Epipolar Geometry from Profiles under Circular Motion

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

Abstract—This paper addresses the problem of motion estimation from profiles (also known as apparent contours) of an object rotating on a turntable in front of a single camera. Its main contribution is the development of a practical and accurate technique for solving this problem from profiles alone, which is precise enough to allow for the reconstruction of the shape of the object. No correspondences between points or lines are necessary, although the method proposed can be used equally when these features are available without any further adaptation. Symmetry properties of the surface of revolution swept out by the rotating object are exploited to obtain the image of the rotation axis and the homography relating epipolar lines in two views in a robust and elegant way. These, together with geometric constraints for images of rotating objects, are then used to obtain first the image of the horizon, which is the projection of the plane that contains the camera centers, and then the epipoles, thus fully determining the epipolar geometry of the image sequence. The estimation of the epipolar geometry by this sequential approach (image of rotation axis—homography—image of the horizon—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, is carried out as a simple one-dimensional optimization problem. The initialization of the parameters is trivial and completely automatic for all stages of the algorithm. After the estimation of the epipolar geometry, the Euclidean motion is recovered using the fixed intrinsic parameters of the camera obtained either from a calibration grid or from self-calibration techniques. Finally, the spinning object is reconstructed from its profiles using the motion estimated in the previous stage. Results from real data are presented, demonstrating the efficiency and usefulness of the proposed methods.

Index Terms—Structure and motion, epipolar geometry, profiles, apparent contours, circular motion.

1 INTRODUCTION

Methods for motion estimation and 3D reconstruction from point or line correspondences in a sequence of images have achieved a high level of sophistication with impressive results [35], [21], [14]. Nevertheless, if corresponding points are not available, the current techniques cannot be applied. That is exactly the case when the scene being viewed is composed by nontextured smooth surfaces, and, in this situation, the dominant feature in the image is the profile or apparent contour of the surface [22]. Besides, even when point correspondences can be established, the profile still offers important clues for determining both motion and shape and, therefore, it should be used whenever available.

The first attempts to estimate motion from profiles date back to Rieger, in 1986 [30], who introduced the concept of a frontier point, interpreted as “centers of spin” [sic] of the image motion. The paper dealt with the case of fronto-parallel orthographic projection. This idea was further developed by Porrill and Pollard [29], who recognized the frontier point as a fixed point on the surface, corresponding to the intersection of two consecutive contour generators [7]. The connection between the epipolar geometry and the frontier points was established in [17] and an algorithm for motion estimation from profiles under circular motion was introduced in [6]. Related works also include [2], where a technique based on registering the images using a planar curve was first developed. This method was implemented in [11], which also showed results of reconstruction from the estimated motion. The work in [19] presents a method where the affine approximation is used to bootstrap the full projective case.

This work presents a method for estimating the motion of an object rotating around a fixed axis from information provided by its profiles alone. It makes use of symmetry properties [41], [12], [15] of the surface of revolution swept out by the rotating object to overcome the main difficulties and drawbacks present in other methods which have attempted to estimate motion from profiles, namely, 1) the need for a very good initialization for the epipolar geometry and an unrealistic demand for a large number of epipolar tangencies [6], [2], [1] (here, as few as two epipolar tangencies are needed), 2) restriction to linear motion [31] (whereas circular motion is a more practical situation), or 3) the use of an affine approximation [39] (which may be used only for shallow scenes).

An interesting comparison can be made between the work presented here and [14]. Both papers tackle the same problem, but while in [14] hundreds of points were tracked and matched for each pair of adjacent images, it will be shown here that a solution can be obtained even when only two epipolar tangencies are available with at least comparable results.
1.1 Outline of the Paper

This paper begins by describing a method to obtain the image of the rotation axis and the coordinates of a special vanishing point, used in the parameterization of the fundamental matrix under circular motion, from symmetry properties of the profile of the surface of revolution swept out by an object placed on a turntable. These provide the homography component of the fundamental matrix in a plane plus parallax [9], [36] representation. The epipolar constraint is then used to estimate the epipoles for each pair of images in the sequence. These epipoles should be collinear and the line containing them corresponds to the horizon. Due to noise, this alignment will not be verified and a line is fitted to the cloud of epipoles to provide an estimate for the horizon. Once this estimate is available, the epipolar constraint is again employed to recompute the epipoles with a minimal parameterization specialized to circular motion [38]. The epipoles are now constrained to lie on the horizon, providing an accurate estimate for the epipolar geometry of each pair of images in the sequence. Intrinsic parameters, either computed from a self-calibration algorithm or precomputed by any standard calibration technique, can then be used together with the fundamental matrices to determine the camera motion.

Section 2 reviews the symmetry properties exhibited by the image of a surface of revolution. Section 3 establishes the relationship between this transformation and the epipolar geometry and also presents two useful parameterizations of the fundamental matrix. These parameterizations allow the estimation of the epipoles to be carried out as independent one-dimensional searches, avoiding local minima and greatly reducing the computational complexity of the estimation. Section 4 presents the algorithm for motion recovery and the implementation of the algorithm for real data is shown in Section 5, which also makes comparisons with previous works. Experimental results using the estimated motion for reconstruction are shown in Section 6.

2 Symmetry in the Image of a Surface of Revolution

An object rotating around a fixed axis sweeps out a surface of revolution [15]. Symmetry properties 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. 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 v and the set of points on a line l fixed is called a perspective collineation with center (or pole) v and axis l. A homology is a perspective collineation whose center and axis are not incident (otherwise, the perspective collineation is called an elation). Let a be a point mapped by a homology onto a point a'. It is easy to show that the center of the homology, v, and the points a and a' are collinear. Let qa be the line passing through these points and va be the intersection of qa and the axis l. If a and a' are harmonic conjugates with respect to v and va, i.e., their cross-ratio is one, the homology is said to be a harmonic homology (see details in [32], [10] and also Fig. 1a). The matrix W representing a harmonic homology with center v and axis l in homogeneous coordinates is given by

$$W = I - 2 \frac{vI^T}{vI}.$$  (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.

The profile of a surface of revolution exhibits a special symmetry property, which can be described by a harmonic homology [24]. The next theorem gives a formal definition for this property.

**Theorem 1.** The profile of a surface of revolution S viewed by a pinhole camera is invariant to the harmonic homology with an axis given by the image of the rotation axis 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 rotation axis and the camera center.

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

**Lemma 1.** Let T : \(\Gamma' \rightarrow \Gamma\) be a harmonic homology with axis \(\Gamma\) and center \(v\) on the plane \(\Gamma'\) and let \(H : \Gamma' \rightarrow \Gamma\) be a bijective 2D homography between the planes \(\Gamma'\) and \(\Gamma\). Then, the transformation \(W = HTH^{-1}\), \(\Gamma' \rightarrow \Gamma\) is a harmonic homology with axis \(I = H^{-1}I'\) and center \(v = Hv'\).
Proof. Since $H$ is bijective, $H^{-1}$ exists. Then,

$$W = H \left( I - 2 \frac{v^T}{v^T v} \right) H^{-1}$$

$$= I - 2 \frac{v^T}{v^T v}$$

since $v^T v = v^T v$. □

The following corollary is a trivial consequence of Lemma 1.

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

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

Without loss of generality, assume that the axis of rotation of the surface of revolution $S$ is coincident with the $y$-axis of a right-handed orthogonal coordinate system. Considering a particular case of Theorem 1 in which the pinhole camera $P$ is given by $P = [I|t]$, where $t = [0 0 \alpha]^T$, for any $\alpha > 0$, symmetry considerations show that the profile $s$ of $S$ will be bilaterally symmetric with respect to the image of the $y$-axis [28], [26], which corresponds to the line $q_\alpha = [1 0 0]^T$ in (homogeneous) image coordinates.

**Proof of Theorem 1 (particular case).** Since $s$ is bilaterally symmetric about $q_\alpha$, there is a transformation $T$ that maps each point of $s$ onto its symmetric counterpart, given by

$$T = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$ (3)

However, as any bilateral symmetry transformation, $T$ is also a harmonic homology, with axis $q_\alpha$ and center $u_x = [1 0 0]^T$ since

$$T = I - 2 \frac{u_x q_\alpha^T}{u_x q_\alpha}.$$ (4)

The transformation $T$ maps the set $s$ onto itself (although the points of $s$ are not mapped onto themselves by $T$, but onto their symmetric counterparts) and, thus, $s$ is invariant to the harmonic homology $T$. 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 this plane is $U_x = [1 0 0 0]^T$, whose image is $u_x$. □

Let $P$ be an arbitrary pinhole camera. The camera $P$ can be obtained by rotating $\hat{P}$ about its optical center by a rotation $R$ and transforming the image coordinate system of $\hat{P}$ by introducing the intrinsic parameters represented by the calibration matrix $K$ [13]. Let $KR = H$. Thus, $P = H[I|t]$ and the point $u_x$ in space with image $u_x$ in $\hat{P}$ will project to a point $v_x = Hu_x$ in $P$. Analogously, the line $q_\alpha$ in $\hat{P}$ will correspond to a line $l_\alpha = H^{-T}q_\alpha$ in $P$. In the general case, it is now possible to derive the proof of Theorem 1.

**Proof of Theorem 1 (general case).** Let $s$ be the profile of the surface of revolution $S$ obtained from the camera $P$. Thus, the counterdomain of the bijection $H$ acting on the profile $s$ is $s$ (or $Hs = s$), and using Lemma 1, the transformation $W = HTH^{-1}$ is a harmonic homology with center $v_x = Hu_x$ and axis $l_\alpha = H^{-T}q_\alpha$. Moreover, from Corollary 1, $WHs = HTs$ or $Ws = HTs$. From the particular case of Theorem 1, it is known that the profile $s$ will be invariant to the harmonic homology $T$, so $Ws = Hs = s$. □

When the camera is pointing directly toward the axis of rotation, the transformation that maps $s$ onto its symmetric counterpart will be reduced to a skewed symmetry [20], [27], [5], which corresponds to a particular case of the harmonic homology in which the pole is at infinity. It is given by

$$S = \frac{1}{\cos(\theta - \rho)} \begin{bmatrix} -\cos(\theta + \rho) & -2\sin\phi\cos\rho & 2d\cos\phi \\ -2\sin\phi\cos\rho & \cos(\theta + \rho) & 2d\cos\phi \\ 0 & 0 & \cos(\theta - \rho) \end{bmatrix}.$$ (5)

where $l_\alpha = [\cos \rho \sin \rho - d]^T$ is the image of the rotation axis, with $d = w_0 \cos \rho + v_0 \sin \rho$, where $(v_0, w_0)$ is camera principal point. The angle $\phi$ gives the orientation of the lines of symmetry, which are the lines joining each point to its skew-symmetric counterpart (see Fig. 1b). The transformation $S$ has 3 degrees of freedom (dof).

If the camera also has zero skew and aspect ratio 1, the transformation is further reduced to a bilateral symmetry, given by

$$B = \begin{bmatrix} -\cos 2\rho & -\sin 2\rho & 2d\cos\rho \\ -\sin 2\rho & \cos 2\rho & 2d\sin\rho \\ 0 & 0 & 1 \end{bmatrix}.$$ (6)

The transformation now has only 2 dof since the lines of symmetry are orthogonal to $l_\alpha$. A graphical representation of the bilateral symmetry, skewed symmetry, and harmonic homology is shown in Fig. 1.

### 3 Parameterizations of the Fundamental Matrix

#### 3.1 Fundamental Matrix 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 [38], [14], which can be expressed explicitly in terms of fixed image features under circular motion (image of rotation axis, pole, and horizon, jointly holding 5 dof) and the relative angle of rotation (1 dof). A simpler derivation of this result will be shown here. Moreover, a novel parameterization based on the harmonic homology will be introduced, providing a connection between the geometry of the complete sequence (harmonic homology) with the geometry of a single pair of images (fundamental matrix).
Consider the pair of camera matrices $P_1$ and $P_2$, given by
\[
P_1 = [I|t], \quad P_2 = [R_y(\theta)|t],
\]
where
\[
t = [0 \quad 0 \quad 1]^T \quad \text{and} \quad R_y(\theta) = \begin{bmatrix}
\cos \theta & 0 & \sin \theta \\
0 & 1 & 0 \\
-\sin \theta & 0 & \cos \theta
\end{bmatrix}.
\]

for $\theta \neq 0$. Let $\hat{F}$ be the fundamental matrix relating $P_1$ and $P_2$. From (7) and (8), it is easy to see that
\[
\hat{F} = -\sin \theta \begin{bmatrix} 0 & \cos \theta - 1 & 0 \\ \cos \theta - 1 & 0 & \sin \theta \\ 0 & -\sin \theta & 0 \end{bmatrix} = -\sin \theta \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta - 1 & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.
\]

Let $U_x$, $U_y$, and $U_z$ be the points at infinity in the $x$, $y$, and $z$ direction, respectively, in world coordinates. Projecting these points using the camera $P_1$, we obtain the vanishing points $u_x$, $u_y$, and $u_z$ given by
\[
u_x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad u_y = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad u_z = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.
\]

The image of the horizon is the line $q_h$, and the image of the rotation axis is the line $q_e$, where
\[
q_e = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad q_h = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.
\]

Substituting (11) and (12) into (10), the desired parameterization is obtained:
\[
\hat{F} = -\sin \theta \left[ u_x \times + \tan\frac{\theta}{2} (q_e q_h^T + q_h q_e^T) \right].
\]

The factor $-\sin \theta$ can be eliminated since the fundamental matrix is defined only up to an arbitrary scale. Assume now that the cameras $P_1$ and $P_2$ are transformed by a rotation $R$ about their optical centers and the introduction of a set of intrinsic parameters represented by the calibration matrix $K$.

The new pair of cameras, $P_1$ and $P_2$, is related to $P_1$ and $P_2$ by
\[
P_1 = HP_1 \quad \text{and} \quad P_2 = HP_2,
\]
where $H = KR$. The fundamental matrix $F$ of the new pair of cameras $P_1$ and $P_2$ is given by
\[
F = H^{-1}FH^{-1} = \det(H)[v_x]_x + \tan\frac{\theta}{2} (l_h l_h^T + l_e l_e^T),
\]
where $v_x = Hu_x$, $l_h = H^{-1}q_h$, and $l_e = H^{-1}q_e$. Since the fundamental matrix is defined only up to a scale factor, (15) can be rewritten as
\[
F(\theta) = [v_x]_x + \kappa \tan\frac{\theta}{2} (l_h l_h^T + l_e l_e^T),
\]
where $\kappa = 1/\det(H)$. The notation $F(\theta)$ was used in (16) to emphasize that, for a given circular motion sequence, the parameters $v_x$, $l_e$, $l_h$, and $\kappa$ are fixed and the fundamental matrices associated with any pair of cameras in the sequence differs only in the value of $\theta$.

### 3.2 Parameterization via Planar Harmonic Homology

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 $P_1$ and $P_2$. If $P_1$ and $P_2$ point toward the axis of rotation and have zero skew and aspect ratio 1, their epipoles $e_1$ and $e_2$ will be symmetric with respect to the image of the rotation axis, or $e_2 = We_1$. According to Fig. 2. In a general situation, the epipoles will simply be related by the transformation $e_2 = We_1$. It is then straightforward to show that the
corresponding epipolar lines $l_1$ and $l_2$ are related by $l_2 = W^{-1}l_1$. This means that the pair of epipoles can be represented with only two parameters once $W$ is known. From (1), it can be seen that $W$ has only 4 dof.

In [25], [40], it has been shown that any fundamental matrix $F$ can be parameterized as $F = [e_2]_x W$, where, from (16), $e_2 = v_2 - \kappa \tan \frac{\theta}{2} [l_1]_x l_0$. Therefore, $F$ has only 6 dof: Four to determine $W$ and two to fix $e_2$, in agreement with [38]. Note that, in the case of skewed symmetry and bilateral symmetry, the dof of the fundamental matrix will be reduced to five and four, respectively, corresponding to the decrease in the dof of the symmetry transformation. A full account of the dof of the fundamental matrix under different configurations is given in Table 1.

From (17), it can be seen that the transformation $W$ corresponds to a plane induced homography (see [18]). This means that the registration of the images can be done by using $W$ instead of a planar contour as proposed in [2], [11]. It is known that different choices of the plane that induces the homography in a plane plus parallax parameterization of the fundamental matrix, such as the one in (17), will result in different homographies, although they will all generate the same fundamental matrix since

$$F = [e_2]_x W,$$

(17)

where, from (16), $e_2 = v_2 - \kappa \tan \frac{\theta}{2} [l_1]_x l_0$. Therefore, $F$ has only 6 dof: Four to determine $W$ and two to fix $e_2$, in agreement with [38]. Note that, in the case of skewed symmetry and bilateral symmetry, the dof of the fundamental matrix will be reduced to five and four, respectively, corresponding to the decrease in the dof of the symmetry transformation. A full account of the dof of the fundamental matrix under different configurations is given in Table 1.

From (17), it can be seen that the transformation $W$ corresponds to a plane induced homography (see [18]). This means that the registration of the images can be done by using $W$ instead of a planar contour as proposed in [2], [11]. It is known that different choices of the plane that induces the homography in a plane plus parallax parameterization of the fundamental matrix, such as the one in (17), will result in different homographies, although they will all generate the same fundamental matrix since

$$F = [e_2]_x W = [e_2]_x [W + e_2 b^T] \quad \forall b \in \mathbb{R}^3.$$

(18)

The three-parameter family of homographies $[W + e_2 b^T]$ parameterized in $b$ has a one-to-one correspondence with the set of planes in $\mathbb{R}^3$. The particular plane that induces the planar homology $W$ is given in the next theorem:

**Theorem 2.** The planar homology $W$ relating the cameras $P_1$ and $P_2$ is induced by the plane $\mathcal{E}$ that contains the axis of rotation and bisects the segment joining the optical centers of the cameras.

**Proof.** The existence and uniqueness of $\mathcal{E}$ satisfying the hypothesis of the theorem are trivial. Let $x_1 = [1 \ 0 \ 0]^T$, $x_2 = [0 \ 1 \ 0]^T$, and $x_3 = [0 \ 0 \ 1]^T$. Without loss of generality, let

$$P_1 = KR[I \mid x_3] \quad \text{and} \quad P_2 = KR[R_y^\theta \mid x_3],$$

(19)

where $K$ is the matrix of intrinsic parameters of $P_1$ and $P_2$, $R$ is the rotation matrix relating the orientation of the coordinate system of $P_1$ to the world coordinate system, and $R_y^\theta$ is a rotation by $\theta$ about the $y$-axis of the world coordinate system, i.e.,

$$R_y^\theta = \begin{bmatrix}
\cos \theta & 0 & \sin \theta \\
0 & 1 & 0 \\
-\sin \theta & 0 & \cos \theta
\end{bmatrix}.$$

(20)

Therefore, $\forall \alpha, \beta \in \mathbb{R}$, the point

$$X = [-\alpha \sin(\theta/2), \beta, \alpha \cos(\theta/2)]^T$$

lies on $\mathcal{E}$. Projecting $X$ using $P_1$ and $P_2$, one obtains

$$u_1 = KR(X + x_4) \quad \text{and} \quad u_2 = KR(R_y^\theta X + x_3).$$

Since

$$R_y^\theta X = \begin{bmatrix}
\alpha \sin \theta \cos(\theta/2) - \alpha \cos \theta \sin(\theta/2) \\
\beta \\
\alpha \sin \theta \sin(\theta/2) + \alpha \cos \theta \cos(\theta/2)
\end{bmatrix},$$

(21)

$$= \begin{bmatrix}
\alpha \sin(\theta/2) \\
\beta \\
\alpha \cos(\theta/2)
\end{bmatrix} = \begin{bmatrix}
-1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix} X,$$

or $R_y^\theta X = (I - 2x_1x_1^T)X$, we have

$$u_2 = KR[(I - 2x_1x_1^T)X + x_3],$$

or

$$u_2 = (I - 2KRx_1x_1^T R^{-1}K^{-1})u_1.$$

It can be shown [26] that $KRx_1 = v_2$ and $x_1^T R^{-1}K^{-1} = I_3^T$, and, thus, $u_2 = Wu_1$.

A graphical representation of the result in Theorem 2 is shown in Fig. 3.
3.3 Epipolar Geometry and Profiles of Surfaces

So far, the discussion of epipolar geometry and parameterizations of the fundamental matrix under circular motion is applicable to arbitrary image features, such as points or profiles, as it does not consider particular aspects of the issue when the only image features available are profiles. This section briefly reviews some of the main geometricizations of the fundamental matrix under circular motion is applicable to arbitrary image features, such as points or profiles, as it does not consider particular aspects of the issue when the only image features available are profiles. This section briefly reviews some of the main geometric components of epipolar geometry from profiles.

Consider a surface $S$ of type $C^1$ viewed by two pinhole cameras $P_1$ and $P_2$. The following definitions apply [8]:

- A contour generator associated with the surface $S$ and the camera $P_1$ corresponds to the space curve $C \subset S$ such that for all points $C \in C$ the line passing through the optical center of $P_1$ and $C$ is tangent to $S$ at $C$.
- The image of the contour generator associated with the camera $P_1$ is a profile or apparent contour.
- If two contour generators associated with the surface $S$ and the cameras $P_1$ and $P_2$ intersect, the points of intersection are denoted frontier points.
- The epipolar plane $\Pi$ defined by the optical centers of the two cameras $P_1$ and $P_2$ and an associated frontier point $X_i$ is tangent to the surface $S$ at $X_i$.
- The epipolar lines corresponding to the epipolar plane $\Pi$ are tangent to their associated profiles and are called epipolar tangents.

The tangency points of corresponding epipolar tangents are the images of the same point on the surface $S$, namely, the frontier point. The above definitions are illustrated in Fig. 4.

4 Motion Estimation

4.1 Estimation of the Harmonic Homology

Consider an object that undergoes a full rotation around a fixed axis. The envelope $\varepsilon$ of its profiles is found by overlapping the images of the sequence and applying a Canny edge detector [4] to the resultant image. This envelope corresponds to the image of a surface of revolution and, thus, it is harmonically symmetric. The homography $W$ related to $\varepsilon$ is then found by sampling $N$ points $x_i$ along $\varepsilon$ and optimizing the cost function

$$f_W(v_\varepsilon, l) = \sum_{i=1}^{N} \text{dist}(\varepsilon, W(v_\varepsilon, l)x_i)^2,$$

where $\text{dist}(\varepsilon, W(v_\varepsilon, l)x_i)$ is the orthogonal distance between the curve $\varepsilon$ and the transformed sample point $W(v_\varepsilon, l)x_i$. The estimation of $W$ is summarized in Algorithm 1.

\begin{algorithm}
\caption{Estimation of the harmonic homology $W$.}
\begin{algorithmic}
\State overlap the images in the sequence;
\State extract the envelope $\varepsilon$ of the profiles using a Canny edge detector;
\State sample $N$ points $x_i$ along $\varepsilon$;
\State initialize the axis of symmetry $l$ and the vanishing point $v_\varepsilon$ using bitangents;
\While {not converged}
\State transform the points $x_i$ using $W$;
\State compute the distances between $\varepsilon$ and the transformed points $W(v_\varepsilon, l)x_i$;
\State update $l$ and $v_\varepsilon$ to minimize the function in (22);
\EndWhile
\end{algorithmic}
\end{algorithm}

The initialization of the line $l$, and the point $v_\varepsilon$ can be made very close to the global minimum by automatically locating one or more pairs of corresponding bitangents on the envelope. Given two bitangents $l(p_1, p_2)$ and $l(q_1, q_2)$ on the two sides of the profile $\varepsilon$ with bitangent points $p_1$, $p_2$ and $q_1$, $q_2$, respectively (see Fig. 5), the intersection of the two bitangents ($l(p_1, p_2)$, $l(q_1, q_2)$) and the intersection of the diagonals ($l(p_1, q_1)$, $l(q_1, p_2)$) give two points defining a line that can be used as an estimate of $l$. An estimate for the vanishing point $v_\varepsilon$ is given by the point of intersection of the lines $l(p_1, q_1)$ and $l(p_2, q_2)$. The initialization of $l$ and $v_\varepsilon$ 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.

4.2 Estimation of the Horizon

After obtaining a good estimation of $W$, one can then search for epipolar tangencies between pairs of images in the sequence using the parameterization given by (17). To obtain a pair of corresponding epipolar tangents in two images, it is necessary to find a line tangent to one profile which is transformed by $W^{-T}$ to a line tangent to the profile in the other image (see Fig. 6). The search for corresponding tangents may be carried out as a one-dimensional optimization problem. The single parameter is the angle $\delta$ that defines the orientation of the epipolar line $l_1$ in the first image and the cost function is given by

$$f_1(\delta) = \text{dist}(W^{-T}l_1(\delta), l_2(\delta)), \tag{23}$$

where $\text{dist}(W^{-T}l_1(\delta), l_2(\delta))$ is the distance between the transferred line $l_2 = W^{-T}l_1$ and a line $l_2^\parallel$ parallel to $l_2$ and tangent to the profile in the second image. Typical values of $\delta$ lie between $-0.5$ rad and $0.5$ rad or $-30^\circ$ and $30^\circ$. The shape of the cost function (23) for the profiles in Fig. 6 can be seen in Fig. 7.

The epipoles can then be computed as the intersection of epipolar lines in the same image. After obtaining this first estimate for the epipoles, the image of the horizon can then be found by robustly fitting a line $l_h$ to the initial set of epipoles, such that $l_h^T v_x = 0$. Fig. 11 shows a typical output of Algorithm 2, together with the horizon $l_h$ fitted to the epipoles.

**Algorithm 2** Estimation of the horizon.

1. extract the profiles of the images using a Canny edge detector; fit B-splines to the top and bottom of the profiles;
2. for each selected pair of images do
   1. initialize $\delta$;
   2. while not converged do
      1. find $l_1, l_2$ and $l_2^\parallel$;
      2. compute the distance between $l_2$ and $l_2^\parallel$;
      3. update $\delta$ to minimize the function in (23);
   3. end while
   4. compute epipoles;
3. end for
4. fit the horizon $l_h$ to the cloud of epipoles.

An alternative method to compute the epipoles is to register the profiles using the homology $W$, eliminating the effects of rotation on the images and, then, apply any of the methods in [2], [31], [11], in a plane plus parallax approach. However, no advantage has been obtained by doing so since to use this method it is necessary to find a common tangent between two profiles, which involves a search at least as complex as the one in Algorithm 2.

4.3 Estimation of the Epipoles Constrained to the Horizon

After estimating the horizon, the only missing term in the parameterization of the fundamental matrix shown in (16) is the scale factor

$$\lambda = \kappa \tan \theta / 2. \tag{24}$$

This parameter can be found, again, by a one-dimensional search that minimizes the geometric error of transferred
epipolar lines, as shown in Fig. 8. Therefore, two distinct parameterizations of the fundamental matrix are used: Equation (17) to obtain the cloud of epipoles and the horizon and (16) to recompute the position of the epipoles constrained to lie on the horizon.

4.4 Limitations of the Algorithm
There are some limitations on the applicability of the algorithms presented here:

- **Density of the sequence of images.** If the number of images in the sequence is too small or the angle of rotation between successive snapshots is too large, the envelope of the profiles no longer approximates the profile of a surface of revolution and, therefore, Algorithm 1 will fail to correctly estimate the image of the rotation axis and the pole. In practice, this problem does not arise if the angles of rotation in a closed sequence are below $20^\circ$. This problem can be overcome by performing a simultaneous search for the harmonic homology and the rotation angles, at the expense of increasing the number of search parameters and therefore the complexity of the optimization.

- **Symmetry of the object.** If the object placed on the turntable is rotationally symmetric and its axis of symmetry coincides with the axis of rotation of the turntable, Algorithm 2 will fail. To understand this problem, consider the alternative formulation of Algorithm 2 in which the epipoles are computed by first registering the images by using the harmonic homology and then computing the epipoles as the intersection of common tangents to the profiles. Under the conditions described above, the registration of the profiles will not produce any effect since the image of a surface of revolution with the same rotation axis as the turntable is invariant to the harmonic homology. Moreover, the profiles will coincide and any tangent to one of the profiles will be a common tangent to the pair of profiles. Therefore, the position of the epipole will be undetermined. To avoid this problem, it is enough to reposition the symmetric object over the turntable so that its symmetry axis does
not coincide anymore with the turntable axis. The further the two axes are, the better. Of course, the placement of the object must not be so distant from the center of the turntable as to remove it from the field of view. In the experiments shown in this paper using a vase and a head model, which are nearly rotationally symmetric in the regions of interest (the top and the bottom of the objects), it was verified that the problem disappears if the axes are separated by a distance of about 50 pixels.

5 IMPLEMENTATION AND EXPERIMENTAL RESULTS

The algorithms described in the previous section were tested using two sets of images from a vase and a head model, respectively (see Fig. 9). Both sets consisted of 36 images, with the turntable rotated by an angle of 10° between successive snapshots. The fact that the angle was fixed was not used either in the estimation of the epipolar geometry or in the reconstruction to be shown in Section 6. For the vase sequence, the symmetry transformation associated to the envelope of its profiles was assumed to be a harmonic homology $W$, whereas, for the head sequence, the transformation was modeled as a skewed symmetry $S$. The choice of the simpler model for the head sequence was motivated by the fact that the camera was nearly pointing toward the axis of the turntable and, therefore, the skewed symmetry transformation could be used. Of course, there would have been no problem in adopting the more complex model. To obtain $W$ and $S$, Algorithm 1 was implemented with 100 evenly spaced sample points along each envelope ($N = 100$). Initializations were done by using bitangents. Less than 10 iterations of the Levenberg-Marquardt algorithm were necessary, with derivatives computed by finite differences. The final positions of the rotation axes can be seen in Fig. 10.
In the implementation of Algorithm 2, 70 pairs of images were selected by uniformly sampling the indexes of the images in each sequence and the resultant estimate of the epipoles for the vase sequence is shown in Fig. 11, which also shows the horizon $l_h$ found by a robust fit. To get $l_h$, a minimization of the median of the squares of the residuals was used, followed by removal of outliers and orthogonal least-squares regression using the remaining points (inliers). The epipolar geometry was then reestimated with the epipoles constrained to lie on $l_h$. Once the epipolar geometry was obtained, precomputed intrinsic parameters were used to convert the fundamental matrices into essential matrices [13] and these were then decomposed to provide the camera motion and orientation. The resulting camera configurations are presented in Fig. 12.

Fig. 12. Camera configurations for the vase (left) and head (right) sequences.

Fig. 13. Estimated angles of rotation between successive views for the vase (left) and head (right) sequences, with RMS errors 0.19° and 0.23°, respectively.
The object was rotated on a manual turntable with resolution of 0.01°, but the real precision achieved is highly dependent on the skill of the operator. The RMS errors in the estimated angles were 0.19° and 0.23° for the vase and head sequence, respectively (see Fig. 13), demonstrating the accuracy of the estimation.

Fig. 14. Details of the reconstruction of the vase by triangulating the profiles, using the method described in [39]. The model was built using 1,224 triangles.

Fig. 15. Details of the reconstruction of the vase by voxel carving, using the method described in [33], showing the triangular mesh and texture mapping. The model was built using 23,144 triangles.
It is interesting to compare this result with the ones shown in [14, p. 166] for the “Head,” “Freiburg,” and “Dinosaur” sequences, where the average number of point matches per image pair varies from 137 to 399, depending on the sequence. It should be stressed that only two epipolar tangents were used for each pair of images in the experiments presented in this paper, with comparable results.

6 RECONSTRUCTION FROM IMAGE PROFILES

The algorithm for motion estimation introduced here can be used even when point correspondences can be established. On the other hand, methods as the ones in [35], [14], and [21] cannot deal with situations where profiles are the only available features in the scene. Earlier attempts to solve the problem of reconstruction from image profiles under known motion include [16], [37], [7], and state of the art algorithms can be found in [34], [3], [39]. We use a simple method based on triangulation to reconstruct the model using the estimated motion. Examples using voxel-carving [33], [23] are also shown.

Details of the 3D reconstruction of the objects are shown in Fig. 14, Fig. 15, and Fig. 16. Although no ground truth is available for a quantitative evaluation, it can be seen that the reconstructions are faithful to the images of the objects. As reported in [7], errors in the camera orientation of a few mrad can render the reconstruction useless, therefore confirming the accuracy of the technique introduced here.

7 SUMMARY AND CONCLUSIONS

This paper introduced a novel technique for motion estimation from image profiles. It does not make use of expensive search procedures, such as bundle adjustment, although it naturally integrates data from multiple images. The method is mathematically sound, practical, and highly accurate. From the motion estimation to the model reconstruction, no point tracking is required and it does not depend on having point correspondences beforehand.

The convergence to local minima, a critical issue in most nonlinear optimization problems, is avoided by a divide-and-conquer approach which keeps the size of the problem manageable. Moreover, a search space with lower dimension results in fewer iterations before convergence. The quality of model reconstructed is remarkable, in particular, if one considers that only the least possible amount of information has been used. Since then the method has been applied to a large number of sequences, including human heads and sculptures.

7.1 Future Work

An interesting problem is the development of an optimal solution for the problem of structure and motion from profiles—so far, there is no equivalent of a bundle adjustment algorithm capable of dealing with profiles. An important step in filling this gap is to develop a model for the effect that image noise has on the detection of profiles, and only when such error model becomes available can a maximum-likelihood estimator for structure and motion from profiles be developed.

ACKNOWLEDGMENTS

Paulo R.S. Mendonça would like to acknowledge the financial support of CAPES, Brazilian Ministry of Education, grant BEX1165/96-8.

REFERENCES


Kwan-Yee K. Wong received the BEng degree in computer engineering, with first class honors from the Chinese University of Hong Kong in 1998 and the MPhil degree in computer vision (information engineering) from the University of Cambridge in 2000. Currently, he is a PhD student under the supervision of Professor Roberto Cipolla in the Speech, Vision and Robotics (SVR) Group, Department of Engineering, University of Cambridge. His research interests include camera calibration, motion estimation from silhouette and surface reconstruction, and representation.

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. His research interests include stochastic optimization, camera self-calibration, projective geometry, structure and motion from image sequences, and geometry-based computer vision.