PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The visual motion of curves and surfaces

The Royal Society

Phil. Trans. R. Soc. Lond. A 1998 **356**, 1103-1121
doi: 10.1098/rsta.1998.0213

The Royal Society

The visual motion of curves and surfaces

By Roberto Cipolla

Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK

For smooth curved surfaces the dominant image feature is the *apparent contour*, or outline. This is the projection of the *contour generator*, the locus of points on the surface which separate visible and occluded parts. The contour generator is dependent of the local surface geometry and the viewpoint. Each viewpoint will generate a different contour generator. This paper addresses the problem of recovering the three-dimensional shape and motion of curves and surfaces from image sequences of apparent contours.

For *known* viewer motion the visible surfaces can then be reconstructed by exploiting a spatio-temporal parametrization of the apparent contours and contour generators under viewer motion. A natural parametrization exploits the contour generators and the epipolar geometry between successive viewpoints. The *epipolar parametrization* leads to simplified expressions for the recovery of depth and surface curvatures from image velocities and accelerations and known viewer motion.

The parametrization is, however, degenerate when the apparent contour is singular since the ray is tangent to the contour generator and at *frontier points* when the epipolar plane is a tangent plane to the surface. At these isolated points the epipolar parametrization can no longer be used to recover the local surface geometry. This paper reviews the epipolar parametrization and shows how the degenerate cases can be used to recover surface geometry and *unknown* viewer motion from apparent contours of curved surfaces. Practical implementations are outlined.

Keywords: epipolar geometry; apparent contour; cusps; frontier points

1. Introduction

Structure and motion from image sequences of point features has attracted considerable attention and a large number of algorithms exist to recover both the spatial configuration of the points and the motion compatible with the views. A key component of these algorithms is the recovery of the *epipolar geometry* between distinct views (Luong & Faugeras 1996). The structure and motion problem for curves and curved surfaces is more challenging. For curved surfaces the dominant image feature is the *apparent contour* which is the projection of the curve on the surface (*contour generator*) dividing visible and occluded parts. The contour generator is dependent on viewpoint and local surface geometry (via tangency and conjugacy constraints) and each viewpoint will generate a different contour generator. The image curves are therefore projections of different space curves and there is no correspondence between points on the curves in the two images.

The family of contour generators generated under continuous viewer motion can be used to represent the visible surface. Giblin & Weiss (1987) and Cipolla & Blake (1992) have shown how the spatio-temporal analysis of deforming image apparent

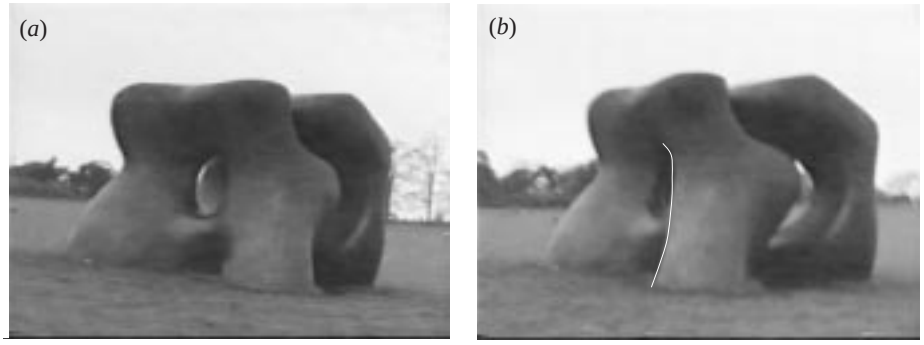

Figure 1. The apparent contours of a curved surface from two different viewpoints. The apparent contours from different viewpoints are projections of different surface curves. The apparent contour can be singular, seen here as the visible apparent contour ending abruptly (*b*).

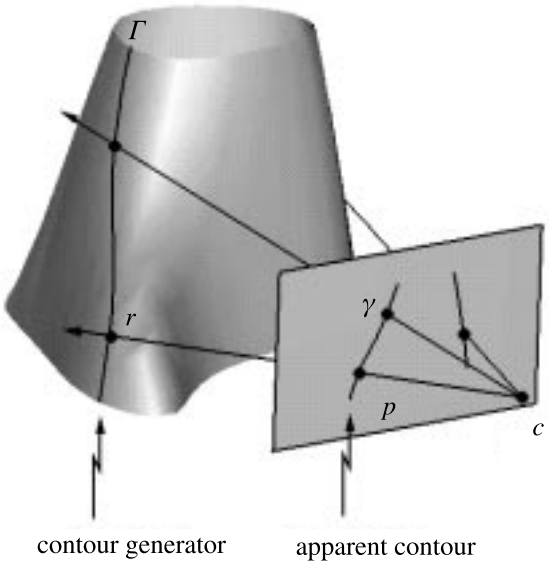

Figure 2. Viewing geometry and parametrization of the surface. For each viewpoint, $\boldsymbol{c}$, the family of rays which are tangent to the surface define the contour generator, $\boldsymbol{\Gamma}$. The image of the contour generator is called the apparent contour, $\boldsymbol{\gamma}$. A surface point, $\boldsymbol{r}$, has position $\boldsymbol{c}+\lambda\boldsymbol{p}$ where λ is the distance from the viewing centre along a ray with direction given by the unit vector $\boldsymbol{p}$.

contours or outlines enables computation of local surface curvature along the corresponding contour generator on the surface. To perform the analysis, however, a spatio-temporal parametrization of image-curve motion is needed, but is underconstrained. The *epipolar* parametrization is most naturally matched to the recovery of surface curvature. In this parametrization (for both the spatio-temporal image and

the surface), *correspondence* between points on successive snapshots of an apparent contour and contour generator is set up by matching along epipolar lines and epipolar planes respectively. The parametrization leads to simplified expressions for the recovery of depth and surface curvature from image velocities and accelerations and known viewer motion.

There are, however, several cases in which this parametrization is degenerate and so can not be used to recover the local surface geometry. The first case of degeneracy occurs at a point of the surface-to-image mapping when a ray is tangent not only to the surface but also to the contour generator. This will occur when viewing a hyperbolic surface patch along an asymptotic direction. For a transparent surface this special point on the contour generator will appear as a *cusp* on the apparent contour. For opaque surfaces, however, only one branch of the cusp is visible and the contour ends abruptly (Koenderink & Van Doorn 1982).

The other case of degeneracy of the epipolar parametrization occurs when contour generators from subsequent viewpoints intersect to form an envelope. This occurs when the epipolar plane is also a tangent plane to the surface. These isolated surface points are called *frontier points* (Giblin & Weiss 1994). The surface can not be reconstructed at these points by the epipolar parametrization since the contour generator is locally stationary. However, the frontier points correspond to real, fixed feature points on the surface which are visible in two views. They can be used to recover the viewer motion.

This paper addresses the problem of recovering the three-dimensional (3D) shape and motion of curves and surfaces from image sequences of apparent contours. As with point features, the epipolar geometry plays an important role in both the recovery of the motion and in the reconstruction of the surface.

2. Reconstruction under known viewer motion

(*a*) *Viewing geometry*

Consider a smooth surface M. For each vantage point, $\boldsymbol{c}$, the sets of points, $\boldsymbol{r}$, on the surface for which the visual ray is tangent to M can be defined. This is called the *contour generator*, Γ, and is the set of points $\boldsymbol{r}$ for which

$$(\boldsymbol{r}-\boldsymbol{c})\cdot\boldsymbol{n}=0, \tag{2.1}$$

where $\boldsymbol{n}$ is the unit normal to the surface at $\boldsymbol{r}$ (figure 2). The contour generator is usually (but not always, see § 3) a smooth curve on the surface separating the visible from the occluded parts and can be parametrized using say s as a parameter.

The image, $\boldsymbol{\gamma}$, of the contour generator, $\boldsymbol{\Gamma}$, is called the apparent contour or outline and is the intersection of the set of rays which are tangent to the surface and the imaging surface. Without loss of generality and for mathematical simplicity, we consider perspective projection onto the unit sphere. An apparent contour point, $\boldsymbol{p}$ (a unit vector specifying the direction of the visual ray), satisfies

$$\boldsymbol{r}=\boldsymbol{c}+\lambda\boldsymbol{p}, \tag{2.2}$$

$$\boldsymbol{p}\cdot\boldsymbol{n}=0, \tag{2.3}$$

where λ is the distance along the ray to the surface point, $\boldsymbol{r}$, from the position of the centre of projection $\boldsymbol{c}$.

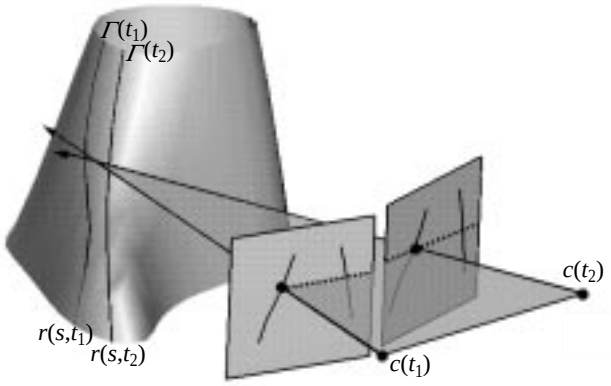

Figure 3. Epipolar parametrization. The surface is parametrized locally by the contour generators from successive viewpoints and *epipolar* curves defined by the intersection of the pencil of epipolar planes and the surface. The parametrization is only possible if the contour generators are not singular and are transverse to the epipolar curves. The parametrization is degenerate at cusp and frontier points.

(*b*) *Spatio-temporal parametrization*

Each viewpoint will generate a different contour generators on the surface M. A moving monocular observer with position at time t given by $\boldsymbol{c}(t)$, will generate a one parameter family of contour generators, indexed by time, $\Gamma(t)$. It is natural to attempt a parametrization of M which is 'compatible' with the motion of the camera centre, in the sense that contour generators are parameter curves. We want there to exist a regular (local) parametrization of M of the form $(s,t) \rightarrow \boldsymbol{r}(s,t)$, where the set of points $\boldsymbol{r}(s,t_0)$, for fixed t_0 (i.e. the s-parameter curve), is the contour generator from viewpoint $\boldsymbol{c}(t_0)$. The set of points $\boldsymbol{p}(s,t_0)$ is the corresponding apparent contour on the unit sphere at the viewpoint; the actual apparent contour points in space are $\boldsymbol{c}(t_0)+\boldsymbol{p}(s,t_0)$. Note that (2.2) and (2.3) become

$$\boldsymbol{r}(s,t) = \boldsymbol{c}(t) + \lambda(s,t)\boldsymbol{p}(s,t), \tag{2.4}$$

$$\boldsymbol{p}(s,t)\cdot\boldsymbol{n}(s,t) = 0. \tag{2.5}$$

However, the spatio-temporal parametrization of the apparent contours, $\boldsymbol{p}(s,t)$, and the surface, $\boldsymbol{r}(s,t)$, is not unique. The choice of the t-parameter curves, $\boldsymbol{p}(s_0,t)$ and $\boldsymbol{r}(s_0,t)$, for fixed s_0, is under-constrained.

(*c*) *Epipolar parametrization*

A natural choice of parametrization is the *epipolar* parametrization. In this parametrization the *correspondence* between points on successive snapshots of an apparent contour and contour generator are set up by matching along epipolar planes. Namely the corresponding ray in the next viewpoint (in an infinitesimal sense), is chosen so that it lies in the epipolar plane defined by the viewer's translational motion and the ray in the first viewpoint (figure 3).

Figure 4. Estimating surface curvatures from three discrete views. Points are selected on image contours in the first view (t_0), indicated by crosses A and B for points on an extremal boundary and surface marking respectively. For the *epipolar parametrization* of the surface corresponding features lie on epipolar lines in the second and third view (t_1 and t_2). Measurement of the three rays lying in an epipolar plane can be used to estimate surface curvatures. Point B can be considered as a degenerate surface point with infinite curvature.

The epipolar parametrization is defined by (Cipolla & Blake 1992)

$$\boldsymbol{r}_t \wedge \boldsymbol{p} = 0 \tag{2.6}$$

and leads to the following *epipolar matching* condition,

$$[\boldsymbol{p}_t, \boldsymbol{c}_t, \boldsymbol{p}] = 0, \tag{2.7}$$

such that the t-parameter curves are defined to lie instantaneously in the epipolar plane defined by the ray and direction of translation. The parametrization leads to simplified expressions for the recovery of depth and surface curvature. By simple manipulation of equations (2.4) to (2.7) and their spatial and temporal derivatives (denoted below by subscripts s and t) it is easy to show that the local surface geometry can be recovered from spatio-temporal derivatives (up to second order) of the apparent contours and the *known* viewer motion as follows:

1. The orientation of the surface normal, $\boldsymbol{n}$, can be recovered from a single view from the vector product of the ray direction, $\boldsymbol{p}$ and the tangent to the apparent contour, $\boldsymbol{p}_s$:

$$\boldsymbol{n} = \frac{\boldsymbol{p} \wedge \boldsymbol{p}_s}{|\boldsymbol{p} \wedge \boldsymbol{p}_s|}. \tag{2.8}$$

2. If the contour generator is smooth at $\boldsymbol{r}$ then its tangent is in a *conjugate* direction (with respect to the second fundamental form) to the visual ray $\boldsymbol{p}$ and

$$\boldsymbol{p} \cdot \boldsymbol{n}_s = 0. \tag{2.9}$$

3. Under viewer motion and the epipolar parametrization, a given point on a contour generator, $\boldsymbol{r}$, will *slip* over the surface with velocity given by $\boldsymbol{r}_t$ and which depends on the distance, λ, and surface curvature (normal curvature in direction of the ray):

$$\boldsymbol{r}_t = -\left(\frac{\boldsymbol{c}_t \cdot \boldsymbol{n}}{\lambda \kappa^t}\right) \boldsymbol{p}. \tag{2.10}$$

Note that the velocity is inversely proportional to the surface curvature and is zero in the limiting case of viewing a space curve or crease. The latter can be simply treated as apparent contours with infinite curvature along the ray direction.

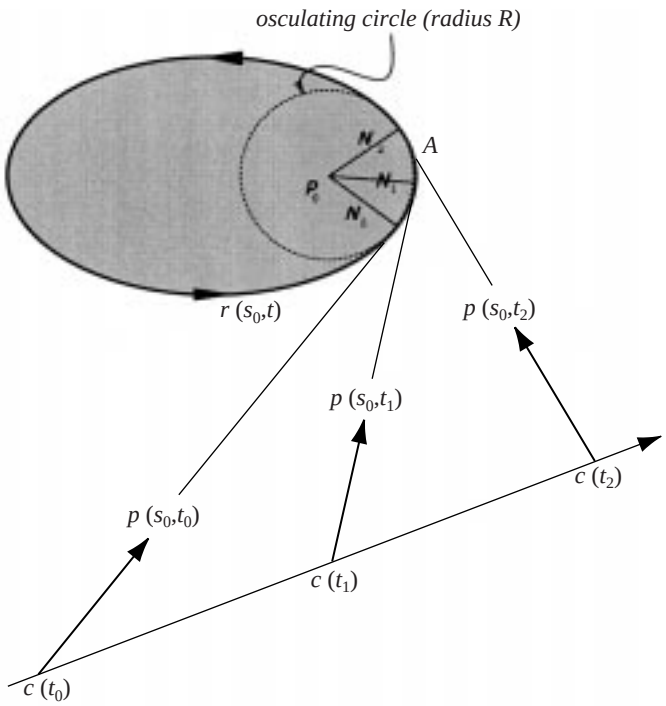

Figure 5. The epipolar plane. Each view defines a tangent to $\boldsymbol{r}(s_0, t)$. For linear camera motion and epipolar parametrization the rays and $\boldsymbol{r}(s_0, t)$ lie in a plane. If $\boldsymbol{r}(s_0, t)$ can be approximated locally by its osculating circle, it can be uniquely determined from measurements in three views. For curvilinear motion the epipolar geometry is continuously changing and the epipolar curve is no longer planar.

4. Depth (distance along the ray, λ) can be computed from the deformation of the apparent contour, ($\boldsymbol{p}_t$), under known viewer motion (translational velocity $\boldsymbol{c}_t$):

$$\lambda = -\frac{\boldsymbol{c}_t \cdot \boldsymbol{n}}{\boldsymbol{p}_t \cdot \boldsymbol{n}}. \tag{2.11}$$

5. The Gaussian curvature at a point on the apparent contour, K, can be recovered from the depth, λ, the normal curvature κ^t along the line of sight and the *geodesic* curvature of the apparent contour, κ^p:

$$K = \frac{\kappa^p \kappa^t}{\lambda}. \tag{2.12}$$

Since the normal section in the direction of the ray must always be convex at a point on the apparent contour, the sign of the Gaussian curvature is determined by the sign of the curvature of the apparent contour. Convexities, concavities and inflections of an apparent contour corresond to elliptic, hyperbolic and parabolic surface points respectively (Koenderink 1984).

Figures 4–6 illustrate the epipolar parametrization and the reconstruction of a strip of surface at the contour generator. Details of the camera calibration and the detection and tracking of the apparent contours with B-spline snakes can be found in Cipolla & Blake (1992) and Boyer & Berger (1997).

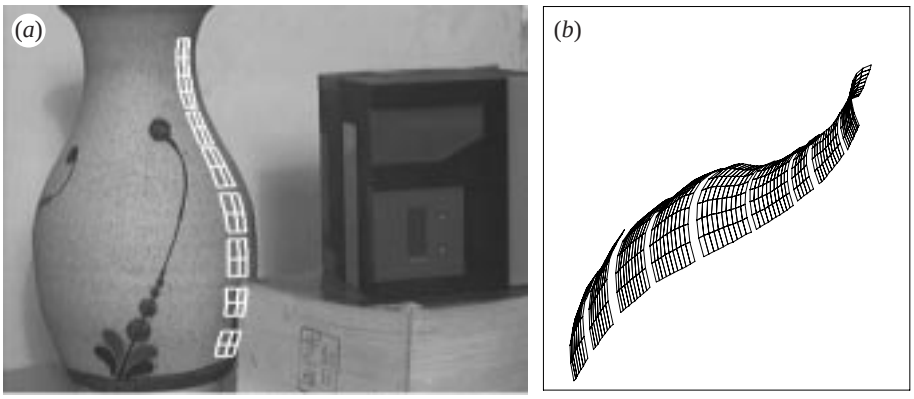

Figure 6. Recovery of surface strip in vicinity of apparent contour. The surface is recovered as a family of s-parameter curves, the contour generators, and t-parameter curves, portions of the *osculating* circles measured in each epipolar plane. The strip is shown projected into the image of the scene from a different viewpoint and after extrapolation (Cipolla & Blake 1992).

3. Degenerate cases of the epipolar parametrization

There are two possible cases where degeneracy of the parametrization arises. These occur when the contour generators and epipolar curves are singular or not transverse, i.e. $\{\boldsymbol{r}_s, \boldsymbol{r}_t\}$ fails to form a basis for the tangent plane of the surface:

$$\boldsymbol{r}_s \wedge \boldsymbol{r}_t = \boldsymbol{0}. \tag{3.1}$$

The first case occurs when $\boldsymbol{r}$ is a hyperbolic surface point viewed along an asymptotic direction. The apparent contour is singular (a cusp is generated), seen as a contour-ending for opaque surfaces. The second case occurs when $\boldsymbol{r}$ is a *frontier point*. The epipolar plane (spanned by the velocity vector $\boldsymbol{c}_t(t)$ of the viewpoint and the ray $\boldsymbol{p}$) coincides with the tangent plane to M and the contour generators form an envelope on M.

(*a*) *Singular apparent contours*

The most common case of degeneracy occurs when viewing a hyperbolic surface point along an *asymptotic* direction (figure 7). The ray grazes the surface with 3-point contact and is tangent to the contour generator, $\boldsymbol{r}_s \wedge \boldsymbol{p} = \boldsymbol{0}$. For a transparent surface this special point on the contour generator (the cusp generator point) will appear as a *cusp* on the apparent contour with $\boldsymbol{p}_s = \boldsymbol{0}$. For opaque surfaces, however, only one branch of the cusp is visible and the apparent contour ends abruptly (Koenderink & Van Doorn 1982).

Cusps (or contour-endings as they appear for opaque surfaces) are stable phenomena and they persist under viewer motion. In any generic view of a curved surface we expect to see smooth apparent contours with a finite number of singularities (cusp or contour-ending) and T-junctions. Under viewer motion the generic events that are visible consist of cusps being created or annihilated in pairs. These are described in Koenderink & Van Doorn (1976) and Koenderink (1990) and illustrated in figures 8 and 9.

The epipolar parametrization can no longer be used to recover the depth and surface curvature at cusp points. In fact the cusp points have a component of motion

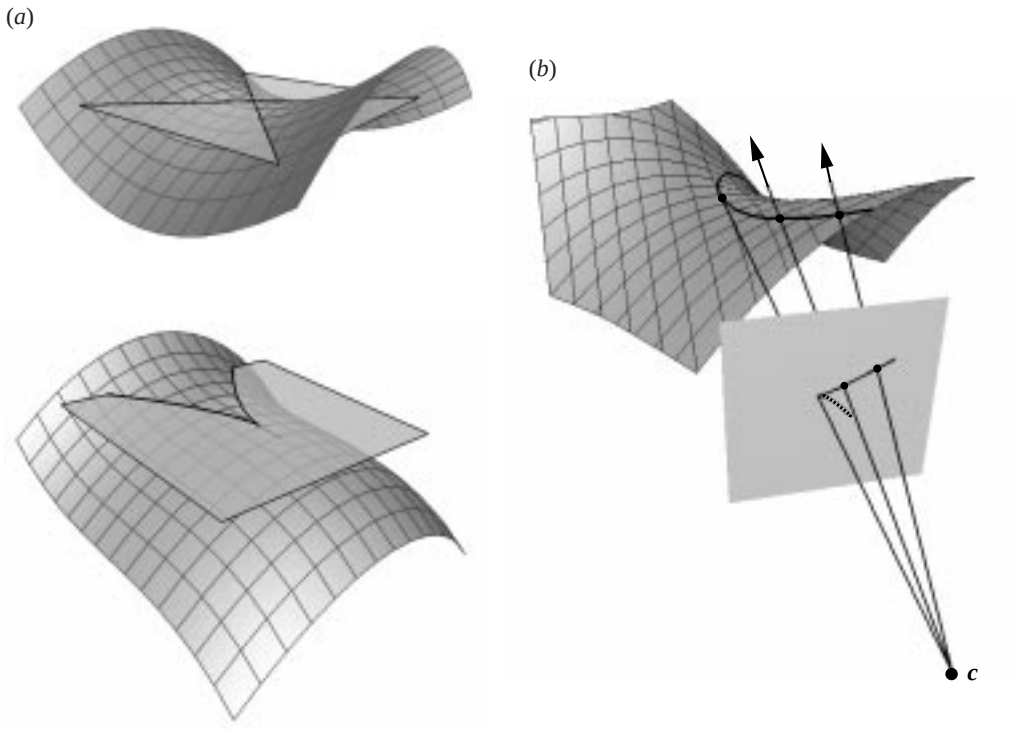

Figure 7. (*a*) The asymptotic directions of a hyperbolic and parabolic surface point. (*b*) The visual ray has 3-point contact with the surface when viewing a hyperbolic surface point along one of its asymptotic directions. The contour generator and ray are parallel and a cusp is generated in the apparent contour. Only one branch of the cusp is visible for an opaque surface and the apparent contour ends abruptly.

out of the epipolar plane and so can not be localized on corresponding epipolar lines in the different views. The cusp point in the image can, in principle, be localized and tracked under viewer motion, and if its trajectory is smooth and can be parametrized by t, $\boldsymbol{p}(t)$, it can be used to induce an alternative parametrization of the surface in the vicinity of the cusp generator (Cipolla *et al*. 1997).

At a cusp, the cuspidal tangent is given by $\boldsymbol{p}_{ss}$ (since $\boldsymbol{p}_s = \boldsymbol{0}$) and the surface normal, depth and Gaussian curvature can be recovered as follows:

$$\boldsymbol{n} = \frac{\boldsymbol{p} \wedge \boldsymbol{p}_{ss}}{|\boldsymbol{p} \wedge \boldsymbol{p}_{ss}|}, \tag{3.2}$$

$$\lambda = -\frac{\boldsymbol{c}_t \cdot \boldsymbol{n}}{\boldsymbol{p}_t \cdot \boldsymbol{n}}, \tag{3.3}$$

$$K = \frac{-(\boldsymbol{p}_t \cdot \boldsymbol{n})^4}{[\boldsymbol{p}, \boldsymbol{c}_t, \boldsymbol{p}_t]^2}. \tag{3.4}$$

Note that the Gaussian curvature can be recovered from first-order derivatives only. Compare this to a normal apparent contour point that requires second-order spatio-temporal derivatives of viewpoint and apparent contour.

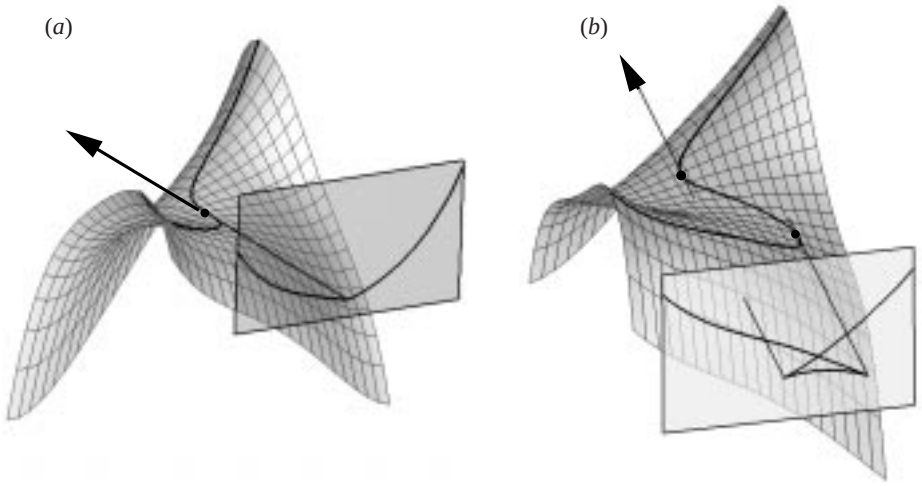

Figure 8. (*a*) The swallowtail transition. A smooth apparent contour develops a pair of cusps and a T-junction under viewer motion. For opaque surfaces only one branch of one of the cusps is visible and the smooth apparent contour comes to a sudden end. (*b*) The transition occurs at an asymptotic direction of a *flecnodal surface point*. The ray has 4-point contact with the surface, (*a*).

(*b*) *Frontiers and epipolar tangencies*

The remaining case of degeneracy of the epipolar parametrization occurs for epipolar planes (spanned by the direction of translation and the ray) which coincide with tangent planes to the surface. This will occur at a finite set of points on the surface where the surface normal $\boldsymbol{n}$ is perpendicular to the direction of translation:

$$\boldsymbol{c}_t \cdot \boldsymbol{n} = 0. \tag{3.5}$$

From (2.10) we see that the contour generator is locally stationary ($\boldsymbol{r}_t = \boldsymbol{0}$). In fact (see figure 10) consecutive contour generators will intersect at points where the epipolar plane is tangent to the surface.

The points of contact on the surface are called *frontier points* because for continuous motion the locus of intersections of consecutive contour generators in an infinitesimal sense define a curve on the surface which represents the boundary of the visible region swept out by the contour generators under viewer motion.

For larger discrete motions the contour generators defined by the discrete viewpoints also intersect at points on the surface where the epipolar plane is tangent to the surface. This is easily seen if we consider the motion to be linear. $\boldsymbol{c}_t$ is then a *constant* vector, and the frontier point on the surface at time t satisfies the frontier condition at subsequent times. The frontier degenerates to a point on the surface. In the discrete case the frontier points are defined by the condition

$$\Delta\boldsymbol{c} \cdot \boldsymbol{n} = 0, \tag{3.6}$$

where $\Delta\boldsymbol{c} = \boldsymbol{c}(t_2) - \boldsymbol{c}(t_1)$ and $\boldsymbol{n}$ is the surface normal at the point in which in the two contour generators for each viewpoint intersect.

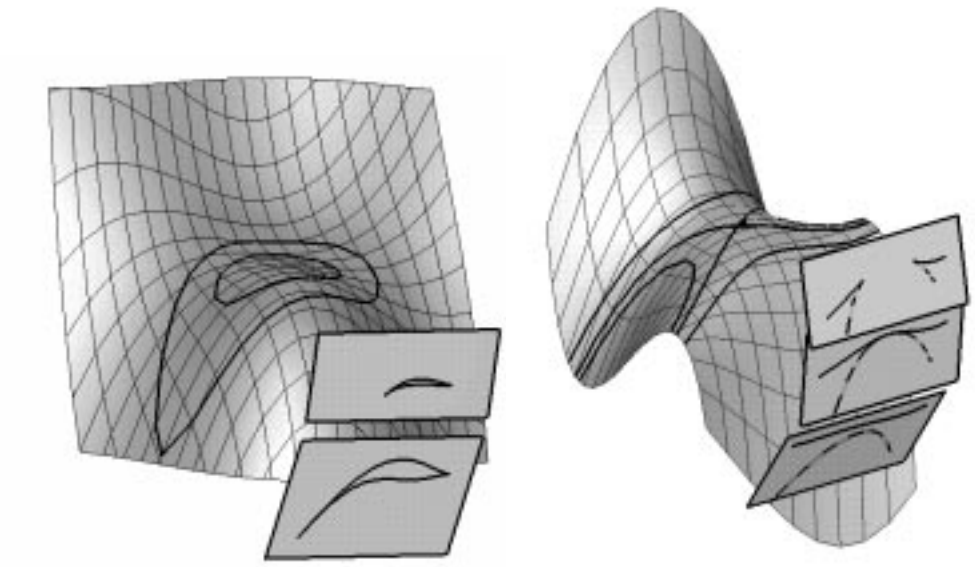

Figure 9. The 'lips' and 'beaks' visual events. The contour generators are singular (an isolated point in the lips transition; a crossing in the case of the beaks transition) when viewing a parabolic surface point along its asymptotic direction. A pairs of cusps are generated under viewer motion.

The frontier point projects to a point on the apparent contour in both views such that its tangent passes through the epipole. It is an epipolar tangency point since the tangent plane is also the epipolar plane.

The surface curvature can not be recovered by the epipolar parametrization at these points since the contour generator is locally stationary. However, frontier points correspond to real, fixed feature points on the surface which are visible in both views, once detected they can be used to provide a constraint on viewer motion. In fact they can be used in the same way as points in the recovery of the epipolar geometry via the epipolar constraint.

4. Recovery of viewer motion

(*a*) *Parametrization of the fundamental matrix*

The epipolar geometry between two uncalibrated views is completely determined by seven independent parameters: the position of the epipoles in the two views $(u_e, v_e, 1)^{\mathrm{T}}$ and $(u'_e, v'_e, 1)^{\mathrm{T}}$ and the three parameters of the homography relating the pencil of epipolar lines in view 1 to those in view 2,

$$\tau' = -\frac{h_2\tau + h_1}{h_4\tau + h3}, \tag{4.1}$$

where τ and τ' represent the directions of a pair of corresponding epipolar lines in the first and second images respectively. The transformation is fixed by three pairs of epipolar line correspondences (Luong & Faugeras 1996).

The epipolar geometry can be conveniently specified by the fundamental matrix, F (3×3 matrix defined up to an arbitrary scale and of rank two), such that the

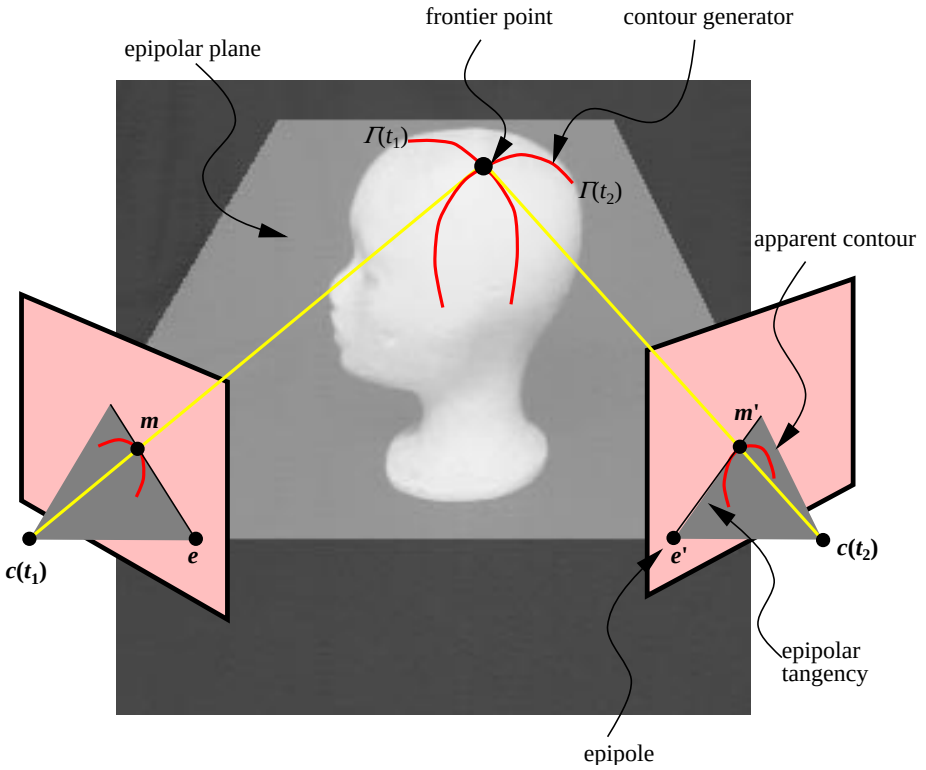

Figure 10. A frontier point appears the intersection of two consecutive contour generators and is visible in both views. The frontier point projects to a point on the apparent contour which is an epipolar tangency point.

image coordinates (projective representation) of a pair of corresponding points, $\boldsymbol{m}$ and $\boldsymbol{m}'$, must satisfy the epipolar constraint:

$$\boldsymbol{m}'^{\mathrm{T}}\boldsymbol{F}\boldsymbol{m} = 0 \tag{4.2}$$

and where the left and right epipoles ($\boldsymbol{e}$ and $\boldsymbol{e}'$) are given by the null space of F and F^{T} respectively.

This gives the following parametrization of the fundamental matrix:

$$\boldsymbol{F} = \begin{pmatrix} h_1 & h_2 & -u_e h_1 - v_e h_2 \\ h_3 & h_4 & -u_e h_3 - v_e h_4 \\ -u'_e h_1 - v' e h_3 & -u'_e h_2 - v' e h_4 & u_e u'_e h_1 + v_e u'_e h_2 + u_e v' e h_3 + v_e v' e h_4 \end{pmatrix}.$$

(*b*) *Finding the epipoles*

Under pure translation the epipolar geometry is completely determined by the position of the epipole in a single view. The position of the epipole is the same in both views if the intrinsic parameters do not change and the epipolar lines have the same directions and are *auto-epipolar*. The bitangents at two consecutive apparent contours are epipolar tangencies and hence the projection of frontier points. The intersection of at least two distinct tangencies (epipolar lines) determines the position of the epipole. See figure 13 (Sato & Cipolla 1998).

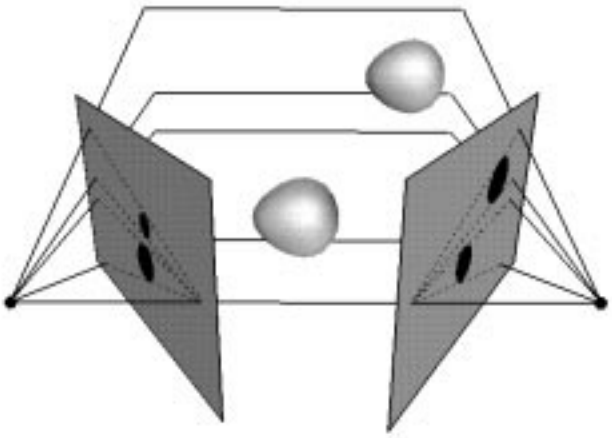

Figure 11. Epipolar geometry and epipolar tangencies under arbitrary motion.

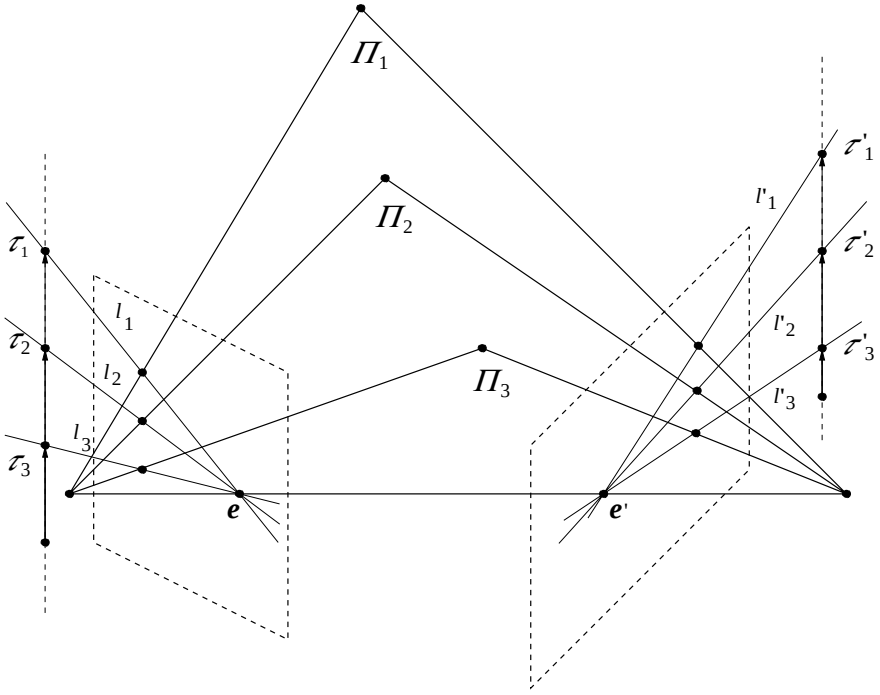

Figure 12. The epipolar geometry of an uncalibrated stereo pair of images is completely specified by the image positions of the epipoles and three pairs of corresponding epipolar lines. The projective parameters τ and τ' represent the intersection of the epipolar line and the line at infinity. The directions in the two views are related by a homography.

The solution is no longer trivial in the case of arbitrary motion with rotation. There is in fact no closed form solution since the epipoles are needed to define the epipolar tangency points (and frontier points) and these are needed to determine the epipoles.

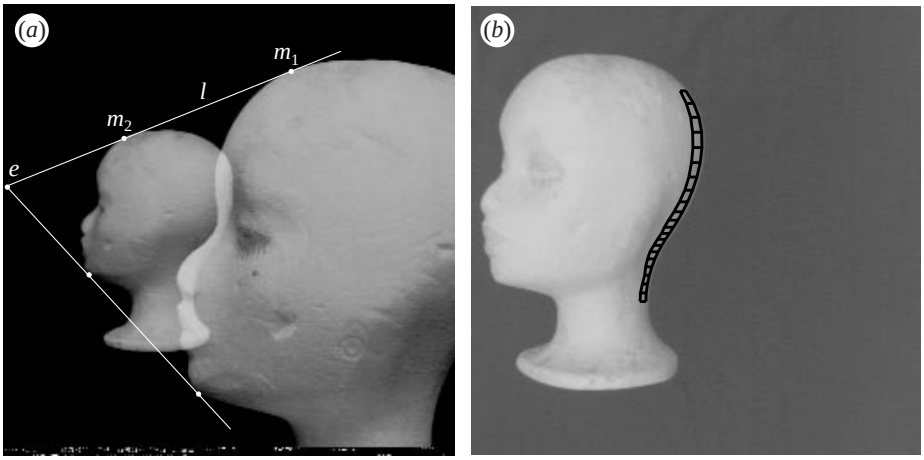

Figure 13. Under pure translation the frontier/epipolar tangency point moves along the epipolar line since the position of the epipole and the direction of the epipolar lines do not change. From a minimum of two bitangents of the apparent contour in two views (a) it is possible to recover the epipole, $\boldsymbol{e}$, and hence to reconstruct the strip of visible surface at the apparent contour (b).

(c) *Optimization*

The solution proceeds as a search and optimization problem to find the position of the epipoles in both views such that the epipolar tangencies in the first view are related to the set of epipolar tangencies in the second view by a one-dimensional homography (Cipolla *et al*. 1995).

A suitable cost function is needed. A geometric criterion (distance) is used in the estimation of the fundamental matrix from point correspondenes and can also be used in the case of curves. The geometric distance is computed as the sum over all tangency points of the square of the distance between the image point and the corresponding epipolar line from the tangency point in the other view.

The key to a successful implementation is to ensure that the search space is reduced and that the optimization begins from a good starting point using approximate knowledge of the camera motion or point correspondences. The solution proceeds as follows:

1. Start with an initial guess or estimate of the epipoles in both views.
2. Compute the epipolar tangencies, $\boldsymbol{m}(\boldsymbol{e})$ and $\boldsymbol{m}'(\boldsymbol{e}')$, in both views respectively. These are points on the apparent contours with tangents passing through the epipole.
3. Estimate the elements of the homography between the pencil of tangencies in both views. This can be done linearly by minimizing
$$\sum_i (h_4\tau_i\tau_i' + h_3\tau_i' - h_2\tau_i - h_1)^2 \tag{4.3}$$
by the method of least squares over all pairs of correspondences (τ and τ').
4. The fundamental matrix is now given by the parametrization above and the distance criterion (i.e. sum of squared distances between tangency point and

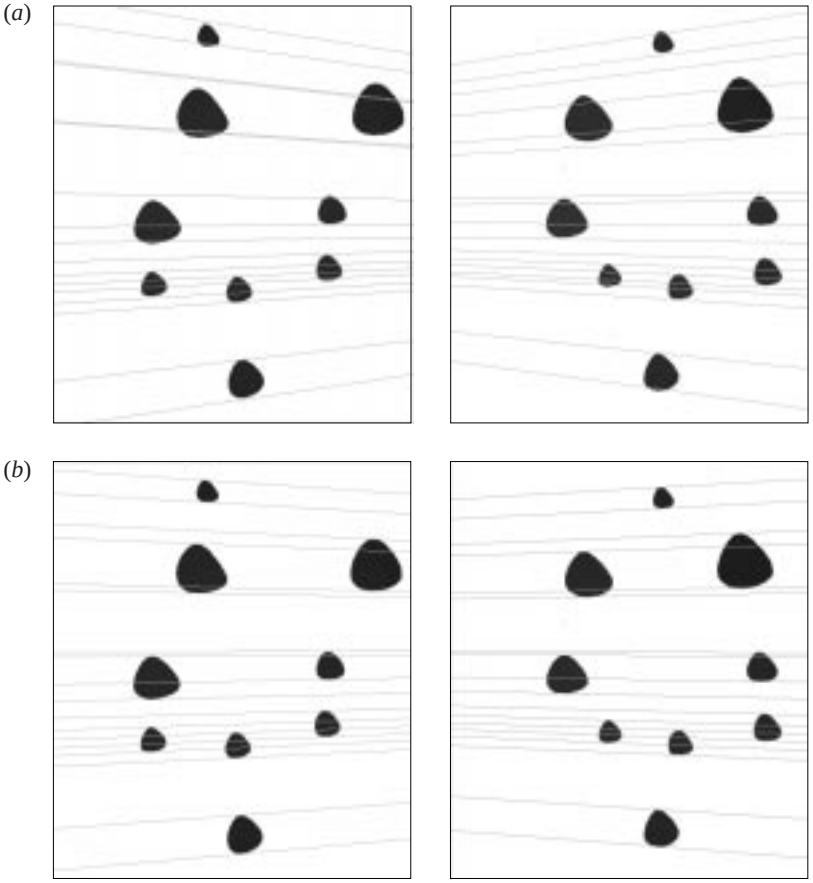

Figure 14. Starting point for optimization (a). An initial guess of the position of the epipoles is used to determine a finite set of epipolar tangencies in both views and the homography relating the two set of epipolar lines. For each tangency point the corresponding epipolar line is drawn in the other view. The distances between epipolar lines and tangecy points is used to search for the correct positions of the epipoles. Convergence to local minimum after five iterations (b). The epipolar lines are now tangent to apparent contours in both views.

corresponding epipolar line) can be computed:

$$C = \sum_i \left(\frac{1}{(\boldsymbol{F}\boldsymbol{m}_i)_1^2 + (\boldsymbol{F}\boldsymbol{m}_i)_2^2} + \frac{1}{(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{m}'_i)_1^2 + (\boldsymbol{F}^{\mathrm{T}}\boldsymbol{m}'_i)_2^2} \right) (\boldsymbol{m}_i'^{\mathrm{T}}\boldsymbol{F}\boldsymbol{m}_i)^2.$$

5. Minimize the distance by the conjugate gradient method. The search space is restricted to the four coordinates of the epipoles only. This requires the first-order partial derivatives of the cost function with respect to the coordinates of the epipoles which can be computed analytically but are more conveniently estimated by numerical techniques.

At each iteration of the algorithm, steps 1 to 4 are repeated, and the positions of the epipoles are refined. The search is stopped when the root-mean-square distance converges to a minimum (usually less than 0.1 pixels). It is of course not guaranteed to find a unique solution.

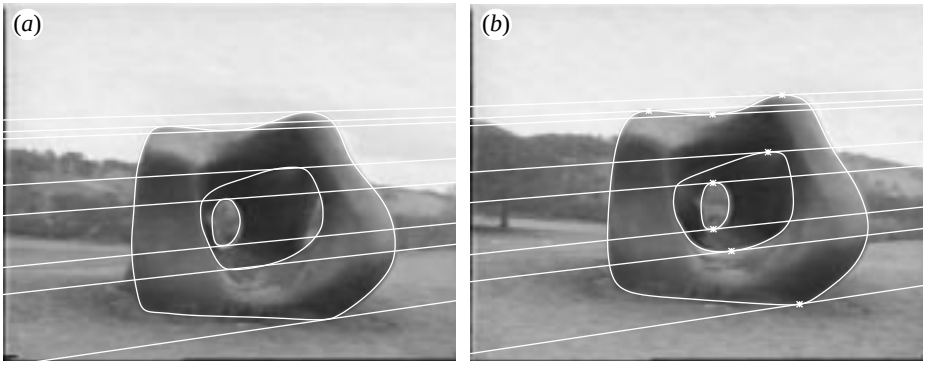

Figure 15. Local minimum obtained by iterative scheme to estimate the epipolar geometry from eight epipolar tangencies.

Several of the experiments were carried out with simulated data (with noise) and known motion (figure 14). The apparent contours were automatically extracted from the sequence by fitting B-splines to the image edge data (Cham & Cipolla 1996). Five to ten iterations each for four different initial guesses for the position of the epipole were sufficient to find the correct solution to within an root-mean-square error of 0.1 pixel per tangency point.

Figure 15 shows an example with real data whose apparent contours are detected and automatically tracked using B-splines snakes. A solution is found very quickly which minimizes the geometric distances but as with all structure from motion algorithms, a limited field of view and small variation in depths result in a solution which is sensitive to image localization errors.

5. Conclusions and future work

The recovery of the epipolar geometry between views is a key part of any algorithm to recover the 3D structure and motion compatible with the views. The structure and motion problem for curves and surfaces is more challenging since the apparent contours are viewpoint dependent and the correspondence of points between the two viewpoints in not given.

We have shown how the viewer motion can be recovered from the outlines (apparent contours) of curved surfaces by searching for epipolar tangency points. The results of initial experiments using these algorithms have been promising but the performance of the algorithm remains to be fully evaluated and it is still unclear whether the extraction of the motion leads to a unique solution. After computing the motion the epipolar geometry can be exploited to parametrize the apparent contours and to recover the visible surface.

We have only used apparent contours in the motion estimate. In practice one would use a combination of image features to estimate motion and could then use the apparent contours to reconstruct the surface. An important test of the usefulness of the proposed theories will be the accuracy of the reconstruction of an arbitrarily curved surface from uncalibrated viewer motion.

I acknowledge the support of the EPSRC. The research on following cusps and frontiers was carried out at the Isaac Newton Institute for Mathematical Sciences and the Department of

Engineering, Cambridge in collaboration with Karl Åström, Andrew Blake, Gordon Fletcher, Peter Giblin, Stefan Rahmann and Jun Sato. The figures of visual events were produced with the Liverpool Surfaces Modelling Package.

References

Åström, K., Cipolla, R. & Giblin, P. J. 1996 Generalised epipolar constraint. In *Proc. 4th European Conf. on Computer Vision, Cambridge, England*, vol. 2, pp. 97–108. Lecture Notes in Computer Science, vol. 1065. Springer.

Boyer, E. & Berger, M. O. 1997 3D Surface reconstruction using occluding contours. *Int. J. Computer Vision* **22**, 219–233.

Cham, T. J. & Cipolla, R. 1996 MDL-based curve representation using B-spline active contours. In *Proc. British Machine Vision Conf., Edinburgh*, pp. 363–372.

Cipolla, R. & Blake, A. 1992 Surface shape from the deformation of apparent contours. *Int. J. Computer Vision* **9**, 83–112.

Cipolla, R., Åström, K. & Giblin, P. J. 1995 Motion from the frontier of curved surfaces. In *Proc. IEEE 5th Int. Conf. on Computer Vision*, pp. 269–275. Boston.

Cipolla, R., Fletcher, G. J. & Giblin, P. J. 1997 Following cusps. *Int. J. Computer Vision* **23**, 115–129.

Faugeras, O. D., Luong, Q.-T. & Maybank, S. J. 1992 Camera self-calibration: theory and experiments. In *Proc. 2nd European Conf. on Computer Vision, Santa Margherita Ligure, Italy, May 1992*, pp. 321–334. Lecture Notes in Computer Science, vol. 588. Springer.

Giblin, P. J. & Weiss, R. 1987 Reconstruction of surfaces from profiles. In *Proc. 1st Int. Conf. on Computer Vision, London*, pp. 136–144.

Giblin, P. J. & Weiss, R. 1994 Epipolar fields on surfaces. In *Proc. 3rd European Conf. on Computer Vision, Stockholm, May 1994*, vol. 1, pp. 14–23. Lecture Notes in Computer Science, vol. 800. Springer.

Giblin, P. J., Pollick, F. E. & Rycroft, J. E. 1994 Recovery of an unknown axis of rotation from the profiles of a rotating surface. *J. Opt. Soc. Am.* A **11**, 1976–1984.

Koenderink, J. J. 1984 What does the occluding contour tell us about solid shape. *Perception* **13**, 321–330.

Koenderink, J. J. 1990 *Solid shape*. MIT Press.

Koenderink, J. J. & Van Doorn, A. J. 1976 The singularities of the visual mapping. *Biol. Cyber.* **24**, 51–59.

Koenderink, J. J. & Van Doorn, A. J. 1982 The shape of smooth objects and the way contours end. *Perception* **11**, 129–137.

Longuet-Higgins, H. C. 1981 A computer algorithm for reconstructing a scene from two projections. *Nature* **293**, 133–135.

Luong, Q.-T. & Faugeras, O. D. 1996 The fundamental matrix: theory, algorithms, and stability analysis. *Int. J. Computer Vision* **17**, 43–76.

Porrill, J. & Pollard, S. 1991 Curve matching and stereo calibration. *Image Vision Computing* **9**, 45–50.

Rieger, J. H. 1986 Three dimensional motion from fixed points of a deforming profile curve. *Opt. Lett.* **11**, 123–125.

Sato, J. & Cipolla, R. 1998 Affine reconstruction of curved surfaces from uncalibrated views of apparent contours. In *Proc. IEEE 6th Int. Conf. on Computer Vision, Bombay*, pp. 715–720.

Vaillant, R. & Faugeras, O. D. 1992 Using extremal boundaries for 3D object modelling. *IEEE Trans. Pattern Recognition Machine Intell.* **14**, 157–173.

Discussion

H. C. Longuet-Higgins (*Laboratory of Experimental Psychology, University of Sussex, Brighton, UK*). From what has been said, one might think that the problems are solved, but one cannot help feeling that there are many 'nasties' in the real world which will defy a neat analysis. An example would be an irregular scene such as a hedge, from which it is difficult to extract tangents. How restrictive are the conditions under which this analysis holds?

R. Cipolla. We must be able to fit B-splines to the curves and surfaces that we are looking at and be able to give them a geometric representation. They must therefore be projections of some geometric entities.

P. Giblin (*Department of Mathematical Sciences, University of Liverpool, UK*). That is also only half the problem. In most cases we do not know if the recovery is unique. Only in the special cases of circular motion and parallel projection is there any rigorous theory which can show uniqueness. All my attempts at proving uniqueness in the general case have failed. One is very likely to get stuck in local minima.

R. Cipolla. Yes, indeed it is much worse than extracting structure from motion from points. With curves the problem is compounded.

P. H. S. Torr (*Department of Engineering Science, University of Oxford, UK*). Given that the work details recovery of the motion using the occluding contour or outline of smooth or differentiable surfaces, how is it determined whether or not an apparent contour observed arises from a smooth surface or a crease?

R. Cipolla. The key idea here is that the parametrization depends upon the epipolar geometry, which is the same for creases, edges of polyhedral objects or curved surfaces. So, we automatically get some measure of curvature; if it is a crease, it will have infinite Gauss curvature.

O. Faugeras (*INRIA, France*). I have a question relating to the numerical stability of the process. Is there, for the occluding contours, any analogue to the critical surfaces that we know to exist for the fundamental matrix F, and what can we say about the stability of the reconstruction?

R. Cipolla. Currently, the main instability we experience is due to poor localization. However, I am sure that it is the case that there exist some surfaces for which degeneracy results if reconstruction is attempted for contour generators lying on such surfaces.

P. Giblin. That is a very good and natural question, but we currently don't know anything about the possible existence of such surfaces.

A. Fitzgibbon (*Department of Engineering, University of Oxford, UK*). What is the minimum structure needed for reconstruction?

R. Cipolla. The requirements are the same as those for reconstruction from points, i.e. for pure translation we would need two epipolar tangencies, for weak perspective, four, and for uncalibrated perspective, seven.

Figure 16. (*a*) A cusped apparent contour, with the occluded part dashed. (*b*) Two cusps and a T-junction which can merge into one at a 'swallowtail transition', when the viewline has an extra degree of contact with the surface (4 point contact altogether).

G. Hunter (*St Mary's University College, Twickenham, UK*). Has anyone considered the corresponding problem for binocular vision? Although the mathematics would become more complicated, one would get some impression of depth.

R. Cipolla. Jean Ponce has worked on the trinocular case. But what I have presented here is actually the motion from stereo views. It is only when it comes to reconstruction that one needs the information from three views or a knowledge of the infinitesimal accelerations.

A. Hopper (*The Olivetti & Oracle Research Laboratory, Cambridge, UK*). What is the stability with respect to recognizable targets? For example, if some target was painted on the object, would this make it easier for the algorithms, or would it make no difference?

R. Cipolla. It is obviously much easier to track and reconstruct targets. Often with targets it is not just points which are important, commercial companies use reflective spheres. In this case we should still be looking at the motion of apparent contours but the distances are often such as to make it reasonable to consider only the centroids of the contours to track.

H. C. Longuet-Higgins. Is it possible to distinguish cusps from T-junctions produced by occlusions?

P. Giblin. The two cases are different but related. I can best explain by means of a drawing (figure 16). When two cusps are about to merge and disappear in a 'swallowtail transition' there is indeed a T-junction close to the positions of the two cusps. Otherwise at a cusp there is an actual ending of the contour, without the presence of a second branch as in the T-junction.

K. Stark (*Technical University, Dresden, Germany*). How is the minimum number of B-spline control points decided upon?

R. Cipolla. For a fixed number of control points one always wants to minimize the error between the B-spline and the edge data. In general then, one would approximate a curve segment by straight lines and then for each line increase the number of control points until the error no longer decreases. Some recent work by T.-J. Cham has looked at this problem of choosing the optimum number and position of control points using the MDL criterion.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

M. Isard (*Robotics Group, University of Oxford, UK*). Does the original image have to be chosen carefully to get the right number of features? For example, if the features come and go, there will be too few control points.

R. Cipolla. Yes, there are big problems using B-splines as snakes. Basically, one has to keep on reinitializing on events. It is always easy to add more control points but very hard to cope with major topological changes.

T. Ihle (*Technical University, Dresden, Germany*). I have a question concerning the detection of frontier points. In the case of translation and rotation of the camera we cannot use the bitangency algorithm to detect the frontier points. In another case of a screw transformation the frontier points may not be defined between the two images. Can these two different cases be distinguished and exploited?

P. Giblin. We have problems if we don't know the camera motion. In this case a search (with an initial guess at the epipole) is required each time—rather an inelegant solution.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES