Estimation of Epipolar Geometry from Apparent Contours: Affine and Circular Motion Cases

Paulo R. S. Mendonça and Roberto Cipolla
Department of Engineering
University of Cambridge
Trumpington Street, Cambridge, UK, CB2 1PZ
{prdsm2,cipolla}@eng.cam.ac.uk

Abstract

This paper addresses the problem of estimating the epipolar geometry from apparent contours in two special cases: under weak perspective and for circular motion. An appropriate parametrization of the fundamental matrix is introduced for both cases, as well as suitable cost functions for the estimation of the epipoles. The algorithm used in the affine approximation proved to be robust and accurate under several conditions. The circular motion case turned out to be much more difficult, but for a wide baseline the method introduced here is successful. For small viewing angles the technique is too sensitive to noise to be used in practice. Nevertheless, circular motion with small baseline can be well modeled by an affine camera system, and this approximation should be used in this circumstance.

1. Introduction

The recovering of the epipolar geometry of image sequences is the first step towards reconstruction of a 3D model. In the case of point features, correspondences are relatively easy to establish, and the fundamental matrix can be computed by several techniques [11]. When the scene being viewed consists of smooth surfaces, there is no obvious matching of points. In this case the the main feature is the *apparent contour*, which is the contour of the projected image of the surfaces.

2. Theoretical Background

This section summarizes results presented in [5] which are relevant to this work.

Under *known* viewer motion the epipolar parametrization [7, 14, 3], shown in fig. 1, can be used to determine the geometry of the surface in the vicinity of the contour gener-
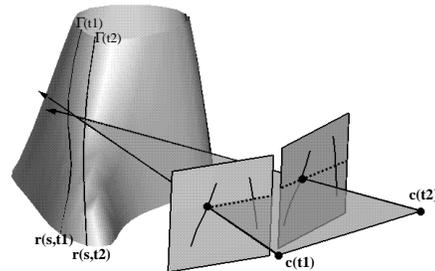

Figure 1. Epipolar parametrisation. The surface is parametrised locally by the contour generators from successive viewpoints and epipolar curves defined by the intersection of the pencil of epipolar planes and the surface.

ation, which is the curve on the surface that projects at the apparent contour.

2.1. Degeneracy of The Epipolar Parametrization

In points where the epipolar plane is tangent to the surface, called *frontier points*, the epipolar parametrization is *degenerate* [10]: the contour generator is locally stationary ($c_t.n = 0$). The frontier points correspond to the intersection of two successive contour generators, and thus are visible in successive images, as depicted in fig. 2. If the correspondence between frontier points could be established, the problem of estimation of the epipolar geometry would turn to be same situation as the one for points. Unfortunately, the determination of the frontier points already depends on the knowledge of the epipolar geometry, and so numerical techniques must be employed.

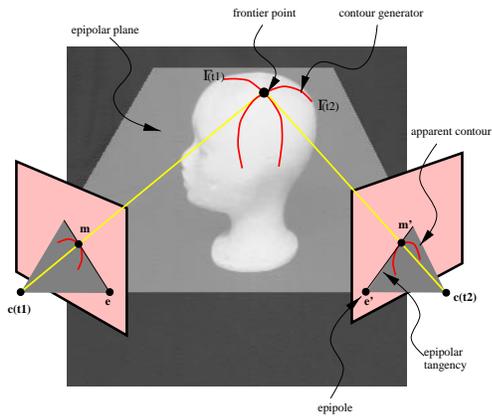Consider the following minimal parametrization of the

Figure 2. A frontier point appears in the intersection of two consecutive contour generators and is visible in both views. The frontier point projects to a point on the apparent contour which is an epipolar tangency point.

fundamental matrix:

$$\mathbf{F} = \begin{bmatrix} \mathbf{H} & \mathbf{Hu} \\ \mathbf{u'}^{\mathrm{T}}\mathbf{H} & \mathbf{u'}^{\mathrm{T}}\mathbf{Hu} \end{bmatrix}, \qquad (1)$$

where

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 \\ h_3 & h_4 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} u \\ v \end{bmatrix}, \mathbf{u'} = \begin{bmatrix} u' \\ v' \end{bmatrix} \qquad (2)$$

with $\mathbf{u}$ and $\mathbf{u'}$ being the coordinates of the epipoles. Given an initial guess for the epipoles, the matrix $\mathbf{H}$ can be estimated from the pencils of epipolar tangencies in both images [6]. This fixes the fundamental matrix under the parametrization above. The distances between the apparent contours and the transfered epipolar lines coming from the corresponding contours in the other image provide a geometric distance criteria [11], as shown in fig. 3, given by

$$C = \sum_i \left(\frac{(\mathbf{m'}_i^T \mathbf{F}\mathbf{m}_i)^2}{(\mathbf{F}\mathbf{m}_i)_1^2 + (\mathbf{F}\mathbf{m}_i)_2^2} + \frac{(\mathbf{m'}_i^T \mathbf{F}\mathbf{m}_i)^2}{(\mathbf{F}^T \mathbf{m'}_i)_1^2 + (\mathbf{F}^T \mathbf{m'}_i)_2^2} \right), \quad (3)$$

to be minimized by searching for the epipoles. Details of the procedure and a similar idea can be found in [2, 5]. The point to be stressed is that the algorithm is quite prone to converge to local minima, and needs at least 8 tangencies in each image, a requirement not satisfied by most images of real scenes.

3. The Affine Case

When the field of view is narrow or the depth variation is small compared with the distance from the camera to the scene, the epipoles will be far from the image centre, and the epipolar lines will be approximately parallel.
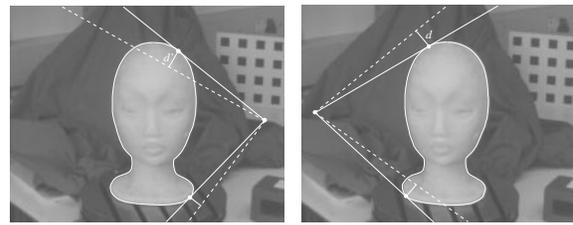

Figure 3. Illustration of the cost function to be minimised in the motion estimation algorithm. From the initial guess of the epipoles the homography is determined, and epipolar tangencies are transferred from one image to the other. The length d is the distance from a tangency point in the first image and an epipolar line obtained by the transfer of an epipolar tangency from the second image. The distance d' is found in the same way, interchanging the roles of the images. The cost function is then the sum $\sum_i (d_i^2 + d_i'^2)$ for each matching pair i of putative epipolar tangencies.

This kind of situation suggests the use of an *affine camera model* [12] and *affine epipolar geometry* [13], that assumes that the epipoles will be at infinity, and reduces the degrees of freedom of the fundamental matrix, which will display the format shown:

$$\mathbf{F} = \begin{bmatrix} 0 & 0 & c \\ 0 & 0 & d \\ a & b & e \end{bmatrix}. \qquad (4)$$

There are two circumstances when the *affine fundamental matrix* may be used. The first is when the affine model can be used to describe the cameras. If the principal point of a camera is (u_0, v_0), the variation of depth in the scene is ΔZ and the mean distance of the features of the scene to the camera is Z_{mean}, the difference of the image of a point taken from a projective camera (u, v) and its image at the affine camera, $(u_{\mathrm{a}}, v_{\mathrm{a}})$ is given by

$$(u - u_{\mathrm{a}}, v - v_{\mathrm{a}}) = (u - u_0, v - v_0)\Delta Z/Z; \qquad (5)$$

When the field of view is narrow, the terms $u - u_0$ and $v - v0$ will be small. In this case, or when the depth variation of the scene is much smaller than its mean depth, e. g. $\Delta Z/Z < 0.1$, the error due to the affine approximation is negligible.

Other favorable situation for the use of the affine fundamental matrix is when the motion is restricted to translation orthogonal to the optical ray and cyclorotation. In this case the affine fundamental matrix can be used even if the affine camera model is inadequate. It is important to notice that a rotation by a small angle around a distant axis is a good approximation for such motion.
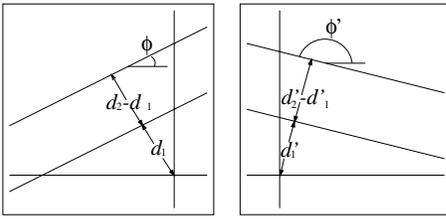
Figure 4. Geometric interpretation of the parametrisation of the affine fundamental matrix presented in (6). The directions of the epipoles on each image are given by the angles ϕ and ϕ', and the parameters α and α' are proportional to the differences of distances $d_2' - d_1'$ and $d_2 - d_1$, respectively (see (7)).

As scale factors are not important, the affine fundamental matrix has only four degrees of freedom, and can be linearly computed from 4 point correspondences. Each epipole, as being at infinity, is described by a single parameter, corresponding to the a direction in the image plane. This observation suggests another parametrisation for the fundamental matrix, where the directions of the epipoles are made explicit. If ϕ and ϕ' are the directions of the epipoles in the first and second images, the affine fundamental matrix can be expressed as

$$\mathbf{F} = \begin{bmatrix} 0 & 0 & \alpha' \sin \phi' \\ 0 & 0 & -\alpha' \cos \phi' \\ -\alpha \sin \phi & \alpha \cos \phi & \sqrt{1 - \alpha^2 - \alpha'^2} \end{bmatrix}, \quad (6)$$

where the parameters α and α' are related to the distances between epipolar lines on each image. The geometric interpretation of the parameters α and α' can be seen in fig. 4. It is easy to show that they are proportional to the distance between epipolar lines, or, in the notation of fig. 4,

$$\begin{bmatrix} \alpha \\ \alpha' \end{bmatrix} = \frac{\begin{bmatrix} d_1' - d_2' \\ d_2 - d_1 \end{bmatrix}}{\sqrt{(d_2 d_1' + d_1 d_2')^2 + (d_1' - d_2')^2 + (d_1 - d_2)^2}}. \quad (7)$$

In the affine case the epipolar tangencies will be parallel lines, with directions given by the corresponding epipole, and, as in the projective case, the epipolar tangencies will touch the apparent contours at corresponding points. Since the number of degrees of freedom of the affine fundamental matrix is 4, this will also be the number of epipolar tangencies necessary for its computation. Thus, the algorithm for computation of the epipolar geometry from apparent contours in the affine case is described as follows:

1. Initialise estimative for the directions of both epipoles.

2. Determine correspondences $\mathbf{m}_i$ and $\mathbf{m}_i'$ from epipolar tangencies consistent with the directions of the epipoles.

3. Compute the affine fundamental matrix from the epipoles and the correspondences. This must be done by using the parametrisation given in (6), so the directions of the epipoles will be kept.

4. Minimise the the sum of geometric distances from the tangent points on the contours to the corresponding epipolar lines. The search is restricted to the directions of the epipoles, and the cost function is the same as given by (3).
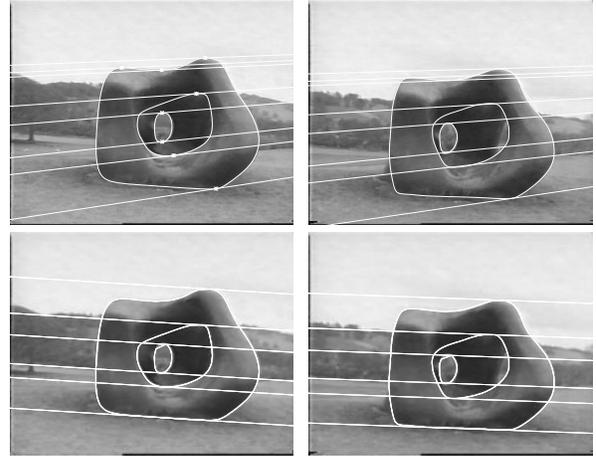
3.1. Experimental Results.

Figure 5. Estimation of the epipolar geometry from the apparent contours. The top pair of images shows epipolar tangencies estimated using the algorithm described in [5]. The bottom pair shows the results of the algorithm presented here.

The algorithm was tested in the images shown in fig. 5, with the directions of the epipoles initialised at $0°$. The contours were extracted by fitting B-splines to image edge data [4]. There is clear discrepancy between the results found by the general algorithm and the ones found by the algorithm for the affine case. Since there is no ground truth, the doubt will remain. However, the epipolar lines found by the algorithm using affine approximation are consistent with a planar motion parallel to the ground, since the epipolar lines are parallel to the image of the physical horizon line. This is a very reasonable guess for the actual motion used in the acquisition of the images. Also, the bottom epipolar line may give a clue that the affine approximation is indeed a good estimative. If the angles of the epipolar lines had opposite signals in respect the horizontal axis, the points put under correspondence by the epipolar geometry would be at opposite places in the bottom of the sculpture, which is, obviously, an incorrect solution. So the angles must have the

same signal. Moreover, the actual values found were such that correspondent tangent points are located at the very bottom of the sculpture, which is also consistent with a rotation motion around the sculpture.

Consider now images taken from a triplet of cameras where the relative position of subsequent cameras is fixed. In this case, the epipolar geometry of adjacent cameras will also be fixed, and if two epipolar tangencies are available, a pair of estimative for the directions of the epipoles will produce, for images showing two frontier points, 4 correspondences, or 2 correspondences for each pair of images. A typical example of such configuration is when the camera undergoes a rotation at fixed angles around an arbitrary axis. This idea was explored in the experiments that follow. The algorithm previously described was implemented
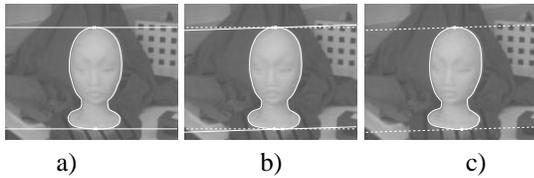

Figure 6. Images and contours used in the computation of the affine fundamental matrix. The rotation angle between successive snapshots is $5°$. The continuous lines and $\times$ marks are corresponding epipolar lines and epipolar tangencies from images a) and b), and the dashed lines a and circles are corresponding epipolar lines and epipolar tangencies from images b) and c), after convergence of the algorithm.

for the images shown in fig. 6. Convergence was achieved after 4 to 5 iterations of both Davidon-Fletcher-Powell or Levenberg-Marquadt optimisation methods. To evaluate the results, a calibration grid substituted the dummy and the same rotation was applied (see fig. 7). The epipolar lines on the right image correctly match the points correspondent to the ones marked on left, showing the quality of the estimated epipolar geometry.

4. Circular Motion

A very common situation for model acquisition is placing an object over a turn-table and taking snapshots at successive angles. There are successful methods for model reconstruction under this particular situation [8] when point features are available. However, for a smooth object, without landmarks on its surface, point correspondences are difficult to obtain. This is obviously a case for techniques based on apparent contours.
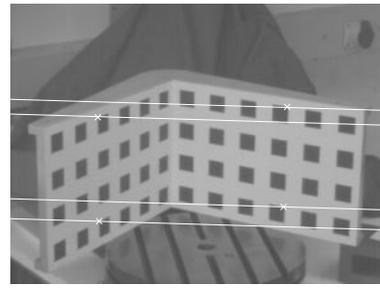

Figure 7. Evaluation of the estimated affine epipolar geometry. The fundamental matrix estimated from the apparent contours of the rotating dummy was tested in a calibration grid under the same motion. The epipolar lines corresponding to the matches of the marked points are shown, and the distance between the lines and points is always less than one pixel, showing the accuracy of the estimation.

There are two main image features in a stereo pair of images obtained under camera rotation: the rotation axis and the horizon line [1, 9]. If the camera intrinsic parameters are kept fixed during the rotation, the projection of the rotation axis will be a line fixed pointwise in the pair of images. The horizon line is the projection of the plane where the rotation plane, i. e., the plane orthogonal to the rotation axis that contains the camera centres. Since the rotation plane contains the camera centres, it also contains the line that passes through the camera centres. Thus, the epipoles must lie on the horizon line.

4.1. Determination of the Epipolar Geometry

The determination of the epipoles and the image of the rotation axis, henceforth called *rotation line*, is enough to fix the epipolar geometry. Since the rotation line is fixed pointwise, the epipoles and three points in the rotation line determine an homography of the epipolar lines from one image to the other. This homography, together with the epipoles, fixes the fundamental matrix and thus the epipolar geometry of the system. So, in the circular motion case the fundamental matrix has 6 degrees of freedom: 4 for the coordinates of the two epipoles plus 2 for the equation of the rotation line.

The relation between the epipoles, the rotation line and the homography is represented in fig. 8. If the images are rotated in such way that the horizon line is horizontal, the homography has a simpler expression, given by

$$\tau' = \frac{\tau}{k - \tau(1+k)\cot\theta}, \tag{8}$$

where $k = a/b$, θ is the angle between the rotation line and the horizon line and τ' and τ are the tangents of the angles ψ' and ψ as shown in fig. 8. The parameters a and b are the distance from e and e' to the intersection of the rotation and horizon lines, as shown in fig. 8. Hence, it is possible to
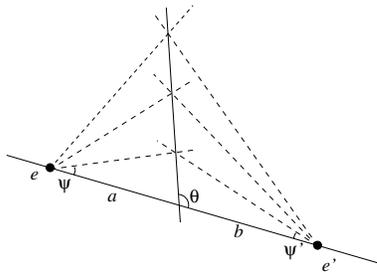

Figure 8. The homography shown in (8) relates the tangents τ and τ' of the angles ψ and ψ'.

parametrise the fundamental matrix in the case of rotation motion with only 6 parameters: 4 for the coordinates of the epipoles, that fix the horizon line, 1 for the intersection of the rotation line with the horizon line and 1 for the angle between the rotation line and the horizon line.

The planes intersecting the camera centres and tangent to the surface being viewed by the cameras define epipolar lines in both images. If these epipolar lines are superimposed in a single image, they must intersect in a point at the rotation line. So, assuming that the epipoles are known, two of such tangent planes would fix the rotation line and thus the homography (see fig. 9). The position of the epipoles is controlled by four degrees of freedom (the coordinates of the epipoles), and four more tangent planes would provide enough information to determine the epipolar geometry of the pair of cameras, in the case of rotation. If more tangent planes cannot be determined but another image from a third camera, whose position relative to the second camera is the same as the relative position of the second camera to the first one, the new tangency points and the second and third camera centres will determine another pair of tangent planes and correspondent epipolar lines between images two and three. Proceeding with this, at the fourth image one would have came up with the six necessary tangent planes, and the epipolar geometry would be fixed. It is necessary that adjacent cameras be related by the same rotation and the intrinsic parameters are kept constant so the epipoles will not move. Intermediate solutions would also be possible, e. g., three tangencies and three images with adjacent cameras related by the same rotation. In general we must have $n(i-1) \geq 6$, where n is the number of tangencies in each image and i is the number of images in order to find the epipolar geometry from apparent contours in the case of rotation. We can now summarise the algorithm for estimation
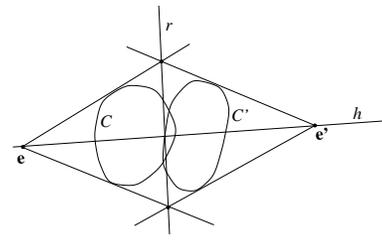

Figure 9. Relation between the epipoles (e and e'), apparent contours (C and C'), horizon line (h) and rotation line (r). Superimposing two adjacent images of a rotating surface, the horizon line can be determined as the line joining the epipoles. The intersection of correspondent epipolar tangencies must lie on the image of the rotation axis, which is thus fixed by two pairs of epipolar tangencies. This is enough to fix the homography relating pencils of epipolar lines between the images, according to (8).

of the epipolar geometry in the circular motion case:

1. Initialise the epipoles at random positions.
2. Superimpose adjacent images and determine epipolar tangencies.
3. Fit by least-squares a line to the set of points generated by the intersection of each pair of correspondent epipolar lines. The sum of the square distance of each point to the fitted line is the cost function which will be then minimised through a search for the correct position for the epipoles.

Preliminary Experiments. In this first experiment with synthetic data, 4 images of an ellipsoid were taken at successive positions related by a rotation of $2\pi/5\,\mathrm{rad}$ around a fixed axis. The epipoles were initialised at coordinates $e_0 = (4000, 200)$ and $e'_0 = (-3000, 400)$, far from the true values $e = (2752.8, 0)$ and $e' = (-2752.8, 0)$. The epipolar tangencies passing through the epipoles were computed, and for each pair of adjacent cameras, their intersection was computed. A line was fitted to the 6 points so produced, and the sum of the square distance in pixels of each point to the fitted line is the cost function to be minimised. The gradient of this function with respect to the coordinates of the epipoles was estimated numerically, and the Davidon-Fletcher-Powell or Levenberg-Marquadt methods were used to optimise the position of the epipoles. In fig. 10 the initial and final set of intersection points is presented, together with the correspondent estimative for the rotation line.

In fig. 11 the initial and final epipolar configuration are shown. Each picture presents the superposition of two adja-

cent images of the ellipsoid and the correspondent tangent lines for each epipole and contour. The final values found
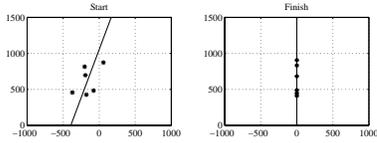

Figure 10. Intersection of epipolar lines at first and last iteration of the algorithm. The fitted line is the estimative of the rotation line.
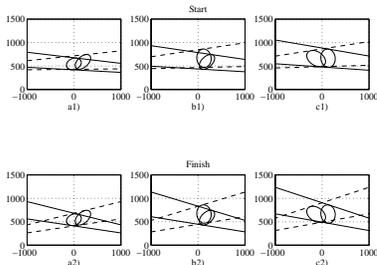

Figure 11. Initial (above) and final (below) epipolar configurations for four images of an ellipsoid. Each figure shows the superposition of adjacent images.

for the coordinates of the epipoles were $e = (2752.9, 0.0)$ and $e' = (2752.8, 0.0)$, very close, by any criteria, of the true values. The experiment was reproduced with noise levels up to one pixel, and the measure *Em* [11] for the relative error in the position of the epipoles, given by

$$Em = \frac{1}{2} \left(\frac{\|e - e_0\|}{\min(\|e\|, \|e_0\|)} + \frac{\|e' - e'_0\|}{\min(\|e'\|, \|e'_0\|)} \right), \quad (9)$$

was always under 0.2.

Other experiments, with different values of the baseline angle, β, were also done, and, *for large baseline angle*, e. g. $\beta > \pi/10$, the results were similar. For smaller values of β, though, the sensitivity to noise of the algorithm increased unacceptably. However, this is a favourable situation for the use of the affine approximation, as pointed in sec. 3, that thus should be used instead.

5. Conclusions and Future Goals

This article presents some techniques for estimation of the epipolar geometry of image sequences from apparent contours under affine approximation and circular motion, situations that often arise in applications. The affine approximation showed robustness and accuracy even in cases where the eccentricity of the apparent contours is small, as in fig 6, which is a very challenging problem. When compared to the algorithm presented in [5], the affine approximation is also simpler and demands less tangent points for its computation, exactly as it happens in the case of point features.

The circular motion case is still under development. The cost function based in the fitting of a putative rotation axis to the intersection of epipolar tangencies is very insensitive to variations in the position of the epipoles except when the eccentricity of the contours is large. However, when the baseline is small the affine approximation can be safely used. A topic proposal for further investigation is the evaluation of the estimation of the epipolar geometry by reconstructing of the surfaces.

References

[1] M. Armstrong, A. Zisserman, and R. Hartley. Self-calibration from image triplets. In *Proc. 4th European Conf. on Computer Vision*, pages 3–16, 1996.

[2] K. Åström, R. Cipolla, and P.J. Giblin. Generalised epipolar constraint. In *Proc. 4th European Conf. on Computer Vision*, volume 2, pages 97–108, 1996.

[3] E. Boyer and M. O. Berger. 3D surface reconstruction using occluding contours. *Int. Journal of Computer Vision*, 22(3):219–233, 1997.

[4] T.J. Cham and R. Cipolla. MDL-based curve representation using b-spline active contours. In *Proc. British Machine Vision Conference*, pages 363–372, 1996.

[5] R. Cipolla. The visual motion of curves and surfaces. *Phil. Trans. Royal Soc. London A*, 356:1103–1121, 1998.

[6] 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.

[7] R. Cipolla and A. Blake. Surface shape from the deformation of apparent contours. *Int. Journal of Computer Vision*, 9(2):83–112, 1992.

[8] 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, 1988.

[9] 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.

[10] P.J. Giblin and R.S. Weiss. Epipolar fields on surfaces. In J-O. Eklundh, editor, *Proc. 3rd European Conf. on Computer Vision*, volume I, pages 14–23. Springer–Verlag, 1994.

[11] Q-T. Luong and O. Faugeras. The fundamental matrix: Theory, algorithm, and stability analysis. *Int. Journal of Computer Vision*, 17:43–75, 1996.

[12] J.L. Mundy and A.Zissermann editors. *Geometric Invariance in Computer Vision*. MIT Press, 1992.

[13] L. S. Shapiro, A. Zisserman, and M. Brady. 3d motion recovery via affine epipolar geometry. *Int. Journal of Computer Vision*, 16:147–182, 1995.

[14] R. Vaillant and O.D. Faugeras. Using extremal boundaries for 3D object modelling. *IEEE Trans. Pattern Analysis and Machine Intell.*, 14(2):157–173, 1992.