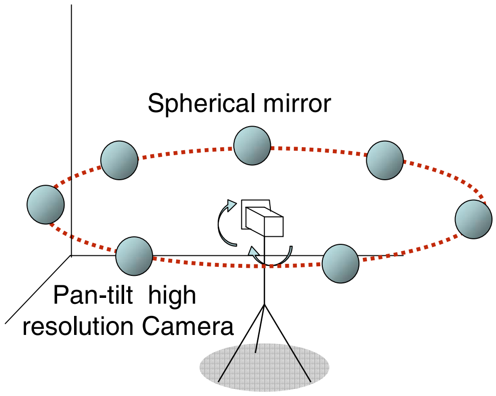

Figure 8. Experimental environment.

IEEE1394 camera (Pointgrey research: Scorpion) that can capture 1600 × 1200 pixel images at 7.5 frames per second, and a spherical mirror whose diameter is 101.6mm. In this experiment, the spherical mirror roughly circles around the camera on the horizontal plane as shown in Figure 8. The sphere is tracked by the camera manually, and 32 images containing the spherical mirror are captured as shown in Figure 9. The mirror sphere moving along an arbitrary path is projected onto the input images. Figure 10 illustrates the relationship between the positions of the spherical mirror at different moments in the camera coordinate system. The three axes colored red, blue, and green in Figure 10 indicate the x, y, and z axes, respectively. Finally, the results of 3D reconstruction are shown in Figure 11. Since the positions of a limited number of feature points are estimated, the reconstructed model has a simple structure. However, we can confirm that the appearance of the generated 3D model is correct. In this experiment, the calculation time of processes which include mirror detection, feature tracking and 3D model reconstruction is a few minutes with a desktop PC (CPU P4 3.2GHz, Memory 1GB).

4. Conclusion

This paper has proposed a 3D reconstruction method using only reflection images of a moving spherical mirror observed in images captured by a high resolution camera. While the spherical mirror moves on an arbitrary path, the camera captures it continuously. Because only one fixed camera is used to capture a real scene in the proposed method, the biggest advantage of our method is the feasibility without complex calibration of the extrinsic camera parameters.

To improve the performance of our proposed method, the image resolution of the spherical mirror in an image (i.e., the size of the mirror in the image) should be higher. Employing a pan-tilt-zoom camera allows us to continuously observe a moving mirror as large as possible. Accordingly, we are developing a 3D reconstruction system by integrating the proposed method and a pan-tilt-zoom camera.

References

[1] D. D. Morris and T. Kanade: "A Unified Factorization Algorithm for Points, Lines Segments and Planes with Uncertainty Models," Proc. 6th Int. Conf. on Computer Vision, pp. 696–702, 1998.

[2] N. Yokoya, T. Shakunaga and M. Kanbara: "Passive Range Sensing Techniques: Depth from Images," IEICE Trans. Information and Systems, Vol. E82-D, No. 3, pp. 523–533, 1999.

[3] T. Sato, M. Kanbara, N. Yokoya and H. Takemura: "Dense 3-D Reconstruction of an Outdoor Scene by Hundreds-baseline

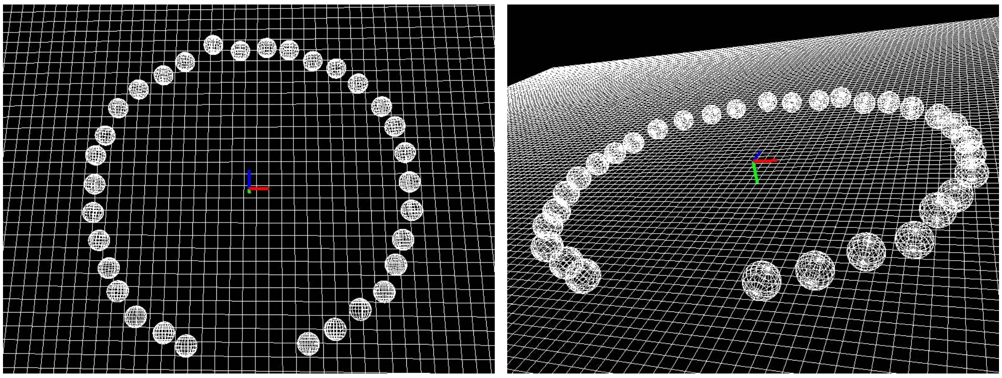

Figure 10. Estimation results of the spherical mirror positions.

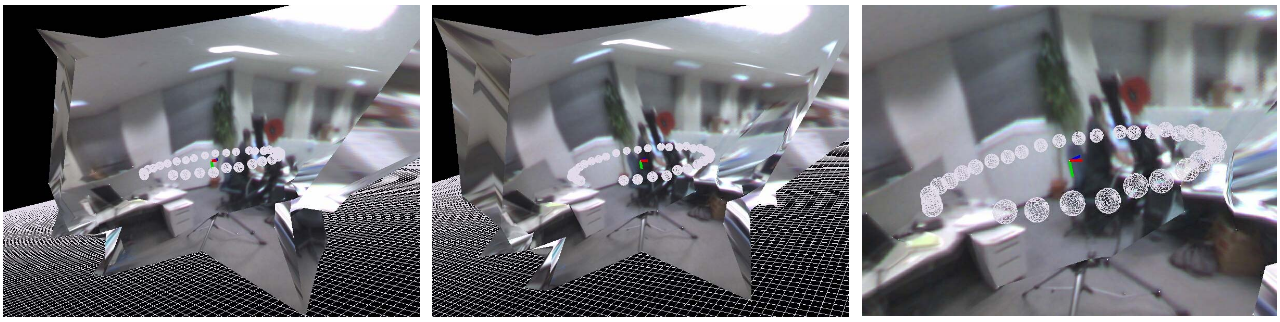

Figure 11. Results of 3D reconstruction.

Stereo Using a Hand-held Video Camera," Int. Jour. of Computer Vision, Vol. 47, No. 1-3, pp. 119–129, 2002.

[4] S. T. Barnard and M. A. Fischler: "Computational Stereo," ACM Computing Survays, Vol. 14, No. 4, pp. 553–572, 1982.

[5] M. Okutomi and T. Kanade: "Multiple-baseline Stereo," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 15, No. 4, pp. 353–363, 1993.

[6] T. Sato and N. Yokoya: "New Multi-baseline Stereo by Counting Interest Points," Proc. Canadian Conf. on Computer and Robot Vision, pp. 96–103, 2005.

[7] G. Roth and A. Whitehead: "Using Projective Vision to Find Camera Positions in an Image Sequence," Proc. 13th Int. Conf. on Vision Interface, pp. 87–94, 2000.

[8] C. Tomasi and T. Kanade: "Shape and Motion from Image Streams under Orthography: A Factorization Method," Int. Jour. of Computer Vision, Vol. 9, No. 2, pp. 134–154, 1992.

[9] M. Pollefeys, R. Koch, M. Vergauwen, A. A. Deknuydt and L. J. V. Gool: "Three-dimentional Scene Reconstruction from Images," Proc. SPIE, Vol. 3958, pp. 215–226, 2000.

[10] S. Savarese, L. F. Fei and P. Perona: "What Do Reflections Tell Us about the Shape of a Mirror?," Proc. Symposium on Applied Perception in Graphics and Visualization, pp. 115–118, 2004.

[11] M. Oren and S. K. Nayar: "A Theory of Specular Surface Geometry," Int. Jour. of Computer Vision, Vol. 24, No. 2, pp. 105–124, 1996.

[12] B. Micusik and T. Pajdla: "Autocalibration and 3D Reconstruction with Non-Central Catadioptric Cameras," Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp. 58–65, 2004.

[13] S. Savarese, M. Chen and P. Perona: "Local Shape from Mirror Reflections," Int. Jour. of Computer Vision, Vol. 64, No. 1, pp. 31–67, 2005.

[14] M. Halstead, B. Barsky, S. Klein and R. Mandell: "Reconstructing Curved Surfaces from Specular Reflection Patterns Using Spline Surface Fitting of Normals," Proc. SIGGRAPH'96, pp. 335–342, 1996.

[15] T. Wada and T. Matsuyama: "Appearance Sphere: Background Model for Pan-Tilt- Zoom Camera," Proc. Int. Conf. on Pattern Recognition, Vol. A, pp. 718–722, 1996.

[16] Z. Zhang: "A Flexible New Technique for Camera Calibration," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 22, No. 11, pp. 1330–1334, 2000.

[17] N. Guil and L. Zapata: "Lower Order Circle and Ellipse Hough Transform," Pattern Recognition, Vol. 30, No. 10, pp. 1729–1744, 1997.

[18] C. J. Harris and M. Stephens: "A Combined Corner and Edge Detector," Proc. 4th Alvey Vision Conf., pp. 147–151, 1988.