Robust Estimation of Surface Curvature from Deformation of Apparent Contours

by

Andrew BLAKE and Roberto CIPOLLA
Robotics Research Group

October 1989
Report No. OUEL 1797/89

University of Oxford
Department of Engineering Science
Parks Road
Oxford OX1 3PJ
England

Tel. +44 (0)865 273000

Robust Estimation of Surface Curvature from Deformation of Apparent Contours

Andrew BLAKE Roberto CIPOLLA
Robotics Research Group
University of Oxford

October, 1989

Abstract

Surface curvature along extremal boundaries is potentially useful information for navigation, grasping and object identification tasks. Previous theories have shown that such qualitative information about curvature can be obtained from a static view. Furthermore it is known that, for orthographic projection, under planar viewer-motion, quantitative curvature information is available from spatio-temporal derivatives of flow. This theory is extended here to arbitrary curvilinear viewer-motion and perspective projection.

We show that curvatures can actually be computed this way in practice, but that they are highly sensitive to errors in viewer-motion estimates. Intuitively, *relative* or *differential* measurements of curvature might be far more robust. Rather than measuring the absolute deformation of an apparent contour, differential quantities depend on the rate at which surface features are swept over an extremal boundary as the viewer moves. It is shown that, theoretically, such differential quantities are indeed far less sensitive to uncertainty in viewer-motion. *Ratios* of differential measurements are less sensitive still. In practice sensitivity is reduced by about two orders of magnitude. We believe this represents a significant step in the development of practical techniques for robust, qualitative 3D vision.

1 Introduction

The deformation of an apparent contour (the silhouette of a smooth surface) under viewer-motion is a potentially rich source of geometric information for navigation, motion-planning and object-recognition. Barrow and Tenenbaum [*Barrow78*] pointed out that surface orientation along an extremal boundary [1] can be computed directly from image data. Koenderink [*Koenderink84*] related the curvature of an apparent contour to the intrinsic curvature of the surface (Gaussian curvature); the sign of Gaussian curvature is equal to the sign of the curvature of the contour. Convexities, concavities and inflections of an apparent contour indicate, respectively, convex, hyperbolic and parabolic surface points. Giblin and Weiss [*Giblin87*] have extended this by adding viewer motions to obtain quantitative estimates of Gaussian and mean curvature. A surface can be reconstructed from the envelope of all its tangent planes, which in turn are computed directly from the family of silhouettes of the surface, obtained under planar motion of the viewer. By assuming that the viewer follows a *great circle* of viewer directions around the object they restricted the problem of analysing the envelope of tangent planes (a 2-parameter family) to the less general one of computing the envelope of a family of lines in a plane. Their algorithm was tested on noise-free, synthetic data (on the assumption that extremal boundaries had been distinguished from other image contours) demonstrating the reconstruction of a planar curve under orthographic projection.

This paper extends these theories further, to the general case of arbitrary non-planar camera motion under perspective projection. The Gaussian curvature of a surface, at a point on its silhouette, can be computed given some known local motion of the viewer. Curvature is computed from spatio-temporal derivatives (up to second order) of image-measurable quantities. The theory can, of course, be applied to detect extremal boundaries and distinguish them from surface markings or discontinuities. Experiments show that, with adequate viewer-motion calibration, itself computed from visual data [*Tsai87*], it is possible to obtain curvature measurements of useful accuracy.

A consequence of the theory, representing an important step towards qualitative vision, concerns the robustness of *differential* measurements of curvature at two nearby points. Intuitively it is relatively difficult to judge, moving around a smooth, featureless object, whether its silhouette is extremal or not — whether the Gaussian curvature along the contour is bounded or not. This judgment is much easier to make for objects with feature-rich surfaces. Under small viewer-motions, features are "sucked" over the extremal boundary, at a rate which depends on surface curvature. Our theory reflects this intuition exactly. It is shown that differential measurements of curvature across two adjacent points are entirely immune

[1] Consider a smooth rigid body for which the tangent plane is defined at every point of the surface. For each vantage point all the rays through the vantage point that are tangent to the surface can be constructed. They touch the object along a smooth curve on its surface called the *extremal boundary* [*Barrow78*] but which is alternatively known as the *contour generator* [*Marr77*], the *rim* [*Koenderink84*] or the *critical set* of a surface whose visual mapping is singular [*Giblin87*]. Its projection is known as the apparent contour, profile, outline or silhouette.

to uncertainties in the viewer's rotational velocity. This is somewhat related to earlier results showing that differential measurements of this kind are important for depth measurement from optic flow [*LHiggins80,Weinshall88*] and for curvature measurements from stereoscopically viewed highlights [*Blake88*]. Furthermore, they are relatively immune to uncertainties in translational motion in that, unlike single-point measurements, they are independent of the viewer's acceleration. Only dependence on velocity remains. Experiments show that this theoretical prediction is borne out in practice. Differential curvature measurements prove to be more than an order of magnitude less sensitive than single-point measurements to errors in viewer-motion calibration. There is some theoretical evidence that *ratios* of differential curvature measurements are less sensitive by about another order of magnitude. In our experiments absolute measurements of curvature were so sensitive that they became unreliable for viewer motion errors of 0.5mm in position and 1mrad in orientation. For ratios of differential measurements of curvature the corresponding figures were about 50mm and 70mrad respectively.

2 Theoretical framework

The following theory relates the geometry of apparent contours and their deformation under viewer-motion to the geometry of the observed object. The use of spherical projection (rather than planar), which has previously proven to be a powerful tool in structure-from-motion [*Maybank85*], makes it feasible to extend the theory of Giblin and Weiss [*Giblin87*] to allow for perspective. A significant technical issue is the choice of spatio-temporal parameterisation. There is no unique parameterisation because the mapping between contour-generators (extremal boundaries) at successive instants is underdetermined. This is essentially a form of the "aperture problem".

2.1 General properties of the extremal boundary and its projection

Generically (that is, given stability under arbitrarily small perturbations such as a small excursion of the vantage point) the extremal boundary is a smooth space curve, in general non-planar, whose tangent is everywhere defined. However, its projection (the apparent contour) may not necessarily be smooth everywhere. For generic situations there may exist a finite number of rays for which the extremal boundary and the view direction are aligned. At these points the apparent contour has cusps. For opaque surfaces only one branch of the cusp is visible, however, corresponding to a contour-ending [*Koenderink82,Koenderink84*].

2.2 Surface Geometry

Consider a point P on the extremal boundary of a smooth, curved surface in R^3 and parameterised locally by a vector valued function $\mathbf{r}(s,t)$. The parametric representation can be considered as covering the surface with 2 families of curves: $\mathbf{r}(s,t_0)$, and $\mathbf{r}(s_0,t)$ where s_0, t_0

are fixed for a given curve in the family. A one-parameter family of views is indexed by the time parameter t and s, t are defined so that the s-parameter curve, $\mathbf{r}(s, t_0)$, is an extremal boundary for a particular view t_0. A t-parameter curve $\mathbf{r}(s_0, t)$ can be thought of as the 3D locus of points grazed by a light-ray from the viewer, under viewer-motion. Given a starting point $s = s_0$, $t = t_0$, such a locus is not uniquely defined because the correspondence, as the viewer moves, between "successive" (in an infinitessimal sense) extremal boundaries is not unique. Hence there is considerable freedom to choose a parameterisation $\mathbf{r}(s, t)$.

Local surface geometry is specified in terms of the basis $\{\mathbf{r}_s, \mathbf{r}_t\}$ for the tangent plane ($\mathbf{r}_s$ and $\mathbf{r}_t$ denote $\partial\mathbf{r}/\partial s$ and $\partial\mathbf{r}/\partial t$ respectively). The surface normal is a unit vector $\mathbf{n}$ (figure 1) satisfying

$$\mathbf{r}_s.\mathbf{n} = 0 \tag{1}$$

$$\mathbf{r}_t.\mathbf{n} = 0. \tag{2}$$

Surface curvature is represented by the first and second fundamental forms [*DoCarmo76*] [*Millman77*]. The first fundamental form (with coefficients represented by a matrix $\mathbf{I}$)

$$\mathbf{I} = \begin{bmatrix} \mathbf{r}_t.\mathbf{r}_t & \mathbf{r}_t.\mathbf{r}_s \\ \mathbf{r}_s.\mathbf{r}_t & \mathbf{r}_s.\mathbf{r}_s \end{bmatrix} \tag{3}$$

is used to express the length of any infinitessimal element in the tangent plane. The second fundamental

$$\mathbf{II} = \begin{bmatrix} \mathbf{r}_{tt}.\mathbf{n} & \mathbf{r}_{ts}.\mathbf{n} \\ \mathbf{r}_{st}.\mathbf{n} & \mathbf{r}_{ss}.\mathbf{n} \end{bmatrix} \tag{4}$$

form quantifies the "bending away" of the surface from the tangent plane. The geometry of the surface is uniquely determined locally up to a rigid motion in R^3 by these two quadratic forms and our aim is therefore to make these and other intrinsic descriptions (such as Gaussian curvature) explicit from image-measurable quantities.

2.3 Imaging model

A general model of the imaging device is to consider it as determining the unit vector, $\mathbf{T}$ of an incoming ray. This is effectively a spherical pin-hole camera of unit radius (figure 3). Its simplicity arises from the fact that there are no special points on the image surface, whereas the origin of the perspective plane is special and the consequent loss of symmetry tends to complicate mathematical arguments.

For a given vantage position t_0 the image of the extremal boundary is a continuous family of rays $\mathbf{T}(s, t_0)$ emanating from the camera's optical centre which are tangent to the surface:

$$\mathbf{T}.\mathbf{n} = 0. \tag{5}$$

The moving observer at position $\mathbf{v}(t)$ sees a 2 parameter family of apparent contours which we represent as a family of unit view vectors $\mathbf{T}(s, t)$. As before with $\mathbf{r}(s, t)$, the parameterisation

is underdetermined (an "aperture problem") – to be fixed later. At any given position, the observer measures the direction $\mathbf{T}$ of a ray from image measurements and knowledge of the orientation of the camera co-ordinate system relative to the world frame. The ray touches the surface at point P (figure 2), whose position vector is

$$\mathbf{r}(s,t) = \mathbf{v}(t) + \lambda \mathbf{T}(s,t), \tag{6}$$

where λ is the distance along the ray to the point P. The tangent to the extremal boundary is

$$\mathbf{r}_s = \lambda_s \mathbf{T} + \lambda \mathbf{T}_s \tag{7}$$

and which must lie in the tangent plane of the surface:

$$\mathbf{r}_s.\mathbf{n} = 0.$$

From (1),(5) and (7), we see that

$$\mathbf{T}_s.\mathbf{n} = 0 \tag{8}$$

and the orientation of the surface normal is (figure 3)

$$\mathbf{n} = \frac{\mathbf{T} \wedge \mathbf{T}_s}{|\mathbf{T} \wedge \mathbf{T}_s|}. \tag{9}$$

This is a generalisation to perspective projection of the well-known result of Barrow and Tenenbaum [*Barrow78,Barrow81*].

2.4 Conjugate relationship of view vector and extremal boundary

It is well known that, for orthographic projection, the view direction and the extremal boundary are in *conjugate* directions with respect to the second fundamental form [*Koenderink76*] [*Koenderink84*]. This is immediately apparent in the current framework since the second fundamental form has the property that

$$II(\mathbf{r}_s, \mathbf{T}) = -\mathbf{T}.\mathbf{n}_s$$

which, by differentiating (5) and substituting (8), is zero.

$$\mathbf{T}.\mathbf{n}_s = 0 \tag{10}$$

The view direction and the extremal boundary are *conjugate* but not in general perpendicular. They are perpendicular only if the ray $\mathbf{T}$ is along a *principal* direction.

2.5 Recovery of depth

Depth can be computed from the 1st derivative of the deformation of the apparent contour under known viewer motion [*Bolles87*]. The contact point of a ray, P, moves with velocity

$$\mathbf{r}_t = \mathbf{v}_t + \lambda_t \mathbf{T} + \lambda \mathbf{T}_t \tag{11}$$

Since $\mathbf{r}_t$ lies in the tangent plane of the surface at P:

$$\mathbf{r}_t.\mathbf{n} = 0$$

so that depth (distance along the ray, λ) is given by

$$\lambda = -\frac{\mathbf{v}_t.\mathbf{n}}{\mathbf{T}_t.\mathbf{n}} \tag{12}$$

This formula is an infinitessimal analogue of triangulation with stereo cameras. The numerator is analogous to baseline and the denominator to disparity. The result also holds for a rigid space curve or an occluding boundary – not surprisingly since we claimed that second but not first derivatives of $\mathbf{T}$ are sensitive to curvature.

2.6 Recovering the first and second fundamental forms

Given an arbitrary choice of $\mathbf{T}_t$, $\mathbf{n}_t$ or $\mathbf{r}_t$, the coefficients of the first and second fundamental forms can be expressed in terms of image-measurable quantities, $\mathbf{T}$ and its spatio-temporal derivatives. The elements of the first fundamental form can be determined from (7), (11) and (12). Taking also their partial derivatives and dot products with $\mathbf{n}$ and using (8) and (10) determines the coefficients of the second fundamental form:

$$\begin{aligned}
\mathbf{r}_{ss}.\mathbf{n} &= \lambda\mathbf{T}_{ss}.\mathbf{n} = -\lambda\mathbf{T}_s.\mathbf{n}_s = \lambda\mathbf{T}.\mathbf{n}_{ss} \\
\mathbf{r}_{tt}.\mathbf{n} &= \lambda\mathbf{T}_{tt}.\mathbf{n} + 2\lambda_t\mathbf{T}_t.\mathbf{n} + \mathbf{v}_{tt}.\mathbf{n} = \lambda\mathbf{T}.\mathbf{n}_{tt} - \mathbf{v}_{tt}.\mathbf{n} - 2\mathbf{v}_t.\mathbf{n}_t \\
\mathbf{r}_{st}.\mathbf{n} &= \lambda\mathbf{T}_{st}.\mathbf{n} + \lambda_s\mathbf{T}_t.\mathbf{n} = \lambda\mathbf{T}.\mathbf{n}_{st} - \mathbf{v}_t.\mathbf{n}_s
\end{aligned} \tag{13}$$

2.7 Choice of Parameterisation

We can consider the problem of choosing a parameterisation as an "aperture problem" for contours on the spherical perspective image ($\mathbf{T}(s,t)$) or on the Gauss sphere ($\mathbf{n}(s,t)$), or between space curves on the surface ($\mathbf{r}(s,t)$) (figure 4).

A natural parameterisation is the "epipolar parameterisation" defined by

$$\mathbf{r}_t \wedge \mathbf{T} = 0 \tag{14}$$

or, equivalently,

$$(\mathbf{r}_t \wedge \mathbf{T}) \wedge \mathbf{T} = 0 \tag{15}$$

for which the tangent-plane basis vectors $\mathbf{r}_s$ and $\mathbf{r}_t$ are in conjugate directions. Substituting for $\mathbf{r}_t$ from (11)

$$(\mathbf{v}_t \wedge \mathbf{T}) \wedge \mathbf{T} + \lambda(\mathbf{T}_t \wedge \mathbf{T}) \wedge \mathbf{T} = 0 \tag{16}$$

leading to the "matching" condition

$$\mathbf{T}_t = \frac{(\mathbf{v}_t \wedge \mathbf{T}) \wedge \mathbf{T}}{\lambda}. \tag{17}$$

Points on different contours are "matched" by moving along great-circles on the image sphere with poles defined by the direction of the viewer's velocity $\mathbf{v}_t$ (figure 5). This induces a *correspondence* on the surface between extremal boundaries from different viewpoints. If the motion is linear corresponding points on the image sphere will lie on an epipolar great-circle. This is equivalent to Epipolar Plane matching in stereo. For a general motion, however, the epipolar structure rotates continuously as the direction of $\mathbf{v}_t$ changes and the space curve, $\mathbf{r}(s_0, t)$, generated by the movement of a contact point will be non-planar.

There are two possible cases of degeneracy of the parameterisation, occurring when $\{\mathbf{r}_s, \mathbf{r}_t\}$ fails to form a basis for the tangent plane:

1. $\mathbf{r}_t = 0$
   The contour is not an extremal boundary but a 3D rigid space curve (surface marking or discontinuity in depth or orientation). The parameterisation degrades gracefully and hence this condition poses no special problems.

2. $\mathbf{r}_s \wedge \mathbf{r}_t = 0$ and $\mathbf{r}_t \neq 0$
   For a special viewing geometry in which the view direction $\mathbf{T}$ is along an asymptotic direction of a hyperbolic patch on the surface, the tangent to the extremal boundary and the view direction are aligned. This is because the asymptotic directions at a point on a hyperbolic patch on the surface are self-conjugate [*DoCarmo76*].[2] A cusp is generated in the projection of the extremal boundary which appears in the image as a contour-ending [*Koenderink82,Koenderink84*].

2.8 Determination of local curvature

Substituting the matching constraint of (17) into (11) and using (5) we obtain:

$$\mathbf{r}_t = (\lambda_t + \mathbf{T}.\mathbf{v}_t)\mathbf{T} \tag{18}$$

$$\mathbf{r}_{tt}.\mathbf{n} = (\lambda_t + \mathbf{T}.\mathbf{v}_t)\mathbf{T}_t.\mathbf{n} \tag{19}$$

The physical interpretation of equation (18) is that the contact (grazing) point on an extremal boundary moves/slips along the line of sight $\mathbf{T}$ with a speed given by $|\mathbf{r}_t|$.

The normal curvature κ^t, of the space curve, $\mathbf{r}(s_0, t)$, at P is given by:

$$\kappa^t = \frac{\mathbf{r}_{tt}.\mathbf{n}}{|\mathbf{r}_t|^2} \tag{20}$$

$$= \frac{\mathbf{T}_t.\mathbf{n}}{(\lambda_t + \mathbf{T}.\mathbf{v}_t)} \tag{21}$$

The numerator is analogous to stereo disparity (as appeared in the denominator of the depth formula (12)) and depends only on the distance of the contact point and the "stereo baseline".

[2] For an asymptotic direction $II(\mathbf{T}, \mathbf{T}) = -\mathbf{T}.\mathbf{n}_t = \mathbf{T}_t.\mathbf{n} = 0$ and the view direction is self-conjugate. The tangent to the extremal boundary and the view direction are also conjugate. These conditions are satisfied if the view direction and the tangent to the extremal boundary are parallel.

The denominator is the speed of the contact point and is therefore inversely proportional to the curvature (normal) of the space curve generated as the viewer moves in time. It "clings" to points with high curvature and speeds up as the curvature is reduced. This property will be used later to distinguish surface markings or creases from extremal boundaries.

Gaussian curvature at a point is defined to be the product of the *principal* curvatures [*DoCarmo76*]. With the epipolar parameterisation , Gaussian curvature can also be expressed as a product of two curvatures: the normal curvature κ^t of the space curve generated by the motion of a contact-point and the geodesic curvature of the image (under spherical projection) of the extremal boundary scaled by inverse-depth. In general the Gaussian curvature can be expressed as the ratio of the determinants of the matrices of coefficients of the second and first fundamental forms:

$$\kappa_{gauss} = \frac{|II|}{|I|} \tag{22}$$

It is easy to show that, as a consequence of choosing the epipolar parameterisation, the second fundamental form is diagonal. The off-diagonal components are both equal to

$$\mathbf{r}_{ts}.\mathbf{n} = -\mathbf{r}_t.\mathbf{n}_s = -|\mathbf{r}_t|\mathbf{T}.\mathbf{n}_s = 0.$$

Consequently the expression for Gaussian curvature is simplified. From (22),

$$\begin{aligned} \kappa_{gauss} &= \frac{(\mathbf{r}_{tt}.\mathbf{n})(\mathbf{r}_{ss}.\mathbf{n})}{(\mathbf{r}_t.\mathbf{r}_t)(\mathbf{r}_s.\mathbf{r}_s) - (\mathbf{r}_t.\mathbf{r}_s)^2} \\ &= \left(\frac{\mathbf{r}_{tt}.\mathbf{n}}{\mathbf{r}_t.\mathbf{r}_t}\right)\left(\frac{\mathbf{r}_{ss}.\mathbf{n}}{\mathbf{r}_s.\mathbf{r}_s}\right) / \left(1 - (\mathbf{T}.\frac{\mathbf{r}_s}{|\mathbf{r}_s|})^2\right) \qquad (23) \\ &= \kappa^t\kappa^s / \left(1 - (\mathbf{T}.\frac{\mathbf{r}_s}{|\mathbf{r}_s|})^2\right) \qquad (24) \end{aligned}$$

where κ^t and κ^s are the normal curvatures of the space curves, $\mathbf{r}(s_0, t)$ and $\mathbf{r}(s, t_0)$ respectively, and the denominator is a scaling factor - the square of the sine of the angle between the view vector $\mathbf{T}$ and the tangent $\mathbf{r}_s$ to the extremal boundary.

Gaussian curvature can be expressed even more simply in terms of the curvature in the image – more precisely the geodesic curvature [3], κ^p, of the apparent contour on the projection sphere, $\mathbf{T}(s_0, t)$.

$$\kappa^p = \frac{\mathbf{T}_{ss}.\mathbf{n}}{|\mathbf{T}_s|^2} \tag{25}$$

Using equation (7) to express $\mathbf{T}_s$ in terms of surface basis vectors $\mathbf{r}_t$ and $\mathbf{r}_s$, the geodesic curvature of the apparent contour can be expressed as:

$$\kappa^p = \frac{\lambda^2\mathbf{T}_{ss}.\mathbf{n}}{(\mathbf{r}_s.\mathbf{r}_s) - (\mathbf{T}.\mathbf{r}_s)^2} \tag{26}$$

[3]The curvature, κ, and the Frenet-Serret normal, N, for a space curve $\mathbf{T}(s)$ are given by [*Faux79, p103*] : $\kappa\mathbf{N} = (\mathbf{T}_s \wedge \mathbf{T}_{ss}) \wedge \mathbf{T}_s / |\mathbf{T}_s|^4$. The normal curvature is the magnitude of the component of $\kappa\mathbf{N}$ in the direction of the surface normal (here $\mathbf{T}$ since $\mathbf{T}(s, t_0)$ is a curve on the image sphere); the geodesic curvature is the magnitude of the component in a direction perpendicular to the curve tangent ($\mathbf{T}_s$) and the surface normal.

It can also be written in terms of the extremal boundary's normal curvature:

$$\kappa^p = \lambda \frac{\kappa^s}{1 - (\mathbf{T}.\frac{\mathbf{r}_s}{|\mathbf{r}_s|})^2} \tag{27}$$

since from equation (13):

$$\begin{aligned} \kappa^s &= \frac{\mathbf{r}_{ss}.\mathbf{n}}{\mathbf{r}_s.\mathbf{r}_s} \\ &= \frac{\lambda \mathbf{T}_{ss}.\mathbf{n}}{\mathbf{r}_s.\mathbf{r}_s} \end{aligned} \tag{28}$$

Equation (27) shows that an apparent contour is smooth except where the parameterisation is degenerate (see earlier). Its curvature has the same sign as the normal curvature κ^s of the extremal boundary. A similar result was derived by Brady et al for orthographic projection [*Brady85*] and Verri and Yuille [*Verri85*] for perspective projection.

Now equations (23) and (27) can be used to derive Koenderink's simple expression for Gaussian curvature:

$$\kappa_{gauss} = \frac{\kappa^p \kappa^t}{\lambda} \tag{29}$$

We can express this purely in terms of image measurables and their spatio-temporal derivatives, up to second order:

$$\kappa_{gauss} = \left(\frac{\mathbf{T}_{ss}.\mathbf{n}}{|\mathbf{T}_s|^2}\right)\left(\frac{1}{\lambda}\right)\left(\frac{\mathbf{T}_t.\mathbf{n}}{\lambda_t + \mathbf{T}.\mathbf{v}_t}\right) \tag{30}$$

where

$$\lambda_t + \mathbf{T}.\mathbf{v}_t = \frac{(\mathbf{T}_{tt}.\mathbf{n})(\mathbf{v}_t.\mathbf{n}) + 2(\mathbf{T}.\mathbf{v}_t)(\mathbf{T}_t.\mathbf{n})^2 - (\mathbf{v}_{tt}.\mathbf{n})(\mathbf{T}_t.\mathbf{n})}{(\mathbf{T}_t.\mathbf{n})^2} \tag{31}$$

and

$$\lambda = -\frac{\mathbf{v}_t.\mathbf{n}}{\mathbf{T}_t.\mathbf{n}} \tag{32}$$

Since the sign of κ^t and the depth λ are always positive at an extremal boundary, a corollary of equation (29) is the well known result that the sign of Gaussian curvature is equal to the sign of the curvature of the projection [*Koenderink84*,*Brady85*].

2.9 Experimental Results: Determining curvatures from absolute measurements

Figure 6 shows 3 views from a sequence of a scene taken from a camera mounted on a moving robot-arm whose motion has been accurately calibrated from visual data [*Tsai87*].

Using a numerical method for estimating surface curvatures from 3 discrete views (see Appendix) we estimate the radius of curvature R for both a point on an extremal boundary of a cup,B, and a surface marking, A.

	measured (mm)	actual (mm)	error (mm)
surface marking A	1.95	0.0	1.95
extremal boundary B	45.7	44.4	1.3

Table 2. Estimating surface curvatures from 3 distinct views of a point on a surface marking and a point on an extremal boundary.

The radius of the cup was measured using calipers as 44.4 $\pm$ 2mm. A surface marking can be considered as a point with infinite curvature and hence ideally will have a radius of curvature of 0.0mm. The estimated curvatures agree with the actual curvatures. However, the results are very sensitive to perturbations in the motion parameters. Figures 7a shows how an error of 1mm (in a translation of 55.3mm) in determining the position of the camera centre for view 2 produces an error of about 190% in the estimate of curvature for the point on the extremal boundary. Similarly the estimate of curvature is very sensitive to errors in the orientation of the camera (figure 8a). In particular a perturbation in rotation about an axis defined by the epipolar plane of 1mrad (in a rotation of 178 mrad) produces an error of 70% in the estimate of the radius of curvature.

3 Differential measurement of curvature

We have seen that although it is perfectly feasible to compute curvature from the observed deformation of an apparent contour, the result is highly sensitive to motion calibration. This may be acceptable for a moving camera mounted on a precision robot-arm or when a grid is in view so that accurate visual calibration can be performed. In such cases it is feasible to determine motion to the accuracy of around 1 part in 1000 that is required. However, when only crude estimates of motion are available another strategy is called for. It is sometimes possible in such a case to use the crude estimate to bootstrap a more precise visual egomotion computation [*Harris87*]. However this requires an adequate number of identifiable corner features, which may not be available in an unstructured environment. Moreover, if the estimate is too crude the egomotion computation may fail; it is notoriously ill-conditioned [*Tsai84*] and is in danger of converging to an incorrect solution.

The alternative approach is to seek qualitative measurements of geometry that are much less sensitive to error in the motion estimate. In this section we show that differential measurements of curvature have just this property. Differential measurements based on two points are insensitive to errors in rotation and in translational acceleration. Typically, the two features might be one point on an extremal boundary and one fixed surface point. The surface point has infinite curvature and therefore acts simply as a stable reference point for the measurement of curvature at the extremal boundary. Intuitively the reason for the insensitivity of differential curvature is that it is a function only of the *differential* motion of image features. Thus global additive errors in motion measurement have no effect.

Consider two visual features whose projections on the image sphere are $\mathbf{T}(s_i, t)$, $i = 1, 2$ which we will abbreviate to $\mathbf{T}^i$, $i = 1, 2$. Think of them as two points on extremal boundaries, which trace out curves $\mathbf{r}(s_1, t)$ and $\mathbf{r}(s_2, t)$ with normal curvatures κ^{t1} and κ^{t2} as the viewer moves. If one of the points, say $\mathbf{T}^1$, is actually a surface feature then it has infinite curvature

$- 1/\kappa^{t1} = 0$. The two features are considered to be instantaneously spatially coincident, that is, initially,

$$\mathbf{T}(s_1, t) = \mathbf{T}(s_2, t).$$

Moreover they are assumed to lie at a common depth λ. In practice, of course, the feature pair will not coincide exactly. The first temporal derivatives of $\mathbf{T}^i$ are dependent only on image position, viewer velocity and rotation and depth:[4]

$$\mathbf{T}_t^i = \frac{(\mathbf{v}_t \wedge \mathbf{T}^i) \wedge \mathbf{T}^i}{\lambda} \tag{33}$$

and hence, instantaneously, $\mathbf{T}_t^1 = \mathbf{T}_t^2 = \mathbf{T}_t$. Second order temporal derivatives are, (rearranging equation (21) and using (31)):

$$\mathbf{T}_{tt}^i.\mathbf{n} = \frac{1}{\lambda}\left[-\frac{(\mathbf{T}_t^i.\mathbf{n})^2}{\kappa^{ti}} + 2(\mathbf{T}^i.\mathbf{v}_t)(\mathbf{T}_t^i.\mathbf{n}) - \mathbf{v}_{tt}.\mathbf{n}\right] \quad i = 1, 2. \tag{34}$$

Now we must define two differential quantities. The *differential curvature* $\Delta\kappa^t$ of the feature pair is defined by

$$\frac{1}{\Delta\kappa^t} = \frac{1}{\kappa^{t1}} - \frac{1}{\kappa^{t2}}. \tag{35}$$

The *relative view vector* is defined to be

$$\delta(t) = \mathbf{T}(s_2, t) - \mathbf{T}(s_1, t) \tag{36}$$

Now, taking the difference of equation (34) for $i = 1, 2$ leads to a relation between these two differential quantities:

$$\delta_{tt}.\mathbf{n} = -\frac{(\mathbf{T}_t.\mathbf{n})^2}{\lambda}\frac{1}{\Delta\kappa^t} \tag{37}$$

which, from (12),

$$= \frac{(\mathbf{v}_t.\mathbf{n})^2}{\lambda^3}\frac{1}{\Delta\kappa^t}. \tag{38}$$

From this equation we can obtain differential curvature $\Delta\kappa^t$ as a function of depth λ, viewer velocity $\mathbf{v}_t$, and the 2nd temporal derivative of δ. Absolute measurement of curvature (34) depended also on the viewer's translational acceleration $\mathbf{v}_{tt}$. Uncertainty from practical measurements (based on finite differences, for example) of the lower derivative will, of course, be much reduced. Hence the relative measurement should be much more robust. Moreover, because δ is a relative measurement on the projection sphere it is unaffected by errors in viewer rotation, unlike the image vectors $\mathbf{T}^i$ which occur in the absolute measurement of curvature.

[4] If we choose the reference frame to be the instantaneous camera co-ordinate system we can express $\mathbf{T}$ and $\mathbf{T}_t$ in terms of an image position vector, $\mathbf{Q}$, and image velocities (optic flow) $\mathbf{Q}_t$. Namely $\mathbf{T} = \mathbf{Q}$ and $\mathbf{T}_t = \mathbf{Q}_t + \omega \wedge \mathbf{Q}$. Equation (33) reduces to the equation of motion and structure from motion from optic flow [*Maybank85*].

In the case that $\mathbf{T}^1$ is known to be a fixed surface reference point, with $1/\kappa^{t1} = 0$, then

$$\Delta\kappa^t = \kappa^{t2}$$

so that the differential curvature $\Delta\kappa^t$ constitutes an estimate, now much more robust, of the normal curvature κ^{t2} at the extremal boundary point $\mathbf{T}^2$. Of course this can now be used in equation(29) to obtain a robust estimate of gaussian curvature.

Our experiments confirm this. Figures 7b and 8b show that the sensitivity of the differential curvature to error in position and rotation computed between points A and B is reduced by an order of magnitude from 190%/mm and 70%/mrad to 17%/mm and 8%/mrad respectively.

Further robustness can be obtained by considering the ratio of differential curvatures. Ratios of two-point differential curvature measurements appear, in theory, to be completely insensitive to viewer motion. This is because in a *ratio* of $\Delta\kappa^t$ measurements for two *different* point-pairs, terms depending on absolute depth λ and velocity $\mathbf{v}_t$ are cancelled out in equation (38). This result corresponds to the following intuitive idea. The rate at which surface features rush towards or away from an extremal boundary is proportional to the (normal) curvature there. The constant of proportionality is some function of viewer-motion and depth; it can be eliminated by considering only ratios of curvatures.

In fact there *is* a hidden dependence on viewer-velocity since the epipolar parameterisation has been used and that depends on $\mathbf{v}_t$ in the "matching equation" (equation(17)). Furthermore, in practice there is a residual sensitivity because the features involved cannot coincide exactly, as the theory requires. Nonetheless our results show a striking decrease in sensitivity again, for *ratios* of differential curvatures, compared with the differential curvatures on their own (figures 7c and 8c). Sensitivities are again reduced by an order of magnitude to 1.5%/mm and 1.1%/mrad respectively.

4 Conclusion

It appears that, in practice, given just one known surface reference point, highly robust relative curvature measurements can be made at an arbitrary number of other points on apparent contours (one reference point can be used as the first partner $\mathbf{T}^1$ in any number of point pairs). These can be used either to measure curvature along an extremal boundary, or to confirm the hypothesis that a point is on a fixed surface feature.

Suppose no known surface feature has been identified in advance. Can the robust relative measurements be made to bootstrap themselves without an independent surface reference? One possibility is that relative (two-point) curvature measurements could be obtained first for a whole set of points (taking many pairings within the set). That fixes curvatures at those points up to one global additive constant. Then look for a *mode* in the set of relative curvature values. That mode, if it exists, is very likely to correspond to the fixed features.

(This strategy would, of course, go wrong for non-generic objects like cylinders which have constant curvature along their extremal boundaries.

We have shown that ratios of differential measurements are theoretically robust. We have shown furthermore that this robustness is exhibited in practice. The next topic for investigation is the analysis of sensitivities both for differential curvature measurements and their ratios. In particular, we will attempt to predict how sensitivity degrades with separation of points in point-pairs.

Appendix: Numerical method for curvature estimation

Computing surface curvatures from equations (30) and (31) is a non-trivial problem since it involves the estimation of second order spatio-temporal derivatives. Numerical analysis for estimating curvature and its uncertainty and sensitivity to errors in image measurement and camera motion is in general difficult. This involves a tradeoff between truncation error and measurement error. We here describe a simple numerical method for estimating surface curvatures from three discrete views for the case of linear camera motion. Numerical analysis for the general case remains to be done.

Consider 3 views taken at times t_0, t_1, and t_2 from camera positions $\mathbf{v}(t_0)$, $\mathbf{v}(t_1)$ and $\mathbf{v}(t_2)$ respectively. From each view and for a point on an image contour, measurements are made of view vector $\mathbf{T}$, surface normal $\mathbf{n}$ and camera position $\mathbf{v}$.

For linear motion and epipolar parameterisation the view vectors and the contact point locus, $\mathbf{r}(s_0, t)$, lie in a plane – the epipolar plane. The 3 view vectors are tangents to $\mathbf{r}(s_0, t)$. They do not, in general, define a unique curve. They may, however, constrain its curvature. If we assume the curvature of the curve $\mathbf{r}(s_0, t)$ is locally constant it can be approximated (locally) as part of a circle of radius R (the reciprocal of curvature) and with centre at $\mathbf{p}$ such that (figure 9):

$$\mathbf{r}(s_0, t) = \mathbf{p} + R\mathbf{N}(t) \tag{39}$$

where $\mathbf{N}$ is the curve normal (which is perpendicular to the view vector and in the case of epipolar parameterisation lies in the epipolar plane and hence is defined by only 2 components) in the Frenet–Serret sense for each view.

Since the view vectors $\mathbf{T}(t)$ are tangent to the curve we can express (39) in terms of image measurables, N(t), and *a priori* unknown quantities $\mathbf{p}$ and R:

$$\begin{aligned} (\mathbf{r}(s_0, t) - \mathbf{v}(t)).\mathbf{N}(t) &= 0 \\ (\mathbf{p} + R\mathbf{N}(t) - \mathbf{v}(t)).\mathbf{N}(t) &= 0 \end{aligned} \tag{40}$$

These quantities can be uniquely determined from measurements in 3 distinct views. For convenience we use subscripts to label the measurements made for each view.

$$\begin{aligned} \mathbf{p}.\mathbf{N}_1 + R &= \mathbf{v}_1.\mathbf{N}_1 \\ \mathbf{p}.\mathbf{N}_2 + R &= \mathbf{v}_2.\mathbf{N}_2 \\ \mathbf{p}.\mathbf{N}_3 + R &= \mathbf{v}_3.\mathbf{N}_3 \end{aligned} \tag{41}$$

Equations (41) are linear equations in 3 unknowns (2 components of $\mathbf{p}$ in the epipolar plane and the radius of curvature, R) and can be solved by standard techniques.

For surface normal, $\mathbf{n}$ (measured from the image from (9)), the normal curvature of $\mathbf{r}(s_0, t)$ is given by:

$$\kappa_t = \frac{1}{R}\mathbf{N}.\mathbf{n} \tag{42}$$

and the Gaussian curvature can then be estimated from (29).

References

[*Barrow78*] H.G. Barrow and J.M. Tenenbaum. *Recovering Intrinsic Scene Characteristics from Images.* A.I Center Technical Report 157, SRI International, 1978.

[*Barrow81*] H.G. Barrow and J.M. Tenenbaum. Interpreting line drawings as three-dimensional surfaces. *Artificial Intelligence*, 17:75–116, 1981.

[*Blake88*] A. Blake and G. Brelstaff. Geometry from specularities. In *Proc. 2nd Int. Conf. on Computer Vision*, pages 394–403, 1988.

[*Bolles87*] R.C. Bolles, H.H. Baker, and D.H. Marimont. Epipolar-plane image analysis: an approach to determining structure. *International Journal of Computer Vision*, vol.1:7–55, 1987.

[*Brady85*] M. Brady, J. Ponce, A. Yuille, and H Asada. Describing surfaces. *Computer Graphics Image Processing*, 32:1–28, 1985.

[*DoCarmo76*] M.P. DoCarmo. *Differential Geometry of Curves and Surfaces.* Prentice-Hall, 1976.

[*Faux79*] I.D. Faux and M.J. Pratt. *Computational Geometry for Design and Manufacture.* Ellis-Horwood, 1979.

[*Giblin87*] P. Giblin and R. Weiss. Reconstruction of surfaces from profiles. In *Proc. 1st Int. Conf. on Computer Vision*, pages 136–144, London, 1987.

[*Harris87*] C.G. Harris. Determination of ego - motion from matched points. In *3rd Alvey Vision Conference*, pages 189–192, 1987.

[*Koenderink76*] J.J. Koenderink and A.J. Van Doorn. The singularities of the visual mapping. *Biological Cybernetics*, 24:51–59, 1976.

[*Koenderink82*] J.J. Koenderink and A.J. Van Doorn. The shape of smooth objects and the way contours end. *Perception*, 11:129–137, 1982.

[*Koenderink84*] J.J. Koenderink. What does the occluding contour tell us about solid shape? *Perception*, 13:321–330, 1984.

[*LHiggins80*] H.C. Longuet-Higgins and K. Pradzny. The interpretation of a moving retinal image. *Proc.R.Soc.Lond.*, B208:385–397, 1980.

[*Marr77*] D. Marr. Analysis of occluding contour. *Proc. Royal Society, London*, 197:441–475, 1977.

[*Maybank85*] S.J. Maybank. The angular velocity associated with the optical flow field arising from motion through a rigid environment. *Proc. Royal Society, London*, A401:317–326, 1985.

[*Millman77*] R.S. Millman and G.D. Parker. *Elements of Differential Geometry.* Prentice-Hall, NJ, 1977.

[*Tsai84*] R.Y. Tsai and T.S. Huang. Uniqueness and estimation of three-dimensional motion parameters of a rigid objects with curved surfaces. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6(1):13–26, 1984.

[*Tsai87*] R.Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, RA-3(4):323–344, 1987.

[*Verri85*] A. Verri and A Yuille. *Perspective Projection Invariants.* Technical Report AIM-832, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, 1985.

[*Weinshall88*] D. Weinshall. Qualitative depth from vertical and horizontal binocular disparities, in agreement with psychophysical evidence. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 159–164, 1988.

Acknowledgments

The authors acknowledge discussions with Professor Mike Brady, Dr Andrew Zisserman, Dr Peter Giblin, Dr David Murray and Dr David Forsyth. Roberto Cipolla acknowledges the support of the IBM UK Scientific Centre.

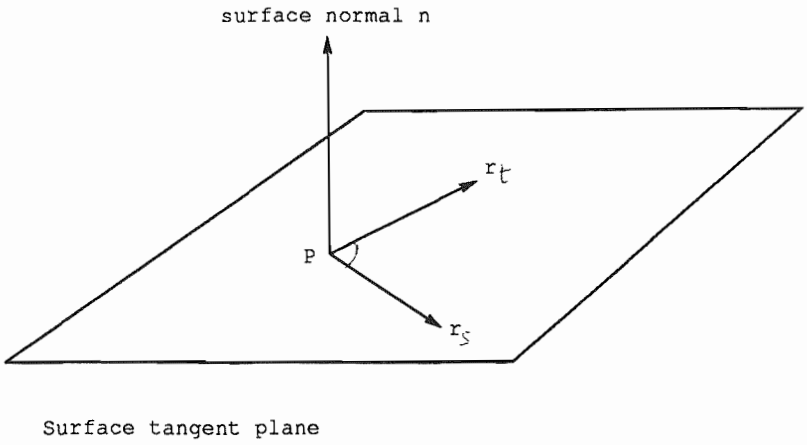


Figure 1. Surface Geometry.

P lies on a smooth surface which is parameterised by a vector valued function $\mathbf{r}(s,t)$. The 2 tangent vectors $\mathbf{r}_s$ and $\mathbf{r}_t$ are a local basis for the tangent plane and are perpendicular to the surface normal $\mathbf{n}$.

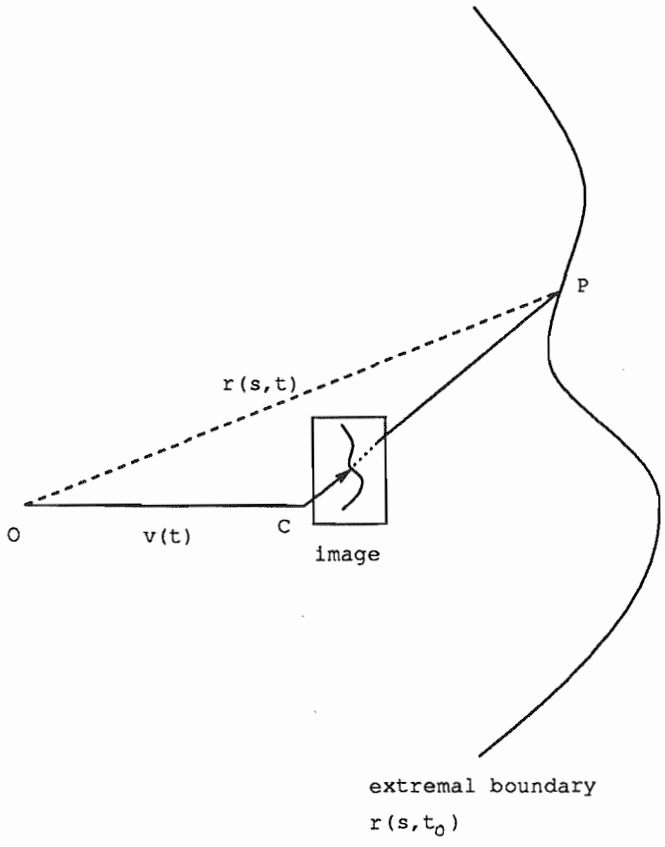


Figure 2. Surface and Viewing Geometry.

The image defines the direction of a ray (unit vector $\mathbf{T}$) which grazes the surface at P. The distance along the ray from the camera centre to the contact point, CP, is λ; the position of the camera centre is $\mathbf{v}(t)$ and a point on the surface has position $\mathbf{r}(s,t)$.

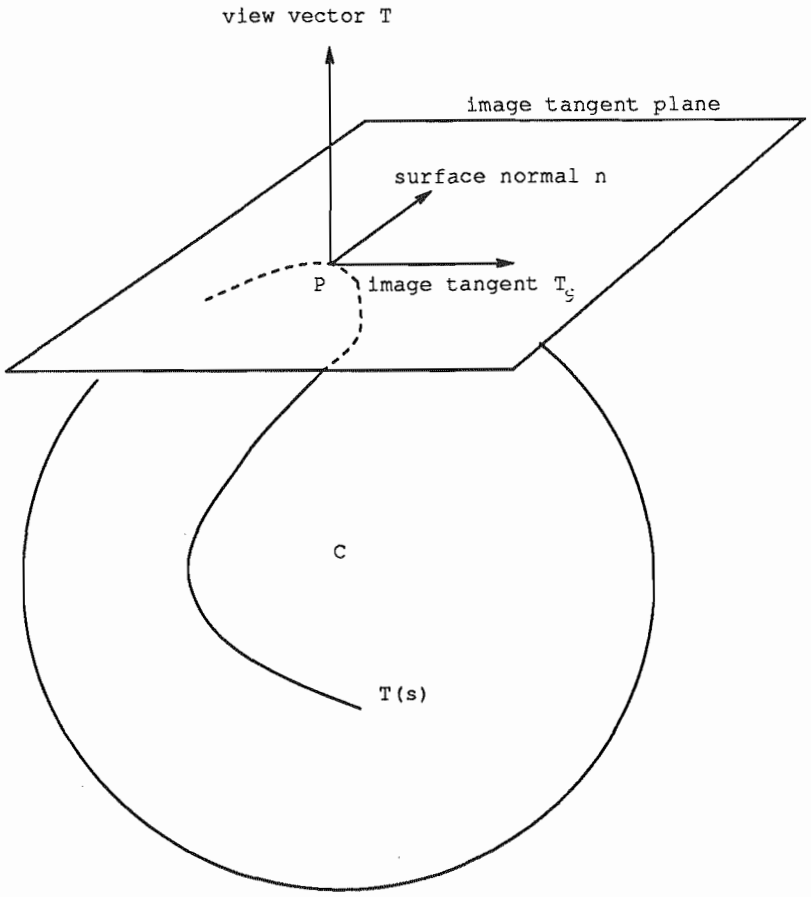


Figure 3. Spherical Perspective Image Geometry

The view vector **T** defined by the image of a point on an extremal boundary and the spherical perspective image tangent $\mathbf{T}_s$ lie in the tangent plane of the surface at P in perpendicular directions. Their measurement can be used to recover the orientation of the surface normal.

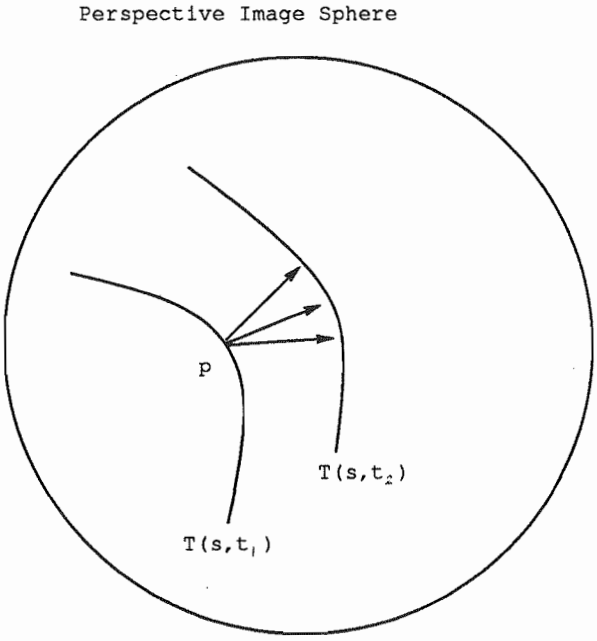


Figure 4. Ambiguity of Parameterisation.

It is impossible to identify a point on the image of an extremal boundary from 1 viewpoint $\mathbf{T}(s,t_1)$, uniquely with a point on the image from a different viewpoint, $\mathbf{T}(s,t_2)$, since the 2 extremal boundaries are different space curves. The choice of constant s parameter curves $\mathbf{T}(t)$ is arbitrary. Each choice defines a different parameterisation of the surface.

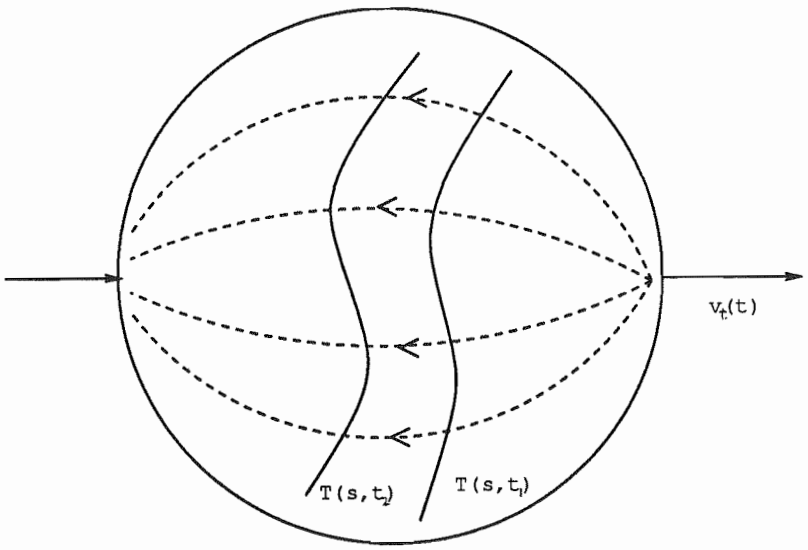
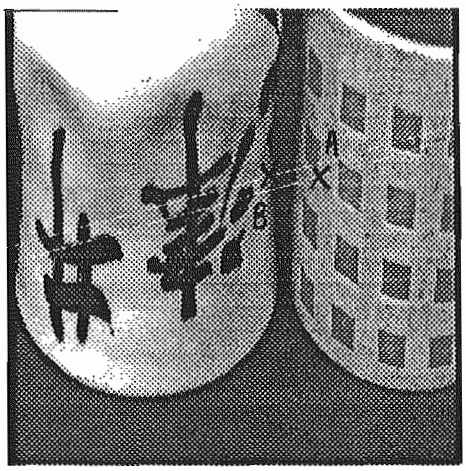


Figure 5. Epipolar Parameterisation

t-parameter curves between the images of an extremal boundary from different viewpoints are defined by moving along great-circles on the image sphere with poles defined by the direction of motion $\mathbf{v}_t$.

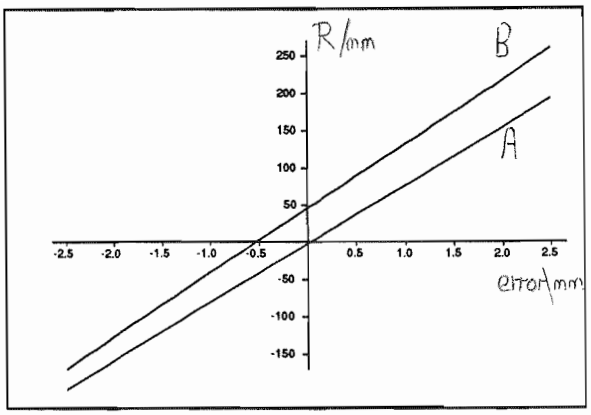
a). Points are selected on image contours in view 1, indicated by crosses A and B for points on a surface marking and extremal boundary respectively.



b). View 2. Corresponding features lie on epipolar lines.

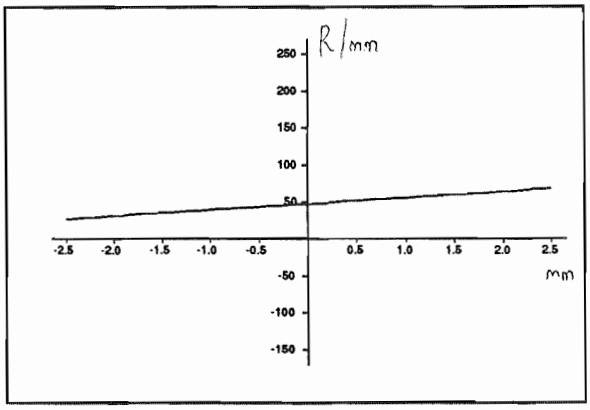
c). Corresponding features in view 3.

Figure 6. Estimating surface curvatures from 3 views



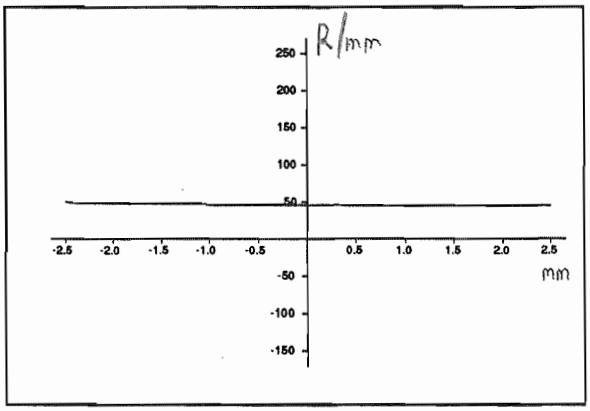
a) Sensitivity of curvature estimated from absolute measurements

The radius of curvature (mm) for both a point on a surface marking (A) and a point on an extremal boundary (B) is plotted against error in the position (mm) of the camera for view 2. The estimation is very sensitive and a perturbation of 1mm in position produces an error of 190% in the estimated radius of curvature for the point on the extremal boundary.



b) Sensitivity of *differential curvature*

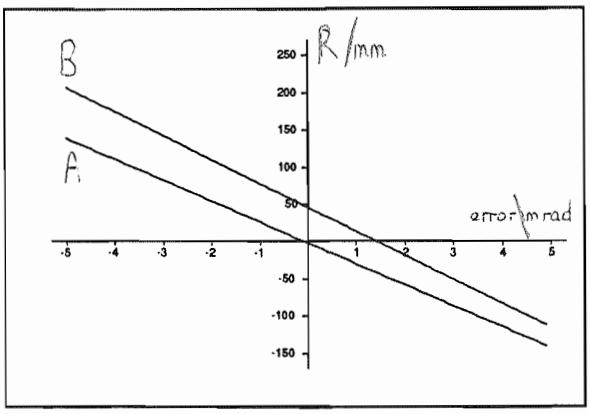
The difference in radii of curvature between a point on the extremal boundary and the nearby surface marking is plotted against error in the position of the camera for view 2. The sensitivity is reduced by an order of magnitude to 17% per mm error.



c) Sensitivity of ratio of *differential curvatures*

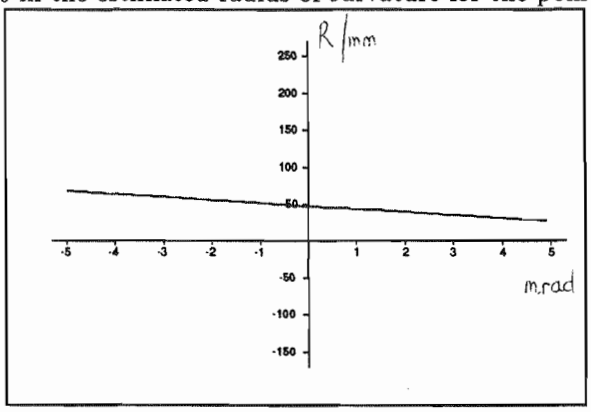
The ratio of *differential curvatures* measurements made between 2 points on an extremal boundary and the same nearby surface marking is plotted against error in the position of the camera for view2. The sensitivity is further reduced by an order of magnitude to 1.5% error for a 1mm error in position. The vertical axis is scaled by the actual curvature for comparision with figures 7a and 7b

Figure 7. Sensitivity of curvature estimate to errors in the camera position for view 2.



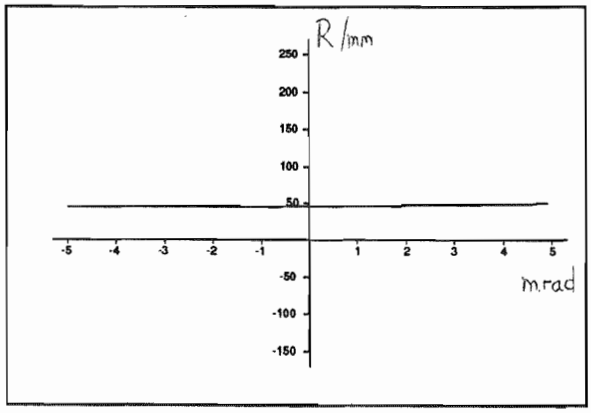
a) Sensitivity of curvature estimated from absolute measurements

The radius of curvature (mm) for both a point on a surface marking (A) and a point on an extremal boundary (B) is plotted against error in the rotation (mrads). The estimation is very sensitive and a perturbation of 1mrad in rotation produces an error of 70% in the estimated radius of curvature for the point on the extremal boundary.



b) Sensitivity of *differential curvature*

The difference in radii of curvature between a point on the extremal boundary and the nearby surface marking is plotted against error in the rotation between view 1 and view 2. The sensitivity is reduced by an order of magnitude to 8% per mrad error.



c) Sensitivity of ratio of *differential curvatures*

The ratio of *differential curvatures* measurements made between 2 points on an extremal boundary and the same nearby surface marking is plotted against error in the rotation The sensitivity is further reduced by an order of magnitude to 1.1% error for a 1mrad error in rotation. The vertical axis is scaled by the actual curvature for comparision with figures 8a and 8b

Figure 8. Sensitivity of curvature estimate to errors in the orientation of the camera for view 2 about an axis defined by the epipolar plane.

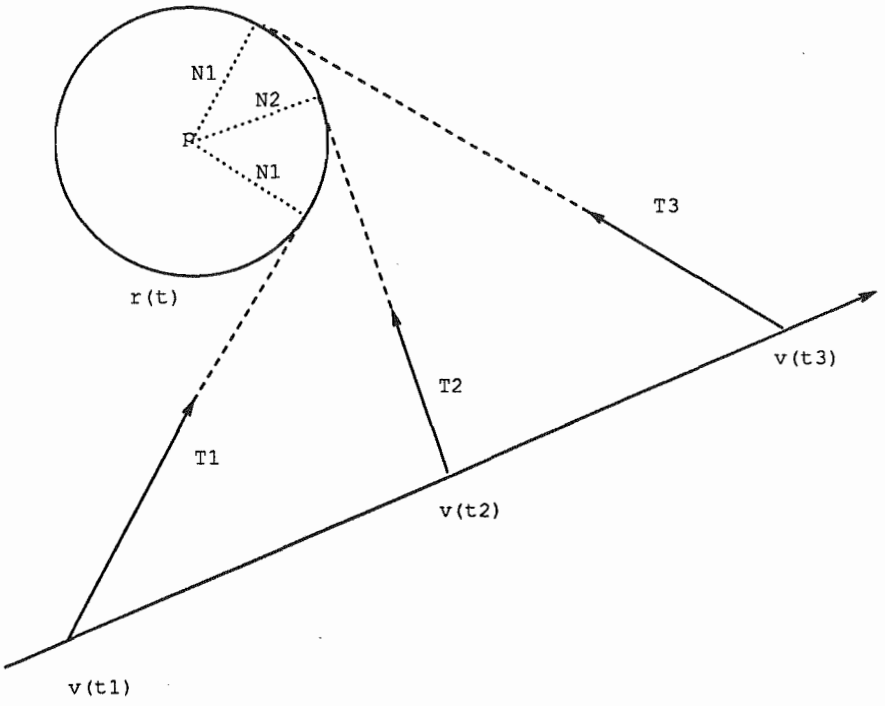


Figure 9. Discrete motion of an extremal boundary

Each view defines a tangent to $\mathbf{r}(s_0, t)$. For linear camera motion and epipolar parameterisation the view vectors and $\mathbf{r}(s_0, t)$ lie in a plane. If $\mathbf{r}(s_0, t)$ can be approximated locally as a circle, it can be uniquely defined from measurements in 3 views.