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

Apparent contours: an outline

The Royal Society

Phil. Trans. R. Soc. Lond. A 1998 **356**, 1087-1102
doi: 10.1098/rsta.1998.0212

Apparent contours: an outline

By Peter Giblin

Department of Mathematical Sciences, University of Liverpool, Liverpool L69 3BX, UK

A point in space viewed from a camera centre $\boldsymbol{c}$ yields a point in the image. A curve in space generally yields a curve in the image. A surface M in space, the boundary of a three-dimensional region, yields a region in the image whose boundary edge is the apparent countour of M from the given viewpoint $\boldsymbol{c}$. More precisely, supposing M has a well-defined tangent plane at every point, those points $\boldsymbol{r}$ of M where the tangent plane passes through $\boldsymbol{c}$ form a curve on M called the *contour generator*. This projects in the image into the *apparent contour*, also called the profile or outline, of M. When the camera centre moves, $\boldsymbol{c} = \boldsymbol{c}(t)$, the contour generator moves over M and the changing apparent contours carry a great deal of information about the surface. In fact, provided cameras are calibrated and camera motion is known, there is in theory enough information to reconstruct a region of M that is swept out by the moving contour generators. In this paper I outline this now 'classical' (1987) result, and some of its more recent variants. It can happen, however, that the contour generators do not sweep out a region but create a boundary, or *frontier* on M, with the contour generators all on one side of this frontier. This situation can be used in principle to recover *motion* as well as the structure of M. I shall describe recent work which seeks to say exactly what the difference is between projections of a space curve, apparent contours of a surface, and apparent contours which yield a frontier. That is, I shall try to describe in a sense all the available information.

Keywords: contour generator; apparent contour; frontier; profile; reconstruction; motion

1. Introduction

This paper is concerned with projections of objects in 3-space to a two-dimensional image. In order to keep the discussion within reasonable bounds, I shall concern myself entirely with so-called *perspective projection*. The model for this is that we have a *world point* $\boldsymbol{r}$, a *camera centre* $\boldsymbol{c}$ and a unit radius *image sphere* S centred at $\boldsymbol{c}$, or an *image plane* P in some way associated with the centre $\boldsymbol{c}$. Rays of light pass from $\boldsymbol{r}$ towards $\boldsymbol{c}$ and are intercepted by S or P to provide image points. We are often interested just in the direction of the ray from $\boldsymbol{c}$ to $\boldsymbol{r}$; as a vector in world coordinates this is denoted $\boldsymbol{p}$. The *unit vector* $\boldsymbol{p}$ can be interpreted geometrically as the vector joining $\boldsymbol{c}$ to the image point in S: it gives the direction in space of the 'visual ray' from $\boldsymbol{c}$ to $\boldsymbol{r}$. Thus we have the fundamental relationship

$$\boldsymbol{r} = \boldsymbol{c} + \lambda \boldsymbol{p}, \tag{1.1}$$

where λ is the distance, or depth, of the point $\boldsymbol{r}$ relative to the camera centre $\boldsymbol{c}$. As a 'world point', the image in S is $\boldsymbol{c} + \boldsymbol{p}$. This is illustrated in figure 1. If we use an image plane P, then the image point in P will have world vector $\boldsymbol{c} + \mu \boldsymbol{p}$, where

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

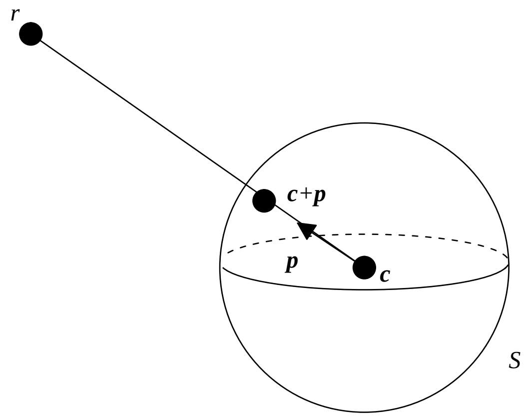

Figure 1. A camera centre $\boldsymbol{c}$, object point $\boldsymbol{r}$ and image point in direction $\boldsymbol{p}$ in the image sphere S.

μ is chosen so that this point lies in P. Of course we would expect to set up some coordinate system in P with which to measure image points. Likewise, in the image sphere S, we may have a system of coordinates which does not coincide with world coordinates, but differs from world coordinates by a rotation R. If $\boldsymbol{c}$ is moving, then R can be expected to depend on the time t.

There are three markedly different cases which will be considered here. The 'scene', that is the collection of world points $\boldsymbol{r}$ which are imaged in S or P, might consist of isolated points in space (dimension 0); it might consist of one of more curves (dimension 1), or it might consist of one or more surfaces (dimension 2). In the latter case the image in S or P of all scene points will be a region of S or P. I shall use only information about the 'apparent contour', which consists roughly speaking of the boundary edges of the image of the whole scene. See figure 2. If we think of a shadow being cast by the object on the image sphere or plane, light rays travelling towards $\boldsymbol{c}$, then it is, roughly speaking, the shadow boundary which is of interest here. This boundary, the *apparent contour*, is also known as the *profile*, or *outline*; hence the title of this paper.

The problems with which I shall be concerned are as follows:

1. Suppose we are given a sequence, or continuous family, of camera centres $\boldsymbol{c}$, and a corresponding sequence of images of a static scene, remembering that for surfaces we take only the occluding boundary. What can we deduce about the scene?

2. What exactly is it that makes one set of images of a scene possible and another set of images impossible?

Cipolla (this volume) considers the possibility of recovering the motion of the camera, too, from measurements in the image. This is much more 'unlikely' in the case of apparent contours of surfaces, since it seems that every image depicts a different part of the object.

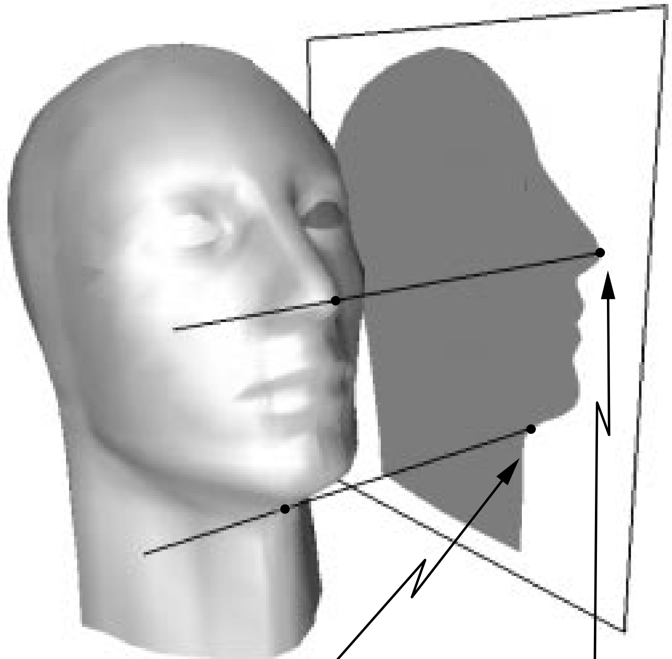

Figure 2. A surface casting a shadow whose boundary is the apparent contour. The apparent contour consists in general of all points where the tangent plane to the surface passes through the camera centre. Here the camera centre is not shown, but is where the two displayed visual rays intersect.

2. A point

Let us look at the images of a single point $\boldsymbol{r}$ in space. Consider a moving camera centre $\boldsymbol{c}(t)$, producing a moving image point. In the coordinate system set up above, this means that

$$\boldsymbol{r} = \boldsymbol{c}(t) + \lambda(t)\boldsymbol{p}(t). \tag{2.1}$$

Here, $\lambda(t)$ will measure the depth from $\boldsymbol{c}$ to the world point $\boldsymbol{r}$.

Given $\boldsymbol{r}$ and $\boldsymbol{c}(t)$, we can determine $\boldsymbol{p}(t)$ from (2.1): $\boldsymbol{p}$ is the unit vector in the direction $\boldsymbol{r} - \boldsymbol{c}(t)$. But what about the converse: what must be true about a camera motion $\boldsymbol{c}(t)$ and an image motion $\boldsymbol{p}(t)$, in order that there exist a *constant* $\boldsymbol{r}$ and a function $\lambda(t)$ with (2.1) satisfied? Geometrically, the condition is obvious: all the rays from $\boldsymbol{c}(t)$ in directions $\boldsymbol{p}(t)$ must be concurrent in one point in space!

The condition '$\boldsymbol{r}$ = const.' can be broken up into two parts: first, the rays must form a 'developable surface'. This means that neighbouring rays are very nearly coplanar; to be precise, the vectors $\boldsymbol{c}_t, \boldsymbol{p}$ and $\boldsymbol{p}_t$ are always coplanar:

$$[\boldsymbol{c}_t, \boldsymbol{p}, \boldsymbol{p}_t] = 0. \tag{2.2}$$

See figure 3. This may be called an 'infinitesimal bilinear constraint' since it depends on two neighbouring rays. In fact, it is the condition needed to solve equation (2.1)

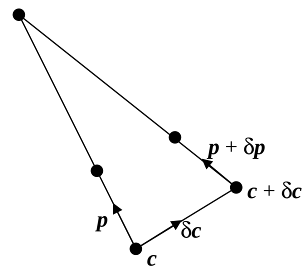

Figure 3. When the visual rays from camera centres $\boldsymbol{c}(t)$ in directions $\boldsymbol{p}(t)$ form a developable surface, this means that neighbouring rays are coplanar. This is equivalent to $[\boldsymbol{p}, \boldsymbol{p}+\delta\boldsymbol{p}, \delta\boldsymbol{c}] = 0$, which becomes $[\boldsymbol{p}, \boldsymbol{p}_t, \boldsymbol{c}_t] = 0$ on dividing by δt and going to the limit.

up to first order. Thus we try to solve

$$(\lambda + (t-t_0)\lambda_t)(\boldsymbol{p} + (t-t_0)\boldsymbol{p}_t) = \boldsymbol{r} - (\boldsymbol{c} + (t-t_0)\boldsymbol{c}_t),$$

λ, $\boldsymbol{c}$, $\boldsymbol{p}$ and their derivatives being evaluated at t_0, and the $(t-t_0)^2$ term obtained by multiplying out the left-hand side being ignored. Comparing coefficients of 1 and $t - t_0$, and writing I for the 3×3 identity matrix and vectors as columns,

$$\begin{pmatrix} I & \boldsymbol{c} & \boldsymbol{p} & 0 \\ 0 & \boldsymbol{c}_t & \boldsymbol{p}_t & \boldsymbol{p} \end{pmatrix} \begin{pmatrix} \boldsymbol{r} \\ -1 \\ -\lambda \\ -\lambda_t \end{pmatrix} = 0.$$

The matrix is 6×6 and we require it to have determinant 0. This is easily seen to be equivalent to the condition that the 3×3 matrix with columns $\boldsymbol{c}_t$, $\boldsymbol{p}_t$, $\boldsymbol{p}$ has determinant 0, which is (2.2).

The second condition ensures that the surface formed by the rays is in fact the very special kind of developable surface which is a 'cone': all rays pass through a point in space. In fact this can be formulated as a second-order condition

$$(\boldsymbol{c}_{tt}\cdot\boldsymbol{p}_t + \boldsymbol{c}_t\cdot\boldsymbol{p}_{tt})(\boldsymbol{p}_t\cdot\boldsymbol{p}_t) - 2(\boldsymbol{c}_t\cdot\boldsymbol{p}_t)(\boldsymbol{p}_t\cdot\boldsymbol{p}_{tt}) = (\boldsymbol{c}_t\cdot\boldsymbol{p})(\boldsymbol{p}_t\cdot\boldsymbol{p}_t)^2. \tag{2.3}$$

This looks more complicated, but in fact it can be linked precisely to an 'infinitesimal trilinear constraint' where in effect what we do is to solve (2.1) *up to order 2* at each value of t. This is essentially using 'three consecutive values of t' (to get the second-order terms) as in a standard trilinear constraint where three values with finite separation are used. That is, we try to solve

$$\begin{aligned}(\lambda + (t-t_0)\lambda_t + \tfrac{1}{2}(t-t_0)^2\lambda_{tt})(\boldsymbol{p} + (t-t_0)\boldsymbol{p}_t \\ + \tfrac{1}{2}(t-t_0)^2\boldsymbol{p}_{tt}) = \boldsymbol{r} - (\boldsymbol{c} + (t-t_0)\boldsymbol{c}_t + \tfrac{1}{2}(t-t_0)^2\boldsymbol{c}_{tt})\end{aligned}$$

at each time instant t_0, terms of degree higher than 2 in t being ignored. Of course this *includes* the earlier condition where we solved to degree 1 in t. It can be shown (Åström & Giblin 1998) that solving to degree 2 is equivalent to solving (2.1) for a constant $\boldsymbol{r}$ and function $\lambda(t)$. We have

Theorem 2.1. *The following are equivalent:*

1. *camera motion* $\boldsymbol{c}(t)$ *and image* $\boldsymbol{p}(t)$ *give a constant point* $\boldsymbol{r}$ *in space;*
2. *equations (2.2) and (2.3) hold for all* t*;*
3. *equation (2.1) can be solved at each time* t_0 *up to degree 2 in* $t - t_0$.

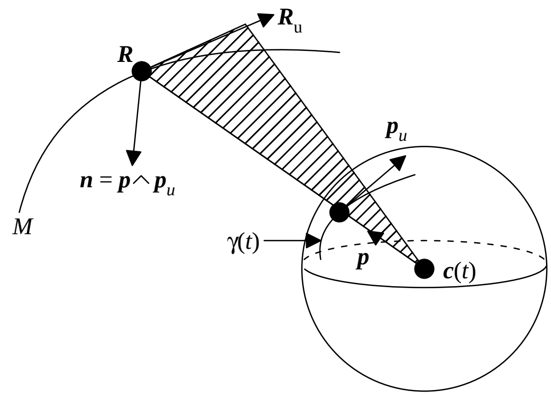

Figure 4. A space curve M and its image $\gamma(t)$ in the image sphere centre $\boldsymbol{c}(t)$. The tangent $\boldsymbol{R}_u$ to M at $\boldsymbol{R}$ projects to the tangent $\boldsymbol{p}_u$ to $\gamma(t)$. The vectors $\boldsymbol{p}$, $\boldsymbol{p}_u$ and $\boldsymbol{R}_u$ lie in a plane which contains the visual ray from $\boldsymbol{c}(t)$ to $\boldsymbol{R}$.

3. Space curve

Suppose now that we have an object in 3-space which is a non-singular space curve M parametrized as $\boldsymbol{R}(u)$, where u is a regular parameter; i.e. $\boldsymbol{R}'(u)$ is never zero. Obviously, the image γ of M in a single image sphere is also a curve; it will be non-singular except that if the viewline from $\boldsymbol{R}(u)$ to the camera centre $\boldsymbol{c}$ is along the tangent line $\boldsymbol{R}'(u)$, then the image will have a cusp. When the camera centre moves on a trajectory $\boldsymbol{c}(t)$, there is an image $\gamma(t)$ for each t, parametrized, say, as $\boldsymbol{p}(u,t)$. Overall there is a two-parameter family of viewlines and requiring that the viewline is tangent to M imposes two conditions, so we do in general expect to see cusps in the image but only at isolated moments of time. Alternatively, the tangent lines to M form in space a surface (the 'tangent developable' of M) and whenever the camera trajectory crosses this surface a cusp will appear in the image.

We have

$$\boldsymbol{R}(u) = \boldsymbol{c}(t) + \lambda(u,t)\boldsymbol{p}(u,t),$$

for a function λ which is the depth or distance from $\boldsymbol{c}(t)$ to the object point $\boldsymbol{R}(u)$. Since $\boldsymbol{R}_u = \lambda_u \boldsymbol{p} + \lambda \boldsymbol{p}_u$, the tangent to M is in the plane of the viewline and the tangent to $\gamma(t)$. The normal to this plane is $\boldsymbol{n} = \boldsymbol{p} \wedge \boldsymbol{p}_u$. See figure 4.

We ask the following question. What is it about $\boldsymbol{p}$ and $\boldsymbol{c}$ which ensures that what is 'out there' producing this sequence of images is actually a space curve? Let us start with a family of curves $\gamma(t)$ in the image sphere, say $\boldsymbol{p}(s,t)$, where s is a regular parameter on $\gamma(t)$. Given also a camera motion $\boldsymbol{c}(t)$, then we construct an object 'out there' as follows, using $\boldsymbol{r}(s,t)$ now to emphasize the dependence on two variables:

$$\boldsymbol{r}(s,t) = \boldsymbol{c}(t) + \lambda(s,t)\boldsymbol{p}(s,t). \tag{3.1}$$

Defining $\boldsymbol{n} = \boldsymbol{p} \wedge \boldsymbol{p}_s$, we find $\boldsymbol{r}_s \cdot \boldsymbol{n} = 0$ automatically and, whether it is a curve or surface 'out there', we shall have $\boldsymbol{r}_t \cdot \boldsymbol{n} = 0$ (compare § 4), which we solve to find λ:

$$\lambda \boldsymbol{p}_t \cdot \boldsymbol{n} = -\boldsymbol{c}_t \cdot \boldsymbol{n}. \tag{3.2}$$

This gives us λ so long as $\boldsymbol{p}_t \cdot \boldsymbol{n} \neq 0$. If what we construct *is* a space curve, then certainly $\boldsymbol{r}_s$ and $\boldsymbol{r}_t$ must be parallel since they are both along the unique tangent

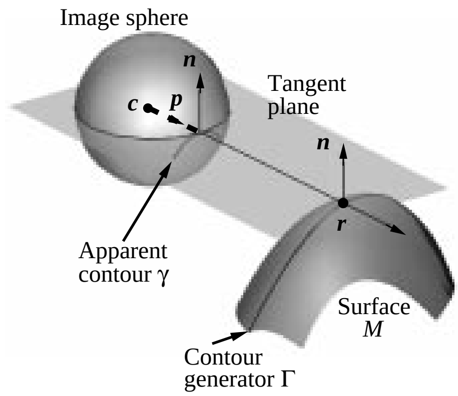

Figure 5. A surface M, contour generator Γ, and corresponding apparent contour γ. The tangent planes to M at points of Γ all pass through the camera centre $\boldsymbol{c}$, and the tangent plane to M at $\boldsymbol{r}$ contains the visual ray direction $\boldsymbol{p}$ and the tangent direction to γ.

to the space curve. Since both these vectors are by construction perpendicular to $\boldsymbol{n}$, there is only one condition for them to be parallel, and this comes to

$$\lambda_s(\boldsymbol{c}_t \cdot \boldsymbol{p}_s + \lambda \boldsymbol{p}_t \cdot \boldsymbol{p}_s) = \lambda(\lambda_t + \boldsymbol{c}_t \cdot \boldsymbol{p})\boldsymbol{p}_s \cdot \boldsymbol{p}_s. \tag{3.3}$$

This can be turned into a condition depending only on $\boldsymbol{p}$ and $\boldsymbol{c}$ by finding λ_s and λ_t. When $\boldsymbol{p}_t \cdot \boldsymbol{n} \neq 0$ we find that λ_s and λ_t are given by

$$\lambda_s(\boldsymbol{p}_t \cdot \boldsymbol{n})^2 = \boldsymbol{c}_t \cdot \boldsymbol{n}(\boldsymbol{p}_{st} \cdot \boldsymbol{n} + [\boldsymbol{p}_t, \boldsymbol{p}, \boldsymbol{p}_{ss}]) - [\boldsymbol{c}_t, \boldsymbol{p}, \boldsymbol{p}_{ss}]\boldsymbol{p}_t \cdot \boldsymbol{n}, \tag{3.4}$$

$$\begin{aligned}\lambda_t(\boldsymbol{p}_t \cdot \boldsymbol{n})^2 &= \boldsymbol{c}_t \cdot \boldsymbol{n}(\boldsymbol{p}_{tt} \cdot \boldsymbol{n} + [\boldsymbol{p}_t, \boldsymbol{p}, \boldsymbol{p}_{st}]) \\ &\quad - (\boldsymbol{c}_{tt} \cdot \boldsymbol{n} + [\boldsymbol{c}_t, \boldsymbol{p}_t, \boldsymbol{p}_s] + [\boldsymbol{c}_t, \boldsymbol{p}, \boldsymbol{p}_{st}])\boldsymbol{p}_t \cdot \boldsymbol{n}.\end{aligned} \tag{3.5}$$

It can be shown that (3.3) is necessary and sufficient for the object reconstructed from $\boldsymbol{p}$ and $\boldsymbol{c}$ using (3.1) and (3.2) to be a space curve, at any rate assuming $\boldsymbol{p}_t \cdot \boldsymbol{n}$ is never zero. See Appendix A. So, assuming this, we have:

Theorem 3.1. *The necessary and sufficient conditions that camera motion* $\boldsymbol{c}(t)$ *and image* $\boldsymbol{p}(s,t)$ *should give a space curve are that (3.3) holds, where the terms* λ, λ_s, λ_t *are given by (3.2), (3.4) and (3.5).*

For another approach to space curves, see Papadopoulo & Faugeras (1996) and Cipolla (1991, § 4.5).

4. Surface

The full image of a surface M in the image sphere is of course a region and not a curve. We concentrate here on the *contour generator* Γ and the *apparent contour* γ of M relative to a camera centre $\boldsymbol{c}$. Write $\boldsymbol{n}$ for the normal to the surface at $\boldsymbol{r}$. Then

$$\Gamma = \{\boldsymbol{r} \in M : (\boldsymbol{r} - \boldsymbol{c}) \cdot \boldsymbol{n} = 0\},$$
$$\gamma = \{\boldsymbol{p} : \boldsymbol{p} = (\boldsymbol{r} - \boldsymbol{c})/\|\boldsymbol{r} - \boldsymbol{c}\| \text{ where } \boldsymbol{r} \in \Gamma\}.$$

Thus Γ consists of those points $\boldsymbol{r}$ of M where the tangent plane passes through $\boldsymbol{c}$ and γ consists of the corresponding image vectors $\boldsymbol{p}$. See figure 5.

The cone of rays through $\boldsymbol{c}$ and γ is tangent to M along Γ. As $\boldsymbol{c}$ moves along a trajectory, this cone sweeps through space, always being tangent to M and in fact creating M in space as an 'envelope'. Thus we confidently expect the apparent contours γ to give us enough information to reconstruct at any rate part of M. The part of M covered by contour generators can be called the 'visible part of M', here meaning that it is that part which from some viewpoint $\boldsymbol{c}$ is seen as an apparent contour point.

It has been known for some time (Giblin & Weiss 1987; Cipolla 1991; Cipolla & Blake 1992) that, given a camera trajectory $\boldsymbol{c}(t)$ and a parametrized family of non-singular apparent contours $\boldsymbol{p}(s,t)$, the surface M with normal vector $\boldsymbol{n}(s,t)$ at $\boldsymbol{r}(s,t)$ can be reconstructed from the following equations:

$$\boldsymbol{r}(s,t) = \boldsymbol{c}(t) + \lambda(s,t)\boldsymbol{p}(s,t); \tag{4.1}$$

$$\boldsymbol{n}(s,t) = \boldsymbol{p}(s,t) \wedge \boldsymbol{p}_s(s,t); \tag{4.2}$$

$$\lambda(s,t)\boldsymbol{p}_t(s,t) \cdot \boldsymbol{n}(s,t) = -\boldsymbol{c}_t(s,t) \cdot \boldsymbol{n}(s,t). \tag{4.3}$$

Here, (4.1) says that the object point $\boldsymbol{r}$ is along the visual ray from $\boldsymbol{c}$ in direction $\boldsymbol{p}$. Differentiating (4.1) with respect to s and using $\boldsymbol{r}_s \cdot \boldsymbol{n} = 0$ shows that $\boldsymbol{p}_s \cdot \boldsymbol{n} = 0$, i.e. that the tangent $\boldsymbol{p}_s$ to the apparent contour is in the tangent plane to M at $\boldsymbol{r}$. Now $\boldsymbol{p}$, $\boldsymbol{p}_s$ are independent so long as $\boldsymbol{p}_s \neq \boldsymbol{0}$, that is so long as the apparent contour is non-singular. Hence we get (4.2); see also figure 5. Equation (4.3) is obtained by differentiating (4.1) with respect to t and using $\boldsymbol{r}_t \cdot \boldsymbol{n} = 0$.

Thus by evaluating λ from (4.2) and (4.3), then substituting in (4.1), the surface points are obtained from $\boldsymbol{p}$ and $\boldsymbol{c}$ alone. Note that this assumes several things:

1. the camera centres $\boldsymbol{c}$ are known in world coordinates;

2. the image directions $\boldsymbol{p}$ from $\boldsymbol{c}$ are known in world coordinates;

3. $\boldsymbol{p}_t \cdot \boldsymbol{n}$ is never zero.

The first two of these say that we are working with calibrated cameras and known camera motion. It is easy to correct the formulae if the image coordinates $\boldsymbol{q}$ are merely rotated from world coordinates $\boldsymbol{p}$ by

$$\boldsymbol{p}(s,t) = R(t)\boldsymbol{q}(s,t) \tag{4.4}$$

for a known rotation $R(t)$ depending on time. For we then have

$$\boldsymbol{p}_t = R(t)\boldsymbol{q}_t + \Omega(t) \wedge R(t)\boldsymbol{q}(s,t), \tag{4.5}$$

where Ω is a vector in the direction of the 'instantaneous axis of rotation', with length equal to the 'instantaneous angular velocity' of the camera coordinate frame with respect to world coordinates. Substituting for $\boldsymbol{p}_t$ from (4.5) into (4.3) gives an expression for λ in terms of the rotated coordinate system. If we 'start the clock' at a moment when $R =$ identity, then the above equation becomes $\boldsymbol{p}_t = \boldsymbol{q}_t + \Omega \wedge \boldsymbol{q}$ at that moment.

The requirement above that $\boldsymbol{p}_t \cdot \boldsymbol{n} \neq 0$ is clearly of a different kind, and requires more explanation. It is closely tied up with the geometry of the apparent contours and the contour generators, as I shall try to explain below. In fact, points where $\boldsymbol{p}_t \cdot \boldsymbol{n} = 0$ turn out to have special importance when we try to deduce *camera motion and rotation* R from the movement of apparent contours. There will be more on this in Roberto Cipolla's paper (this volume; see also Åström *et al.* 1997; Cipolla *et al.* 1995).

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

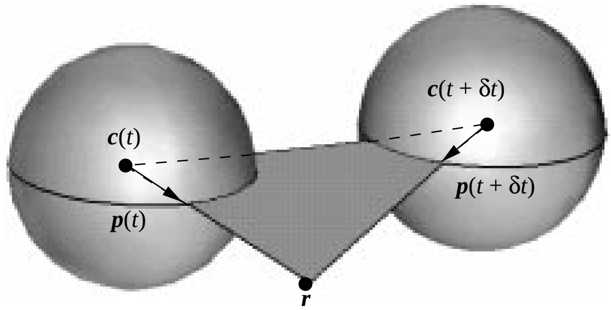

Figure 6. Viewing a single point $\boldsymbol{r}$ from two positions $\boldsymbol{c}(t)$ and $\boldsymbol{c}(t+\delta t)$, the epipolar plane spans the base-line between camera centres and the visual ray through $\boldsymbol{c}(t)$ and $\boldsymbol{r}$. This plane meets the second image sphere in a great circle and $\boldsymbol{p}(t+\delta t)$ must lie on this circle.

(*a*) *Epipolar correspondence*

Some of this discussion is adapted from Cipolla & Giblin (1998), but see also Cipolla & Blake (1992).

If we look at the same point $\boldsymbol{r}$ from two viewpoints, $\boldsymbol{c}(t)$ and $\boldsymbol{c}(t+\delta t)$, then the plane containing $\boldsymbol{c}(t)$, $\boldsymbol{c}(t+\delta t)$ and $\boldsymbol{r}$ is an *epipolar plane*. If we know only $\boldsymbol{c}(t)$, $\boldsymbol{c}(t+\delta t)$ and the image vector $\boldsymbol{p}(t)$ of $\boldsymbol{r}$ in the first view, then we still know the epipolar plane, since $\boldsymbol{r}$ is on the line through $\boldsymbol{c}(t)$ in the direction $\boldsymbol{p}(t)$, and we also know that the image of $\boldsymbol{r}$ in the second view must be on the great circle where the epipolar plane meets the image sphere centred at $\boldsymbol{c}(t+\delta t)$. See figure 6.

When we view a surface there is no precisely corresponding idea, since we are rarely seeing the same point twice. However, there is the following suggestive analogy. In figure 7*a*, we show a sequence of visual rays from camera centres

$$\boldsymbol{c}(t-\delta t), \boldsymbol{c}(t), \boldsymbol{c}(t+\delta t),$$

each of which is nearly tangent to the surface, i.e. meets the surface at two nearby points. Each visual ray determines the next: join the next camera centre to the second point of intersection of the current ray with the surface. Any two consecutive visual rays, for example those through $\boldsymbol{c}(t)$ and $\boldsymbol{c}(t+\delta t)$, are then coplanar (since they intersect). The plane containing these two rays is analogous to the epipolar plane in the point case above.

As the visual rays become more nearly tangent to the surface, they become tangent to a curve on the surface, which is called an *epipolar curve*. The tangents to the epipolar curve are thus always along visual rays. This is illustrated in figure 7*b*. Epipolar curves are studied in detail in Giblin & Weiss (1995). The *epipolar planes* in this 'continuous' situation are spanned by a visual ray and the base-line joining two 'infinitely close' camera centres, that is by a visual ray and the tangent line $\boldsymbol{c}_t(t)$ to the path of centres. Thus

$$\text{The normal to the epipolar plane is } \boldsymbol{c}_t \wedge \boldsymbol{p}.$$

If $\boldsymbol{c}_t$ is parallel to $\boldsymbol{p}$, that is, if the camera motion is instantaneously towards the surface point $\boldsymbol{r}$, then *the epipolar plane is indeterminate.*

When we parametrize contour generators, or the corresponding apparent contours, by a parameter s which is constant along the epipolar curves in the surface (or their corresponding images in the image sphere or image plane) we say that we are using

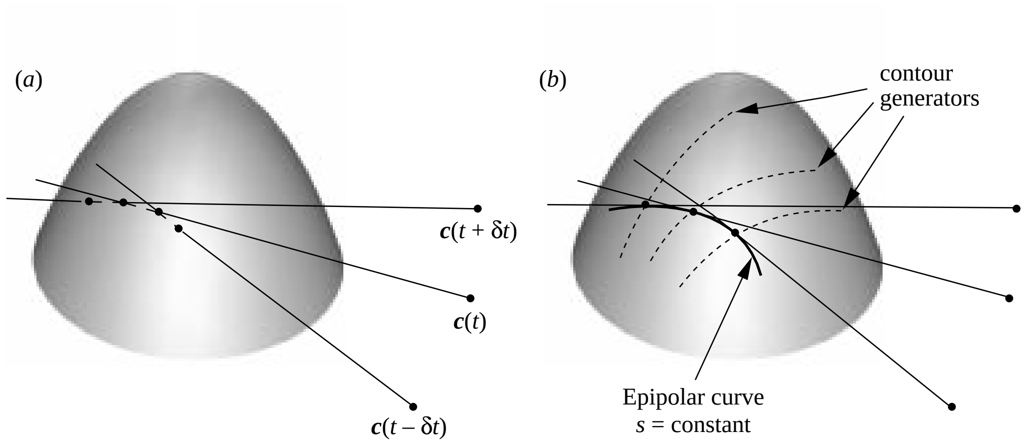

Figure 7. (a) Several visual rays almost tangent to the surface, with each ray determining the next. (b) What happens when the visual rays become tangent to the surface. They are tangent to a curve on the surface called the epipolar curve. This crosses the contour generators on the surface. Epipolar curves and contour generators generally provide a coordinate grid on the surface.

the *epipolar parametrization*. Thus the surface is locally parametrized $\boldsymbol{r}(s, t)$ where t = const. gives the contour generators, and s = const. gives the epipolar curves. The apparent contour for $t = t_0$ is parametrized $\boldsymbol{p}(s, t_0)$. We then have

$$\text{For the epipolar parametrization } \boldsymbol{r}_t(s,t) \text{ is parallel to } \boldsymbol{p}(s,t) \qquad (4.6)$$

for all values of s, t. Of course, this assumes that the above procedure does in fact produce a parametrization of M. See § 4 *b* below.

The epipolar parametrization makes a number of formulae much simpler, and is a natural parametrization to use when tracking (see Cipolla & Blake 1992; Cipolla & Giblin 1998, ch. 2).

(b) *Epipolar tangencies (frontiers)*

It can happen that the attractive procedure of taking contour generators and epipolar curves to form a coordinate system on M simply does not work. This can be for several reasons:

1. The contour generators can be smooth curves but fail to form part of a coordinate grid at all on M. This is what happens along the *frontier*; see below.

2. *The vector* $\boldsymbol{r}_s$ *can be parallel to* $\boldsymbol{p}$, in which case we cannot require that $\boldsymbol{r}_t$ is also parallel to $\boldsymbol{p}$ since for a valid parametrization $\boldsymbol{r}(s, t)$ we must have $\boldsymbol{r}_s$, $\boldsymbol{r}_t$ independent. This is the case of a *cusp on the apparent contour*. We shall not go into this case here; it is analysed in detail in Cipolla *et al*. (1997).

3. The contour generators can be *singular curves* on M. This happens at a 'lips' or 'beaks' point (see, for example, Koenderink 1990, p. 458; Cipolla & Giblin 1998, ch. 4).

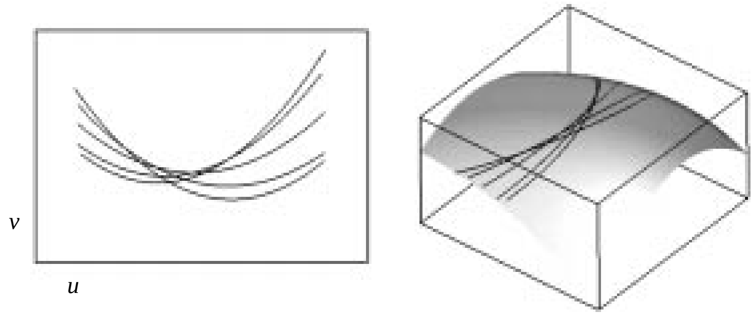

Figure 8. Envelope of contour generators (left) in the parameter plane of M and (right) on M itself. The frontier points on M are envelope points where two 'very nearby' contour generators meet. The contour generators sweep out the 'visible region', which is to the right and below the frontier in the picture of M.

4. We note that another degeneracy of a slightly different kind occurs if we are in fact viewing a *space curve* rather than a surface. In that case we have $\boldsymbol{r}_s$ parallel to $\boldsymbol{r}_t$ for all s and t. Compare § 3.

When contour generators are smooth but fail to form part of a coordinate grid on M, they are forming an 'envelope', as shown in figure 8. This figure shows the contour generators in parameter space for the surface and also on the surface itself. The envelope of contour generators is called the *frontier* relative to the given camera motion, since, at any rate locally, it divides the region of M which is covered by contour generators from the region which is empty of them. Figure 9 shows another example, with the camera trajectory also shown. It also shows another interpretation of the situation via the *spatio-temporal surface* $\tilde{M}$. This is formed by taking contour generators $\Gamma(t)$ for different times t in the (u, v) parameter plane for M and lifting the one for time t to height t in (u, v, t)-space. The surface $\tilde{M}$ 'lies above' only the part of the parameter plane of M which is actually swept out by contour generators—i.e. the 'visible part' of M. When the contour generators form an envelope, then $\tilde{M}$ folds over the envelope curve. This is shown in figure 9*b*.

In fact we have the following (Cipolla & Giblin 1998, ch. 4). The set-up is that we are given a camera trajectory $\boldsymbol{c}(t)$ and a family $\boldsymbol{p}(s, t)$ of non-singular apparent contours of a surface M. The vectors $\boldsymbol{p}$ are unit vectors.

Theorem 4.1. *The following are equivalent, and when any of them occur we say that the point $\boldsymbol{r}$ of M is a frontier point. Note that s, t will not then be local coordinates on M.*

(i) *$\boldsymbol{r}$ is an envelope point of the contour generators on M (this says that $\boldsymbol{r}_s$ and $\boldsymbol{r}_t$ are parallel).*

(ii) $\boldsymbol{c}_t \cdot \boldsymbol{n} = 0$.

(iii) $\boldsymbol{p}_t \cdot \boldsymbol{n} = 0$.

(iv) *$\boldsymbol{p}$ is an envelope point of the apparent contours (this says $\boldsymbol{p}_s$ is parallel to $\boldsymbol{p}_t$ there).*

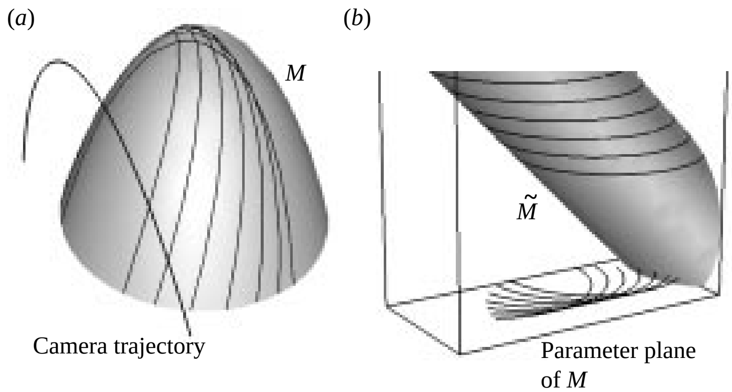

Figure 9. (a) A camera trajectory and contour generators forming an envelope on the surface M. (b) The spatio-temporal surface $\tilde{M}$ with the horizontal curves corresponding to the same contour generators, and their projection to the parameter plane of M, forming an envelope there. The envelope points in the parameter plane are under the 'fold curve' of $\tilde{M}$, which consists of points of $\tilde{M}$ where the tangent plane is vertical.

(v) *Provided $\boldsymbol{c}_t$ is not parallel to $\boldsymbol{p}$, the epipolar plane spanned by $\boldsymbol{c}_t$ and the visual ray is tangent to the surface at $\boldsymbol{r}$. (The excluded case occurs when camera motion is directly towards the surface point.)*

(vi) *The projection from the spatio-temporal surface $\tilde{M}$ down to M is 'folded', that is the tangent plane to $\tilde{M}$ contains the t direction.*

The equivalence of theorem 4.1(ii) and (iii) follows from (3.2), bearing in mind that our surfaces will be finite (λ never infinite) and the camera centre will never hit the surface (λ never 0). Figure 10 illustrates theorem 4.1(v), showing an epipolar plane tangency where the camera centres are separated by a finite distance. The epipolar tangency at a frontier point is the limiting case of this as the centres $\boldsymbol{c}_1$, $\boldsymbol{c}_2$ tend to coincidence.

According to theorem 4.1(iv), frontier points can be recognized by looking at the image, but this assumes that the image is known in world coordinates. When the image is not known, then frontier points are much harder to recognize. It is very useful to be able to recognize them since, according to theorem 4.1(ii), the normal to the apparent contour is perpendicular to $\boldsymbol{c}_t$ at such points, in other words we have a constraint on $\boldsymbol{c}_t$. When the camera coordinate system is rotated with respect to world coordinates (using $\boldsymbol{q}$ instead of $\boldsymbol{p}$ coordinates, as in (4.4) above), we know that the normal and $\boldsymbol{c}_t$, as measured in our camera coordinate system, are perpendicular. Clearly this constraint on camera motion is useful to have. The first person to notice that such a constraint is possible appears to be Rieger (1986).

Unfortunately, the rotated coordinates version of theorem 4.1(iii) is not so simple. From (4.5) the condition $\boldsymbol{p}_t \cdot \boldsymbol{n} = 0$ becomes

$$R(\boldsymbol{q}_t) \cdot \boldsymbol{n} + [\Omega, R\boldsymbol{q}, \boldsymbol{n}] = 0.$$

If we use the convention that at $t = 0$ we have $R =$ identity, then the R can be omitted from this equation at $t = 0$. Alternatively, if we measure $\boldsymbol{n}$ in the rotated coordinate frame too, calling the result $\boldsymbol{N}$, then, for a vector $\Psi = R^{-1}\Omega$, we have

$$\boldsymbol{q}_t \cdot \boldsymbol{N} = -[\Psi, \boldsymbol{q}, \boldsymbol{N}] = \Psi \cdot \boldsymbol{T} \tag{4.7}$$

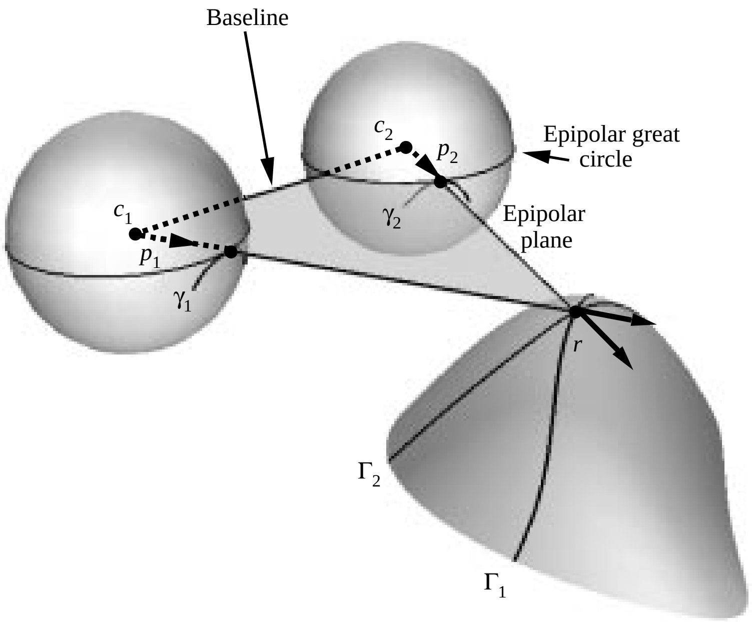

Figure 10. Two contour generators Γ_1, Γ_2 which intersect in a point $\boldsymbol{r}$ near to the frontier. This is an epipolar tangency with camera centres separated by a finite distance. The epipolar tangency in the text is the limiting case, which corresponds with $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ tending to coincidence. The intersection of Γ_1 and Γ_2 becomes a frontier point—a point of the envelope of contour generators. There are corresponding points $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ on the apparent contours γ_1, γ_2.

for any time t, where $\boldsymbol{T}$ is the unit tangent to the apparent contour. This is known as a *generalized epipolar constraint*.

Cipolla (this volume) shows how to use this idea, in work which generalizes Rieger (1986), Giblin *et al*. (1994), Åström *et al*. (1997), and Cipolla *et al*. (1995), to constrain camera motion with real images.

5. Surface case: necessary and sufficient conditions

Finally, we turn to the problem of identifying exactly what conditions on the camera trajectory $\boldsymbol{c}(t)$ and apparent contours $\boldsymbol{p}(s,t)$ must be present in order that, using the reconstruction method above, based on (4.1), we obtain a surface with a frontier (envelope of contour generators) on it. What is special about the reconstruction of a surface near the frontier is that contour generators intersect, that is points $\boldsymbol{r}$ of M near the frontier are 'seen' twice. In the image, there are two points $\boldsymbol{p}_1, \boldsymbol{p}_2$ which are both images of $\boldsymbol{r}$. See figure 10, and also figure 8. The region near the frontier is 'doubly covered' by the apparent contours, as figure 9*b* also illustrates.

We have seen that for frontier points, in world coordinates, the apparent contours form an envelope (theorem 4.1(iii)). But, given only $\boldsymbol{p}$ and $\boldsymbol{c}$, we cannot deduce that the contour generators form an envelope: counter-examples appear in Fletcher (1996) and Fletcher & Giblin (1998). We must make use of the structure away from the frontier points, and we turn to that now.

(*a*) *Discrete constraints in the surface case*

Let us assume that we are just given a family of smooth curves $\boldsymbol{p}(s,t)$ in the image sphere and a camera motion $\boldsymbol{c}(t)$. Let $\boldsymbol{n} = \boldsymbol{p} \times \boldsymbol{p}_s$. In order to use the reconstruction process (4.1) at all we must assume that

> there exists a smooth function $\lambda(s,t)$ such that $\boldsymbol{p}_t \cdot \boldsymbol{n} + \lambda \boldsymbol{c}_t \cdot \boldsymbol{n} = 0$ for all s, t.

Of course this is no condition away from $\boldsymbol{p}_t \cdot \boldsymbol{n} = 0$, but it does impose restrictions where $\boldsymbol{p}_t \cdot \boldsymbol{n} = 0$: for example, $\boldsymbol{c}_t \cdot \boldsymbol{n} = 0$ there too.

Consider the single equation,

$$[\boldsymbol{p}(s,t), \boldsymbol{p}_s(s,t), \boldsymbol{c}(t+u) - \boldsymbol{c}(t)] = 0, \tag{5.1}$$

which is one equation in three variables s, t, u. Equation (5.1) is locally soluble for s as a function of u, t close to a solution where $[\boldsymbol{p}, \boldsymbol{p}_{ss}, \boldsymbol{c}(t+u) - \boldsymbol{c}(t)] \neq 0$, which says that $\boldsymbol{p}_{ss}$ does not lie in the plane spanned by $\boldsymbol{p}$ and $\boldsymbol{p}_s$, i.e. the apparent contour does not have a geodesic inflexion. When we are looking at apparent contours of a surface this says that the surface point is not parabolic.

Let us write

$$s = s(u,t), \qquad \tilde{\boldsymbol{p}}(u,t) = \boldsymbol{p}(s(u,t),t), \tag{5.2}$$

but from now on drop the ˜ so that we write just $\boldsymbol{p}(u,t)$. Thus we have reparametrized the apparent contours, using u instead of s, and

$$[\boldsymbol{p}(u,t), \boldsymbol{p}_u(u,t), \boldsymbol{c}(t+u) - \boldsymbol{c}(t)] = 0, \tag{5.3}$$

identically in u, t. Note that this uses the fact that the t variable remains unchanged under the reparametrization, so that $\boldsymbol{p}_s$ becomes $\boldsymbol{p}_u$.

We assume that

$$[\boldsymbol{p}(u,t), \boldsymbol{p}(-u,t+u), \boldsymbol{c}(t+u) - \boldsymbol{c}(t)] = 0 \tag{5.4}$$

holds identically, regarding this as a constraint.

Theorem 5.1. *From the above assumption, we can deduce*

$$\boldsymbol{r}_t = 2\boldsymbol{r}_u \text{ along } u = 0; \qquad \boldsymbol{r}(u,t) = \boldsymbol{r}(-u,t+u) \text{ for all } u, t.$$

Thus (5.4) is sufficient (also in fact necessary) for apparent contours and camera motion to yield a surface with frontier, corresponding to $u = 0$.

This says that the constructed surface $\boldsymbol{r} = \boldsymbol{c} + \lambda \boldsymbol{p}$ really is a doubly covered surface with boundary along $u = 0$: apparent contour points with parameters (u,t) and $(-u, t+u)$ give the same surface point, so that the parameter space (u,t) is 'folded over' before being mapped onto the surface. Note that with the special parametrization by u and t, we have $\boldsymbol{r}_t = 2\boldsymbol{r}_u$ along the frontier. We expect $\boldsymbol{r}_u$ and $\boldsymbol{r}_t$ to be parallel there, by theorem 4.1(i). The fact that $\boldsymbol{c}_t \cdot \boldsymbol{n} = 0$ along $u = 0$ follows from (5.3) by dividing the third term by u and letting $u \to 0$. The examples of Fletcher (1996) and Fletcher & Giblin (1998) arise because the '$\boldsymbol{r}_u || \boldsymbol{r}_t$' and '$\boldsymbol{c}_t \cdot \boldsymbol{n} = 0$' curves have become separated.

Figure 11 illustrates the geometrical meaning of the reparametrization by u and t that has been carried out. What the theorem says is that, once this is done, (5.4) contains all the constraints derivable from camera motion and apparent contours.

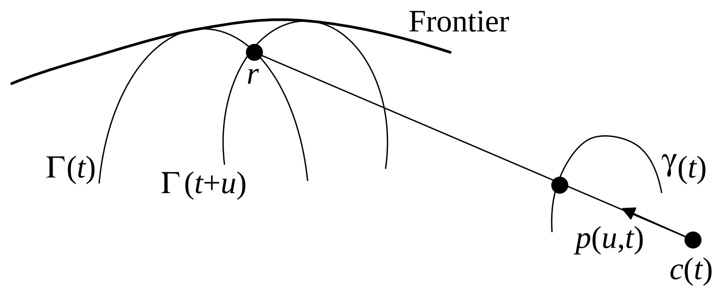

Figure 11. A special parametrization defined by intersecting contour generators. A point on $\gamma(t)$ is given the parameter u when the corresponding surface point $\boldsymbol{r}$ also lies on $\Gamma(t+u)$.

See Appendix A for a proof of the theorem.

I am very grateful to my co-authors over the past decade or so, especially Richard Weiss, Andrew Zisserman, Roberto Cipolla and Kalle Åström. The work described here on 'necessary and sufficient conditions' is joint work with Kalle Åström, partly supported by a grant from the London Mathematical Society. I also acknowledge with gratitude EPSRC grants GR/F 45745 and GR/H 59855, as well as European grants VIVA and VANGUARD. Most of the computer graphics uses Richard Morris's Liverpool Surfaces Modelling Package (LSMP) (Morris 1997), and I am grateful to Gordon Fletcher for help with the figures.

Proofs of some results above

Proof of theorem 3.1. Suppose $\boldsymbol{r}_s || \boldsymbol{r}_t$ for all s, t. Then there is a function $\mu(s,t)$ such that $\mu \boldsymbol{r}_s = \boldsymbol{r}_t$. We seek to find the curves in s, t space which map to fixed points of the object curve in $\boldsymbol{R}^3$. That is, we seek a function $u = \theta(s,t)$ so that the constant values of u give constant values in space of $\boldsymbol{r}(s,t)$. The new parameter space (u,t) has $t =$ const. corresponding to profiles and $u =$ const. corresponding to fixed points in the object.

Let $u = \theta(s,t)$ solve for s as $s = \sigma(u,t)$. Then

$$\frac{\partial}{\partial t}\boldsymbol{r}(\sigma(u,t),t)$$

is to be identically zero. This gives $\boldsymbol{r}_s \sigma_t + \boldsymbol{r}_t = 0$, so we want $\sigma_t(u,t) = -\mu(\sigma(u,t),t)$. But $s = \sigma(\theta(s,t),t)$ so

$$\sigma_t(\theta(s,t),t) = -\frac{\theta_t(s,t)}{\theta_s(s,t)},$$

and we want to find θ with

$$\theta_t(s,t) = \mu(s,t)\theta_s(s,t) \tag{1}$$

for all s, t. However, θ is far from unique, since it just has to be constant at points s, t giving a fixed point in space. In fact we can specify $\theta(s,0)$ for all s: this fixes the u-value at all points of a particular profile and so should fix it for all points s, t. And indeed the differential equation (1) has a unique solution, at least locally, with the initial condition $\theta(s,0) = u_0(s)$ for a given (strictly increasing or decreasing) function u_0. ■

Proof of theorem 5.1. As usual we define $\boldsymbol{n}(u,t) = \boldsymbol{p}(u,t) \times \boldsymbol{p}_u(u,t)$. We first want to check that, for all u, t,

$$\boldsymbol{n}(u,t) \text{ is parallel to } \boldsymbol{n}(-u,t+u).$$

To make sure that we do not miss any exceptional cases, we will do this carefully. Note that by (5.3) and its companion

$$[\boldsymbol{p}(-u,t+u),\boldsymbol{p}_u(-u,t+u),\boldsymbol{c}(t)-\boldsymbol{c}(t+u)]=0,$$

we have, writing $\boldsymbol{v}$ for $\boldsymbol{c}(t+u)-\boldsymbol{c}(t)$, that there exist real numbers α, β, γ, δ such that

$$\begin{aligned}\boldsymbol{v}&=\alpha\boldsymbol{p}(u,t)+\beta\boldsymbol{p}_u(u,t)\\&=\gamma\boldsymbol{p}(-u,t+u)+\delta\boldsymbol{p}_u(-u,t+u).\end{aligned}$$

Hence $\boldsymbol{p}(u,t)\times\boldsymbol{v}=\beta\boldsymbol{n}(u,t)$ and $\boldsymbol{p}(-u,t+u)\times\boldsymbol{v}=\delta\boldsymbol{n}(-u,t+u)$. On the other hand, using (5.4), we have

$$\boldsymbol{v}=\boldsymbol{c}(t+u)-\boldsymbol{c}(t)=\xi\boldsymbol{p}(u,t)+\eta\boldsymbol{p}(-u,t+u) \tag{2}$$

for some real numbers ξ, η. Thus

$$\boldsymbol{p}(u,t)\times\boldsymbol{v}=\eta\boldsymbol{p}(u,t)\times\boldsymbol{p}(-u,t+u),$$

is parallel to

$$\boldsymbol{p}(-u,t+u)\times\boldsymbol{v}=-\xi\boldsymbol{p}(u,t)\times\boldsymbol{p}(-u,t+u).$$

We can now deduce that the two normal vectors are parallel, assuming only that neither of the two $\boldsymbol{p}$ vectors is along $\boldsymbol{v}$, i.e. that the line joining the camera centres does not point directly along either of the viewlines.

It is a straightforward matter now to deduce that $\boldsymbol{r}(u,t)=\boldsymbol{r}(-u,u+t)$, as follows. We have (away from $u=0$), and writing $\boldsymbol{n}$ for a vector in the same direction as the two normals above (it does not matter whether they have the same length),

$$\lambda(u,t)=-\frac{\boldsymbol{c}_t(t)\cdot\boldsymbol{n}}{\boldsymbol{p}_t(u,t)\cdot\boldsymbol{n}},\qquad \lambda(-u,t+u)=-\frac{\boldsymbol{c}_t(t+u)\cdot\boldsymbol{n}}{\boldsymbol{p}_t(-u,t+u)\cdot\boldsymbol{n}}.$$

Differentiating (2) with respect to t and dotting with $\boldsymbol{n}$ we get

$$\boldsymbol{c}_t(t+u)\cdot\boldsymbol{n}-\boldsymbol{c}_t\cdot\boldsymbol{n}=\xi\boldsymbol{p}_t(u,t)\cdot\boldsymbol{n}+\eta\boldsymbol{p}_t(-u,t+u)\cdot\boldsymbol{n},$$

while differentiating with respect to u and dotting with $\boldsymbol{n}$ gives

$$\boldsymbol{c}_t(t+u)\cdot\boldsymbol{n}=\eta\boldsymbol{p}_t(-u,t+u)\cdot\boldsymbol{n}.$$

It follows that $\eta=-\lambda(-u,t+u)$ and $\xi=\lambda(u,t)$. It now follows immediately that

$$\boldsymbol{r}(u,t)-\boldsymbol{r}(-u,t+u)=\boldsymbol{c}(t)+\xi\boldsymbol{p}(u,t)-\boldsymbol{c}(t+u)+\eta\boldsymbol{p}(-u,t+u)=0$$

by (2). ■

References

Åström, K. & Giblin, P. 1998 Necessary and sufficient conditions on image motion. (In preparation.)

Åström, K., Cipolla, R. & Giblin, P. 1997 Generalised epipolar constraint. Preprint, Technical University of Lund.

Cipolla, R. 1991 Active visual inference of surface shape. Ph.D. thesis, University of Oxford. (Also *Springer Lecture Notes on Computer Science*, vol. 1016, 1996.)

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

Cipolla, R. & Giblin, P. J. 1998 *Visual motion of curves and surfaces.* Cambridge University Press. (Submitted.)

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

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

Fletcher, G. 1996 Geometrical problems in computer vision. Ph.D. thesis, University of Liverpool, UK.

Fletcher, G. J. & Giblin, P. J. 1998 Apparent contours, surfaces and motion. In *Proc. Int. Workshop on Computer Vision and Applied Geometry, Sophus Lie Institute, Nordfjordeid, Norway, August 1995.* (In the press.)

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

Giblin, P. J. & Weiss, R. S. 1995 Epipolar curves on surfaces. *Image Vision Computing* **13**, 33–44. (Short version: Epipolar fields on surfaces. In *Proc. ECCV, Stockholm 1994.* Springer Lecture Notes on Computer Science, vol. 800 (1994), pp. 14–23.)

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. 1990 *Solid shape.* Boston, MA: MIT Press.

Morris, R. J. 1997 The use of computer graphics for solving problems in singularity theory. In *Visualization and mathematics* (ed. H.-C. Hege & K. Polthier), pp. 53–66. Heidelberg: Springer.

Papadopoulo, T. & Faugeras, O. 1996 *Computing structure and motion of general 3D curves from monocular sequences of perspective images.* Springer Lecture Notes in Computer Science, vol. 1065, pp. 696–708.

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