

# Silhouette-Based Camera Calibration from Sparse Views under Circular Motion

Po-Hao Huang

Shang-Hong Lai

Computer Vision Laboratory

Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan  
{even, lai}@cs.nthu.edu.tw

## Abstract

*In this paper, we propose a new approach to camera calibration from silhouettes under circular motion with minimal data. We exploit the mirror symmetry property and derive a common homography that relates silhouettes with epipoles under circular motion. With the epipoles determined, the homography can be computed from the frontier points induced by epipolar tangencies. On the other hand, given the homography, the epipoles can be located directly from the bi-tangent lines of silhouettes. With the homography recovered, the image invariants under circular motion and camera parameters can be determined. If the epipoles are not available, camera parameters can be determined by a low-dimensional search of the optimal homography in a bounded region. In the degenerate case, when the camera optical axes intersect at one point, we derive a closed-form solution for the focal length to solve the problem. By using the proposed algorithm, we can achieve camera calibration simply from silhouettes of three images captured under circular motion. Experimental results on synthetic and real images are presented to show its performance.*

## 1. Introduction

Camera calibration from uncalibrated image sequences without using any specific calibration patterns can be classified into two camps; namely, the feature-based and silhouette-based approaches. In the feature-based approach, structure from motion (SfM) algorithm [20] determines the camera information and the 3D structure of the object simultaneously from the feature correspondences. In the silhouette-based approach, the cameras are first calibrated from the object silhouettes and then the volumetric description, intersection [1] or the image-based visual hull (IBVH) [2] techniques are applied for object modeling.

It takes more efforts to calibrate cameras in general motion from silhouettes [18, 19] than feature points. Boyer [19] proposed a quantitative and qualitative criterion for measuring the silhouette consistency from viewing cones

(or named visual cones), but the estimated results heavily depend on the initial parameters. Sinha et al. [18] proposed a camera calibration method from dynamic silhouettes. The key step is to robustly compute the epipolar geometry from two views. Therefore, dynamic silhouettes are needed for providing enough constraints.

Additional constraints on camera motion, such as circular motion, can facilitate the camera calibration. It may seem to be a restriction to assume the camera motion to be circular motion. However, in real applications, such as 3D digitization in digital museum, 3D object reconstruction from a circular motion sequence is a practical and widely used approach for generating 3D models and a number of 3D reconstruction methods focused on circular motion have been proposed, including the feature-based [3,4,5,6,7] and silhouette-based [8,9,12,13,14,15] methods.

In this paper, we deal with the problem of camera calibration from silhouettes under circular motion. Different from previous silhouette-based methods based on surface of revolution (SoR), the proposed approach can be applied with very few sparsely located views to calibrate the cameras. We introduce the concept of mirror symmetry property [16,17] into circular motion and derive the property that images under circular motion share a common homography that relates silhouettes with epipoles. With this property, epipoles can be located directly from the silhouettes and then the remaining camera parameters can be recovered.

In addition, in the degenerate case, when the optical axes of all cameras intersect at one point, for a reference camera, the projection of x-axis direction in an image, or named x-axis vanishing point, is an infinite point, which influences the calibration process. For this degenerate case, we derive a closed-form solution to determine the focal length by using the epipoles between three images captured under circular motion.

This paper is focused on the problem of camera calibration from silhouettes instead of features, thus it is particularly useful when it is difficult to establish accurate feature correspondences from texture-less, transparent, translucent, or reflective objects, such as jadeite material.

The remainder of this paper is organized as follows. We review some previous works in section 2. Section 3

describes the circular motion geometry. In section 4, we present how to apply the mirror symmetry property to formulate the common homography. The proposed camera calibration algorithm from silhouettes under circular motion is described in section 5. Experimental results on both synthetic and real data are given in section 6. Finally, we conclude this paper in section 7.

## 2. Previous works

For camera calibration under circular motion, the feature-based [3,4,5,6,7] and silhouette-based [8,9,12,13,14,15] approaches have been proposed.

In the feature-based approaches, Fitzgibbon et al. [3] computed the fundamental matrices and trifocal tensors to uniquely determine the rotation angles. The reconstruction is determined up to a two-parameter family. Jiang et al. [4,5] further developed a method that avoids the computation of multi-view tensors to recover the circular motion geometry by either fitting conics to tracked points in at least five images or computing a plane homography from minimally two points in four images. Cao et al. [6] aimed at the problem of varying focal lengths under circular motion with a constant but unknown rotation angle. With the invariance of the essential matrix, the ratio of varying focal lengths can be determined. Zhong and Hung [7] proposed a circular projective reconstruction method.

In the silhouette-based approaches, Mendonca and Cipolla [8] addressed the problem of estimating the epipolar geometry from contours in the affine and circular motion cases. For the circular motion case, the relation between the epipolar tangencies and the image of rotation axis is used to define a cost function. Under the assumption of constant rotation angle, they iteratively minimize the cost function to determine the two common epipoles between adjacent images. Huang and Lai [15] further extended this method to determine the focal length. In [9], Mendonca et al. exploited the symmetry properties of the SoR swept out by the rotating object to obtain an initial guess of image invariants, followed by several one-dimensional searching steps to obtain the epipolar geometry. However, it still requires the knowledge of the camera intrinsic parameters in advance to recover the rotation angles and the object structure. Zhang et al. further extended this method to achieve auto-calibration [12] and formulated the circular motion as 1D camera geometry to achieve more robust motion estimation [13].

Most of the silhouette-based methods are based on the SoR to obtain an initial guess of image invariants, thus it is infeasible to deal with sparsely located images. Wong et al. [10,11] presented a method based on two epipolar tangencies between image pairs to deal with this problem, but it requires the camera intrinsic parameters to be known

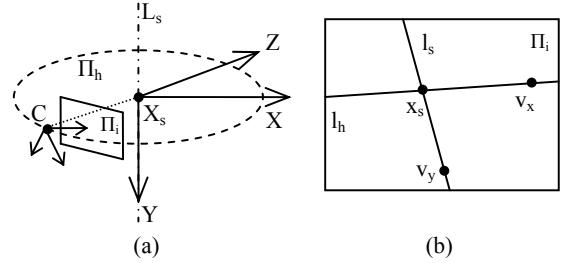


Figure 1. (a) The geometry of circular motion. (b) The image invariants under circular motion

in advance and needs to solve a high-dimensional optimization problem with non-trivial initialization. The method described in [15] could deal with this problem to achieve auto-calibration. Nevertheless, it is restricted to the constant interval angle assumption. Recently, Hernandez et al. [14] defined the silhouette coherence that uses the entire visible silhouettes, not only the tangencies, to measure the silhouette consistency. Cameras are calibrated by maximizing the silhouette coherence and good results could be obtained when the errors of the initial interval angles are within 15 degrees. However, a roughly initial guess for camera parameters is still needed.

## 3. The geometry under circular motion

The geometry of circular motion can be depicted in Figure 1(a). Let the world coordinate system be centered at the circle center \$X\_s\$ and a reference camera C be placed on the negative Z-axis and rotate about Y-axis (\$L\_s\$). If the camera intrinsic parameters and pose are kept constant under the circular motion, the image \$\Pi\_i\$ will contain image invariants [3,4,5,12] of the geometry as shown in Figure 1(b). In this figure, line \$l\_h\$ is the image of the camera plane \$\Pi\_h\$ and line \$l\_s\$ is the projection of \$L\_s\$. Three points, \$v\_x\$, \$v\_y\$, and \$x\_s\$ are the vanishing points of X-axis, Y-axis and Z-axis, respectively. The point \$x\_s\$ is also the intersection of \$l\_s\$ and \$l\_h\$.

In mathematical expression, the camera projection matrix P can be written as:

$$P = KR[R_y(\theta)|-T] \quad (1)$$

where K is camera intrinsic parameter, \$R=[r\_1 \ r\_2 \ r\_3]\$ is the camera pose, T is the camera position in world coordinate and \$R\_y(\theta)\$ is the rotation matrix about \$L\_s\$ with angle \$\theta\$.

The three vanishing points can be expressed as:

$$[v_x \ v_y \ x_s] \sim KR = K[r_1 \ r_2 \ r_3] \quad (2)$$

where symbol “~” denotes the equivalence relation in the homogenous coordinate.

Therefore, the image of \$L\_s\$ and the projection of camera plane \$\Pi\_h\$ are related to these invariant points as follows:

$$l_h = x_s \times v_x, \quad l_s = x_s \times v_y \quad (3)$$

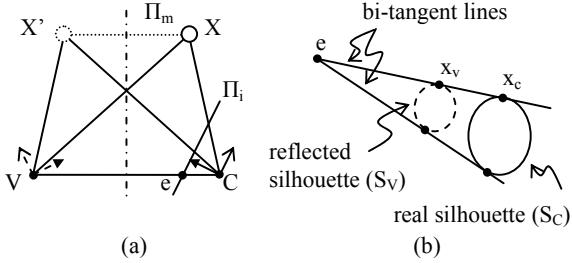


Figure 2. (a) The bird's eye view of the two view geometry generate by a plane mirror. (b) Epipole and bi-tangent lines.

## 4. Homography from mirror symmetry

In this section, we describe how to apply the mirror symmetry property to derive the plane homography that relates the object contours with epipoles.

### 4.1. Epipolar geometry from mirror reflection

As described in [16,17], the corresponding geometry established by placing a plane mirror in the scene can be illustrated as a bird's eye view in Figure 2(a). In the figure,  $\Pi_m$  is the mirror plane, X is an object point and its reflection  $X'$  could be imagined as a virtual point at the opposite side of the mirror plane. C represents the real camera with the associated image plane denoted by  $\Pi_i$ . Similarly, there exists a virtual camera V on the other side of mirror and its projection onto  $\Pi_i$ , which is the epipole denoted by e. Because of mirror reflection, the projection of  $X'$  onto camera C can be regarded as the projection of X onto camera V. Therefore, the two view geometry for a camera with a plane mirror is established.

For simplicity, we denote an object silhouette in the image plane of camera C as  $S_c$ . In this setting, the image plane  $\Pi_i$  contains two silhouettes, real  $S_c$  and reflected  $S_v$ , as depicted in Figure 2(b). The major property is that the epipole e can be located by the outer tangent lines of the two silhouettes,  $S_c$  and  $S_v$ , and the tangent points, for instance  $x_c$  and  $x_v$ , will be considered as the corresponding points (see [16,17] for details).

### 4.2. Circular motion and mirror reflection

The geometry under the setting of plane mirrors is similar to the circular motion. To be more specific, the relationship is illustrated in Figure 3. In Figure 3(a),  $C_1$  and  $C_2$  denote two instances of the camera under circular motion. Imagine that there is a plane mirror  $\Pi_m$ , which is parallel to and passes through the rotation axis, to be placed in the middle of two camera centers. According to the mirror symmetry property, there is a virtual camera  $V_2$ , which is the reflection of camera  $C_1$  according to  $\Pi_m$ , centered at the same position of camera  $C_2$ . Therefore, we can relate the epipole  $e_{C2}$ , which is the projection of  $C_2$  onto

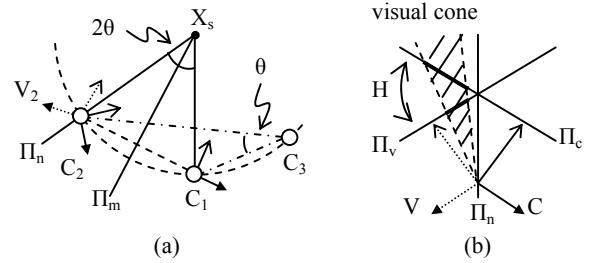


Figure 3. The geometry relationship between the circular motion and the mirror reflection.

$C_1$ , with silhouettes  $S_{C1}$  and  $S_{V2}$  by the following equation:

$$e_{C2}^T \cdot (x_{C1} \times x_{V2}) = 0 \quad (4)$$

where  $x_{C1}$  and  $x_{V2}$  is a pair of corresponding points from the tangencies of silhouettes  $S_{C1}$  and  $S_{V2}$ .

Similarly, camera  $V_2$  can be regarded as the reflection of camera  $C_2$  according to another mirror plane  $\Pi_n$ . To be more specific, we enlarge this part in Figure 3(b). In this figure,  $\Pi_c$  and  $\Pi_v$  are the image planes of cameras C and V, respectively. Because of the coincident camera center of cameras C and V, silhouettes  $S_c$  and  $S_v$  can be transformed with each other via a non-singular 3x3 homography H [20].

From mathematical derivation, the camera projection matrices for cameras C and V can be expressed as:

$$P_C = KR[R_y(\theta)|-T] \quad (5)$$

$$P_V = KR\Sigma[R_y(\theta)|-T] \quad (6)$$

where  $\Sigma = \text{diag}([-1 \ 1 \ 1])$ .

Suppose the projections of an object point X onto cameras C and V are denoted by  $x_c$  and  $x_v$ , respectively. The relation between the corresponding image points can be expressed as:

$$x_v \sim KR\Sigma R^T K^{-1} x_c = Hx_c \quad (7)$$

The homography H only involves the constant camera intrinsic parameters and camera pose. Thus, all cameras under circular motion share the common homography.

Substituting equation (7) into equation (4) leads to

$$e_{C2}^T \cdot (x_{C1} \times Hx_{C2}) = 0 \quad (8)$$

Similarly, we can derive the following relation

$$e_{C1}^T \cdot (x_{C2} \times Hx_{C1}) = 0 \quad (9)$$

For any two views under circular motion, a pair of correspondence points with two epipoles can provide two linear constraints on H according to equation (8) and (9). With enough constraints, H can be linearly determined by the least-square estimation.

On the other hand, given the homography H, a silhouette can be transformed to its reflected silhouette. Therefore, the epipoles of any two views can be determined from the bi-tangent lines of one of the silhouettes and the transformed silhouette of the other one.

In addition, equation (8) can be rewritten as follow:

$$x_{C1}^T ([e_{C2}]_x H) x_{C2} = 0 \quad (10)$$

Therefore, the fundamental matrix can be expressed as:

$$F = [e_{C2}]_x H \quad (11)$$

The derived equations are similar to those in [9,20]. However, the property is derived from the mirror symmetry property. Therefore, we have a new observation that the epipoles of two views under circular motion can be directly located instead of performing one-dimensional search [9].

### 4.3. Homography and image invariants

With the pole-polar relationship, we can show that

$$l_s = \omega v_x \quad (12)$$

where  $\omega = K^{-T} K^{-1}$  is the image of the absolute conic.

The homography  $H$  derived in the previous section can be further used to relate to the image invariants as follow:

$$H = KR\Sigma R^T K^{-1} = I - 2K\mathbf{r}_1\mathbf{r}_1^T K^{-1} = I - 2\frac{v_x l_s^T}{v_x^T l_s} \quad (13)$$

Therefore, if there is knowledge about  $v_x$  or  $l_s$ , equation (8) and (9) can also be applied to determine the unknown part with a linear solution.

## 5. Camera Calibration

In this section, we discuss how to estimate the camera parameters by using the proposed property and how to determine the focal length under the degenerate case. We assume the optical centers coincide with the image centers, and the camera intrinsic parameter has only one unknown, i.e. the focal length.

### 5.1. Recovery of the homography

If the epipoles are available, the tangent points of silhouettes induced by the epipoles are considered as point correspondences, and equation (8) and (9) are used to compute the homography  $H$ . One way to extract the epipoles from object silhouettes can be found in [8].

If the epipoles are not available, a low-dimensional search of the optimal  $H$  in a bounded region is performed. As given in equation (7),  $H$  contains four parameters which are focal length, denoted by  $f$ , and three rotation angles, denoted by  $\theta_x$ ,  $\theta_y$  and  $\theta_z$ , respectively, of the camera pose. In fact,  $H$  is independent of the rotation angle of  $x$ -axis because of the following relation:

$$R_x(\theta_x)\Sigma R_x(\theta_x)^T = \Sigma \quad (14)$$

Thus, the number of unknowns of  $H$  is reduced to three.

Under circular motion, all cameras are positioned on the same plane, which is  $\Pi_h$  as shown in Figure 1(a), and their projections, which are epipoles, are co-linear. The homography  $H$  can hence be determined by minimizing the

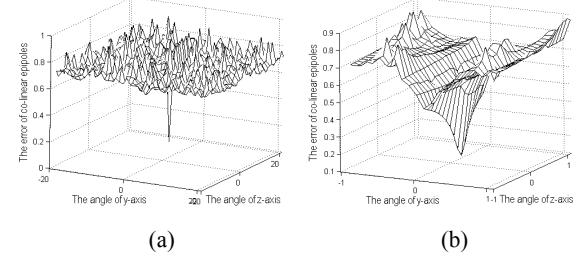


Figure 4. The error function of different parameters of  $H$ .

distance from the epipoles to the co-linear line,  $l_h$ , and the cost function is defined as follow:

$$Cost(f, \theta_y, \theta_z) = \sum_{(s_i, s_j) \in S} (dist(e_{ij}, l_h) + dist(e_{ji}, l_h)) \quad (15)$$

where the function  $dist(\cdot)$  gives the distance from a point to a line, the set  $S$  contains all silhouette pairs and  $e_{ab}$  denotes the epipole which is the projection of camera  $C_b$  onto  $C_a$ .

As described in section 4.2, the epipoles of a pair of images can be located from a given  $H$  and the epipole co-linear line,  $l_h$ , is computed by a least-square line fitting.

Unfortunately, under a given focal length, the error function is complex as shown in Figure 4. Figure 4(a) is the error function of the rotation angles within  $\pm 20$  degrees of the ground truth and Figure 4(b) is that within  $\pm 1$  degree. Thus, the optimization methods will be easily trapped into local minima without very good initialization. However, the values of the unknowns can be bounded within a reasonable range. The optimal  $H$  can be searched discretely in a finite parameter spaces.

In practice, a candidate list of  $H$  with the cost smaller than a predefined threshold, which can be set to the average cost, will be kept because of the noise. The final solution is chosen with the smallest error in the epipolar tangency constraints, whose detailed description can be found in [10].

### 5.2. Recovery of the image invariants

When the homography  $H$  is obtained, the epipoles of a pair of views can be determined. The line  $l_h$  is determined via line fitting of epipoles. The line  $l_s$  is determined via line fitting of the intersections of the corresponding epipolar lines [8]. The point  $x_s$  hence can be computed from the intersection of  $l_h$  and  $l_s$ . The point  $v_x$  is computed from equation (12) given  $f$  and  $l_s$  or from equation (8), (9) and (13) when the epipoles are available.

### 5.3. Recovery of the focal length

If the homography  $H$  is determined via a bounded-search process, the focal length is therefore determined. If it is determined via given epipoles, with the recovered image

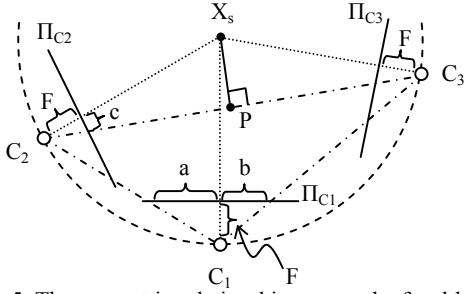


Figure 5. The geometric relationship among the focal length and the epipoles.

invariants,  $v_x$  and  $l_s$  will provide two constraints to the image of the absolute conic as expressed in equation (12). The focal length can therefore be computed by solving two quadratic equations.

#### 5.4. Recovery of the camera pose

From equation (2), camera pose  $R$  can be computed by

$$\begin{aligned} r_1 &= \alpha K^{-1} v_x / \|K^{-1} v_x\|, \alpha = \pm 1 \\ r_3 &= \beta K^{-1} x_s / \|K^{-1} x_s\|, \beta = \pm 1 \\ r_2 &= r_3 \times r_1 \end{aligned} \quad (16)$$

Notice that, the sign of rotation axis has no difference for projection to the image coordinates, but back-projection of image points will lead to a sign ambiguity. This ambiguity can be resolved by back-projecting the epipole, which is obtained from images, and checking the sign with the corresponding camera position. Finally, the overall optimal orthogonalization is done by SVD.

#### 5.5. Recovery of the rotation angle about $L_s$

The inscribed angle theorem of circle states that an inscribed angle is exactly half the corresponding central angle. As shown in Figure 3(a), we have

$$\angle C_1 X_s C_2 = 2 \angle C_1 C_3 C_2 \quad (17)$$

where  $C_3$  is an arbitrary camera on the circle.

In this case, the rotation angle about  $L_s$  between cameras  $C_1$  and  $C_2$  can be obtained by calculating the angle between the back-projection lines of the corresponding epipoles of cameras  $C_1$  and  $C_2$  in the image of camera  $C_3$  as follow:

$$\angle C_1 C_3 C_2 = \text{ang}(K^{-1} e_1, K^{-1} e_2) \quad (18)$$

where  $e_1$  and  $e_2$  are epipoles of  $C_1$  and  $C_2$  in the image of  $C_3$ .

#### 5.6. Degenerate case

When the optical axes intersect at one point, the x-axis vanishing point,  $v_x$ , is an infinite point and can not be determined from the SoR [9,12,13]. Therefore, the constraints arisen from  $v_x$ , such as equation (12), can not be used for determining the focal length.

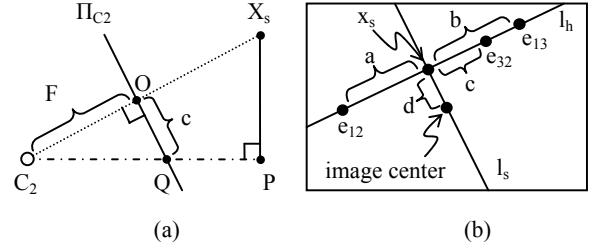


Figure 6. (a) The enlarged figure of  $\Delta XPC_2$  (b) The relationship among epipoles and image entities.

In addition, in this case,  $\theta_y$  is zero, the parameters of  $H$  will hence be reduced to only one, which is  $\theta_z$ .

$$H = KR_z(\theta_z) \Sigma R_z(\theta_z)^T K^{-1} = R_z(\theta_z) \Sigma R_z(\theta_z)^T \quad (19)$$

where  $K = \text{diag}(f, f, f, 1)$  and  $KR_z = R_z K$ .

Thus, the focal length will also be eliminated in  $H$ .

However, from three images, there are constraints on the epipoles and the focal length, as shown in Figure 5, thus a closed-form solution of the focal length can be derived in this case. For the derivation, we introduce some symbol definitions. In Figure 5,  $C_1$  to  $C_3$  are three cameras and  $\Pi_1$  to  $\Pi_3$  are their corresponding image planes. The image plane is placed at a distance  $f$  from the camera center,  $F$  is the distance from camera center to  $x_s$ , and  $e_{ab}$  denotes the epipole that is the projection of camera  $C_b$  on camera  $C_a$ . As depicted in Figure 6(b),  $a$ ,  $b$ ,  $c$  denote the distance from  $x_s$  to  $e_{12}$ ,  $e_{13}$ , and  $e_{23}$ , respectively, and  $d$  is the distance from  $x_s$  to the image center. In Figure 5, assume the world center is placed at circle center  $X_s$  in 2D case and the coordinate of  $C_1$  is  $[0 -r]$ , the coordinate of  $C_2$  and  $C_3$  can be expressed as:

$$C_2 = [2arF/(a^2 + F^2), r(F^2 - a^2)/(a^2 + F^2)] \quad (20)$$

$$C_3 = [2brF/(b^2 + F^2), r(F^2 - b^2)/(b^2 + F^2)] \quad (21)$$

Note that Figure 6(a) is the enlarged figure of  $\Delta XPC_2$  in Figure 5. Since  $\Delta XPC_2$  is similar to  $\Delta QOC_2$ , we have the following relationship:

$$\overline{QO} : \overline{OC_2} = \overline{X_s P} : \overline{PC_2} \quad (22)$$

where

$$\overline{QO} = c, \overline{OC_2} = F, \overline{X_s P} = \sqrt{r^2 - \overline{PC_2}^2}, \overline{PC_2} = \frac{\overline{C_2 C_3}}{2} \quad (23)$$

The distance from  $C_2$  to  $C_3$  can be calculated from equation (20) and (21). Also,  $F$  is the distance from camera center to  $x_s$ , given by

$$F = \sqrt{f^2 + d^2} \quad (24)$$

Therefore, from equation (22), (23) and (24), we can derive the closed-form solution for the focal length  $f$  by using  $a$ ,  $b$ ,  $c$  and  $d$  as follow:

$$f = \sqrt{\pm (bc - ac) - ab - d^2} \quad (25)$$

where  $f$  is chosen as the positive solution.

In addition to the case with a rotationally symmetric object, another degenerate case occurs when the contact points of the epipolar planes with the object lie on the rotation axis. In this case, points  $x_c$  and  $x_v$ , depicted in Figure 2(b), are identical and the derived constraint cannot be used. However, this problem can be avoided by putting the object away from the rotation axis. The expense is to sacrifice image resolution for the object, which is not a problem for a high-resolution camera.

## 6. Experimental results

In this section, both synthetic and real data are used to evaluate the proposed algorithm. For each test image sequence, only the silhouettes, instead of feature correspondences, are used for camera calibration.

### 6.1. Synthetic data

In this part, the Stanford bunny model is used to generate 100 data sets to test the algorithm. The model is projected to each image, with the coordinates of all pixels rounded, and then the convex hull algorithm is applied to find the silhouette. Each test set contains 12 images of size 800x600 pixels with interval angle 30 degrees. The value of focal length and the camera pose are randomly generated. The range of  $f$  is 1500~5000 pixels,  $\theta_x$  is within -10~50 degrees,  $\theta_y$  and  $\theta_z$  are within -5~5 degrees. For each test, the number of images used for the reconstruction is randomly chosen from 3 to 6.

The experimental results are shown in Table 1. In this Table, # is the number of images used and  $\Delta\theta_i$  is the average error of the interval angles. The errors of angles are represented in degrees. The error in focal length is computed by taking the focal length difference divided by the ground truth.

For the experiment with the degenerate case, we generate another 100 data sets based on the same setting as the previous ones except fixing  $\theta_y$  to 0 degree. In this experiment, the number of images used for reconstruction is set to 3 and the estimated results are listed in Table 2.

### 6.2. Real data

In the experiment on real data set, we provide three kinds of image sequences. All sequences are sparsely located under circular motion and are not suitable for producing SoR. We applied the proposed algorithm to estimate the camera parameters and then minimized the epipolar tangency constraints [10] for all pairs of views to refine the camera parameters.

The first real-data experiment is on the Fox sequence, which contains 12 images of size 640x480 pixels and the angle between images is about 30 degrees (may deviate a little bit due to mechanical error). Some example views are

Table 1. Accuracy of the recovered camera parameters

#	err.	$\Delta\theta_x$ (°)	$\Delta\theta_y$ (°)	$\Delta\theta_z$ (°)	$\Delta\theta_f$ (°)	$\Delta f$ (%)
3	avg.	0.96	0.23	0.84	2.61	5.32
	std.	1.23	0.49	1.17	5.71	4.86
4	avg.	0.86	0.17	0.51	1.49	4.85
	std.	1.17	0.14	0.58	1.47	4.73
5	avg.	0.53	0.13	0.38	1.12	3.98
	std.	0.51	0.11	0.34	0.73	3.60
6	avg.	0.46	0.15	0.43	0.93	3.70
	std.	0.39	0.13	0.32	0.55	2.89

Table 2. Accuracy of the recovered camera parameters for the degenerate case

#	err.	$\Delta\theta_x$ (°)	$\Delta\theta_y$ (°)	$\Delta\theta_z$ (°)	$\Delta\theta_f$ (°)	$\Delta f$ (%)
3	avg.	0.56	0	0.15	1.41	2.87
	std.	1.03	0	0.36	1.82	4.56

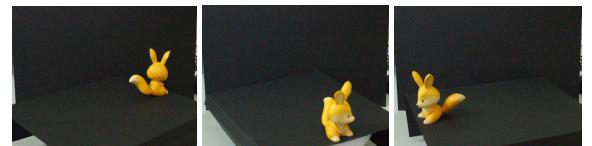


Figure 7. Example images in the Fox sequence

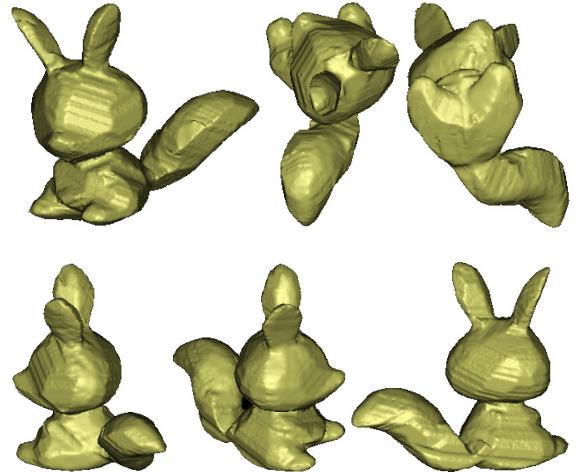


Figure 8. Sample views of the reconstructed Fox model.

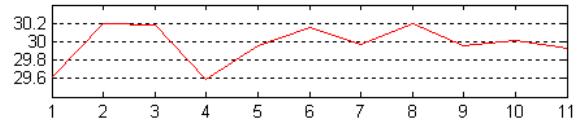


Figure 9. The estimated angles for the Fox sequence.

depicted in Figure 7. In this case, the epipoles are first determined from the method in [8], and then used to compute the homography and the remaining camera parameters. Sample views of the reconstructed 3D model and the estimated angles are shown in Figure 8 and 9, respectively.



Figure 10. Example images in the crystal apple sequence



Figure 11. Sample views of the reconstructed Apple model.

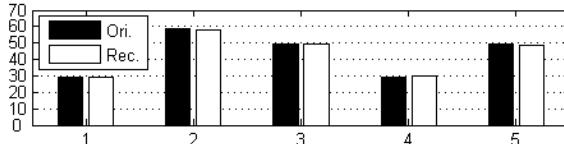


Figure 12. The estimated interval angles between images

The second experiment is on the crystal apple sequence as depicted in Figure 10. Notice that the object is made of glass and it is almost impossible to establish feature correspondences in this case. In this part, six images of size 640x480 pixels with different interval angles are chosen. The estimated results are shown in Figure 11 and 12. Figure 12 shows the ground truth angles in black and the reconstructed values in white. We can see the estimated angles are quite close to the true angles.

Finally, we show the experiment on the horse sequence [11], which contains 14 images of size 640x480 pixels. The sequence was taken under an approximately circular motion and the rectified step makes this sequence become a degenerate case. We chose 3 images, as shown in Figure 13, with their optical axes approximately intersecting at a point, for reconstruction. The comparison of the estimated focal length and the ground truth is listed in Table 3.

In Table 3, ‘GT’ means the ground truth, ‘Ini.’ is the focal length estimated by the proposed method, and ‘Opt.’ is the refined focal length estimate after minimizing the overall epipolar tangency error. The error of the initial estimated focal length is about 5.3% compared to the ground truth. Sample views of the reconstructed 3D model are shown in Figure 14. Because it is reconstructed from 3 views only, the results may seem to be rough. In [11], the camera intrinsic parameters are obtained from off-line calibration. For the proposed method, the camera is auto-calibrated.

### 6.3. Comparison with Other Methods

The proposed method is also compared with the methods proposed in [10] and [12] in the following experiments.



Figure 13. Images chosen from the horse sequence

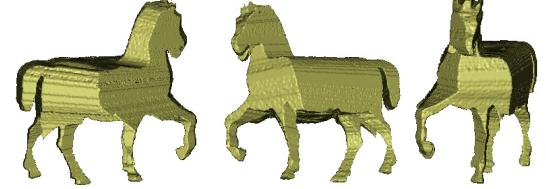


Figure 14. Sample views of the reconstructed horse model.

Table 3. Comparison of the recovered focal length with the ground truth for the degenerate case.

Method	GT	Proposed Ini.	Proposed Opt.
$f$	684.98	723.29	664.85

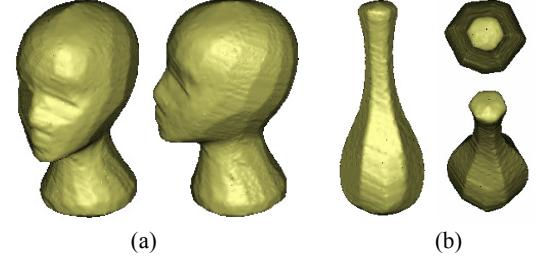


Figure 15. Sample views of the reconstructed (a) Head model and (b) Vase model.

First, the head image sequence in [10] was used for reconstruction. The sample views of the reconstructed model and the estimated angles are shown in Figure 15(a) and 16, respectively. In [10], the focal length was calibrated off-line, while the proposed method auto-calibrated the camera. The RMS error of interval angles in [10] is  $0.2131^\circ$  and it is  $0.1015^\circ$  by using the proposed method.

In addition, we applied the proposed algorithm to four example images of the vase sequence that were directly clipped from [9] and enlarged to the size of 640x480 pixels. The reconstructed model is shown in Figure 15(b) and the comparison of the estimated focal length with [12] and the ground truth is listed in Table 4. As shown in Table 4, our estimated error is about 4% compared to the ground truth. Notice that, we only used 4 clipped and enlarged images to calibrate the camera instead of the original 18 images used in [12].

It is obvious that more correct results could be obtained when more views are used. However, our proposed method can still work well when only very few sparse views are available, while the previous silhouette-based algorithms are not feasible for this kind of sequences.

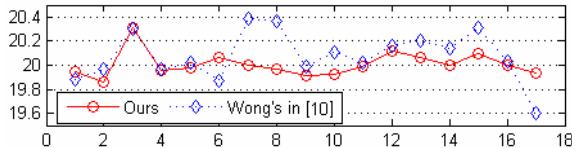


Figure 16. The estimated interval angles between images.

Table 4. Comparison of the recovered focal length with the ground truth and the method by Zhang et al. [12].

Method	GT	Zhang et al.[12]	Proposed Opt.
#	-	18	4
f	2389.8	2370.7	2288.4

## 7. Conclusion

In this paper, we introduced the concept of mirror symmetry property into circular motion and provide a new observation that the epipoles can be located directly from the silhouettes by using the derived homography under circular motion. With this proposed algorithm, the camera can be calibrated from the silhouettes of sparse views, which can be as small as three, under circular motion. This is quite different from the previous silhouette-based methods, which need dense (the interval angle smaller than 20 degrees) and complete (make a turn) image sequences to generate SoR. In addition, we derived a closed-form solution for focal length under the degenerate case when the camera optical axes intersect at one point.

Experimental results on synthetic and real data sets are presented to demonstrate the performance of the proposed algorithm. Although the epipoles and the geometry under circular motion can be well determined for the feature-based methods, the silhouette-based approach has its advantage when the feature correspondences could not be reliably established, which is very common in practice. Previous silhouette-based methods can produce good camera calibration results only with dense views. By using the proposed algorithm, the camera calibration from silhouettes at very sparse views becomes feasible.

## Acknowledgements

This work was supported by National Science Council, Taiwan, under the grant NSC 96-2221-E-007-043.

## References

- [1] G. G. Slabaugh, W. B. Culbertson, T. Malzbender, and R. W. Schafer, A survey of methods for volumetric scene reconstruction from photographs, In Proc. Int. Workshop on Volume Graphics, 81-100, 2001.
- [2] W. Matusik, C. Buehler, R. Raskar, S. J. Gortler and L. McMillan, Image-based visual hulls, In Proc. SIGGRAPH, 369-374, 2000.
- [3] A. W. Fitzgibbon, G. Cross, and A. Zisserman, Automatic 3D model construction for turn-table sequences, In Proc. 3D Structure from Multiple Images of Large-Scale Environments, 155-170, 1998.
- [4] G. Jiang, H.T. Tsui, L. Quan, and A. Zisserman, Single axis geometry by fitting conics, IEEE Trans. Pattern Analysis Machine Intelligence, 25(10):1343-1348, 2002.
- [5] G. Jiang, L. Quan, H.T. Tsui, Circular motion geometry using minimal data, IEEE Trans. Pattern Analysis Machine Intelligence, 26(6):721-731, 2004.
- [6] X. Cao, J. Xiao, H. Foroosh, and M. Shah, Self-calibration from turn-table sequences in presence of zoom and focus, Computer Vision and Image Understanding, 102(3):227-237, 2006.
- [7] H. Zhong and Y. S. Hung, Multi-stage 3D reconstruction under circular motion, Image and Vision Computing, 25(11): 1814-1823, 2007.
- [8] P. R. S. Mendonca, and R. Cipolla, Estimation of epipolar geometry from apparent contours: affine and circular motion cases, In Proc. Computer Vision Pattern Recognition, 1:9-14, 1999.
- [9] P. R. S. Mendonca, K.-Y. K. Wong, and R. Cipolla, Epipolar geometry from profiles under circular motion, IEEE Trans. Pattern Analysis Machine Intelligence, 23(6):604-616, 2001.
- [10] K.-Y. K. Wong, P. R. S. Mendonca, and R. Cipolla, Structure and motion estimation from apparent contours under circular motion, Image and Vision Computing, 20(5):441-448, 2002.
- [11] K.-Y. K. Wong and R. Cipolla, Reconstruction of sculpture from its profiles with unknown camera positions, IEEE Trans. on Image Processing, 13(3):381-389, 2004.
- [12] H. Zhang, G. Zhang, and K.-Y. K. Wong, Auto-calibration and motion recovery from silhouettes for turntable sequences, In Proc. British Machine Vision Conference, 1:79-88, 2005.
- [13] G. Zhang, H. Zhang, and K.-Y. K. Wong, 1D camera geometry and its application to circular motion estimation, In Proc. British Machine Vision Conference, 1:67-76, 2006.
- [14] C. Hernandez, F. Schmitt, and R. Cipolla, Silhouette coherence for camera calibration under circular motion, IEEE Trans. Pattern Analysis Machine Intelligence, 29(2):343-349, 2007.
- [15] P.-H. Huang and S.-H. Lai, Camera calibration from silhouette under incomplete circular motion with a constant interval angle, In Proc. Asian Conference on Computer Vision, 1:106-115, 2007.
- [16] K. Forbes, F. Nicolls, G. de Jager and A. Voigt, Shape-from-silhouette with two mirrors and an uncalibrated camera, In Proc. 9th European Conference on Computer Vision, 2:165-178, 2006.
- [17] P.-H. Huang and S.-H. Lai, Contour-based structure from reflection, In Proc. Computer Vision Pattern Recognition, 1:379-386, 2006.
- [18] S. N. Sinha, M. Pollefeys and L. McMillan, Camera network calibration from dynamic silhouettes, In Proc. Computer Vision and Pattern Recognition, 1:195-202, 2004.
- [19] E. Boyer, On using silhouettes for camera calibration, In Proc. Asian Conference on Computer Vision, 1-10, 2006
- [20] R. Hartley and A. Zisserman, Multiple View Geometry, Cambridge University Press, 2000.