Recovery of Circular Motion from Profiles of Surfaces

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

Department of Engineering, University of Cambridge,
Trumpington Street, Cambridge, CB2 1PZ, UK
{prdsm2, kykw2, cipolla}@eng.cam.ac.uk

Abstract. This paper addresses the problem of motion recovery from image profiles, in the important case of turntable sequences. No correspondences between points or lines are used. Symmetry properties of surfaces of revolution are exploited to obtain, in a robust and simple way, the image of the rotation axis of the sequence and the homography relating epipolar lines. These, together with geometric constraints for images of rotating objects, are used to obtain epipoles and, consequently, the full epipolar geometry of the camera system. This sequential approach (image of rotation axis — homography — epipoles) avoids many of the problems usually found in other algorithms for motion recovery from profiles. In particular, the search for the epipoles, by far the most critical step for the estimation of the epipolar geometry, is carried out as a one-dimensional optimization problem, with a smooth unimodal cost function. The initialization of the parameters is trivial in all three stages of the algorithm. After the estimation of the epipolar geometry, the motion is recovered using the fixed intrinsic parameters of the camera, obtained either from a calibration grid or from self-calibration techniques. Results from real data are presented, demonstrating the efficiency and practicality of the algorithm.

1 Introduction

Points and lines have long been used for the recovery of structure and motion from images of 3D objects. Nevertheless, for a smooth surface the predominant feature in the image is its *profile* or *apparent contour*, defined as the projection of a *contour generator* of the surface. A contour generator corresponds to the set of points on a surface where the normal vector to the surface is orthogonal to the rays joining the points in the set and the camera center (for details, see [3, 4]). If the surface does not have noticeable texture, the profile may actually be the only source of information available for estimating the structure of the surface and the motion of the camera.

The problem of motion recovery from image profiles has been tackled in several works. The concept of *frontier point*, defined as a point on a surface tangent to any plane of the pencil of epipolar planes related to a pair of images, was introduced in [15]. The idea was further developed in [14], where the frontier point was recognized as a fixed point on a surface, created by the intersection of two contour generators. A frontier point projects on its associated images as an *epipolar tangency*. The use of frontier points and epipolar tangencies for motion recovery was first shown in [2]. A parallax based technique, using a reference planar contour was shown in [1], where the images are registered using the reference contour and common tangents are used

to determine the projections of the frontier point. The techniques described above face two main difficulties: the likely non-uniqueness of the solution, due to the presence of local minima, and the unrealistic requirement of having at least 7 corresponding epipolar tangencies available on each image pair. Better results can be achieved when an affine approximation is used, as shown in [11]. In this case the problem can be solved when as few as 4 epipolar tangencies are available, but the application of the method is constrained to situations where the affine approximation is valid.

In the case of circular motion, the envelope of the profiles exhibits symmetry properties that greatly simplify this estimation problem. This is an idea well developed for orthographic projection. In [15] it is shown that, when the image plane is parallel to the axis of rotation, the image of the axis of rotation will be perpendicular to common tangents to the images of the profile. The use of bilateral symmetry to obtain the axis of rotation was first introduced in [13]. The condition of parallelism between the image plane and the axis of rotation was relaxed in [8], but orthographic projection was still used.

In this paper we introduce a novel technique for the estimation of the motion parameters of turntable sequences. It based on symmetry properties of the set of apparent contours generated by the object that undergoes the rotation. In Section 2, a method for obtaining the images of the axis of rotation and a special vanishing point is presented. The algorithm is simple, efficient and robust, and it does not make direct use of the profiles. Therefore, its use can be extended to non-smooth objects, and the quality of the results obtained justifies doing so. Section 3 makes use of the previous results to introduce a parameterization of the fundamental matrix based on the *harmonic homology*. This parameterization allows for the estimation of the epipoles to be carried out as independent one-dimensional searches, avoiding local minima points and greatly decreasing the computational complexity of the estimation. These results are used in Section 4, which presents the algorithm for motion estimation. Experimental results are shown in Section 5, and conclusions and future work are described in Section 6.

2 Theoretical Background

An object rotating about a fixed axis sweeps out a surface of revolution [8]. Symmetry properties [18, 19] of the image of this surface of revolution can be exploited to estimate the parameters of the motion of the object in a simple and elegant way, as will be shown next.

2.1 Symmetry Properties of Images of Surfaces of Revolution

In the definitions that follow, points and lines will be referred to by their representation as vectors in homogeneous coordinates.

A 2D homography that keeps the pencil of lines through a point $\mathbf{u}$ and the set of points on a line $\mathbf{l}$ fixed is called a *perspective collineation* with center $\mathbf{u}$ and axis $\mathbf{l}$. An *homology* is a perspective collineation whose center and axis are not incident (otherwise the perspective homology is called an *elation*). Let $\mathbf{a}$ be a point mapped by an homology onto a point $\mathbf{a}'$. It is easy to show that the center of the homology $\mathbf{u}$, $\mathbf{a}$ and $\mathbf{a}'$ are

collinear. Let $\mathbf{q}_a$ be the line passing through these points, and $\mathbf{v}_a$ the intersection of $\mathbf{q}_a$ and the axis $\mathbf{l}$. If $\mathbf{a}$ and $\mathbf{a}'$ are harmonic conjugates with respect to $\mathbf{u}$ and $\mathbf{v}_a$, i.e., their cross-ratio is one, the homology is said to be a *harmonic homology* (see details in [16, 5]). The matrix $\mathbf{W}$ representing a harmonic homology with center $\mathbf{u}$ and axis $\mathbf{l}$ in homogeneous coordinates is given by

$$\mathbf{W} = \mathbb{I} - 2\frac{\mathbf{u}\,\mathbf{l}^{\mathrm{T}}}{\mathbf{u}^{\mathrm{T}}\mathbf{l}}. \tag{1}$$

Henceforth a matrix representing a projective transformation in homogeneous coordinates will be used in reference to the transformation itself whenever an ambiguity does not arise.

An important property of profiles of surfaces of revolution is stated in the next theorem:

Theorem 1. *The profile of a surface of revolution S viewed by a pinhole camera is invariant to the harmonic homology with axis given by the image of the axis of rotation of the surface of revolution and center given by the image of the point at infinity in a direction orthogonal to a plane that contains the axis of rotation and the camera center.*

The following lemma will be used in the proof of Theorem 1.

Lemma 1. *Let $\mathbf{T} : \Gamma \mapsto \Gamma$ be a harmonic homology with axis $\mathbf{l}$ and center $\mathbf{u}$ on the plane Γ, and let $\mathbf{H} : \Gamma \mapsto \Gamma'$ be a bijective 2D homography. Then, the transformation $\mathbf{W} = \mathbf{H}\mathbf{T}\mathbf{H}^{-1} : \Gamma' \mapsto \Gamma'$ is a harmonic homology with axis $\mathbf{l}' = \mathbf{H}^{-\mathrm{T}}\mathbf{l}$ and center $\mathbf{u}' = \mathbf{H}\mathbf{u}$.*

Proof. Since $\mathbf{H}$ is bijective, $\mathbf{H}^{-1}$ exists. Then

$$\begin{aligned}\mathbf{W} &= \mathbf{H}\left(\mathbb{I} - 2\frac{\mathbf{u}\mathbf{l}^{\mathrm{T}}}{\mathbf{u}^{\mathrm{T}}\mathbf{l}}\right)\mathbf{H}^{-1} \\ &= \mathbb{I} - 2\frac{\mathbf{u}'\mathbf{l}'^{\mathrm{T}}}{\mathbf{u}'^{\mathrm{T}}\mathbf{l}'},\end{aligned} \tag{2}$$

since $\mathbf{u}^{\mathrm{T}}\mathbf{l} = \mathbf{u}'^{\mathrm{T}}\mathbf{l}'$. □

The following corollary is a trivial consequence of Lemma 1:

Corollary 1. *Let $\mathbf{T}$, $\mathbf{H}$ and $\mathbf{W}$ be defined as in Lemma 1. The transformation $\mathbf{H}$ is a* isomorphism *between the structures $(\mathbf{T}, \Gamma)$ and $(\mathbf{W}, \Gamma')$, i.e, $\forall \boldsymbol{\gamma} \in \Gamma$, $\mathbf{H}\mathbf{T}\boldsymbol{\gamma} = \mathbf{W}\mathbf{H}\boldsymbol{\gamma}$.*

An important consequence of Lemma 1 and Corollary 1 is that if a set of points s, e.g., the profile of a surface of revolution, is invariant to a harmonic homology $\mathbf{T}$, the set $\hat{s}$ obtained by transforming s by a 2D projective transformation $\mathbf{H}$ is invariant to the harmonic homology $\mathbf{W} = \mathbf{H}\mathbf{T}\mathbf{H}^{-1}$.

Without loss of generality assume that the axis of rotation of the surface of revolution S is coincident with the y-axis of an right-handed orthogonal coordinate system. Considering a particular case of Theorem 1 where the pinhole camera $\mathbf{P}$ is given by

$\mathbf{P} = [\mathbb{I}\,|\mathbf{t}]$, where $\mathbf{t} = [0\ 0\ \alpha]^{\mathrm{T}}$, for any $\alpha > 0$, symmetry considerations show that the profile s of S will be bilaterally symmetric with respect to the image of y (a proof is presented in the Appendix 1), which corresponds to the line $\mathbf{q}_s = [1\ 0\ 0]^{\mathrm{T}}$ in (homogeneous) image coordinates.

Proof of Theorem 1 (particular case). Since s is bilaterally symmetrical about $\mathbf{q}_s$, there is a transformation $\mathbf{T}$ that maps each point of s on its symmetrical counterpart, given by

$$\mathbf{T} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3}$$

However, as any bilateral symmetry transformation, $\mathbf{T}$ is also a harmonic homology, with axis $\mathbf{q}_s$ and center $\mathbf{v}_x = [1\ 0\ 0]^{\mathrm{T}}$, since

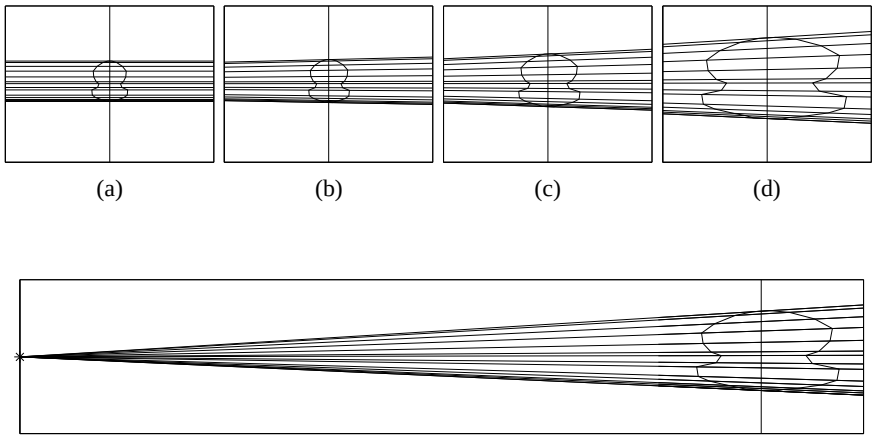
$$\mathbf{T} = \mathbb{I} - 2\frac{\mathbf{v}_x \mathbf{q}_s^{\mathrm{T}}}{\mathbf{v}_x^{\mathrm{T}} \mathbf{q}_s}. \tag{4}$$

The transformation $\mathbf{T}$ maps the set s onto itself (although the points of s are not mapped onto themselves by $\mathbf{T}$, but on their symmetrical counterparts), and thus s is invariant to the harmonic homology $\mathbf{T}$. Since the camera center lies on the z-axis of the coordinate system, the plane that contains the camera center and the axis of rotation is in fact the yz-plane, and the point at infinity orthogonal to the yz-plane is $\mathbf{U}_x = [1\ 0\ 0\ 0]^{\mathrm{T}}$, whose image is $\mathbf{v}_x$. □

Let $\hat{\mathbf{P}}$ be an arbitrary pinhole camera. The camera $\hat{\mathbf{P}}$ can be obtained by rotating $\mathbf{P}$ about its optical center by a rotation $\mathbf{R}$ and transforming the image coordinate system of $\mathbf{P}$ by introducing the intrinsic parameters represented by the matrix $\mathbf{K}$. Let $\mathbf{KR} = \mathbf{H}$. Thus, $\hat{\mathbf{P}} = \mathbf{H}[\mathbf{I}\,|\mathbf{t}]$, and the point $\mathbf{U}_x$ in space with the image $\mathbf{v}_x$ in $\mathbf{P}$ will project as a point $\mathbf{u}_x = \mathbf{H}\mathbf{v}_x$ in $\hat{\mathbf{P}}$. Analogously, the line $\mathbf{q}_s$ in $\mathbf{P}$ will correspond to a line $\mathbf{l}_s = \mathbf{H}^{-\mathrm{T}}\mathbf{q}_s$ in $\hat{\mathbf{P}}$. It is now possible to derive the proof of Theorem 1 in the general case.

Proof of Theorem 1 (general case). Let $\hat{s}$ be the profile of the surface of revolution S obtained from the camera $\hat{\mathbf{P}}$. Thus, the counter-domain of the bijection $\mathbf{H}$ acting on the profile s is $\hat{s}$ (or $\mathbf{H}s = \hat{s}$), and, using Lemma 1, the transformation $\mathbf{W} = \mathbf{HTH}^{-1}$ is a harmonic homology with center $\mathbf{u}_x = \mathbf{H}\mathbf{v}_x$ and axis $\mathbf{l}_s = \mathbf{H}^{-\mathrm{T}}\mathbf{q}_s$. Moreover, from Corollary 1, $\mathbf{WH}s = \mathbf{HT}s$, or $\mathbf{W}\hat{s} = \mathbf{HT}s$. From the particular case of the Theorem 1 it is known that the profile s will be invariant to the harmonic homology $\mathbf{T}$, so $\mathbf{W}\hat{s} = \mathbf{H}s = \hat{s}$. □

The images of a rotating object are the same as the images of a fixed object taken by a camera rotating around the same axis, or by multiple cameras along that circular trajectory. Consider any two of such cameras, denoted by $\mathbf{P}$ and $\mathbf{P}'$. If $\mathbf{P}$ and $\mathbf{P}'$ point towards the axis of rotation, their epipoles $\mathbf{e}$ and $\mathbf{e}'$ will be symmetrical with respect to the image of the rotation axis, or $\mathbf{e}' = \mathbf{Te}$, according to Figure 2. In a general situation, the epipoles will simply be related by the transformation $\mathbf{e}' = \mathbf{We}$. It is then



(a) (b) (c) (d)

(e)

Fig. 1. Lines joining symmetric points with respect to the image of rotation axis $\mathbf{l}_s$ (images are scaled and translated independently for better observation). (a) The optical axis points directly towards the rotation axis. (b) The camera is rotated about its optical center by an angle ρ of 20° in a plane orthogonal to the rotation axis. (c) $\rho = 40^\circ$. (d) $\rho = 60^\circ$. (e) Same as (d), but the vanishing point $\mathbf{v}_x$ is also shown.

straightforward to show that the corresponding epipolar lines $\mathbf{l}$ and $\mathbf{l}'$ are related by $\mathbf{l}' = \mathbf{W}^{-\mathrm{T}}\mathbf{l}$. This means that the pair of epipoles can be represented with only two parameters once $\mathbf{W}$ is known. From (2) it can be seen that $\mathbf{W}$ has only four degrees of freedom (dof). Therefore, the fundamental matrix relating views of an object under circular motion must have only 6 dof, in agreement with [17].

3 Parameterization of the Fundamental Matrix

3.1 Epipolar Geometry under Circular Motion

The fundamental matrix corresponding to a pair of cameras related by a rotation around a fixed axis has a very special parameterization, as shown in [17, 7]. A simpler derivation of this result will be shown here.

Consider the pair of camera matrices $\mathbf{P}_1$ and $\mathbf{P}_2$, given by

$$\begin{aligned} \mathbf{P}_1 &= [\mathbb{I}\,|\,\mathbf{t}\,] \\ \mathbf{P}_2 &= [\mathbf{R}_y(\theta)|\,\mathbf{t}\,], \end{aligned} \tag{5}$$

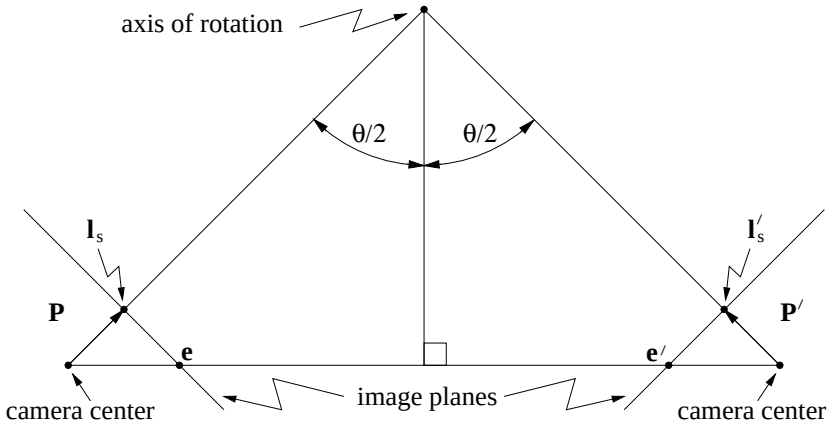


Fig. 2. If the cameras are pointing towards the axis of rotation, the epipoles $\mathbf{e}$ and $\mathbf{e}'$ are symmetric with respect to the image of the axis of rotation.

where

$$\mathbf{t} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^{\mathrm{T}} \text{ and} \tag{6}$$
$$\mathbf{R}_y(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}.$$

Let $\mathbf{F}$ be the fundamental matrix relating $\mathbf{P}_1$ and $\mathbf{P}_2$. From (5) and (6), it is easy to see that

$$\mathbf{F} = \begin{bmatrix} 0 & \cos\theta - 1 & 0 \\ \cos\theta - 1 & 0 & \sin\theta \\ 0 & -\sin\theta & 0 \end{bmatrix} \tag{7}$$
$$= -\sin\theta \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}_{\times} + (\cos\theta - 1) \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \right). \tag{8}$$

Let now $\mathbf{U}_X, \mathbf{U}_Y$ and $\mathbf{U}_Z$ be the points at infinity in the x, y and z direction, respectively, in world coordinates. Projecting these points using the camera $\mathbf{P}_1$, we obtain $\mathbf{u}_x, \mathbf{u}_y$ and $\mathbf{u}_z$ given by

$$\mathbf{u}_x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{u}_y = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } \mathbf{u}_z = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{9}$$

The image of the horizon is the line $\mathbf{q}_\mathrm{h}$, and the image of the screw axis is the line $\mathbf{q}_\mathrm{s}$, where

$$\mathbf{q}_\mathrm{s} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \text{ and } \mathbf{q}_\mathrm{h} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \tag{10}$$

Substituting (9) and (10) in (8), the desired parameterization is obtained:

$$\mathbf{F} = -\sin\theta \left[[\mathbf{u}_x]_\times + \tan\frac{\theta}{2}(\mathbf{q}_\mathrm{s}\mathbf{q}_\mathrm{h}^\mathrm{T} + \mathbf{q}_\mathrm{h}\mathbf{q}_\mathrm{s}^\mathrm{T}) \right]. \tag{11}$$

The factor "$-\sin\theta$" can be eliminated since the fundamental matrix is defined only up to an arbitrary scale. Assume now that the cameras $\mathbf{P}_1$ and $\mathbf{P}_2$ are transformed by a rotation $\mathbf{R}$ about their optical centers and the introduction of a set of intrinsic parameters represented by the matrix $\mathbf{K}$. The new pair of cameras, $\hat{\mathbf{P}}_1$ and $\hat{\mathbf{P}}_2$, is related to $\mathbf{P}_1$ and $\mathbf{P}_2$ by

$$\begin{aligned} \hat{\mathbf{P}}_1 &= \mathbf{H}\mathbf{P}_1 \text{ and} \\ \hat{\mathbf{P}}_2 &= \mathbf{H}\mathbf{P}_2, \end{aligned} \tag{12}$$

where $\mathbf{H} = \mathbf{K}\mathbf{R}$. The fundamental matrix $\hat{\mathbf{F}}$ of the new pair of cameras $\hat{\mathbf{P}}_1$ and $\hat{\mathbf{P}}_2$ is given by

$$\begin{aligned} \hat{\mathbf{F}} &= \mathbf{H}^{-\mathrm{T}}\mathbf{F}\mathbf{H}^{-1} \\ &= \det(\mathbf{H})[\mathbf{v}_x]_\times + \tan\frac{\theta}{2}(\mathbf{l}_\mathrm{s}\mathbf{l}_\mathrm{h}^\mathrm{T} + \mathbf{l}_\mathrm{h}\mathbf{l}_\mathrm{s}^\mathrm{T}), \end{aligned} \tag{13}$$

where $\mathbf{v}_x = \mathbf{H}\mathbf{u}_x$, $\mathbf{l}_\mathrm{h} = \mathbf{H}^{-\mathrm{T}}\mathbf{q}_\mathrm{h}$ and $\mathbf{l}_\mathrm{s} = \mathbf{H}^{-\mathrm{T}}\mathbf{q}_\mathrm{s}$.

3.2 Parameterization via Planar Harmonic Homology

The epipole $\mathbf{e}'$ in the image obtained from the camera $\mathbf{P}_2$ in (5) is given by

$$\mathbf{e}' = \mathbf{u}_x - \tan\frac{\theta}{2}\mathbf{u}_z, \tag{14}$$

which can be obtained from (5). The planar harmonic homography $\mathbf{T}$ relating the symmetric elements in the stereo camera system $\mathbf{P}_1$ and $\mathbf{P}_2$ (e.g. epipoles and pencils of epipolar lines) can be parameterized as

$$\mathbf{T} = \mathbb{I} - 2\frac{\mathbf{u}_x\mathbf{q}_\mathrm{s}^\mathrm{T}}{\mathbf{u}_x^\mathrm{T}\mathbf{q}_\mathrm{s}}. \tag{15}$$

Direct substitution of (14) and (15) in (11) shows that the fundamental matrix can be parameterized by $\mathbf{e}'$ and $\mathbf{T}$ as:

$$\mathbf{F} = [\mathbf{e}']_\times \mathbf{T}. \tag{16}$$

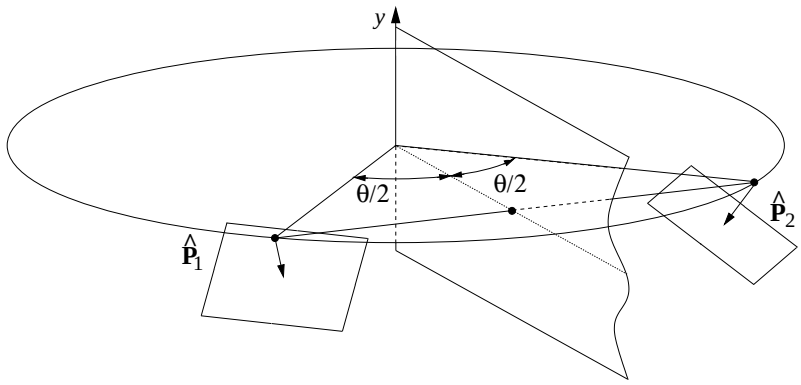


Fig. 3. The harmonic homology is a homography induced by the plane that contains the axis of rotation and bisects the segment joining the camera centers.

Again, it is easy to show that the result does not depend on the transformation $\mathbf{H}$, and the general result becomes

$$\hat{\mathbf{F}} = [\hat{\mathbf{e}}']_\times \mathbf{W}, \text{ with } \hat{\mathbf{e}}' = \mathbf{v}_x - \tan\frac{\theta}{2}\mathbf{v}_z. \tag{17}$$

Thus, we have proved that the transformation $\mathbf{W}$ corresponds to a plane induced homography (see [9]). This means that the registration of the images can be done by using $\mathbf{W}$ instead of a planar contour as proposed in [1, 6]. It is known that different choices of the plane that induces the homography in a *plane plus parallax* parameterization of the fundamental matrix will result in different homographies, although they will all generate the same fundamental matrix, since

$$\hat{\mathbf{F}} = [\hat{\mathbf{e}}']_\times \mathbf{W} = [\hat{\mathbf{e}}']_\times [\mathbf{W} + \hat{\mathbf{e}}'\mathbf{a}^{\mathrm{T}}] \; \forall \mathbf{a} \in \mathbb{R}^3. \tag{18}$$

The three parameter family of homographies $[\mathbf{W} + \hat{\mathbf{e}}'\mathbf{a}^{\mathrm{T}}]$ has a one to one correspondence with the set of planes in $\mathbb{R}^3$. In particular, the homology $\mathbf{W}$ relating the cameras $\hat{\mathbf{P}}_1$ and $\hat{\mathbf{P}}_2$ is induced by a plane Ξ that contains the axis of rotation y and bisects the segment joining the optical centers of the cameras, as shown in Figure 3.

4 Algorithms for Motion Recovery

4.1 Estimation of the Harmonic Homology

Consider an object that undergoes a full rotation around a fixed axis. The envelope ϵ of the profiles is found by overlapping the image sequence and applying a Canny edge detector to the resultant image (Figure 4(b)). The homography $\mathbf{W}$ is then found by sampling N points $\mathbf{x}_i$ along ϵ and optimizing the cost function

$$f_{\mathbf{W}}(\mathbf{v}_x, \mathbf{l}_s) = \sum_{i=1}^{N} \mathrm{dist}(\epsilon, \mathbf{W}(\mathbf{v}_x, \mathbf{l}_s)\mathbf{x}_i)^2, \tag{19}$$

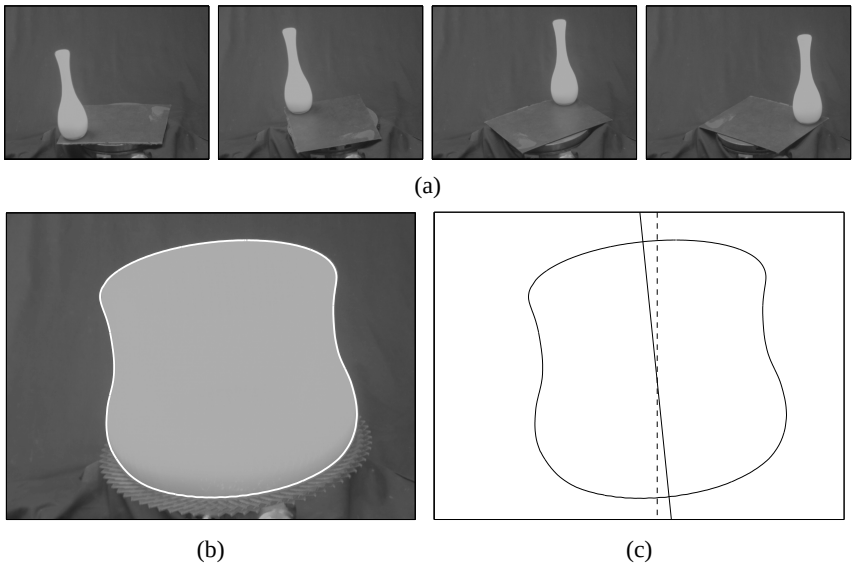


Fig. 4. (a) Image 1, 8, 15 and 22 in the sequence of 36 images of a rotating vase. (b) Envelope of apparent contours produced by overlapping all images in the sequence. (c) Initial guess (dashed line) and final estimation (solid line) of the image of the rotation axis.

Algorithm 1 Estimation of the harmonic homology $\mathbf{W}$.

overlap the images in sequence;
extract the envelope ϵ of the profiles using Canny edge detector;
sample N points $\mathbf{x}_i$ along ϵ;
initialize the axis of symmetry $\mathbf{l}_s$ and the vanishing point $\mathbf{v}_x$;
while not converged **do**
  transfer the points $\mathbf{x}_i$ using $\mathbf{W}$;
  compute the distances between ϵ and the transferred points;
  update $\mathbf{l}_s$ and $\mathbf{v}_x$ to minimize the function in (19);
end while

where $\mathrm{dist}(\epsilon, \mathbf{W}(\mathbf{v}_x, \mathbf{l}_s)\mathbf{x}_i)$ is the distance between the curve ϵ and the transformed sample point $\mathbf{W}(\mathbf{v}_x, \mathbf{l}_s)\mathbf{x}_i$.

The initialization of the line $\mathbf{l}_s$ is trivial, and can be made simply by picking a coarse approximation for the axis of symmetry of ϵ. This can be done via user intervention or by automatically locating one or more pairs of corresponding bitangents. In all practical situations, the camera should be roughly pointing towards the rotation axis, which means that the point $\mathbf{v}_x$ is far (or even at infinity) and at a direction orthogonal to $\mathbf{l}_s$. The estimation of $\mathbf{W}$ is summarized in Algorithm 1.

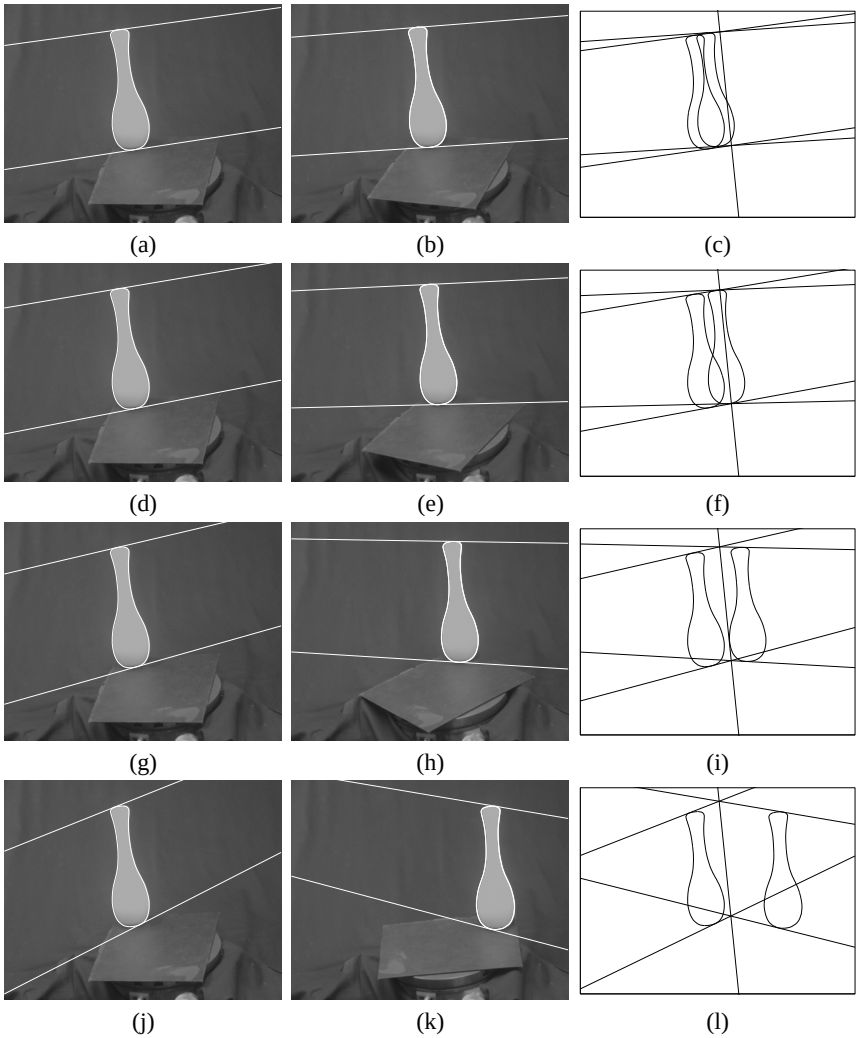


Fig. 5. Five images from a single camera and circular motion after a rotation of 10°, 20°, 40° and 80° are shown in (b), (e), (h) and (k), and the base image at 0° can be seen in (a), (d), (g), (g). The epipolar geometry between image pairs is shown. The overlapping of corresponding pairs can be seen in (c), (f), (i) and (l). Corresponding epipolar lines intersect at the image of the rotation axis, and all epipoles lie on a common horizon.

4.2 Estimation of the Epipoles

After obtaining a good estimation of $\mathbf{W}$, one can then search for *epipolar tangencies* between pairs of images in the sequence. Epipolar tangencies are important for motion estimation from profiles since they are the only correspondences that can be established between image pairs [2]. To obtain a pair of corresponding epipolar tangencies in two

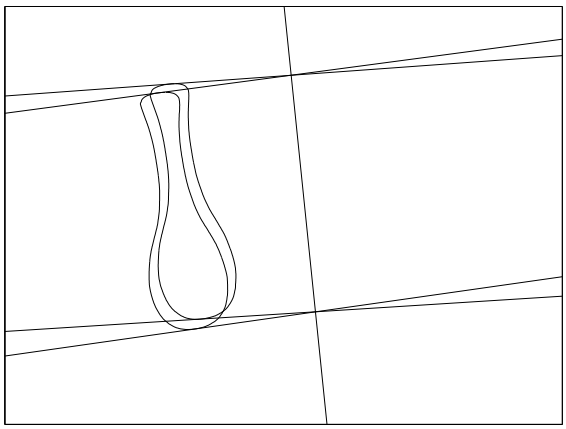


Fig. 6. Corresponding pairs of epipolar tangencies near the top and bottom of two images.

Algorithm 2 Estimation of the orientation of the epipolar lines.

extract the profiles of two adjacent images using Canny edge detector;
fit b-splines to the top and the bottom of the profiles;
initialize α;
while not converge **do**
  find $\mathbf{l}$, $\mathbf{l}'$ and $\mathbf{l}'_{\|}$;
  compute the distance between $\mathbf{l}'$ and $\mathbf{l}'_{\|}$;
  update α to minimize the function in (20);
end while

images, it is necessary to find a line tangent to one profile which is transferred by $\mathbf{W}^{-\mathrm{T}}$ to a line tangent to the profile in the other image (see Figure 6). The search for corresponding tangent lines may be carried out as a one-dimensional optimization problem. The single parameter is the angle α that defines the orientation of the epipolar line $\mathbf{l}$ in the first image, and the cost function is given by

$$f_\alpha = \mathrm{dist}(\mathbf{W}^{-T}\mathbf{l}(\alpha), \mathbf{l}'_{\|}(\alpha)), \tag{20}$$

where $\mathrm{dist}(\mathbf{W}^{-T}\mathbf{l}(\alpha), \mathbf{l}'_{\|}(\alpha))$ is the distance between the transferred line $\mathbf{l}' = \mathbf{W}^{-T}\mathbf{l}$ and a parallel line $\mathbf{l}'_{\|}$ tangent to the profile in the second image. Typical values of α lie between -0.5 rad and 0.5 rad, or -30° and 30°.

Given a pair of epipolar lines near the top and the bottom of a profile, the epipole can be computed as the intersection point of the two epipolar lines, and the fundamental matrix relating the two cameras follows from (17). Using the camera calibration matrix obtained either from a calibration grid or from self-calibration techniques, the essential matrix can be found. The decomposition of the essential matrix gives the relative motion between two cameras.

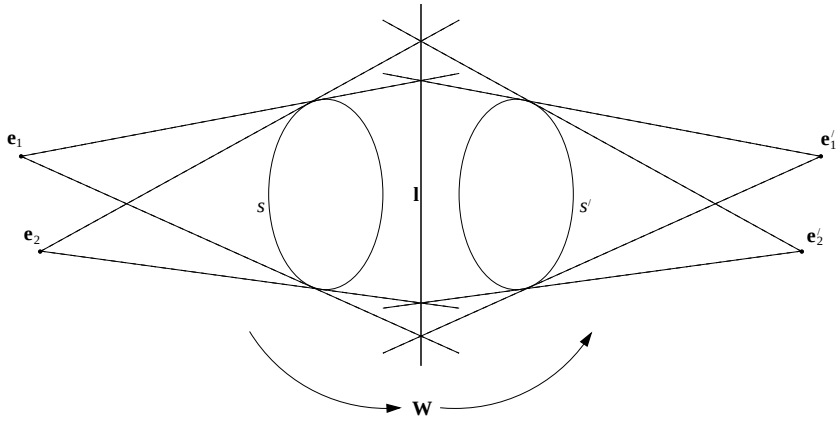


Fig. 7. If the apparent contours are related by the homography $\mathbf{W}$, there will be multiple solutions for the positions of the epipoles. Both pairs $(\mathbf{e}_1, \mathbf{e}_1{}')$ and $(\mathbf{e}_2, \mathbf{e}_2{}')$ are valid epipoles, consistent with the transformation $\mathbf{W}$ (and thus with $\mathbf{l}$) and the contours s and s'.

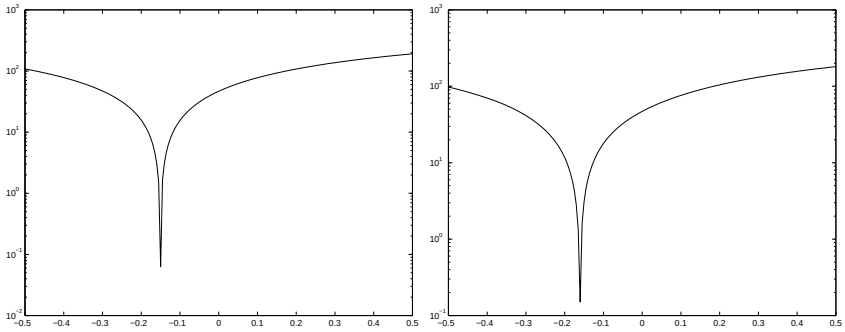
4.3 Critical Configurations

There is a configuration where the algorithm described in Algorithm 2 fails. Let $\mathcal{N}_\mathbf{t}$ and $\mathcal{N}'_\mathbf{t}$ be subsets of two adjacent apparent contours, with $\mathcal{N}_\mathbf{t}$ and $\mathcal{N}'_\mathbf{t}$ related by the homography $\mathbf{W}$ found in Algorithm 1. Any value of α in Algorithm 2 such that the resulting epipolar tangencies are in $\mathcal{N}_\mathbf{t}$ and $\mathcal{N}'_\mathbf{t}$ will minimize the cost function in (19). The proof follows from observing that if α is the orientation of a putative epipolar line with corresponding epipolar tangency in $\mathcal{N}_\mathbf{t}$ in the first contour, the mapping of the epipolar line tangency via $\mathbf{W}$, as required by Algorithm 2, will result in a line tangent to the second contour, as shown in Figure 7. To overcome this problem it is enough then to choose another contour as the first one of the pair where the problem appeared, and proceed with the algorithm.

The ultimate degenerate configuration occurs when the surface being viewed is a surface of revolution (if not completely, at least in the neighbourhood of the frontier points), and the axis of rotation of the turntable is coincident with the axis of rotation of the surface (or the axis of rotation of the rotationally symmetric neighbourhoods). In this case, all the contours are the same, since the contour generator is a fixed curve in space, and the substitution of one contour for another will not make any difference.

5 Implementation and Experimental Results

The algorithms described in the previous session were tested using a set of 36 images of a vase placed on a turntable (see Figure 4(a)) rotated by an angle of 10° between successive snapshots. To obtain $\mathbf{W}$, the Algorithm 1 was implemented with 40 evenly spaced sample points along the envelope ($N = 40$). An approximation for the image of the rotation axis was manually picked by observing the symmetry of the envelope.



(a) (b)

Fig. 8. Plot of the cost function (20) for a pair of images in the sequence. (a)/(b) Cost function for a pair of corresponding epipolar tangencies near the top/bottom of the profile.

Algorithm 3 Motion estimation.

```
estimate motion between IMAGE(1) and IMAGE(2);
update the direction of the axis of rotation;
for i = 3 TO END do
   j = i - 1;
   while motion is bad do
      estimate motion between IMAGE(j) and IMAGE(i);
      j = j - 1;
   end while
   update the direction of the axis of rotation;
end for
```

This provided an initial guess for $\mathbf{l}_s$. The vanishing point $\mathbf{v}_x$ was initialized at infinity, at a direction orthogonal to $\mathbf{l}_s$. The cost function (19) was minimized using the BFGS algorithm [10]. The initial and final configurations can be seen in Figure 4(c).

For the estimation of the motion, the Algorithm 2 was applied for pairs of images to obtain the essential matrix $\mathbf{E}$. The camera calibration matrix was obtained using a calibration grid. The cost function in (20) was minimized using the Golden Section method. This optimization problem is rather simple since the cost function is smooth and unimodal (see Figure 8).

The direction of the axis of rotation was initialized as that obtained from the first pair of images. The quality of each subsequent estimation was checked by comparing the direction of the rotation axis computed from the current pair with the average direction found for all the previous pairs. If the deviation was greater than $10°$, the motion was estimated by using a different combination of images (see Algorithm 3). Such process of quality control is completely automatic.

The remaining problem was to fix the ratio of the norm of the relative translations. Since the camera is performing circular motion, it is easy to show that the relative translations are proportional to $\sin\theta/2$, where θ is the angle of the relative rotation

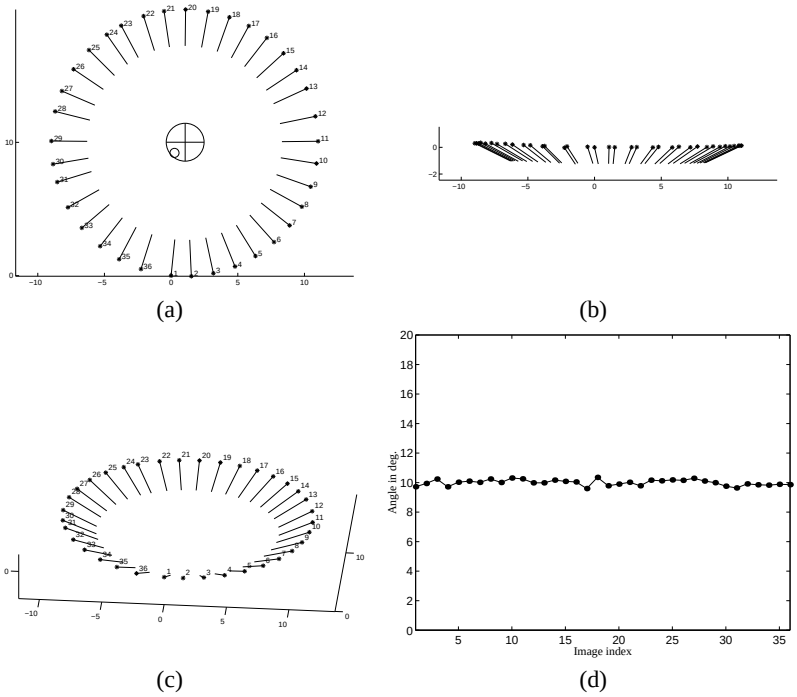


Fig. 9. (a-c) Final configuration of the estimated motion of the cameras. (d) Estimated angles of rotation.

between the two cameras. The resulting camera configurations are presented in Figure 9(a-c). The estimated relative angles between adjacent cameras are accurate, as shown in fig 9(d) and the camera centers are virtually on the same plane and the motion closely follows a circular path.

6 Conclusions and Future Work

This paper introduces a new method of motion estimation by using profiles of a rotating object. No affine approximation has been used and only minimal information (two epipolar tangencies) is required, as long as the object performs a complete rotation. This means that the algorithm can be applied in any practical situation involving circular motion. If more information is available, the estimation problem will be more constrained, and numerical results can be further improved. By proceeding in a divide-and-conquer approach, the difficulties due to initialization and presence of local minima are overcome. The search space in the main loop of the algorithm is one-dimensional, making the technique highly efficient.

Some ideas can be explored to further improve the results presented in this work. A promising approach is to make simultaneous use of the parameterizations shown

in (11) and (17). After estimating the position of the epipoles using Algorithm 2, the horizon line can be found by fitting a line $\mathbf{l}_{\mathrm{h}}$ to the epipoles, such that $\mathbf{l}_{\mathrm{h}}^{\mathrm{T}}\mathbf{v}_x = 0$. This should be done by using a robust method, such as Hough transform or RANSAC. Then, Algorithm 2 can be run again, now with the constraint that all the epipoles must lie on the horizon line. This procedure constrains the cameras to exactly follow a circular path, and integrates information from all images in the estimation of the horizon. This approach has already been proved to produce more accurate results, allowing for high quality reconstructions [12].

Appendix A: Bilateral Symmetry of Images of Surfaces of Revolution

Let S be the surface of revolution parameterized as

$$S = \{\mathbf{S}(\tau,\phi) = [f(\tau)\sin\phi \quad g(\tau) \quad -f(\tau)\cos\phi]^{\mathrm{T}}, (\tau,\phi) \in I_\tau \times I_\phi\}, \tag{21}$$

where $f : \mathbb{R} \supset I_\tau \mapsto \mathbb{R}$ is a differentiable map for which $\exists a > 0$ such that $0 < f(\tau) < a\ \forall\tau \in I_\tau$, and $g : \mathbb{R} \supset I_\tau \mapsto \mathbb{R}$ is a differentiable map for which $\exists b, c$ such that $b < g(\tau) < c\ \forall\tau \in I_\tau$. Also, $\dot{f}^2 + \dot{g}^2 > 0$, where $\dot{f}$ and $\dot{g}$ are the derivatives of the maps f and g. The normal vector at the point $\mathbf{S}(\tau,\phi)$ is given by $\mathbf{n} = \mathbf{S}_\phi \times \mathbf{S}_\tau = f(\tau)[-\dot{g}\sin\phi \quad \dot{f} \quad \dot{g}\cos\phi]^{\mathrm{T}}$, where $\mathbf{S}_\chi$ is the partial derivative of $\mathbf{S}$ with respect to the variable χ. Let $\mathbf{P} = [\mathbb{I}\,|\mathbf{t}\,]$ be the matrix of a pinhole camera, with $\mathbf{t} = [0\ 0\ \alpha]^{\mathrm{T}}$ and $\alpha > a$.

The profile s of S obtained from $\mathbf{P}$ is the projection of the set of points of S where $(\mathbf{S}(\tau,\phi)+\mathbf{t})\cdot\mathbf{n} = 0$. This constraint can be expressed as $g(\tau)\dot{f} - \dot{g}f(\tau) + \alpha\dot{g}\cos\phi = 0$, and for $\tau \in I_\tau$ such that $\dot{g}(\tau) \neq 0$ the resulting expression for $\mathbf{s} \in s$ after removing the dependence on ϕ is given by

$$\mathbf{s}(\tau) = \begin{bmatrix} \pm\frac{f\sqrt{(\alpha\dot{g})^2 - (\dot{g}f - g\dot{f})^2}}{\alpha^2\dot{g} - f(\dot{g}f - g\dot{f})} \\ \frac{g}{\alpha^2\dot{g} - f(\dot{g}f - g\dot{f})} \end{bmatrix} \tag{22}$$

$\forall\tau$ such that $|(\dot{g}f - g\dot{f})/(\alpha\dot{g})| < 1$. Observe that this condition implies that $\alpha^2\dot{g} - f(\dot{g}f - g\dot{f}) \neq 0$, otherwise one would have $|(\dot{g}f - g\dot{f})/(\alpha\dot{g})| = |\alpha/f| > 1$. From (22), one can see that the profile s is bilaterally symmetric about the line $\mathbf{q}_{\mathrm{s}} = [1\ 0\ 0]^{\mathrm{T}}$ (observe the sign "$\pm$").

Acknowledgements

Paulo R. S. Mendonça gratefully acknowledges the financial support of CAPES, Brazilian Ministry of Education, grant BEX1165/96-8. Roberto Cipolla acknowledges the support of the EPSRC and EC project Vigor.

References

1. K. Åström, R. Cipolla, and P. J. Giblin. Generalised epipolar constraints. In B. F. Buxton and R. Cipolla, editors, *Proc. 4th European Conf. on Computer Vision*, volume II, pages 97–108. Springer–Verlag, 1996.
2. R. Cipolla, K. Åström, and P. J. Giblin. Motion from the frontier of curved surfaces. In *Proc. 5th Int. Conf. on Computer Vision*, pages 269–275, 1995.
3. R. Cipolla and A. Blake. Surface shape from the deformation of apparent contours. *Int. Journal of Computer Vision*, 9(2):83–112, 1992.
4. R. Cipolla and P. J. Giblin. *Visual Motion of Curves and Surfaces*. Cambridge University Press, Cambridge, 1999.
5. H. S. M. Coxeter. *Introduction to Geometry*. John Wiley and Sons, New York, second edition, 1969.
6. G. Cross, A. Fitzgibbon, and A. Zisserman. Parallax geometry of smooth surfaces in multiple views. In *Proc. 7th Int. Conf. on Computer Vision*, volume I, pages 323–329, 1999.
7. A. W. Fitzgibbon, G. Cross, and A. Zisserman. Automatic 3D model construction for turn-table sequences. In *3D Structure from Multiple Images of Large-Scale Environments, European Workshop SMILE'98*, Lecture Notes in Computer Science 1506, pages 155–170, 1998.
8. P. J. Giblin, F. E. Pollick, and J. E. Rycroft. Recovery of an unknown axis or rotation from the profiles of a rotating surface. *J. Opt. Soc. America A*, 11:1976–1984, 1994.
9. R. Hartley. Projective reconstruction and invariants from multiple images. *IEEE Trans. Pattern Analysis and Machine Intell.*, 16(10):1036–1041, 1994.
10. D. G. Luenberger. *Linear and Nonlinear Programming*. Addison-Wesley, USA, second edition, 1984.
11. P. R. S. Mendonça and R. Cipolla. Estimation of epipolar geometry from apparent contours: Affine and circular motion cases. In *Proc. Conf. Computer Vision and Pattern Recognition*, volume I, pages 9–14, 1999.
12. P. R. S. Mendonça, K-Y. K. Wong, and R. Cipolla. Camera pose estimation and reconstruction from image profiles under circular motion. In *Proc. 6th European Conf. on Computer Vision*, Dublin, Ireland, 2000. Springer–Verlag.
13. V. S. Nalwa. Line-drawing interpretation: Bilateral symmetry. *IEEE Trans. Pattern Analysis and Machine Intell.*, 11(10):1117–1120, 1989.
14. J. Porrill and S. B. Pollard. Curve matching and stereo calibration. *Image and Vision Computing*, 9(1):45–50, 1991.
15. J. H. Rieger. Three dimensional motion from fixed points of a deforming profile curve. *Optics Letters*, 11:123–125, 1986.
16. J. G. Semple and G. T. Kneebone. *Algebraic Projective Geometry*. Oxford University Press, 1952.
17. T. Vieville and D. Lingrand. Using singular displacements for uncalibrated monocular visual systems. In *Proc. 4th European Conf. on Computer Vision*, volume II, pages 207–216, 1996.
18. A. Zisserman, D. Forsyth, J. Mundy, and C. A. Rothwell. Recognizing general curved objects efficiently. In J.L Mundy and A. Zisserman, editors, *Geometric Invariance in Computer Vision*, chapter 11, pages 228–251. MIT Press, Cambridge, Mass., 1992.
19. A. Zisserman, J. L. Mundy, D. A. Forsyth, J. Liu, N. Pillow, C. Rothwell, and S. Utcke. Class-based grouping in perspective images. In *Proc. 5th Int. Conf. on Computer Vision*, pages 183–188, 1995.