

Camera Calibration from Surfaces of Revolution

Kwan-Yee K. Wong, Paulo R.S. Mendonça, and Roberto Cipolla, *Member, IEEE*

Abstract—This paper addresses the problem of calibrating a pinhole camera from images of a surface of revolution. Camera calibration is the process of determining the intrinsic or internal parameters (i.e., aspect ratio, focal length, and principal point) of a camera, and it is important for both motion estimation and metric reconstruction of 3D models. In this paper, a novel and simple calibration technique is introduced, which is based on exploiting the symmetry of images of surfaces of revolution. Traditional techniques for camera calibration involve taking images of some precisely machined calibration pattern (such as a calibration grid). The use of surfaces of revolution, which are commonly found in daily life (e.g., bowls and vases), makes the process easier as a result of the reduced cost and increased accessibility of the calibration objects. In this paper, it is shown that two images of a surface of revolution will provide enough information for determining the aspect ratio, focal length, and principal point of a camera with fixed intrinsic parameters. The algorithms presented in this paper have been implemented and tested with both synthetic and real data. Experimental results show that the camera calibration method presented here is both practical and accurate.

Index Terms—Camera calibration, surface of revolution, harmonic homology, absolute conic, vanishing point.

1 INTRODUCTION

AN essential step for motion estimation and 3D Euclidean reconstruction, two important tasks in computer vision, is the determination of the intrinsic parameters of cameras. This process, known as *camera calibration*, usually involves taking images of some special pattern with known geometry, extracting the features from the images, and minimizing their reprojection errors. Details of such calibration algorithms can be found in [1], [2], [3], [4], [5, chapter 7] and [6, chapter 3]. These methods do not require direct mechanical measurements on the cameras, and often produce very good results. Nevertheless, they involve the design and use of highly accurate tailor-made calibration patterns, which are often both difficult and expensive to be manufactured.

In this paper, a novel technique for camera calibration is introduced. It relates the idea of calibration from vanishing points [7], [8], [9] to the symmetry properties exhibited in the silhouettes of surfaces of revolution [10], [11], [12], [13], [14]. The method presented here allows the camera to be calibrated from two or more silhouettes of surfaces of revolution (like bowls and vases, etc.), which are commonly found in daily life. The use of such objects has the advantages of easy accessibility and low cost, in contrast to traditional calibration patterns.

This paper is organized as follows: Section 2 gives a literature survey on existing camera calibration techniques. Section 3 defines the camera model used and gives brief reviews on camera calibration from vanishing points and the symmetry properties associated with the silhouettes of

surfaces of revolution. These provide the theoretical background needed for the camera calibration method [15] introduced in Section 4. The algorithms and implementation details are given in Section 5, followed by a discussion of the singular cases and an error analysis in Sections 6 and 7. Section 8 first presents results of experiments conducted on synthetic data, which are used to perform an evaluation on the robustness of the algorithms in the presence of noise. Experiments on real data then show the usefulness of the proposed method. Finally, conclusions are given in Section 9.

2 PREVIOUS WORKS

Classical calibration techniques [16], [17], [18] in photogrammetry involve full-scale nonlinear optimizations with large numbers of parameters. These techniques are able to cope with complex camera models and they produce accurate results, but require a good initialization and are computationally expensive. In [19], Abdel-Aziz and Karara presented the *direct linear transformation* (DLT) technique, which is one of the most commonly used calibration techniques in the field of computer vision. By ignoring lens distortion and treating the coefficients of the 3×4 projection matrix as unknowns, DLT only involves solving a system of linear equations, which can be done by a linear least-squares method. In practice, the linear solution obtained from DLT is usually refined iteratively by minimizing the reprojection errors of the 3D reference points [1], [6]. In [2], [3], Tsai and Lenz introduced the *radial alignment constraint* (RAS) and developed a technique which also accounts for lens distortion.

All the calibration techniques mentioned so far require the knowledge of the 3D coordinates of a certain number of reference points and their corresponding image coordinates. In [7], Caprile and Torre showed that, under the assumption of zero skew and unit aspect ratio, it is possible to calibrate a camera from the vanishing points associated with three mutually orthogonal directions. This idea was further elaborated in [8], [9] to develop practical systems for reconstructing architectural scenes. In contrast to traditional calibration techniques, these methods depend only on the

- K.-Y.K. Wong is with the Department of Computer Science and Information Systems, The University of Hong Kong, Pokfulam Road, Hong Kong. E-mail: kykwong@csis.hku.hk.
- P.R.S. Mendonça is with GE Global Research Center, Schenectady, NY 12301. E-mail: mendonca@crd.ge.com.
- R. Cipolla is with the Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, United Kingdom. E-mail: cipolla@eng.cam.ac.uk.

Manuscript received 30 Apr. 2002; revised 9 Aug. 2002; accepted 16 Aug. 2002.

Recommended for acceptance by L. Quan.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 116448.

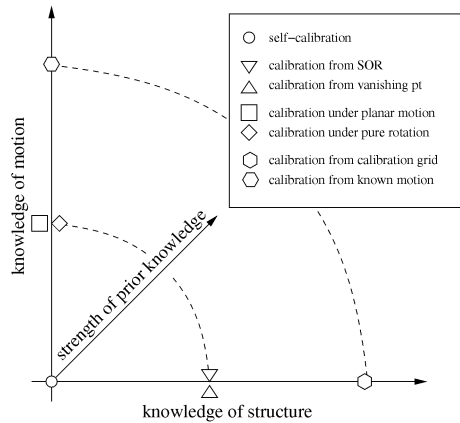


Fig. 1. Different categories of camera calibration techniques.

presence of some special structures, but not on the exact geometry of those structures.

The theory of *self-calibration* was first introduced by Maybank and Faugeras [20], who established the relationship between camera calibration and the epipolar transformation via the *absolute conic* [6]. Implementation of the theory in [20], together with real data experiments, were given by Luong and Faugeras [21] for fixed intrinsic parameters. In [22], Triggs introduced the *absolute quadric* and gave a simpler formulation which can easily incorporate any constraint on the intrinsic parameters. Based on [22], a practical technique for self-calibration of multiple cameras with varying intrinsic parameters was developed by Pollefeys et al. in [23]. Specialized methods of self-calibration have also been derived when the camera motion is restricted to pure rotation [24] or planar motion [25].

The calibration technique introduced in this paper, namely, *calibration from surfaces of revolution*, falls into the same category as calibration from vanishing points (see Fig. 1). Like calibration from vanishing points, which only requires the presence of three mutually orthogonal directions, the technique presented here only requires the calibration target to be a surface of revolution, but the exact geometry of the surface is not important. An important problem not addressed in this paper is how to detect surfaces of revolution in an image, so that they can be used in the calibration algorithm. A practical solution to this has been presented by Zisserman et al. in [13]. The technique developed in that paper is based on the same symmetry properties exploited here, and works well even in the presence of partial occlusion and clustered background.

3 THEORETICAL BACKGROUND

3.1 Camera Model

In this paper, a camera is modeled as a pinhole (perspective) camera and the imaging process can be expressed as

$$\alpha \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{P} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (1)$$

where (X, Y, Z) are the coordinates of a 3D point \mathbf{X} , (u, v) are the image coordinates of the projection of \mathbf{X} and α is an arbitrary scale factor. \mathbf{P} is a 3×4 matrix known as the

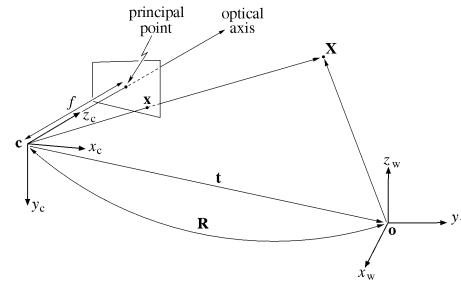


Fig. 2. The extrinsic parameters of a camera represent the rigid body transformation between the world coordinate system (centered at o) and the camera coordinate system (centered at c), and the intrinsic parameters represent the camera internal parameters like focal length, aspect ratio, skew, and principal point.

projection matrix [26], which models the pinhole camera. The projection matrix \mathbf{P} is not a general 3×4 matrix but has a special structure given by [6]

$$\mathbf{P} = \mathbf{K}[\mathbf{R} \ \mathbf{t}], \quad (2)$$

where \mathbf{K} is a 3×3 upper triangular matrix known as the *camera calibration matrix*, \mathbf{R} is a 3×3 rotation matrix and \mathbf{t} is a 3×1 translation vector. \mathbf{R} and \mathbf{t} are called the *extrinsic parameters* [6] of the camera and they represent the rigid body transformation between the camera and the scene (see Fig. 2). The camera calibration matrix \mathbf{K} has the form [6]

$$\mathbf{K} = \begin{bmatrix} f_u & \varsigma & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} af & \varsigma & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3)$$

where f is the *focal length*, $a = f_u/f_v$ is the *aspect ratio* and ς is the *skew* which depends on the angle between the image axes. (u_0, v_0) is called the *principal point* and it is the point at which the optical axis (z_c -axis) intersects the image plane (see Fig. 2). The focal length, aspect ratio, skew, and principal point are referred to as the *intrinsic parameters* [6] of the camera, and *camera calibration* is the process of estimating these parameters. If the image axes are orthogonal to each other, which is often the case, ς will be equal to zero. In practice, the aspect ratio and skew of a camera are often assumed to be one and zero, respectively, to reduce the dimension of the search space in camera calibration. This generally speeds up the calibration process and makes the results more stable. Such an initial estimate of the intrinsic parameters can be further refined later by relaxing the unit aspect ratio and zero skew constraints. A camera is said to be calibrated if its intrinsic parameters are known. If both the intrinsic and extrinsic parameters of a camera are known, then the camera is said to be fully calibrated.

3.2 Calibration from Vanishing Points

In [7], Caprile and Torre showed that under the assumption of zero skew and unit aspect ratio, the principal point of a camera will coincide with the orthocenter of a triangle with vertices given at three vanishing points from three mutually orthogonal directions. Such properties of the vanishing points, together with the symmetry properties associated with the silhouettes of surfaces of revolution, will be used later in Section 4 to derive a simple technique for camera calibration. A simple derivation of Caprile and Torre's result is given below.

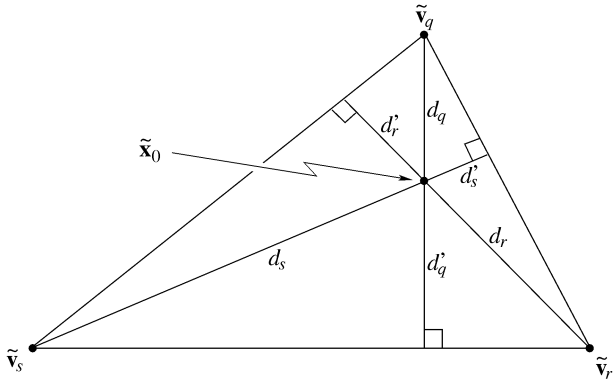


Fig. 3. The principal point $\tilde{\mathbf{x}}_0$ of the camera coincides with the orthocenter of the triangle with vertices given at the vanishing points $\tilde{\mathbf{v}}_q$, $\tilde{\mathbf{v}}_r$, and $\tilde{\mathbf{v}}_s$ associated with three mutually orthogonal directions, and the focal length of the camera is given by $f = \sqrt{d_q d'_q} = \sqrt{d_r d'_r} = \sqrt{d_s d'_s}$.

Consider a pinhole camera with focal length f , unit aspect ratio, zero skew, and principal point $\tilde{\mathbf{x}}_0$. The vector from the camera center to any point $\tilde{\mathbf{x}}$ on the image plane, in camera coordinate system, is given by $[(\tilde{\mathbf{x}} - \tilde{\mathbf{x}}_0)^T \ f]^T$. Let $\tilde{\mathbf{v}}_q$, $\tilde{\mathbf{v}}_r$, and $\tilde{\mathbf{v}}_s$ be three vanishing points associated with three mutually orthogonal directions \mathbf{N}_q , \mathbf{N}_r , and \mathbf{N}_s , respectively. The three vectors from the camera center to $\tilde{\mathbf{v}}_q$, $\tilde{\mathbf{v}}_r$, and $\tilde{\mathbf{v}}_s$ will be mutually orthogonal to each other and, hence,

$$(\tilde{\mathbf{v}}_q - \tilde{\mathbf{x}}_0) \cdot (\tilde{\mathbf{v}}_r - \tilde{\mathbf{x}}_0) + f^2 = 0, \quad (4)$$

$$(\tilde{\mathbf{v}}_r - \tilde{\mathbf{x}}_0) \cdot (\tilde{\mathbf{v}}_s - \tilde{\mathbf{x}}_0) + f^2 = 0, \quad (5)$$

$$(\tilde{\mathbf{v}}_s - \tilde{\mathbf{x}}_0) \cdot (\tilde{\mathbf{v}}_q - \tilde{\mathbf{x}}_0) + f^2 = 0. \quad (6)$$

Subtracting (6) from (4) gives

$$(\tilde{\mathbf{v}}_q - \tilde{\mathbf{x}}_0) \cdot (\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_s) = 0. \quad (7)$$

Equation (7) shows that $\tilde{\mathbf{x}}_0$ lies on a line passing through $\tilde{\mathbf{v}}_q$ and orthogonal to the line joining $\tilde{\mathbf{v}}_r$ and $\tilde{\mathbf{v}}_s$. Similarly, subtracting (4) from (5) and (5) from (6) gives

$$(\tilde{\mathbf{v}}_r - \tilde{\mathbf{x}}_0) \cdot (\tilde{\mathbf{v}}_s - \tilde{\mathbf{v}}_q) = 0, \quad (8)$$

$$(\tilde{\mathbf{v}}_s - \tilde{\mathbf{x}}_0) \cdot (\tilde{\mathbf{v}}_q - \tilde{\mathbf{v}}_r) = 0. \quad (9)$$

Equations (7), (8), and (9) imply that the principal point $\tilde{\mathbf{x}}_0$ coincides with the orthocenter of the triangle with vertices $\tilde{\mathbf{v}}_q$, $\tilde{\mathbf{v}}_r$, and $\tilde{\mathbf{v}}_s$. Besides, (4), (5), and (6) show that the focal length f is equal to the square root of the product of the distances from the orthocenter to any vertex and to the opposite side (see Fig. 3). As a result, under the assumption of zero skew and unit aspect ratio, it is possible to estimate the principal point and the focal length of a camera using vanishing points from three mutually orthogonal directions. A similar derivation was also presented by Cipolla et al. in [8].

3.3 Symmetry in Surfaces of Revolution

As will be shown in the next paragraph, the silhouette of a surface of revolution, viewed under a pinhole camera, will be invariant to a harmonic homology [13]. Such properties of the silhouette can be exploited to calibrate the intrinsic parameters of a camera, as will be shown in Section 4. A simple proof of such symmetry properties is given below which also shows that the axis of the associated harmonic homology is given by the image of the revolution axis, and that the center of the homology is given by the vanishing

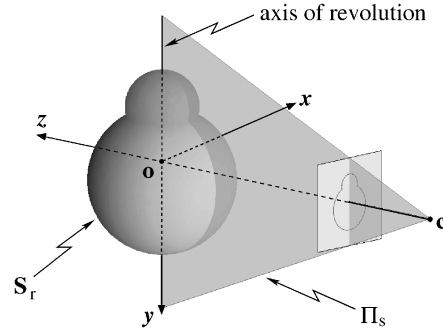


Fig. 4. A surface of revolution S_r , whose axis of revolution coincides with the y -axis, being viewed by a pinhole camera $\hat{\mathbf{P}} = [\mathbb{I}_3 \ -\mathbf{c}]$ centered at $\mathbf{c} = [0 \ 0 \ -d_z]^T$.

point corresponding to the normal direction of the plane containing the axis of revolution and the camera center.

Consider a surface of revolution S_r , whose axis of revolution coincides with the y -axis, being viewed by a pinhole camera $\hat{\mathbf{P}} = [\mathbb{I}_3 \ -\mathbf{c}]$ centered at $\mathbf{c} = [0 \ 0 \ -d_z]^T$, with $d_z > 0$ (see Fig. 4). By symmetry considerations, it is easy to see that the silhouette $\hat{\rho}$ of S_r formed on the image plane will be bilaterally symmetric about the image of the revolution axis $\hat{\mathbf{l}}_s = [1 \ 0 \ 0]^T$. The lines of symmetry (i.e., lines joining symmetric points in $\hat{\rho}$) will be parallel to the normal $\mathbf{N}_x = [1 \ 0 \ 0]^T$ of the plane Π_s that contains the axis of revolution and the camera center, and the vanishing point associated with \mathbf{N}_x is given by $\hat{\mathbf{v}}_x = [1 \ 0 \ 0]^T$. The bilateral symmetry exhibited in $\hat{\rho}$ can be described by the transformation [27], [28]

$$\begin{aligned} \mathbf{T} &= \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \mathbb{I}_3 - 2 \frac{\hat{\mathbf{v}}_x \hat{\mathbf{l}}_s^T}{\hat{\mathbf{v}}_x^T \hat{\mathbf{l}}_s}. \end{aligned} \quad (10)$$

Note that the transformation \mathbf{T} is a *harmonic homology* (see the Appendix and also [29], [30] for details) with axis $\hat{\mathbf{l}}_s$ and center $\hat{\mathbf{v}}_x$, which maps every point in $\hat{\rho}$ to its symmetric counterpart in $\hat{\rho}$. The silhouette $\hat{\rho}$ is thus said to be invariant to the harmonic homology \mathbf{T} (i.e., $\hat{\rho} = \mathbf{T}\hat{\rho}$).

Now, consider an arbitrary pinhole camera \mathbf{P} by introducing the intrinsic parameters represented by the camera calibration matrix \mathbf{K} to $\hat{\mathbf{P}}$ and by applying the rotation \mathbf{R} to $\hat{\mathbf{P}}$ about its optical center. Hence, $\mathbf{P} = \mathbf{KR}[\mathbb{I}_3 \ -\mathbf{c}]$ or $\mathbf{P} = \mathbf{H}\hat{\mathbf{P}}$, where $\mathbf{H} = \mathbf{KR}$. Let \mathbf{x} be the projection of a 3D point \mathbf{X} in \mathbf{P} , hence

$$\begin{aligned} \mathbf{x} &= \mathbf{P}\mathbf{X} \\ &= \mathbf{H}\hat{\mathbf{P}}\mathbf{X}, \\ &= \mathbf{H}\hat{\mathbf{x}} \end{aligned} \quad (11)$$

where $\hat{\mathbf{x}} = \hat{\mathbf{P}}\mathbf{X}$. Equation (11) implies that the 3×3 matrix \mathbf{H} represents a planar homography which transforms the image formed by $\hat{\mathbf{P}}$ into the image formed by \mathbf{P} . Similarly, \mathbf{H}^{-1} transforms the image formed by \mathbf{P} into the image formed by $\hat{\mathbf{P}}$. The silhouette ρ of S_r , formed on the image plane of \mathbf{P} , can thus be obtained by applying the planar homography \mathbf{H} to $\hat{\rho}$ (i.e., $\rho = \mathbf{H}\hat{\rho}$). Let $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}'$ be a pair of

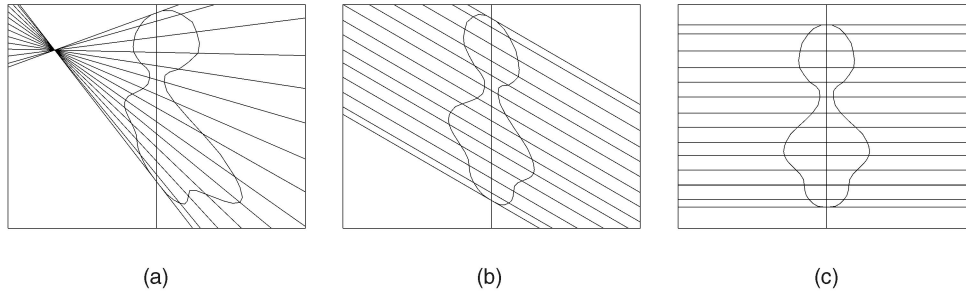


Fig. 5. (a) Silhouette of a surface of revolution under general viewing conditions. The symmetry of the silhouette is described by a harmonic homology defined by the image of the revolution axis and a vanishing point. (b) When the camera is pointing directly towards the axis of revolution, the transformation reduces to a skew symmetry, which is a particular case of the harmonic homology where the vanishing point is at infinity. (c) If the camera also has zero skew and unit aspect ratio, the transformation becomes a bilateral symmetry, in which the vanishing point is at infinity and has a direction orthogonal to the image of the revolution axis.

symmetric points in $\hat{\rho}$, and $\mathbf{x} = \mathbf{H}\hat{\mathbf{x}}$ and $\mathbf{x}' = \mathbf{H}\hat{\mathbf{x}}'$ be their correspondences in ρ . The symmetry between $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}'$ is given by

$$\hat{\mathbf{x}}' = \mathbf{T}\hat{\mathbf{x}}. \quad (12)$$

Substituting $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}'$ in (12) by $\mathbf{H}^{-1}\mathbf{x}$ and $\mathbf{H}^{-1}\mathbf{x}'$, respectively, gives [27], [28]

$$\begin{aligned} (\mathbf{H}^{-1}\mathbf{x}') &= \mathbf{T}(\mathbf{H}^{-1}\mathbf{x}) \\ \mathbf{x}' &= \mathbf{H}\mathbf{T}\mathbf{H}^{-1}\mathbf{x} \\ &= \mathbf{H} \left(\mathbb{I}_3 - 2 \frac{\hat{\mathbf{v}}_x \hat{\mathbf{l}}_s^T}{\hat{\mathbf{v}}_x^T \hat{\mathbf{l}}_s} \right) \mathbf{H}^{-1}\mathbf{x} \\ &= \left(\mathbb{I}_3 - 2 \frac{\mathbf{v}_x \mathbf{l}_s^T}{\mathbf{v}_x^T \mathbf{l}_s} \right) \mathbf{x}, \end{aligned} \quad (13)$$

where $\mathbf{v}_x = \mathbf{H}\hat{\mathbf{v}}_x$, and $\mathbf{l}_s = \mathbf{H}^{-T}\hat{\mathbf{l}}_s$. Note that \mathbf{v}_x is the vanishing point corresponding to the normal direction \mathbf{N}_x in \mathbf{P} , and \mathbf{l}_s is the image of the revolution axis of \mathbf{S}_r in \mathbf{P} . Let $\mathbf{W} = \mathbf{H}\mathbf{T}\mathbf{H}^{-1}$ be the harmonic homology with axis \mathbf{l}_s and center \mathbf{v}_x . Equation (13) shows that \mathbf{W} will map each point in ρ to its symmetric counterpart in ρ and, hence, ρ is invariant to the harmonic homology \mathbf{W} (i.e., $\rho = \mathbf{W}\rho$).

In general, the harmonic homology \mathbf{W} has four degrees of freedom. When the camera is pointing directly towards the axis of revolution, the harmonic homology will reduce to a *skew symmetry* [31], [32], [33], [34], where the vanishing point \mathbf{v}_x is at infinity. The skew symmetry can be described by the transformation

$$\mathbf{S} = \frac{1}{\cos(\phi - \theta)} \begin{bmatrix} -\cos(\phi + \theta) & -2\cos\phi \sin\theta & 2d_1 \cos\phi \\ -2\sin\phi \cos\theta & \cos(\phi + \theta) & 2d_1 \sin\phi \\ 0 & 0 & \cos(\phi - \theta) \end{bmatrix}, \quad (14)$$

where $d_1 = u_0 \cos\theta + v_0 \sin\theta$. The image of the revolution axis and the vanishing point are given by $\mathbf{l}_s = [\cos\theta \ \sin\theta \ -d_1]^T$ and $\mathbf{v}_x = [\cos\phi \ \sin\phi \ 0]^T$, respectively, and \mathbf{S} has only three degrees of freedom. If the camera also has zero skew and unit aspect ratio, the transformation will then become a *bilateral symmetry*, given by

$$\mathbf{B} = \begin{bmatrix} -\cos 2\theta & -\sin 2\theta & 2d_1 \cos\theta \\ -\sin 2\theta & \cos 2\theta & 2d_1 \sin\theta \\ 0 & 0 & 1 \end{bmatrix}. \quad (15)$$

While \mathbf{l}_s will have the same form as in the case of skew symmetry, the vanishing point will now be at infinity and will have a direction orthogonal to \mathbf{l}_s . As a result, \mathbf{B} has only two degrees of freedom. These three different cases of symmetry are illustrated in Fig. 5.

4 CAMERA CALIBRATION

4.1 Vanishing Points and the Harmonic Homology

Consider a surface of revolution \mathbf{S}_r viewed by a pinhole camera $\mathbf{P} = \mathbf{K}[\mathbf{R} \ \mathbf{t}]$. Let ρ be the silhouette of \mathbf{S}_r , \mathbf{l}_s be the image of the revolution axis of \mathbf{S}_r , and \mathbf{v}_x be the vanishing point corresponding to the normal direction \mathbf{N}_x of the plane Π_s that contains the revolution axis of \mathbf{S}_r and the camera center of \mathbf{P} . The silhouette ρ is then invariant to the harmonic homology \mathbf{W} with axis \mathbf{l}_s and center \mathbf{v}_x (see Section 3.3).

Now, consider any two vectors \mathbf{N}_y and \mathbf{N}_z parallel to the plane Π_s and orthogonal to each other, which together with \mathbf{N}_x form a set of three mutually orthogonal directions (see Fig. 6). Under the assumption of zero skew and unit aspect ratio, the vanishing points associated with these three directions can be used to determine the principal point and the focal length of \mathbf{P} , as shown in Section 3.2. By construction, the vanishing points \mathbf{v}_y and \mathbf{v}_z , corresponding to the directions \mathbf{N}_y and \mathbf{N}_z , respectively, will lie on the image of the revolution axis \mathbf{l}_s . Given the harmonic homology \mathbf{W} associated with ρ , with an axis given by the image of the

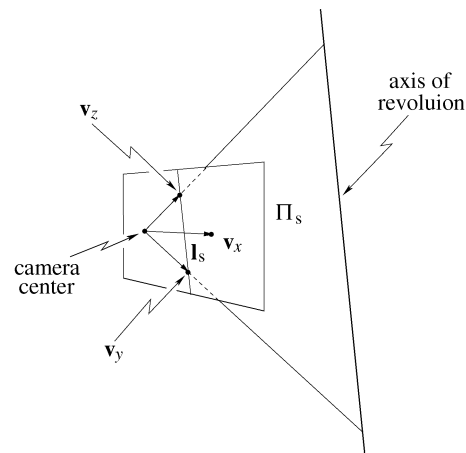


Fig. 6. Three mutually orthogonal directions associated with a surface of revolution.

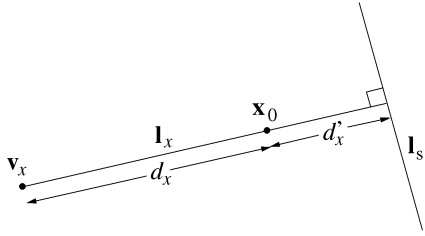


Fig. 7. The vanishing point \mathbf{v}_x and the image of the revolution axis \mathbf{l}_s define a line \mathbf{l}_x on which the principal point \mathbf{x}_0 must lie, and the focal length f is equal to $\sqrt{d_x d'_x}$.

revolution axis \mathbf{l}_s and a center given by the vanishing point \mathbf{v}_x , the principal point \mathbf{x}_0 of \mathbf{P} will therefore lie on a line \mathbf{l}_x passing through \mathbf{v}_x and orthogonal to \mathbf{l}_s , and the focal length f will be equal to the square root of the product of the distances from the principal point \mathbf{x}_0 to \mathbf{v}_x and to \mathbf{l}_s , respectively, (see Fig. 7). As a result, given two or more silhouettes of surfaces of revolution, the principal point can be estimated as the intersection of the lines \mathbf{l}_x and the focal length follows.

4.2 Pole-Polar Relationship and the Absolute Conic

Following the notations in the previous section, consider the equation of the plane Π_s which can be deduced from \mathbf{P} and the image of the revolution axis \mathbf{l}_s , and is given by

$$\Pi_s = \mathbf{P}^T \mathbf{l}_s. \quad (16)$$

By definition, \mathbf{v}_x is the vanishing point corresponding to the normal direction \mathbf{N}_x of the plane Π_s and, hence,

$$\mathbf{v}_x = \mathbf{P} \mathbf{N}_x. \quad (17)$$

Now, let

$$\Omega = \begin{bmatrix} \Pi_3 & \mathbf{0}_3 \\ \mathbf{0}_3^T & 0 \end{bmatrix}$$

be the *absolute dual quadric* [22]. Observe that Π_s can be expressed as $\Pi_s = [\mathbf{n}_x^T \ -d]^T$, where \mathbf{n}_x is the normal direction of Π_s in Cartesian (nonhomogeneous) coordinates. Therefore, $\Omega \Pi_s = [\mathbf{n}_x^T \ 0]^T = \mathbf{N}_x$, which allows (17) to be rewritten as

$$\begin{aligned} \mathbf{v}_x &= \mathbf{P} \Omega \Pi_s \\ &= \mathbf{P} \Omega \mathbf{P}^T \mathbf{l}_s \\ &= \mathbf{K} \mathbf{K}^T \mathbf{l}_s \\ \mathbf{K}^{-T} \mathbf{K}^{-1} \mathbf{v}_x &= \mathbf{l}_s \\ \omega \mathbf{v}_x &= \mathbf{l}_s, \end{aligned} \quad (18)$$

where $\omega = \mathbf{K}^{-T} \mathbf{K}^{-1}$ is the projection of the absolute quadric in \mathbf{P} , known as the *image of the absolute conic*. Equation (18) gives the pole-polar relationship, with respect to the image of the absolute conic, between the vanishing point \mathbf{v}_x of the normal direction of the plane Π_s and the vanishing line \mathbf{l}_s of Π_s [35]. By assuming the skew of \mathbf{P} to be zero, expanding (18) gives

$$\begin{bmatrix} \frac{1}{f_u^2} & 0 & -\frac{u_0}{f_u^2} \\ 0 & \frac{1}{f_v^2} & -\frac{v_0}{f_v^2} \\ -\frac{u_0}{f_u^2} & -\frac{v_0}{f_v^2} & \left(\frac{u_0}{f_u}\right)^2 + \left(\frac{v_0}{f_v}\right)^2 + 1 \end{bmatrix} \mathbf{v}_x = \mathbf{l}_s, \quad (19)$$

where f_u, f_v , and (u_0, v_0) are the intrinsic parameters of \mathbf{P} . It follows that the harmonic homology associated with the silhouette of a surface of revolution will provide two constraints on the four intrinsic parameters of a camera. As a result, under the assumption of fixed intrinsic parameters and zero skew, it is possible to calibrate a camera from two or more silhouettes of surfaces of revolution. Further, under the assumption of unit aspect ratio (i.e., $f_u = f_v$), it can be derived from (19) that the focal length f is equal to the square root of the product of the distances from the principal point (u_0, v_0) to the vanishing point \mathbf{v}_x and to the image of the revolution axis \mathbf{l}_s . These results agree with the analysis of the vanishing points in the previous section.

5 ALGORITHMS AND IMPLEMENTATIONS

5.1 Estimation of the Harmonic Homology

The silhouette ρ of a surface of revolution can be extracted from the image by applying a Canny edge detector [36] (see Fig. 8). The harmonic homology \mathbf{W} that maps each side of the silhouette ρ to its symmetric counterpart is then estimated by minimizing the geometric distances between the original silhouette ρ and its transformed version $\rho' = \mathbf{W} \rho$. This can be done by sampling N evenly spaced points \mathbf{x}_i along the silhouette ρ and optimizing the cost function

$$\text{Cost}_{\mathbf{w}}(\mathbf{v}_x, \mathbf{l}_s) = \sum_{i=1}^N \text{dist}(\mathbf{W}(\mathbf{v}_x, \mathbf{l}_s) \mathbf{x}_i, \rho)^2, \quad (20)$$

where $\text{dist}(\mathbf{W}(\mathbf{v}_x, \mathbf{l}_s) \mathbf{x}_i, \rho)$ is the orthogonal distance from the transformed sample point $\mathbf{x}'_i = \mathbf{W}(\mathbf{v}_x, \mathbf{l}_s) \mathbf{x}_i$ to the original silhouette ρ .

The success of most nonlinear optimization problems requires a good initialization so as to avoid convergence to local minima. This is achieved here by using bitangents of the silhouette [10]. Two points in the silhouette ρ near a

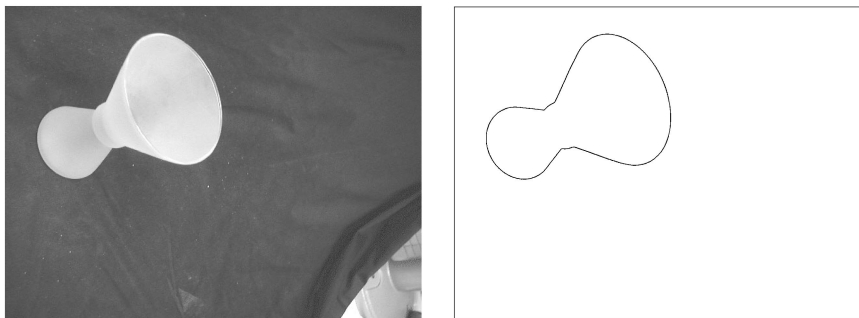


Fig. 8. The silhouette of a surface of revolution (candle holder) extracted by applying a Canny edge detector.

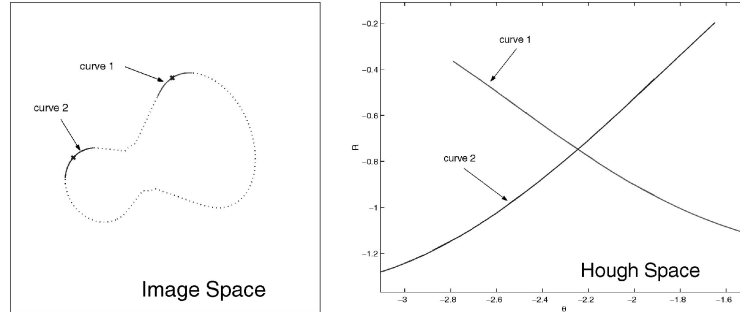


Fig. 9. Two points in the silhouette ρ near a bitangent are selected manually and the neighborhood of each point is transformed into a curve in the Hough space. The bitangent and the bitangent points can then be located by finding the intersection of the two transformed curves in the Hough space.

bitangent are selected manually and the neighborhood of each point is transformed into a curve in the Hough space (see Fig. 9). The bitangent and the bitangent points can then be located by finding the intersection of the two transformed curves in the Hough space (see [37] for details).

Consider two corresponding bitangents l_b and l'_b on the two sides of ρ , with bitangent points x_1, x_2 and x'_1, x'_2 , respectively, (see Fig. 10). Let l_d be the line joining x_1 and x'_2 , and l'_d be the line joining x'_1 and x_2 . The intersection of l_b with l'_b and the intersection of l_d with l'_d define a line which will provide an estimate for the image of the revolution axis l_s . Let l_c be the line joining x_1 and x'_1 , and l'_c be the line joining x_2 and x'_2 . The intersection of l_c with l'_c will provide an estimate for the vanishing point v_x . The initialization of l_s and v_x from bitangents often provides an excellent initial guess for the optimization problem. This is generally good enough to avoid any local minimum and allows convergence to the global minimum in a small number of iterations. Note that bitangents are used here only to provide an initial estimate for l_s and v_x , which will be further refined by optimizing the cost function given in (20). As a result, error in the estimation of the bitangents will not

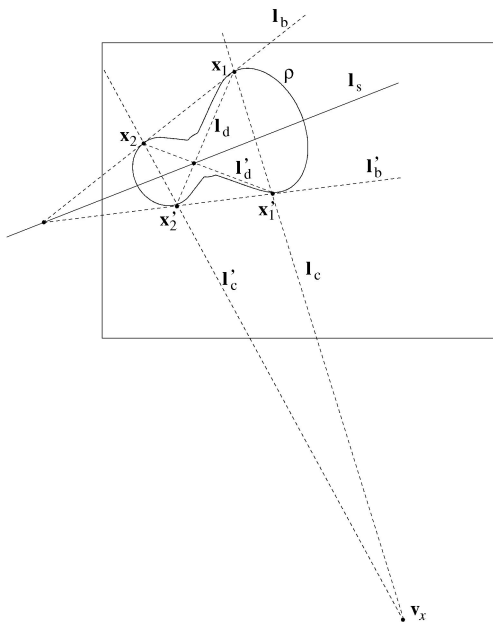


Fig. 10. Initialization of the optimization parameters l_s and v_x from the bitangents and lines formed from the bitangent points.

directly affect the accuracy of the final estimation of the intrinsic parameters. Alternatively, in the absence of any bitangents, l_s can be initialized manually by observing the symmetry in the silhouette, and v_x can be initialized to be a point at infinity having a direction orthogonal to l_s .

The above approach of estimating the harmonic homology is similar to the one presented in [13]. However, the initialization using bitangent points allows for optimizing the full projective model, represented by the harmonic homology, in contrast to the affine approximation used in [13], which corresponds to a skew symmetry transformation. An alternative method, which computes the harmonic homology implicitly, was developed in [38]. The method presented here is much simpler than the one shown in [38], which relies on an unnecessarily complicated prediction, verification and selection scheme. An important point to note is that both the results in [13] and [38] demonstrate that the harmonic homology can be successfully computed in the presence of partial occlusion and clustered background.

5.2 Estimation of the Intrinsic Parameters

5.2.1 Method 1

Under the assumption of zero skew and unit aspect ratio, the line l_x passing through the principal point (u_0, v_0) and the vanishing point v_x will be orthogonal to the image of the revolution axis l_s (see Section 4). Let $v_x = [v_1 \ v_2 \ v_3]^T$ and $l_s = [l_1 \ l_2 \ l_3]^T$. The line l_x can be expressed in terms of v_x and l_s , and is given by

$$l_x = \frac{1}{\sqrt{(l_2 v_3)^2 + (l_1 v_3)^2}} \begin{bmatrix} l_2 v_3 \\ -l_1 v_3 \\ l_1 v_2 - l_2 v_1 \end{bmatrix}. \quad (21)$$

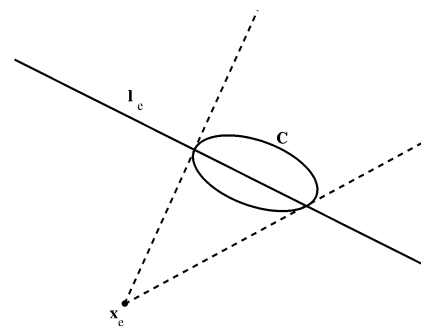


Fig. 11. A conic C will be invariant to any harmonic homology with a center given by any point x_e outside the conic, and an axis given by $l_e = Cx_e$.

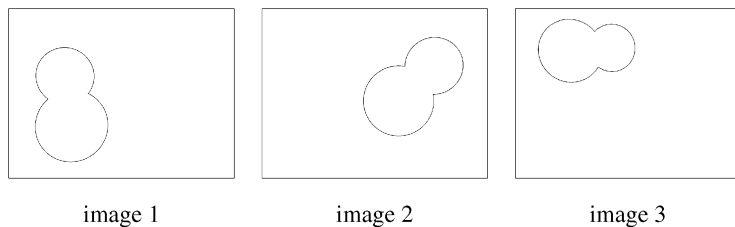


Fig. 15. Silhouettes of the surface of revolution in the three images taken by the synthetic camera with $f = 700$.

6 SINGULAR CASES

The algorithms presented in the previous section have two main steps: the computation of the harmonic homologies, and the subsequent computation of the camera intrinsic parameters using the just computed homologies. Therefore, the only two situations where the algorithms fail are when the homologies cannot be computed, and when the homologies do not provide enough information for the calibration of the camera. The following sections analyze the occurrence of these degeneracies.

6.1 Conic Silhouette

If the silhouette ρ of a surface of revolution is a conic, there will be an infinite number of harmonic homologies to which the silhouette ρ will be invariant. Such a situation results in a singular case for camera calibration from surfaces of revolution.

Consider a conic represented by a 3×3 symmetric matrix \mathbf{C} , such that every point \mathbf{x} on the conic satisfies

$$\mathbf{x}^T \mathbf{C} \mathbf{x} = 0. \quad (30)$$

Given a point \mathbf{x}_e outside the conic \mathbf{C} , two tangents can be drawn from \mathbf{x}_e to \mathbf{C} (see Fig. 11) and the line \mathbf{l}_e passing through the two tangent points is given by

$$\mathbf{l}_e = \mathbf{C} \mathbf{x}_e. \quad (31)$$

Let \mathbf{W}_e be a harmonic homology with axis \mathbf{l}_e and center \mathbf{x}_e , i.e.,

$$\mathbf{W}_e = \mathbb{I}_3 - 2 \frac{\mathbf{x}_e \mathbf{l}_e^T}{\mathbf{x}_e^T \mathbf{l}_e}. \quad (32)$$

Substituting (31) into (32) gives

$$\mathbf{W}_e = \mathbb{I}_3 - 2 \frac{\mathbf{x}_e \mathbf{x}_e^T \mathbf{C}^T}{\mathbf{x}_e^T \mathbf{C} \mathbf{x}_e}. \quad (33)$$

Let \mathbf{x} be a point on \mathbf{C} and $\mathbf{x}' = \mathbf{W}_e \mathbf{x}$, and consider the equation

$$\begin{aligned} \mathbf{x}'^T \mathbf{C} \mathbf{x}' &= (\mathbf{W}_e \mathbf{x})^T \mathbf{C} (\mathbf{W}_e \mathbf{x}) \\ &= \mathbf{x}^T (\mathbf{W}_e^T \mathbf{C} \mathbf{W}_e) \mathbf{x}. \end{aligned} \quad (34)$$

Substituting (33) into (34) gives

$$\begin{aligned} \mathbf{x}'^T \mathbf{C} \mathbf{x}' &= \mathbf{x}^T \left[\left(\mathbb{I}_3 - 2 \frac{\mathbf{x}_e \mathbf{x}_e^T \mathbf{C}^T}{\mathbf{x}_e^T \mathbf{C} \mathbf{x}_e} \right)^T \mathbf{C} \left(\mathbb{I}_3 - 2 \frac{\mathbf{x}_e \mathbf{x}_e^T \mathbf{C}^T}{\mathbf{x}_e^T \mathbf{C} \mathbf{x}_e} \right) \right] \mathbf{x} \\ &= \mathbf{x}^T \left[\left(\mathbb{I}_3 - 2 \frac{\mathbf{C} \mathbf{x}_e \mathbf{x}_e^T}{\mathbf{x}_e^T \mathbf{C} \mathbf{x}_e} \right) \left(\mathbf{C} - 2 \frac{\mathbf{C} \mathbf{x}_e \mathbf{x}_e^T \mathbf{C}^T}{\mathbf{x}_e^T \mathbf{C} \mathbf{x}_e} \right) \right] \mathbf{x} \\ &= \mathbf{x}^T \mathbf{C} \mathbf{x} \\ &= 0. \end{aligned} \quad (35)$$

Equation (35) implies that any point \mathbf{x}_e outside the conic \mathbf{C} and the corresponding line $\mathbf{l}_e = \mathbf{C} \mathbf{x}_e$ will define a harmonic homology \mathbf{W}_e to which the conic \mathbf{C} will be invariant. As a result, if the silhouette of the surface of revolution is a conic, there will not be a unique solution to the optimization problem of the harmonic homology \mathbf{W} associated with the silhouette and, hence, it provides no information on the intrinsic parameters of the camera.

Assume now that the silhouette can be represented as a homogeneous algebraic curve g of degree d [40]. As a result of Bezout's theorem, the silhouette will have $k = 3d(d-2)$ inflection points [41]. If a curve is invariant to a harmonic homology \mathbf{W} , its inflection points, which are projective invariants, will be mapped to each other by \mathbf{W} . If $d \geq 3$, there

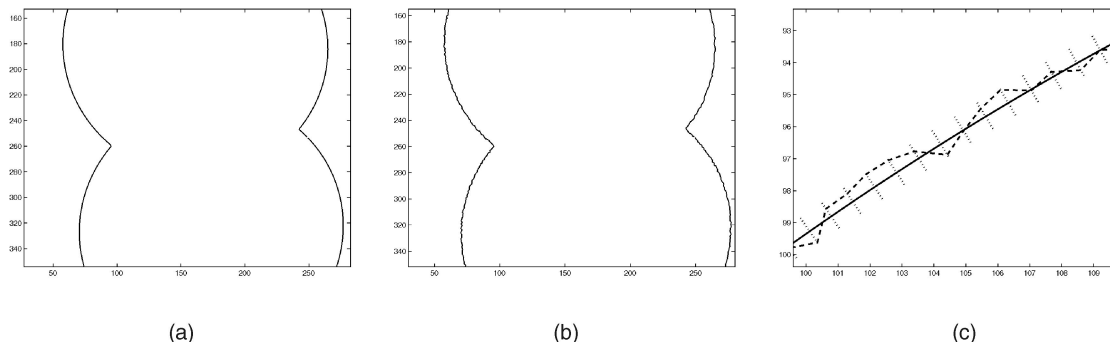


Fig. 16. (a) The original silhouette. (b) The resultant silhouette after uniform random noise of maximum 0.5 pixels being added. (c) The noise-free and noisy silhouettes are represented by solid and dash lines, respectively, and the dotted lines indicate the bounds for noise along the normal direction of each point.

TABLE 1
Experimental Results of Calibration from Silhouettes under Different Focal Lengths and Noise Levels

		Synthetic Camera 1 ($f = 700$)				Synthetic Camera 2 ($f = 1400$)			
		Percentage Errors (%)				Percentage Errors (%)			
noise lv	method	f_u	f_v	u_0	v_0	f_u	f_v	u_0	v_0
0.5	I	1.1921	1.1921	0.6325	0.3354	2.2382	2.2382	0.7371	0.4008
0.5	IIa	1.1516	1.0945	0.6023	0.7591	2.0834	2.0280	0.6254	0.8150
0.5	IIb	1.1254	1.1254	0.5687	0.7462	2.0541	2.0541	0.6108	0.8082
0.7	I	1.7181	1.7181	0.8986	0.4725	3.1868	3.1868	1.1699	0.5610
0.7	IIa	1.7111	1.6250	0.8417	1.0659	2.7825	2.7070	1.0280	1.2551
0.7	IIb	1.6711	1.6711	0.7937	1.0478	2.7423	2.7423	1.0052	1.2454
1.0	I	2.4277	2.4277	1.3113	0.7239	4.3078	4.3078	1.8183	0.9197
1.0	IIa	2.3610	2.2334	1.2908	1.6372	3.6626	3.5731	1.2814	1.5222
1.0	IIb	2.3007	2.3007	1.2184	1.6064	3.6161	3.6161	1.2513	1.5134
1.2	I	3.0194	3.0194	1.7414	0.8761	5.5788	5.5788	1.7734	1.0189
1.2	IIa	2.5415	2.4044	1.5292	2.1493	4.6212	4.5503	1.4714	1.7224
1.2	IIb	2.4749	2.4749	1.4469	2.1164	4.5831	4.5831	1.4401	1.7127
1.5	I	4.5876	4.5876	2.7362	1.3309	6.2149	6.2149	1.9063	1.5326
1.5	IIa	4.0019	3.8079	2.0678	3.1898	5.9250	5.8213	1.7284	2.1655
1.5	IIb	3.9031	3.9031	1.9597	3.1490	5.8700	5.8700	1.6910	2.1504
1.7	I	5.6534	5.6534	3.1541	1.7898	8.0902	8.0902	2.9027	1.4357
1.7	IIa	4.3192	4.1158	2.1550	3.7542	6.4024	6.2320	1.8617	2.3701
1.7	IIb	4.2144	4.2144	2.0343	3.7192	6.3107	6.3107	1.8216	2.3488
2.0	I	6.7864	6.7864	4.3649	3.4956	8.8957	8.8957	2.6345	1.9192
2.0	IIa	5.8438	5.6100	3.1055	5.3920	7.2582	7.1219	2.1700	2.3457
2.0	IIb	5.7052	5.7052	2.9625	5.3566	7.1867	7.1867	2.1304	2.3292
method IIa : method II without the unit aspect ratio constraint									
method IIb : method II with the unit aspect ratio constraint									

will be $k \geq 9$ inflection points, providing at least four matches for computing the harmonic homology. Observe that, this result is valid even when the inflection points are imaginary, since there is no problem in mapping an imaginary inflection point to an also imaginary counterpart by a real harmonic homology. The result also holds when the inflection points have multiplicity greater than one, because then the derivatives of g at the inflection point will also be preserved by \mathbf{W} . As a result, if $d \geq 3$ and the matching of the inflection points is known, the harmonic homology can be determined. If the matching of the inflection points is not known, there will be at most a finite number of solutions, for different choices in the matching. In general, it should be possible to choose the correct \mathbf{W} by validating it against the rest of the curve. The discussion above demonstrates that conics are the only general (i.e., not for a particular view of a particular example) degeneracy in the computation of the harmonic homology when the silhouette is in the class of homogeneous algebraic

curves. It is worthwhile noting that any curve that admits a rational parameterization, such as B-splines and NURBS, can be expressed in homogeneous algebraic form [42] and, therefore, the proof is also valid in these particular cases.

6.2 Vanishing Point at Infinity

When the camera is pointing towards the revolution axis of the surface, the silhouette will exhibit bilateral or skew symmetry (see Section 3.3), and the vanishing point \mathbf{v}_x will be at infinity (i.e., $v_3 = 0$). In this situation, the line \mathbf{l}_x , on which the principal point is constrained to lie, cannot be determined, and this may cause the calibration equations to be under constrained.

To simplify the analysis, consider a camera with zero skew and unit aspect ratio. Assume now that one homology $\mathbf{W}(\mathbf{v}_x, \mathbf{l}_s)$ with $v_3 \neq 0$ is available. If a second homology $\mathbf{W}(\mathbf{v}'_x, \mathbf{l}'_s) \neq \mathbf{W}(\mathbf{v}_x, \mathbf{l}_s)$ is also available, there will then be four distinct possibilities for the computation of the principal point \mathbf{x}_0 and the focal length f :

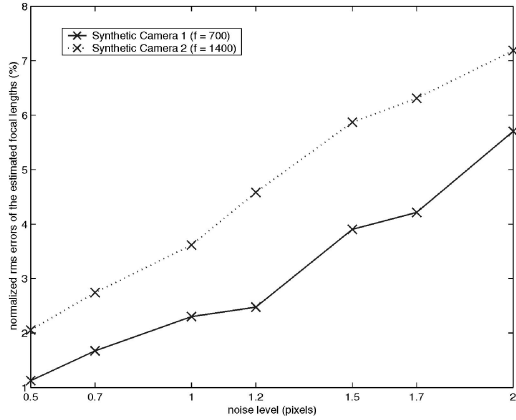


Fig. 17. Normalized rms errors of the estimated focal lengths obtained using method IIb under different noise levels.

1. $v'_3 \neq 0$ and $\mathbf{l}_x \neq \mathbf{l}'_x$: there will be a unique solution for \mathbf{x}_0 , given by $\mathbf{x}_0 = \mathbf{l}_x \times \mathbf{l}'_x$.
2. $v'_3 \neq 0$ and $\mathbf{l}_x = \mathbf{l}'_x$: there will be exactly one solution for \mathbf{x}_0 such that $\mathbf{l}'_x \mathbf{x}_0 = 0$ satisfying

$$\begin{aligned} f^2 &= \text{dist}(\mathbf{x}_0, \mathbf{v}_x) \times \text{dist}(\mathbf{x}_0, \mathbf{l}_s) \\ &= \text{dist}(\mathbf{x}_0, \mathbf{v}'_x) \times \text{dist}(\mathbf{x}_0, \mathbf{l}'_s). \end{aligned}$$

3. $v'_3 = 0$ and $\mathbf{l}_x \neq \mathbf{l}'_x$: note that the principal point is now constrained to lie on \mathbf{l}'_s and there will be a unique solution for \mathbf{x}_0 , given by $\mathbf{x}_0 = \mathbf{l}_x \times \mathbf{l}'_x$.
4. $v'_3 = 0$ and $\mathbf{l}_x = \mathbf{l}'_x$: there will be infinite number of solutions for \mathbf{x}_0 .

The discussion above demonstrates that whenever there are two distinct homologies \mathbf{W} and \mathbf{W}' such that both v_3 and v'_3 are not equal to zero, the computation of the principal point and, therefore, of the focal length is possible. Moreover, even when there is only one homology with $v_3 \neq 0$, the computation of the principal point and the focal length is still possible, as long as the highly unlikely condition that $\mathbf{l}_x = \mathbf{l}'_x$ does not occur. When the camera is pointing towards the revolution axis in all images (i.e., $v_3 = 0 \forall \mathbf{v}_x$), then only the principal point can be estimated.

7 ERROR ANALYSIS

7.1 Proportionality of the Error with the Focal Length

Experiments show that in estimating the harmonic homology \mathbf{W} associated with the silhouette ρ of a surface of revolution, the uncertainty is essentially in the vanishing point \mathbf{v}_x . Since \mathbf{v}_x is, in general, tens of thousands of pixels away from the axis \mathbf{l}_s , its error in a direction orthogonal to \mathbf{l}_s can be neglected in the computation of the principal point and focal length. On the other hand, the error of \mathbf{v}_x in a direction parallel to \mathbf{l}_s will lead to the same error in the estimated principal point \mathbf{x}_0 . This is due to the fact that, under the assumption of zero skew and unit aspect ratio, \mathbf{x}_0 must lie on the line \mathbf{l}_x passing through \mathbf{v}_x and orthogonal to \mathbf{l}_s (see Section 4).

Fig. 12 shows a point \mathbf{x} in ρ which is transformed by \mathbf{W} to its symmetric counterpart \mathbf{x}' in ρ . If \mathbf{v}_x has an error ξ in a direction parallel to \mathbf{l}_s , then the transformed point will have an error ε (see Fig. 12). It is easy to see that ξ and ε are related to each other by

$$\frac{\xi}{\varepsilon} = \frac{d_x + d'_x - w}{w + w'}. \quad (36)$$

Since $d_x = f^2/d'_x$ is much greater than d'_x, w , and w' , and that w and w' have nearly the same value, (36) can be rewritten as

$$\frac{\xi}{\varepsilon} \simeq \frac{f^2}{2d'_x w} = \frac{f}{2 \tan \psi w}, \quad (37)$$

where ψ is the angle between the optical axis and the plane Π_s , and $d'_x = f \tan \psi$. Equation (37) implies that, if ψ, w , and ε are assumed to be approximately constant, then the error ξ of \mathbf{v}_x and, hence, the error of the principal point \mathbf{x}_0 , in a direction parallel to \mathbf{l}_s , will be proportional to f . This might limit the usefulness of the technique to wide angle cameras.

7.2 Validation of the Error Analysis

To validate the error analysis described in the previous section, a simple experiment was performed. Four coplanar points configured into a square were placed in front of a camera with zero skew, unit aspect ratio, and variable focal length. Initially, the image plane of the camera was parallel to the square formed by these four points and the optical axis of

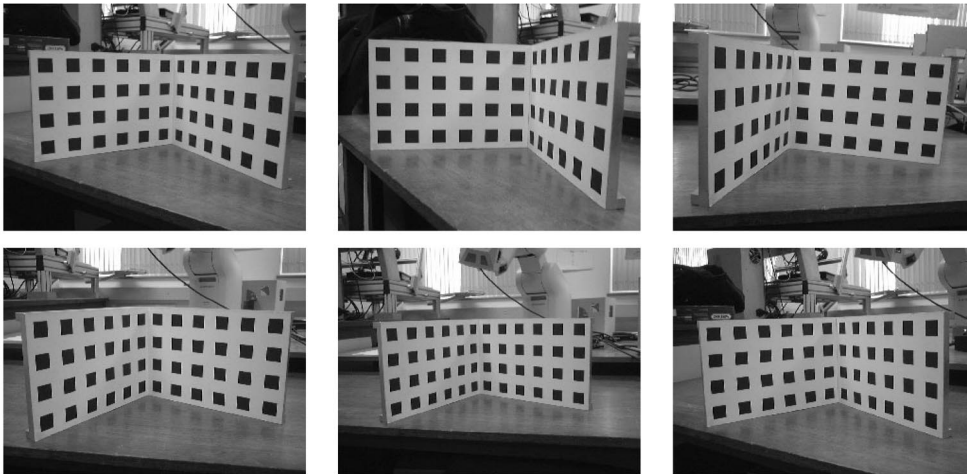


Fig. 18. Six images of a calibration grid taken by the digital camera for calibration.

TABLE 2
Results of Calibration from the Six Images
of the Calibration Grid

assumption: zero skew and unit aspect ratio				
	f_u	f_v	u_0	v_0
mean	687.92	687.92	320.98	230.88
std	1.90	1.90	4.00	3.35
assumption: zero skew				
	f_u	f_v	u_0	v_0
mean	687.70	688.09	320.98	230.88
std	1.83	2.10	4.00	3.34

(all units are in pixels)

the camera passed through the center of this square. The camera was then rotated about its optical center and away from the symmetry plane of the four points, distorting the otherwise bilaterally symmetric image to one invariant to a harmonic homology with the vanishing point at a finite position. Gaussian noise was then added to the coordinates of the projected points and the corresponding harmonic homology was computed. One can easily relate the parameters in this experiment with those in (37): ε is the standard deviation of the noise (i.e., noise level), ψ is the angle by which the camera was rotated from the symmetry plane prior to the projection of the points, and w was taken as the average distance from the points to \mathbf{l}_s . It is important to notice that w changes as the focal length changes, thus introducing a nuisance factor that will hide the true linear dependency between the magnitude of the focal length and the error in the position of the principal point. To compensate for that, the optical center of the camera was translated according to the value of the focal length (i.e., the larger the focal length, the further away from the four points the camera was placed), keeping the average value of w constant. Intuitively, this has the effect of maintaining the size of the calibration object (in this case, the four points) approximately constant in the image, despite the changes in the focal length. This experiment was repeated 100 times for each noise level, and the average norm of the errors in the position of the principal point is shown in Fig. 13, as a function of the focal length. The linear relationship described in (37) clearly holds, and the proportionality of the error with the noise level σ is also apparent.

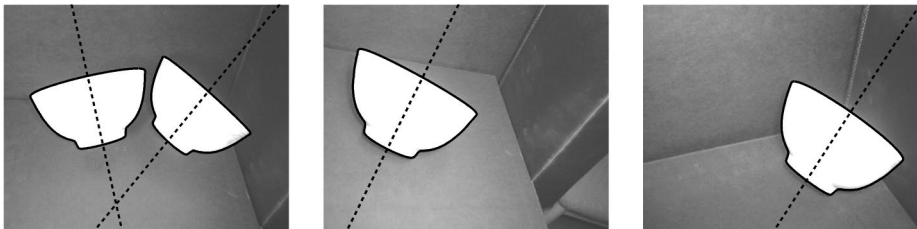


Fig. 19. Three images of two bowls with the extracted silhouettes and estimated images of the revolution axis plotted in solid and dash lines, respectively.

8 EXPERIMENTS AND RESULTS

Experiments on both synthetic and real data were carried out, and the results are presented in the following sections. In both cases, the cameras were assumed to have zero skew.

8.1 SYNTHETIC DATA

8.1.1 Generation of Data

The experimental setup consisted of a surface of revolution viewed by a synthetic camera at three different positions, as shown in Fig. 14. The synthetic images had a dimension of 640×480 pixels and the intrinsic parameters of the synthetic camera were given by the calibration matrix

$$\mathbf{K} = \begin{bmatrix} f & 0 & 320 \\ 0 & f & 240 \\ 0 & 0 & 1 \end{bmatrix} \text{ where } f = 700, 1400. \quad (38)$$

The surface of revolution was composed of two spheres intersecting each other. Each sphere was represented by a 4×4 symmetric matrix \mathbf{Q}_i whose projection was given by [43]

$$\mathbf{C}_{ij} = \left(\mathbf{P}_j \mathbf{Q}_i^{-1} \mathbf{P}_j^T \right)^{-1}, \quad (39)$$

where \mathbf{P}_j was a 3×4 projection matrix and \mathbf{C}_{ij} was a 3×3 symmetric matrix representing the conic, which was the projection of \mathbf{Q}_i in \mathbf{P}_j . The silhouette of the surface of revolution in each image was found by projecting each sphere \mathbf{Q}_i onto the image j as the conic \mathbf{C}_{ij} and finding points on each conic that lie outside the other conic. The silhouettes in the three images taken by the synthetic camera with $f = 700$ are shown in Fig. 15.

In order to evaluate the robustness of the algorithms described in Section 5, uniform random noise was added to each silhouette. Each point in the silhouette was perturbed in a direction normal to the local tangent, and the magnitudes of the noise were smoothed by a Gaussian filter so as to avoid unrealistic jaggedness along the silhouette (see Fig. 16).

8.1.2 Results on Synthetic Data

Experiments on the synthetic data with seven different noise levels were carried out. The seven noise levels were 0.5, 0.7, 1.0, 1.2, 1.5, 1.7, and 2.0 pixels, respectively. The noise level for typical real images ranges from 0.7 to 1.5 pixels and the distortion of the silhouette will be too great to be realistic when the noise level is above 2.0 pixels. For each noise level, 100 experiments were conducted using the algorithms described in Section 5. In the estimation of the harmonic homology, the number of sample points used was 100. Both Methods I and II described in Section 5.2 were used for the computation of the intrinsic parameters.

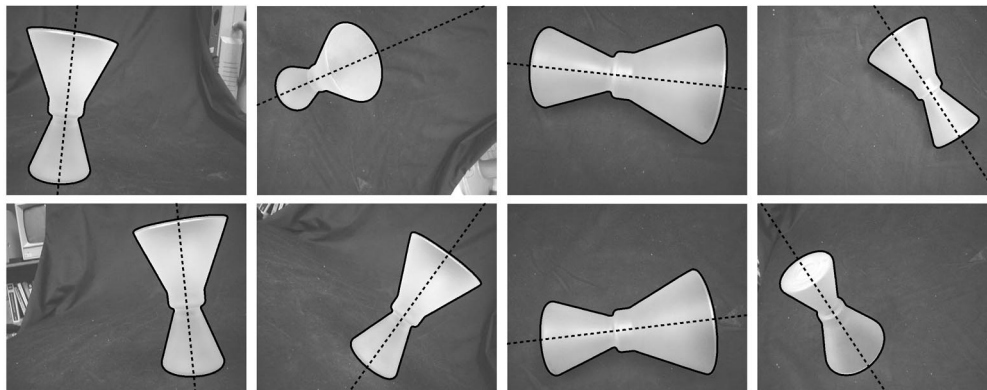


Fig. 20. Eight images of a candle holder with the extracted silhouettes and estimated images of the revolution axis plotted in solid and dash lines, respectively.

The experimental results are presented in Table 1, which shows the root-mean-square (rms) errors of the estimated intrinsic parameters. Note that the values listed in the table have been normalized and are the percentage errors relative to the corresponding ground truth focal lengths. It can be seen from Table 1 that the focal lengths obtained using Method II were better than those obtained using Method I, regardless of whether the unit aspect ratio constraint was imposed or not. For different versions of Method II (i.e., IIa with free aspect ratio and IIb with unit aspect ratio), IIb gave the best results. This is consistent with the well-known fact that the enforcement of known constraints gives more accurate results in camera calibration. Observe that this does not contradict the fact that Method IIa gave better results than Method I for, in this case, different algorithms are being compared. As the noise level increased, the relative errors in the estimated intrinsic parameters increased. Table 1 also shows that the errors increased with the focal length of the camera (see also Fig. 17), and this agrees with the error analysis presented in Section 7. For a noise level of 2.0 pixels, the errors in the estimated focal lengths were less than 6.0 percent and 7.5 percent for the synthetic cameras with $f = 700$ and $f = 1,400$, respectively.

8.2 Real Data

8.2.1 The Ground Truth

The camera used in the real data experiments was a digital camera with a resolution of 640×480 pixels. The ground truth for the intrinsic parameters of the camera was

obtained using a calibration grid. Six images of a calibration grid were taken with the camera at different orientations (see Fig. 18). Corner features were extracted from each image using a Canny edge detector [36] and line fitting techniques. For each image, the camera was calibrated using the DLT technique [19] followed by an optimization which minimized the reprojection errors of the corner features [1], [6]. The results of calibration from the calibration grid are shown in Table 2.

8.2.2 Results on Real Data

Two sets of real images of surfaces of revolution were used for the calibration of the digital camera. The first set consisted of three images of two bowls, which provided four silhouettes for camera calibration (see Fig. 19). The second set consisted of eight images of a candle holder, which provided eight silhouettes for camera calibration (see Fig. 20). The results of calibration from the two image sets are shown in Table 3, and Table 4 shows the percentage errors of the estimated intrinsic parameters relative to the ground truth focal lengths. Similar to the results of the synthetic experiments, the intrinsic parameters obtained in the real data experiments using Method II were better than those obtained using Method I. The focal lengths estimated from both the bowls set and the candle holder set (using Method II with unit aspect ratio constraint) had an error of only around 3 percent relative to the ground truth focal length. Figs. 19 and 20 show the

TABLE 3

Results of Calibration from the Bowls and Candle Holder Sets

image set	method	f_u	f_v	u_0	v_0
bowls	I	727.12	727.12	332.34	227.21
bowls	IIa	706.65	708.55	321.25	246.25
bowls	IIb	708.63	708.63	322.15	244.75
candle holder	I	727.93	727.93	279.21	231.40
candle holder	IIa	708.44	705.47	331.53	231.18
candle holder	IIb	707.40	707.40	330.60	231.54

(all units are in pixels)

TABLE 4
Percentage Errors in the Results of Calibration
from the Two Image Sets

		Percentage Errors (%)			
image set	method	f_u	f_v	u_0	v_0
bowls	I	5.70	5.70	1.65	0.53
bowls	IIa	2.76	2.97	0.04	2.23
bowls	IIb	3.01	3.01	0.17	2.02
candle holder	I	5.82	5.82	6.07	0.08
candle holder	IIa	3.02	2.53	1.53	0.04
candle holder	IIb	2.83	2.83	1.40	0.10

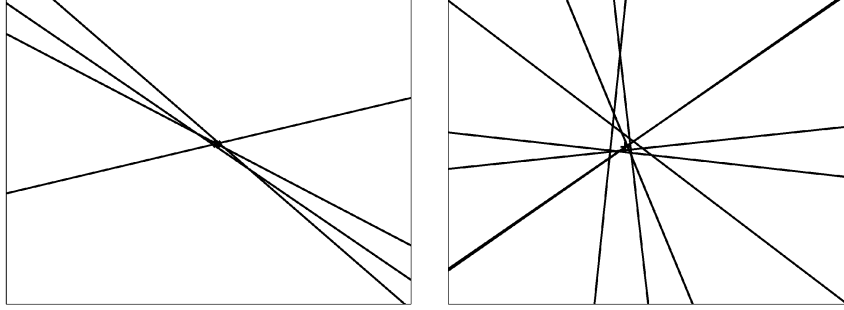


Fig. 21. The solid lines represent the lines l_{xi} passing through the corresponding vanishing point v_{xi} and orthogonal to the corresponding image of the axis revolution l_{si} . Since the principal point x_0 must lie on these lines, it can be estimated as the intersection of two or more lines l_{xi} .

extracted silhouettes and the estimated images of the revolution axis. Fig. 21 shows the lines l_{xi} passing through the corresponding vanishing point v_{xi} and orthogonal to the corresponding image of the revolution axis l_{si} .

From Table 3 and Table 4, it can be seen that the intrinsic parameters estimated from the candle holder set were slightly better than those from the bowls set. This can be explained as the silhouettes in the candle holder set showed much greater perspective effect than those in the bowls set (see Fig. 19 and Fig. 20). Besides, the candle holder set also provided more silhouettes, and hence more constraints, than the bowls set for the estimation of the intrinsic parameters.

9 CONCLUSIONS

By exploiting the symmetry properties exhibited in the silhouettes of surfaces of revolution and relating them to vanishing points, a practical technique for camera calibration has been developed. The use of surfaces of revolution makes the calibration process easier in not requiring the use of any precisely machined device with known geometry, such as a calibration grid. Besides, a surface of revolution can always be generated by rotating an object of any arbitrary shape around a fixed axis. It means that the calibration technique introduced here can be integrated into a motion estimation and model building system for turntable sequences [44], [28].

Despite the fact that strong perspective effect is required, the method introduced here is promising as demonstrated by the experimental results on both synthetic and real data. The focal lengths were estimated with high accuracy, having an error of only around 3 percent with respect to the ground truth. Nonetheless, note that neither of the implementations proposed in Section 5 is statistically optimal, even though the computation of each individual harmonic homology is. The statistically optimal way, which will not be discussed in details here, would be integrating all the information provided by the silhouettes to estimate the intrinsic parameters and the harmonic homologies simultaneously. For a set of N images, this would involve an optimization over a $2N + 4$ parameter space: 4 for the intrinsic parameters, or equivalently, for the image of the absolute conic ω ; and $2N$ for the N vanishing points v_x or, equivalently, for the N images of the revolution axes l_s (as $\omega v_x = l_s$). This might be included in future work.

APPENDIX

DEFINITION OF THE HARMONIC HOMOLOGY

A *perspective collineation* [30], with center x_c and axis l_a , is a collineation which leaves all the lines through x_c and points of l_a invariant. If the center x_c and the axis l_a are not incident, the perspective collineation is called a *homology* [30]; otherwise, it is called an *elation* [30]. Consider a point x which is mapped by a homology with center x_c and axis l_a to the point x' . Let x'_c be the point of intersection between the axis l_a and a line passing through the points x and x' . The homology is said to be harmonic if the points x and x' are harmonic conjugates with respect to x_c and x'_c (i.e., the cross-ratio $\{x_c, x'_c; x, x'\}$ equals -1). The matrix W representing a *harmonic homology* [30] with center x_c and axis l_a , in homogeneous coordinates, is given by $W = \Pi_3 - 2 \frac{x_c l_a^T}{x_c^T l_a}$. More details on harmonic homology can be found in [29], [30].

ACKNOWLEDGMENTS

The authors would like to thank Dr. Tom Drummond, Dr. Andrew Fitzgibbon, and the three anonymous referees for their valuable comments. This work was partially funded by The University of Hong Kong. A preliminary version of this paper was presented at the Ninth IMA Conference on The Mathematics of Surfaces, September 2000, Cambridge, United Kingdom.

REFERENCES

- [1] O.D. Faugeras and G. Toscani, "The Calibration Problem for Stereo," *Proc. Conf. Computer Vision and Pattern Recognition*, pp. 15-20, June 1986.
- [2] R.Y. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," *IEEE Trans. Robotics and Automation*, vol. 3, no. 4, pp. 323-344, Aug. 1987.
- [3] R.K. Lenz and R.Y. Tsai, "Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3D Machine Vision Metrology," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 10, no. 5, pp. 713-720, Sept. 1988.
- [4] Z. Zhang, "A Flexible New Technique for Camera Calibration," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330-1334, Nov. 2000.
- [5] R.I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, UK: Cambridge Univ. Press, 2000.
- [6] O.D. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*. Cambridge, Mass.: MIT Press, 1993.

- [7] B. Caprile and V. Torre, "Using Vanishing Points for Camera Calibration," *Int'l J. Computer Vision*, vol. 4, no. 2, pp. 127-140, Mar. 1990.
- [8] R. Cipolla, T.W. Drummond, and D. Robertson, "Camera Calibration from Vanishing Points in Images of Architectural Scenes," *Proc. British Machine Vision Conf.*, T. Pridmore and D. Elliman, eds., vol. 2, pp. 382-391, Sept. 1999.
- [9] D. Liebowitz and A. Zisserman, "Combining Scene and Auto-Calibration Constraints," *Proc. Seventh Int'l Conf. Computer Vision*, pp. 293-300, Sept. 1999.
- [10] A. Zisserman, D. Forsyth, J.L. Mundy, and C.A. Rothwell, "Recognizing General Curved Objects Efficiently," *Geometric Invariance in Computer Vision*, J.L. Mundy and A. Zisserman, eds., Artificial Intelligence Series, chapter 11, Cambridge, Mass.: MIT Press, pp. 228-251, 1992.
- [11] J. Liu, J.L. Mundy, D.A. Forsyth, A. Zisserman, and C.A. Rothwell, "Efficient Recognition of Rotationally Symmetric Surface and Straight Homogeneous Generalized Cylinders," *Proc. Conf. Computer Vision and Pattern Recognition*, pp. 123-129, June 1993.
- [12] J.L. Mundy and A. Zisserman, "Repeated Structures: Image Correspondence Constraints and 3D Structure Recovery," *Proc. Applications of Invariance in Computer Vision, Second Joint European-US Workshop*, J.L. Mundy, A. Zisserman, and D. Forsyth, eds., vol. 825, pp. 89-106, Oct. 1993.
- [13] A. Zisserman, J.L. Mundy, D.A. Forsyth, J. Liu, N. Pillow, C. Rothwell, and S. Utcke, "Class-Based Grouping in Perspective Images," *Proc. Fifth Int'l Conf. Computer Vision*, pp. 183-188, June 1995.
- [14] R.W. Curwen, C.V. Stewart, and J.L. Mundy, "Recognition of Plane Projective Symmetry," *Proc. Sixth Int'l Conf. Computer Vision*, pp. 1115-1122, Jan. 1998.
- [15] K.-Y.K. Wong, P.R.S. Mendonça, and R. Cipolla, "Camera Calibration from Symmetry," *The Mathematics of Surfaces IX*, R. Cipolla and R. Martin, eds., Cambridge, UK: Springer-Verlag, pp. 214-226, Sept. 2000.
- [16] D.C. Brown, "Close-Range Camera Calibration," *Photogrammetric Eng.*, vol. 37, no. 8, pp. 855-866, Aug. 1971.
- [17] I. Sobel, "On Calibrating Computer Controlled Cameras for Perceiving 3-D Scenes," *Artificial Intelligence*, vol. 5, no. 2, pp. 185-198, June 1974.
- [18] W. Faig, "Calibration of Close-Range Photogrammetry System: Mathematical Formulation," *Photogrammetric Eng. and Remote Sensing*, vol. 41, no. 12, pp. 1479-1486, Dec. 1975.
- [19] Y.I. Abdel-Aziz and H.M. Karara, "Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close-Range Photogrammetry," *Proc. ASP/UI Symp. Close-Range Photogrammetry*, pp. 1-18, Jan. 1971.
- [20] S.J. Maybank and O.D. Faugeras, "A Theory of Self-Calibration of a Moving Camera," *Int'l J. Computer Vision*, vol. 8, no. 2, pp. 123-151, Aug. 1992.
- [21] Q.T. Luong and O.D. Faugeras, "Self-Calibration of a Moving Camera from Point Correspondences and Fundamental Matrices," *Int'l J. Computer Vision*, vol. 22, no. 3, pp. 261-289, Mar. 1997.
- [22] B. Triggs, "Autocalibration and the Absolute Quadric," *Proc. Conf. Computer Vision and Pattern Recognition*, pp. 609-614, June 1997.
- [23] M. Pollefeys, R. Koch, and L.J. van Gool, "Self-Calibration and Metric Reconstruction In spite of Varying and Unknown Intrinsic Camera Parameters," *Int'l J. Computer Vision*, vol. 32, no. 1, pp. 7-25, Aug. 1999.
- [24] L. deAgapito, E. Hayman, and I.D. Reid, "Self-Calibration of Rotating and Zooming Cameras," *Int'l J. Computer Vision*, vol. 45, no. 2, pp. 107-127, Nov. 2001.
- [25] M. Armstrong, A. Zisserman, and R. Hartley, "Self-Calibration from Image Triplets," *Proc. Fourth European Conf. Computer Vision*, B. Buxton and R. Cipolla, eds., vol. 1064, pp. 3-16, Apr. 1996.
- [26] L. G. Roberts, "Machine Perception of Three-Dimensional Solids," *Optical and Electro-Optical Information Processing*, J.T. Tippett, D.A. Berkowitz, L.C. Clapp, C.J. Koester, and A. Vanderburgh, Jr., eds., pp. 159-197, MIT Press, 1965.
- [27] P.R.S. Mendonça, K.-Y.K. Wong, and R. Cipolla, "Recovery of Circular Motion from Profiles of Surfaces," *Vision Algorithms: Theory and Practice*, B. Triggs, A. Zisserman, and R. Szeliski, eds., vol. 1883, pp. 151-167, Sept. 1999.
- [28] P.R.S. Mendonça, K.-Y.K. Wong, and R. Cipolla, "Epipolar Geometry from Profiles under Circular Motion," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 604-616, June 2001.
- [29] J.G. Semple and G.T. Kneebone, *Algebraic Projective Geometry*. Oxford, UK: Clarendon Press, 1998, originally published in 1952.
- [30] H.S.M. Coxeter, *Introduction to Geometry*, second ed. New York: Wiley and Sons, 1989.
- [31] J.R. Kender and T. Kanade, "Mapping Image Properties into Shape Constraints: Skewed Symmetry, Affine-Transformable Patterns, and the Shape-from-Texture Paradigm," *Human and Machine Vision*, J. Beck, B. Hope, and A. Rosenfeld, eds., pp. 237-257, New York: Academic Press, 1983.
- [32] D.P. Mukherjee, A. Zisserman, and J.M. Brady, "Shape from Symmetry—Detecting and Exploiting Symmetry in Affine Images," *Philosophical Trans. Royal Soc. London A*, vol. 351, pp. 77-106, 1995.
- [33] T.J. Cham and R. Cipolla, "Geometric Saliency of Curve Correspondences and Grouping of Symmetric Contours," *Proc. Fourth European Conf. Computer Vision*, B. Buxton and R. Cipolla, eds., pp. 385-398, Apr. 1996.
- [34] J. Sato and R. Cipolla, "Affine Integral Invariants for Extracting Symmetry Axes," *Image and Vision Computing*, vol. 15, no. 8, pp. 627-635, Aug. 1997.
- [35] A. Zisserman, D. Liebowitz, and M. Armstrong, "Resolving Ambiguities in Auto-Calibration," *Philosophical Trans. Royal Soc. London A*, vol. 356, no. 1740, pp. 1193-1211, 1998.
- [36] J. Canny, "A Computational Approach to Edge Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679-698, Nov. 1986.
- [37] M. Wright, A.W. Fitzgibbon, P.J. Giblin, and R.B. Fisher, "Convex Hulls, Occluding Contours, Aspect Graphs and the Hough Transform," *Image and Vision Computing*, vol. 14, no. 8, pp. 627-634, Aug. 1996.
- [38] R. Glacahet, M. Dhome, and J.T. Lapreste, "Finding the Perspective Projection of an Axis of Revolution," *Pattern Recognition Letters*, vol. 12, pp. 693-700, Nov. 1991.
- [39] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, second ed. Cambridge, UK: Cambridge Univ. Press Jan. 1993.
- [40] F. Kirwan, *Complex Algebraic Curves*. Cambridge, UK: Cambridge Univ. Press, 1992.
- [41] J. Plücker, *System der Analytischen Geometrie: Auf neue Betrachtungsweisen Gegründet und Insbesondere eine Ausführliche Theorie der Curven Dritter Ordnung Enthaltend*. Berlin, 1835.
- [42] T.W. Sederberg, D.C. Anderson, and R.N. Goldman, "Implicit Representation of Parametric Curves and Surfaces," *Computer Vision, Graphics, and Image Processing*, vol. 28, no. 1, pp. 72-84, Oct. 1984.
- [43] G. Cross and A. Zisserman, "Quadric Reconstruction from Dual-Space Geometry," *Proc. Sixth Int'l Conf. Computer Vision*, pp. 25-31, Jan. 1998.
- [44] A.W. Fitzgibbon, G. Cross, and A. Zisserman, "Automatic 3D Model Construction for Turn-Table Sequences," *Proc. 3D Structure from Multiple Images of Large-Scale Environments, European Workshop, SMILE '98*, R. Koch and L. Van Gool, eds., vol. 1506, pp. 155-170, June 1998.



Kwan-Yee K. Wong received the BEng degree in computer engineering, with first class honors, from the Chinese University of Hong Kong in 1998. From 1998 to 2001, he studied in United Kingdom at the University of Cambridge, and he received the MPhil and PhD degrees, both in computer vision, from the University of Cambridge in 2000 and 2001, respectively. Since 2001, he has been with the Department of Computer Science and Information Systems at the University of Hong Kong, where he is now an assistant professor. His research interests include camera calibration, surface reconstruction and representation, and motion estimation from image sequences.



Paulo R.S. Mendonça received the BSc degree in electrical engineering from the Federal University of Pernambuco, Brazil, in 1995, and the MSc degree in signal processing from the Federal University of Rio de Janeiro, in 1997. In 2001, he was awarded the PhD degree in computer vision from the University of Cambridge, United Kingdom. He is currently working as a computer scientist for GE Global Research Center. His research interests include stochastic

optimization, camera self-calibration, projective geometry, structure and motion from image sequences, geometry-based computer vision, and medical imaging.



Roberto Cipolla received the BA degree (engineering) from the University of Cambridge, England, in 1984, and the MSEE (electrical engineering) from the University of Pennsylvania in 1985. From 1985 to 1988, he studied and worked in Japan at the Osaka University of Foreign Studies (Japanese language) and the Electro-Technical Laboratory, Tsukuba (visiting scientist), and he obtained the MEng degree (robotics) from the University of Electro-Communications in Tokyo in 1988. In 1991, he was awarded the DPhil (computer vision) degree by the University of Oxford, England, and from 1991–1992, he was a Toshiba fellow and engineer at the Toshiba Corporation Research and Development Center in Kawasaki, Japan. He joined the Department of Engineering, University of Cambridge, in 1992, as a lecturer and a fellow of Jesus College. He became a reader in 1997 and a professor of information engineering in 2000. His research interests are in computer vision and robotics and include recovery of motion and 3D shape of visible surfaces from image sequences, visual tracking and navigation, robot hand-eye coordination, algebraic and geometric invariants for object recognition and perceptual grouping, and novel man-machine interfaces using visual gestures and visual inspection. He is the author of two books, editor of five volumes, and co-author of more than 150 papers. He is a member of the IEEE.

▷ For more information on this or any other computing topic, please visit our Digital Library at <http://computer.org/publications/dlib>.