Shape from specularities: computation and psychophysics

ANDREW BLAKE[1] AND HEINRICH BÜLTHOFF[2]†

[1] *Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, U.K.*
[2] *Center for Biological Information Processing, Department of Brain and Cognitive Sciences, MIT, Cambridge, MA 02139, U.S.A.*

SUMMARY

Images of artificial and natural scenes typically contain many 'specularities' generated by mirror-like reflection from glossy surfaces. Until fairly recently computational models of visual processes have tended to regard specularities as obscuring underlying scene structure. Mathematical modelling shows that, on the contrary, they are rich in local geometric information. Recent psychophysical findings support the notion that the brain can apply that information. Our results concern the inference of 3D structure from 2D shaded images of glossy surfaces. Stereoscopically viewed highlights or 'specularities' are found to serve as cues for 3D local surface-geometry.

1. INTRODUCTION

The idea that human vision exploits physical constraints is, of course, not new. It has been argued vigorously by Marr (1982) and is exemplified by surface continuity and epipolar constraints in theories of stereo vision (Julesz 1971; Marr & Poggio 1979; Mayhew & Frisby 1981). Continuity constraints also underlie certain theories of motion perception (Bülthoff *et al.* 1989; Hildreth 1984; Yuille & Grzywacz 1988) which have survived some psychophysical testing. In the stereoscopic analysis of specularities also it seems that the visual system may exploit, at least partially, some physical constraints. In that case the constraints are expressed in terms of the differential geometry of a surface patch and its interaction with incident light.

Earlier sections of this paper deal, in some detail, with the physical constraints on specular reflection. Later sections describe psychophysical experiments designed to test some of the predictions of the theory. Although the experiments are stereoscopic the theory is developed in terms of visual motion. It is an extension of the theory of Koenderink & van Doorn (1980), which described qualitative changes in the pattern of specularities on a surface under viewer motion. Principally they found that, as the viewer moves, specularities are created and annihilated, in pairs, on (or, in fact, near) parabolic lines. Whilst it is very robust, such a theory is of limited use for short-range motion or stereopsis. Long-range motion of the viewer is necessary before it is at all likely that the creation or annihilation of specularities will be witnessed. However, robust effects also occur between creation and annihilation. The direction and magnitude of the image motion of the specularity are cues for surface geometry. Specular motion is theoretically at its most robust for convexity–concavity in reversible figures. The measured motion is always relative, relative, that is, to the motion of a nearby surface point. It is therefore *parallax* rather than raw image motion that serves as the geometric cue. This reinforces what appears to be a general principle, already shown in the cases of egomotion recovery (Longuet-Higgins & Prazdny 1980) and recovery of surface curvature from apparent contours (Blake & Cipolla, 1990): parallax is a robust cue.

The specular motion model maps naturally to a stereoscopic model via substitution of stereoscopic baseline for viewer motion and of stereoscopic disparity for image motion. It is mathematically convenient to model imaging as a projection onto a sphere, through its centre (Maybank 1985). In that case image motion at a point on the sphere is represented by a vector in the tangent plane. Stereoscopic disparity is therefore a vector difference of positions on the left and right image spheres. The stereoscopic baseline generates a natural coordinate system on the spheres: the epipolar lines form lines of 'longitude'. Each disparity vector then has an 'epipolar' and a 'circumpolar' component. These correspond approximately to the conventional horizontal and vertical components under planar projection, if the eyes are verged on a distant point. Both components are, theoretically, linked to surface shape. We will show that these links are found experimentally too.

It could be argued that specularity is a marginal visual phenomenon since specularities are relatively sparse in images compared with texture edges and other features. Moreover it is associated more with artefacts, relatively recent on an evolutionary time-scale, than with 'natural' objects. Is it really likely, as we claim, that we have developed mechanisms to analyse specularities? In reply, it is worth noting firstly that specularities do commonly occur on (hairless)

† Current address: Department of Cognitive and Linguistic Sciences, Brown University, Box 1978, Providence RI 02912.

Phil. Trans. R. Soc. Lond. B (1991) **331**, 237–252
Printed in Great Britain

faces and that facial recognition is, presumably, important for survival. Wet surfaces, commonly found in nature, are also specular. More significant though, it is not necessary to claim that the ability to deal with specular motion and stereo developed via evolution. The processing of specularities could simply be an extended usage of the parallax mechanism, learnt in a modern environment filled with specular artefacts.

2. THE OPTICAL PRINCIPLE

The basic principle of the 'specular stereo' model derives from familiar concepts of ray-optics (figure 1). The image of a light source, a specularity, appears behind a glossy, convex surface and in front of a concave one, provided both viewer and source are sufficiently distant from the surface. Before we can test its predictive power for human visual phenomena, this principle must be extended in several ways.

(i) It must deal with hyperbolic (saddle) surfaces that are convex in some planes and concave in others.

(ii) The cases of small source and viewer distances must be covered. For example, as the viewer crosses a 'caustic' (Bruce & Giblin 1984) of a concave surface the image of the source vanishes (figure 2). Stereoscopic disparity of the specularity becomes very large as the viewer approaches the caustic and may then change sign as the caustic is crossed.

(iii) the theory must include reflecting surfaces that would make poor optical elements because they are severely astigmatic. In that case the 'depth' of a specularity is not well defined (Koenderink 1976; Longuet-Higgins 1982) because, in general, two rays reflected from a point source off such a surface will be skew. They have no intersection and hence form no virtual image. Rather than depth, 'relative stereoscopic disparity', a vector quantity, is used to characterize the stereoscopic appearance of a specular reflection.

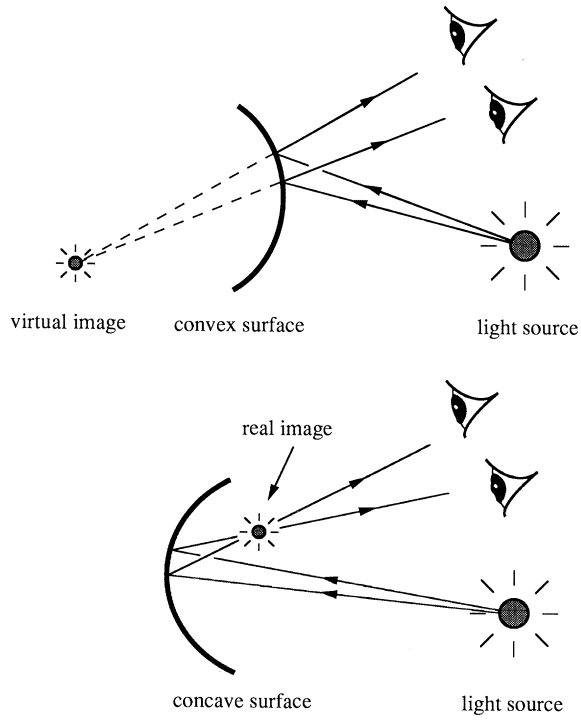


Figure 1. Specular stereo, the basic principle: specularities appear behind a convex mirror but in front of a concave one.

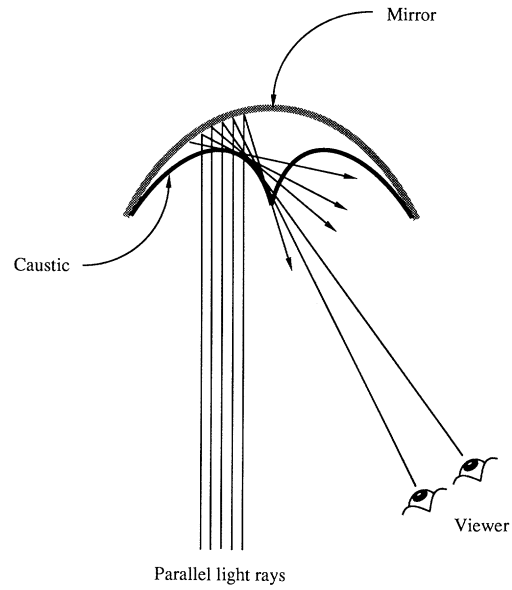


Figure 2. The caustic of a simple cylindrical parabolic mirror. The caustic is the envelope of rays from the point source reflected by the surface. As the viewer crosses the caustic the real image of the source vanishes.

(iv) To make precise predictions, a quantitative model of these phenomena will be needed. First a model is presented that relates the motion of a monocular observer to motion of the specularity and the position, orientation and curvature of the surface. Motion parallax causes the observer to see the specularity in depth, behind or in front of the surface. For example, if the specularity appears behind the surface its image motion will lag behind that of a fixed surface feature. Image motion in the monocular model can then be used to derive a first-order, approximate model for stereoscopic observation of specularities. This is done by regarding stereoscopic disparity as image-motion and stereoscopic baseline as viewer-motion, over some short time Δt. Later in the paper, experimental work on stereoscopic specular phenomena is described. Work on specular motion is in progress.

3. SPECULARITIES AND THE MOVING OBSERVER

Before discussing the full models for specular motion and stereo, in the two succeeding chapters, two simple examples will be given. A monocular observer, moving relative to a glossy surface illuminated by a point source, sees the images of specularities in motion relative to the images of fixed surface points. In the cases of spheres and cylinders, under symmetrical viewing conditions, we can write down simple linear expressions relating the motion parallax of a specularity to the motion of the viewer. The coefficient of the linear relationship depends on distance from the viewer to the object and on surface curvature.

Images will be considered to be formed by projection

onto a unit image-sphere, through a pinhole at its centre. This is done not so much because the retina is roughly spherical but for mathematical simplicity (Maybank, 1985). The parallax $\boldsymbol{\Delta}_t$ of a specularity is its velocity, on the image sphere, relative to some fixed surface feature or marking.

Now consider the case of a glossy sphere of radius a, viewed along a ray passing through the centre of the sphere. Suppose also that the viewer's velocity v_t is orthogonal to the line of sight. Provided that the viewer is not too close to the surface of the sphere, parallax and viewer velocity are related approximately in the following simple way:

$$2\frac{\lambda^2}{a}\boldsymbol{\Delta}_t = -\boldsymbol{v}_t, \tag{1}$$

where λ is the distance from the viewer to the spherical surface, along the line of sight. Several characteristic properties of the behaviour of moving specularities can be observed in this relation.

(i) the magnitude of the parallax diminishes according to an inverse square law, as the object distance λ increases.

(ii) The magnitude of the parallax increases with the radius of the sphere. In the limit of small radius, the sphere can be regarded as a fixed speck, so clearly parallax should become very small. For larger radius, the specularity, appearing to travel around the sphere, covers larger distances and hence generates larger parallax for a given viewer velocity.

(iii) Note that the approximation breaks down for the limit of large radius a. The formula above suggests that parallax should continue to increase as a gets larger and the surface is locally closer and closer to planarity. In fact this does not happen. For approximately planar surfaces the specularity moves 'with' the viewer, rather like the reflection of the moon in a lake, seen from a moving vehicle. This behaviour is correctly predicted by the full model, described later.

(iv) The magnitude of parallax is proportional to the magnitude of viewer velocity.

(v) The parallax vector is parallel to the viewer's velocity only in the special case of the sphere, but not in general.

Now consider a cylinder of radius a with its axis running in a direction $\boldsymbol{u}$ that is orthogonal to the line of sight. Now the relation of parallax to viewer motion is

$$2\frac{\lambda^2}{a}\pi_{\boldsymbol{u}}\boldsymbol{\Delta}_t = -\boldsymbol{v}_t, \tag{2}$$

where $\pi_{\boldsymbol{u}}\boldsymbol{\Delta}_t$ is the component of the parallax vector $\boldsymbol{\Delta}_t$ that is orthogonal to the cylinder axis. Most of the properties that held for the sphere are seen here except that, as in the general case, the parallax vector is no longer necessarily parallel to the motion vector.

4. THE SPECULAR MOTION MODEL

The computational model for specular motion is based on ray-optics and differential geometry. Motion of the viewer generates, via a linear system, apparent motion of specularities. It is clear that the system must be linear: if the velocity of the viewer doubles then all image velocities are also doubled. The coefficients of the linear system depend on viewing and illumination geometry, surface depth and orientation, and surface curvature. These coefficients vary with time as the specularity travels over the surface. Thus an observed apparent motion of a specularity serves as a cue for surface curvature by constraining the coefficients of the linear system. The examples of specular motion from the previous section involved spheres and cylinders viewed along the surface normal. A general model will now be developed which allows for arbitrary local surface shape, and arbitrary positioning of light-source and viewer relative to the surface.

(*a*) *Surface curvature*

Surface curvature can be represented via fundamental forms (matrices of 1st and 2nd derivatives), principal curvatures and directions (do Carmo 1976), or the 'Weingarten map' (Thorpe 1979). The first two representations are tied to a coordinate frame or a particular parameterization of the surface. A particularly convenient coordinate frame (Blake 1985) is one whose axes include the surface normal and the vector orthogonal to both incident and reflected light-rays. However, since we will be using spherical projection, it should be possible to construct purely vector arguments, not tied to any coordinate frame. In that case local surface curvature is concisely represented by means of the Weingarten map W. Consider a moving point $\boldsymbol{r}(t)$ on a surface. The normal vector $\boldsymbol{n}(t)$ to the surface at $\boldsymbol{r}(t)$ changes direction gradually as the point moves. The more highly curved the surface, the greater is the rate of change of direction of $\boldsymbol{n}(t)$. This rate is not merely a scalar but a linear, vector-valued function W of the surface-point motion† $\boldsymbol{r}_t$:

$$\boldsymbol{n}_t = W(\boldsymbol{r}_t). \tag{3}$$

The mapping W is considered to be restricted to the tangent plane at $\boldsymbol{r}$ since $\boldsymbol{n}_t$ and $\boldsymbol{r}_t$ must both lie in the tangent plane:

$$\boldsymbol{n}_t \cdot \boldsymbol{n} = 0, \quad \boldsymbol{r}_t \cdot \boldsymbol{n} = 0.$$

For the two examples mentioned above, the sphere and the cylinder, the Weingarten map is shown in figure 3. For a sphere of radius a the rate of change $\boldsymbol{n}_t$ in $\boldsymbol{n}$, given by the Weingarten map W, is always parallel to the point motion $\boldsymbol{r}_t$:

$$\boldsymbol{n}_t = \frac{1}{a}\boldsymbol{r}_t.$$

The magnitude of the rate of change $\boldsymbol{n}_t$ is in inverse proportion with the radius of the sphere. However, for a cylinder of radius a and with its axis in the direction $\boldsymbol{u}$, $\boldsymbol{n}$ is always orthogonal to the $\boldsymbol{u}$,

$$\boldsymbol{n}_t = \frac{1}{a}\pi_{\mathrm{u}}(\boldsymbol{r}_t)$$

where

$$\pi_{\mathrm{u}}(\boldsymbol{x}) \equiv \boldsymbol{u} \times (\boldsymbol{x} \times \boldsymbol{u})$$

† The derivative of $\boldsymbol{r}$ with respect to time is denoted $\boldsymbol{r}_t$.

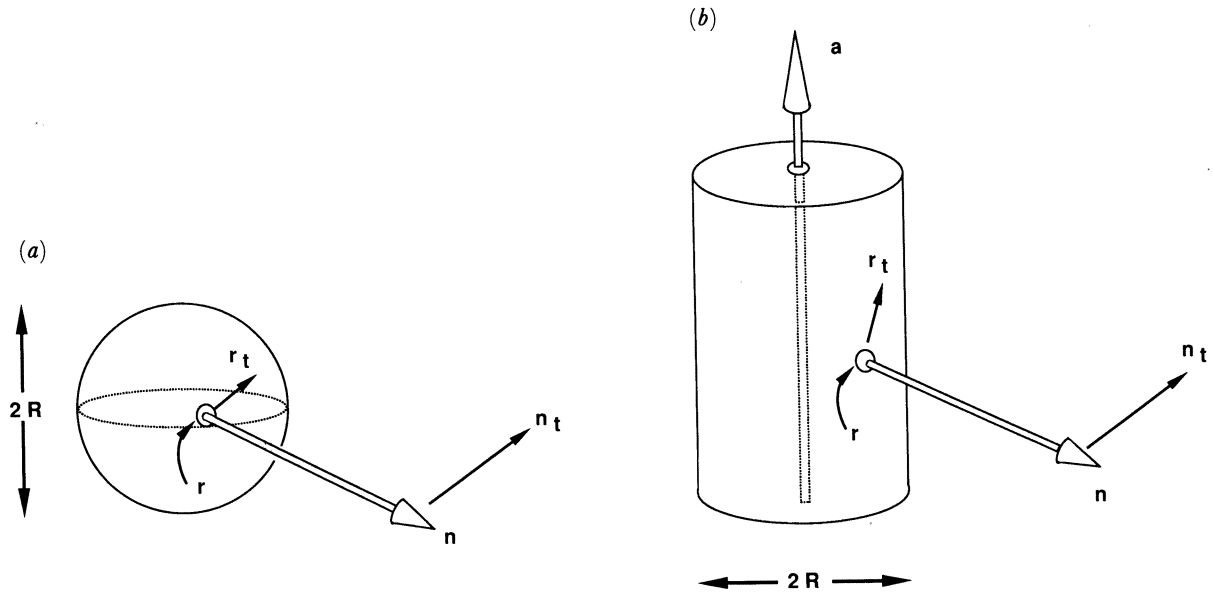


Figure 3. The local curvature at a point $\boldsymbol{r}$ on a surface is concisely represented by the Weingarten map W which specifies the motion $\boldsymbol{n}_t$ of the normal in terms of the motion $\boldsymbol{r}_t$ on the surface. This is illustrated for a sphere (*a*) and a cylinder (*b*).

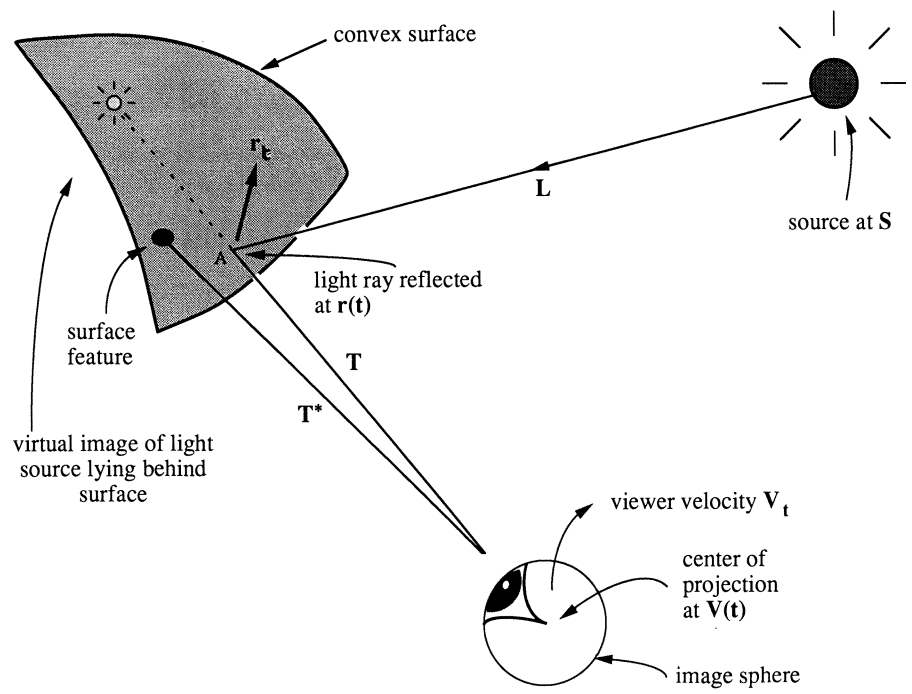


Figure 4. Viewer in motion is at position $\boldsymbol{v}(t)$ at time t and sees a ray from the light source at $\boldsymbol{S}$ reflected from the curved surface at $\boldsymbol{r}(t)$. The reflected ray is in the direction $\boldsymbol{T}(t)$.

denotes the component of a vector $\boldsymbol{x}$ that is orthogonal to $\boldsymbol{u}$. For the cylinder, as for the spherical case, the rate at which the normal $\boldsymbol{n}$ turns is inversely proportional to radius a.

(*b*) *Moving monocular observer*

The geometrical configuration for the specular motion problem is shown in figure 4. At time t, the observer's centre of projection is at position $\boldsymbol{v}(t)$ (refer to figure 4) and the specularity is generated by a ray from the light source at $\boldsymbol{S}$ incident, from direction $\boldsymbol{L}(t)$, at position $\boldsymbol{r}(t)$ on the surface. The reflected ray lies in the direction $\boldsymbol{T}(t)$ and strikes the unit image sphere.

The image of the specularity on the unit image-sphere, allowing for the rotation $R(t)$ of the image-sphere relative to the world frame, is at $R(t)$, $\boldsymbol{T}(t)$ where the rotation operator $R(t)$ represents the attitude of the world-frame relative to the observer's frame at time t. The ray from the surface feature to the centre of projection $\boldsymbol{v}(t)$ is in the direction $\boldsymbol{T}^*(t)$. Again, the corresponding image point, related to the ray direction

by a rotation $R(t)$, is at $R(t)\,\boldsymbol{T}^*(t)$. It serves as a surface reference point for the motion of the specularity. The viewer observes the parallax of the specularity relative to the reference point. Without loss of generality we set $R(0) = I$, the identity transformation, so that the world-frame and the observer's frame are initially aligned. Furthermore it is assumed that the specularity and the reference point initially coincide, so that

$$\boldsymbol{T}(0) = \boldsymbol{T}^*(0). \tag{4}$$

(*c*) *Specular motion constrains surface curvature*

The model consists of the specular motion equation relating the relative motion $\boldsymbol{\Delta}_t$ of the specularity to the viewer motion $\boldsymbol{v}_t$. The model equation is derived in Appendix A., from equations for ray-optics and for differential geometry. The model equation, consisting of coefficients which depend on depth λ and on surface curvature, expressed by the Weingarten map W, is as follows:

$$\lambda^2(2W - k\Pi)\,\mathscr{P}\boldsymbol{\Delta}_t = -\Pi\boldsymbol{v}_t, \tag{5}$$

where the operator Π is defined in terms of operators π_T, π_n for projection orthogonal to $\boldsymbol{T}$ and $\boldsymbol{n}$ respectively:

$$\Pi = \frac{\pi_n \pi_T}{\boldsymbol{T}\cdot\boldsymbol{n}}. \tag{6}$$

(Projection operators of the form π_a were defined earlier. Note that they do not commute: $\pi_a \pi_b \not\equiv \pi_b \pi_a$.) The operator $\mathscr{P}$ is simply a mapping from the tangent plane of the image-sphere to the surface tangent plane, defined by its projection/scaling action on a general vector $\boldsymbol{x}$:

$$\mathscr{P}(\boldsymbol{x}) = \frac{\boldsymbol{n}\times(\boldsymbol{x}\times\boldsymbol{T})}{\boldsymbol{T}\cdot\boldsymbol{n}}. \tag{7}$$

The constant k is an 'effective curvature', the sum of curvatures induced by the finite distances to viewer (λ) and source (μ):

$$k = \frac{1}{\mu} + \frac{1}{\lambda}. \tag{8}$$

The form of the general model (5) is similar to the relationship (1) that was stated earlier for the sphere under symmetrical viewing geometry, but generalized in several ways. In place of the spherical curvature $1/a$ we now have local surface curvature expressed generally in terms of the Weingarten map W. Projection operators $\mathscr{P}$ and Π take care of general viewing geometry, that is, arbitrary positioning of the source and viewer relative to the surface tangent plane. The additional k term in (5) is dominant for surfaces that are almost planar, so that W is negligible. The inclusion of the k term leads to the correct prediction of motion 'with the viewer' for planar surfaces (the reflection of the moon moving over the lake).

5. THE SPECULAR STEREO MODEL

Stereoscopic disparity is a vector quantity, conventionally taken (Mayhew & Longuet-Higgins 1982) to have a horizontal and a vertical component equal to the differences in x, y coordinates of corresponding image points in left and right planar projections. Since we are using spherical projection a modified definition is required, as the difference of two vectors on the image-sphere. Left and right stereoscopic views will be considered as successive views of the monocular observer at times $t = 0$ and $t = \Delta t$ respectively, as in figure 5. Then the disparity of the specularity is

$$\boldsymbol{\Delta} = R(\Delta t)\,\boldsymbol{T}(\Delta t) - \boldsymbol{T}(0), \tag{9}$$

and similarly the disparity of the reference feature is

$$\boldsymbol{\Delta}^* = R(\Delta t)\,\boldsymbol{T}^*(\Delta t) - \boldsymbol{T}^*(0). \tag{10}$$

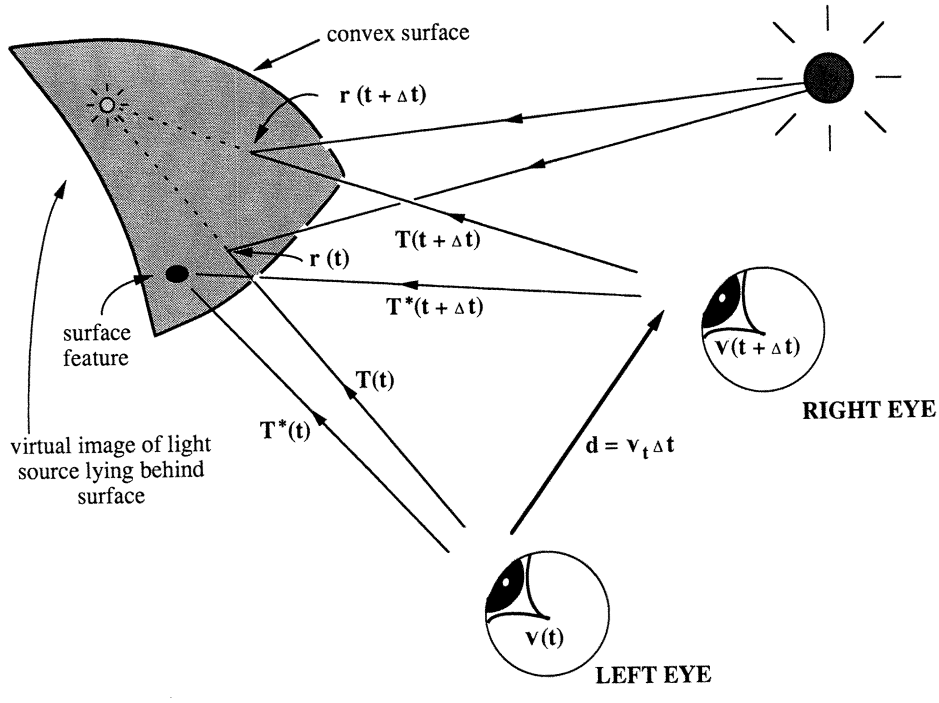


Figure 5. Geometry for specular reflection in stereoscopic views. The right eye is considered as a copy of the left eye after motion over a time interval Δt.

Figure 6. The displacement of the specularity in the right (versus left) image is oblique, whereas that of any nearby surface feature is horizontal. The surface feature is constrained by epipolar geometry but the specularity is not.

The reference feature, being fixed in space, generates a disparity vector $\boldsymbol{\Delta}^*$ that must lie in an 'epipolar plane' (Mayhew & Frisby 1981), the plane containing the baseline vector (figure 5)

$$\boldsymbol{d} = \boldsymbol{v}(\Delta t) - \boldsymbol{v}(0),$$

and the image ray $\boldsymbol{T}^*(0)$. Provided the baseline $\boldsymbol{d}$ is known, the magnitude of the disparity $\boldsymbol{\Delta}^*$ can be used to compute, by triangulation, the distance of the reference feature from the centre of projection. However, the disparity $\boldsymbol{\Delta}$ of the specularity need not lie in the epipolar plane (figure 6) when, as in general, the surface acts as an astigmatic mirror. Its disparity has both an 'epipolar' component and a 'circumpolar' component. We refer to 'epipolar' and 'circumpolar' components of disparity rather than horizontal and vertical ones. These alternative decompositions of the disparity vector are identical only in the special case that optic axes of the two eyes are parallel and perpendicular to the baseline $\boldsymbol{d}$. For typical stereoscopic viewing conditions (modest angle of vergence), however, the two decompositions are approximately the same.

When the baseline is not too large the stereoscopic model can be approximated by using temporal derivatives of quantities from the monocular motion model. The approximations for stereoscopic baseline and disparities, valid to first order when the time interval Δt is short, are:

$$\boldsymbol{d} = \boldsymbol{v}_t \Delta t \tag{11}$$

$$\boldsymbol{\Delta} = (\boldsymbol{T}_t + R_t \boldsymbol{T})\ \Delta t \tag{12}$$

$$\boldsymbol{\Delta}^* = (\boldsymbol{T}_t^* + R_t \boldsymbol{T}^*)\, \Delta t. \tag{13}$$

(*a*) *Relative disparity of specularities*

Surface curvature will be shown, both in theory and later experimentally, to be directly related to the relative disparity of specularities. The relative disparity $\boldsymbol{\Delta}'$ of the specularity is the vector difference between its disparity and that of the reference point:

$$\boldsymbol{\Delta}' = \boldsymbol{\Delta} - \boldsymbol{\Delta}^*. \tag{14}$$

It has both epipolar and circumpolar components each of which turns out to be an important constraint on local surface curvature. The relative disparity of the specularity can be approximated to first order in $\boldsymbol{\Delta}_t$, via equations (12) and (13), in terms of the image velocity of the specularity relative to the reference feature. Thus

$$\boldsymbol{\Delta}' = \boldsymbol{\Delta}_t \Delta t, \tag{15}$$

where

$$\boldsymbol{\Delta}_t = (\boldsymbol{T}_t - \boldsymbol{T}_t^*). \tag{16}$$

Fortuitously, the relative image velocity $\boldsymbol{\Delta}_t$ and hence the relative disparity $\boldsymbol{\Delta}'$ of the specularity are independent (to first order in Δt) of the relative rotation between the left and right image spheres. Relative disparity is therefore (approximately) invariant to rotational error in stereoscopic calibration, for example, errors in estimation of vergence. This proves to be part of the reason for the robustness of relative disparity of specularities (see the following section) as a cue for surface curvature.

(*b*) *The computational model*

In the previous section, the computational model for specular motion was described. From that the specular stereo model can be derived by taking a constant motion $\boldsymbol{v}_t$ over some short time Δt and sweeping out a baseline $\boldsymbol{d} = \boldsymbol{v}_t \Delta t$. It is expressed as a relationship between surface curvature and stereoscopic relative disparity.

(*c*) *The specular stereo model*

The specular stereo model is obtained from the monocular motion model, running over a time-interval Δt. Multiplying (5) by Δt and substituting the baseline $\boldsymbol{d} = \boldsymbol{v}_t \Delta t$ and relative disparity $\boldsymbol{\Delta}' = \boldsymbol{\Delta}_t \Delta t$ yields the specular stereo equation:

$$\lambda^2 (2W - k\Pi)\, \mathscr{P} \boldsymbol{\Delta}' = -\Pi \boldsymbol{d}. \tag{17}$$

Note that the usual inverse-square scaling of disparity with distance appears in the model as the λ^2 term in (17). As expected, the magnitude of relative disparity $\boldsymbol{\Delta}'$ is proportional to the magnitude of the baseline $\boldsymbol{d}$.

The linear operator $(2W - k\Pi)$ in the model has maximum rank 2 because it approximates the mapping between two-dimensional spaces, namely the object surface and the image sphere†. In general the operator has full rank but in special cases, when its rank falls below 2, it is degenerate. The geometrical interpretation of this degeneracy is that the viewer's position $\boldsymbol{v}$ has coincided with a 'caustic' of the surface patch (A. Zisserman, Personal Communication). Now a caustic (Bruce & Giblin 1984) is a surface over which reflected light is concentrated. (It is well known that the inner cylindrical surface of a teacup, for example, concentrates light over a caustic surface which is seen on the bottom of the cup forming two arcs joined by a cusp.) As a monocular viewer crosses a caustic, the apparent image-motion of the specularity approaches infinity and the specularity suddenly disappears altogether. A stereoscopic viewer, close to a caustic, may therefore see a specularity in one way but not in the other.

† In fact the linear operator is the differential of the mapping from the surface patch to the image-sphere.

(*d*) *Approximate model*

Provided the point of reflection $\boldsymbol{r}(t)$ does not lie at a parabolic point of the surface (the Weingarten map is of full rank) and the viewer and source are not too close to the surface the k term is negligible, giving the simplified relationship

$$2\lambda^2 W \mathscr{P} \boldsymbol{\Delta}' = -\Pi \boldsymbol{d}. \tag{18}$$

It can clearly be seen that the relative disparity $\boldsymbol{\Delta}'$ varies in inverse proportion to the magnitude of curvature (W). As an example, consider again the simple case of a spherical surface patch of radius a, viewed along a surface normal, with baseline $\boldsymbol{d}$ perpendicular to the line of sight. In this case the geometric operators $\Pi, \mathscr{P}$ are just identities, the Weingarten map $W = 1/a$. Then the approximate model (18) gives the relative disparity as

$$\boldsymbol{\Delta}' = -\frac{a}{2\lambda^2}\boldsymbol{d}, \tag{19}$$

inversely proportional to the curvature $1/a$. This is very similar to the example of specular motion for a sphere, that was given in section 3. The associated relative depth D (apparent depth of the specularity below the surface) can be shown by straightforward trigonometry to be

$$D = -\lambda^2 |\boldsymbol{\Delta}'|/|\boldsymbol{d}|. \tag{20}$$

Substituting (19) into (20) we find that $D = a/2$, half the radius of the sphere. If the spherical surface patch is concave, so that $W = -1/a$, then the apparent depth is $D = -a/2$ and the specularity appears to be above the surface.

Degeneracy in the approximate model (18) occurs when the Weingarten map W has less than full rank. That occurs precisely when one of the principal curvatures of the surface is zero, that is on parabolic lines. This prediction, valid for the approximate model, that specularities are annihilated and created on parabolic lines has been noted by Koenderink and van Doorn (1976). We have seen however, that the general result is more complicated. Annihilation and creation actually occur as the viewer crosses the caustic (A. Zisserman, personal communication) and the corresponding surface location of the specularity, at the instant of disappearance, is close to but not generally on a parabolic line, becoming close to the parabolic line if the viewer and source are further away from the surface.

(*e*) *Robustness of the specular stereo cue*

Reinforcing the experimental evidence reported below that specularities are significant shape cues, there is a compelling theoretical argument. Specular image motion is robust to uncertainties in viewer motion, and the same is true of specular relative disparities. There are several quite independent senses in which this is true.

(i) *Uncertainty in rotational viewer-motion*

First curvature measurements based on the model (17) are relatively insensitive to motion uncertainty because the relative disparity vector $\boldsymbol{\Delta}'$ is a difference of disparities and was shown, unlike absolute disparities, to be independent of errors in angular velocity of the viewer relative to the surface. This result relies on the assumption that the images $\boldsymbol{T}$ and $\boldsymbol{T}^*$ of specularity and surface reference point respectively are instantaneously coincident. In practice (see, for example, figure 10*a*) the centre of a specularity may not coincide precisely with any surface marking. This small error will give rise to some residual sensitivity to viewer rotation.

(ii) *Robustness to noise from spatial differentiation*

Given the robustness to error in viewer rotation, the remaining motion uncertainty of curvature measurements is no worse than the sensitivity for measurement of the depth λ which appears in the model (17). Curvature measurements from non-specular texture features require the computation of a second spatial derivative λ_{ss} of depth along a contour so that although sensitivity to motion uncertainty is comparable with the specular case, one should expect greatly increased sensitivity to image measurement uncertainty via the second spatial derivative. In the specular case, however, no spatial derivatives need to be computed because differentiation is performed, effectively, by the physical system itself. Ray optics on a curved surface behaves like a differentiation machine.

(iii) *Sensitivity to magnitude of translational motion*

When viewer and source are sufficiently far from the surface the approximate model (18) applies. Within a small surface patch, furthermore, depth λ can be considered to be roughly constant. Under those two assumptions the effect of uncertainty in the magnitude of the viewer motion $\boldsymbol{v}_t$ is confined to a global rescaling of curvatures via the λ^2 (depth-squared) term in the approximate model (18). Thus any ratio of a pair of curvature measurements from the specular stereo model, for adjacent points, is independent of the magnitude of viewer motion, within the limits of our approximations. There is still, of course, a dependence on the direction of the viewer's translational motion $\boldsymbol{v}_t$ via the 'input' to the linear system on the right-hand side of (18). This dependence is well-conditioned however. For example, considering the earlier case of a spherical patch viewed along its surface normal, an error of 5 degrees in the direction of motion would result in errors in curvature terms of less than 10%.

(iv) *Invariance to source position*

No matter how gross the error in the depth λ may be, it remains positive so that motion errors cannot cause a change in the sign of measured curvatures. This leads to the following rule for discrimination of convexity–concavity (Zisserman *et al.* 1989) for which a detailed argument is given in the appendix:

If the epipolar component $\boldsymbol{\Delta}_t \cdot (\pi_T \boldsymbol{v}_t)$ of relative specular image motion $\boldsymbol{\Delta}_t$ is positive then the underlying surface is either convex or hyperbolic. Otherwise, when the epipolar component of $\boldsymbol{\Delta}_t$ is negative, the surface is either concave or hyperbolic.

This rule is independent of all viewer and illumination geometry except that it still requires that the direction of the motion be known. It is derived from the approximate model and hence must be modified for the full model. In that case the epipolar component of $\boldsymbol{\Delta}_t$ is always positive for convex surfaces. For concave surfaces it is usually negative. It may however be positive if the viewer is close enough to the surface. This can happen when the viewer's position $\boldsymbol{v}$ lies between the surface and one of its caustics. However, if viewer and source are far from the surface the approximate model holds, and the convex–concave discrimination rule above holds good.

(v) *Robustness of specular stereo*

The insensitivity results above are presented for the motion case, but apply equally well to stereoscopic viewing. The stereoscopic equivalent of rotational viewer motion is rotation of one eye relative to the other. The equivalent of translational motion is the stereoscopic baseline and the counterpart of image motion is stereoscopic disparity. Thus the rule for convex–concave discrimination in the stereoscopic case depends on the sign of the epipolar component $\boldsymbol{\Delta}' \cdot (\boldsymbol{\pi}_{\mathrm{T}} \boldsymbol{v}_t)$ of the relative disparity $\boldsymbol{\Delta}'$ of the specularity. This requires only that the direction of the stereoscopic baseline is known, which is a plausible assumption.

6. EXPERIMENTAL PREDICTIONS OF THE MODEL

The specular stereo model can be seen in one of two ways. For given surface and viewing geometry, it predicts the linear dependence of the relative motion of the specularity on viewer motion. The relative motion can be computed by inverting the linear operator $(2W - k\Pi)$ in (18), restricted to the surface tangent plane. Alternatively, when relative motion $\boldsymbol{\Delta}_t$ is regarded as a curvature cue, the model represents two constraints on the Weingarten map W, given viewing geometry, depth, source position observer motion. These two ways of viewing the model correspond to two kinds of experiments (Blake & Bülthoff 1990). In the first, independent surface curvature cues are visible and the subject is free to adjust the relative disparity ($\boldsymbol{\Delta}'$) of the specularity. The model makes a prediction of the adjusted disparity. In the second experiment, the stimulus contains an ambiguous surface, relative disparity of the specularity is fixed at one of two values and the subject makes a forced choice between ambiguous interpretations. The model predicts the influence of relative disparity on perceived curvature.

If we assume that the specular stereo model, which describes the physical interaction of light-rays with curved surfaces, also accounts, at least partially, for human perception of glossy, curved surfaces then the following predictions can be made.

(i) In the adjustment mode (1) dense surface texture is available, serving as a strong cue for curvature even in the absence of a specularity. In that case light source direction can be inferred. If the viewer is far from the surface then, unless the light source is itself visible, it is excluded from the cone of view. It must also, therefore, be far from any surface patch that is seen near the centre of the field of view. In that case the approximate specular stereo model is valid so that specular relative disparities are approximately independent of the (unknown) source distance. It is therefore possible that the visual system might be capable of detecting, with some precision, quantitative discrepancies in relative disparity, both epipolar and circumpolar components.

(ii) The model predicts non-zero relative epipolar disparity for specularities on a general curved surface. This means that specularity does not, according to ray-optics, lie on the surface, rather in front or behind. Subjects should therefore perceive a specularity as more realistic if it does not coincide, in depth, with the surface.

(iii) The circumpolar (vertical) component of specular relative disparity is also determined by the model and unlike the circumpolar relative disparity of surface features, is generally non-zero. The model predicts therefore that non-zero circumpolar relative disparity should be a cue to specularity.

(iv) Specular relative disparity is a direct curvature cue whereas disparities of non-specular features serve as curvature cues only after noise-sensitive computation of second derivatives (Rogers & Cagnello 1989). The specular cue should therefore be more robust unless surface texture cues are sufficiently dense to compensate for the ill-conditioning of spatial differentiation.† When both specular and non-specular stereoscopic cues are present, and are mutually inconsistent, it would therefore be expected that the specular curvature cue should dominate. In particular, that dominance should increase as the non-specular texture becomes sparser.

(v) The sign of epipolar relative disparity should be a particularly robust cue, capable of convex–concave disambiguation, regardless of whether information about light-source position is available (experiment 2).

(vi) If the approximate model applies to the stereoscopic analysis of specularities, that is human vision embodies no compensation for finite source and viewer distances, then convex–concave disambiguation may occasionally fail. When the viewer (or source) is close to a concave surface the sign of epipolar relative disparity of the specularity may none the less be positive, and thus act as a cue for convexity. This occurs only on sufficiently shallow, concave surfaces and is therefore probably rather uncommon.

(vii) The monocular shape of the specularity may act as a cue to shape. The reflection of a spherical source, for example, in a curved surface, is approximately an ellipse whose dimensions and orientation depend on surface curvature. Theory predicts (Blake & Brelstaff 1988) that the monocular cue must be subject to (at least) a fourfould ambiguity. The four interpretations include one convex, one concave and two hyperbolic (figure 7).

(viii) It has been shown (Blake & Brelstaff 1988) that this monocular information is complementary to that available from specular stereo. Specular stereo, involving the two components of relative disparity of a

† This explains why, for example, the slightest dent in a polished car door becomes plainly visible when it catches a highlight.

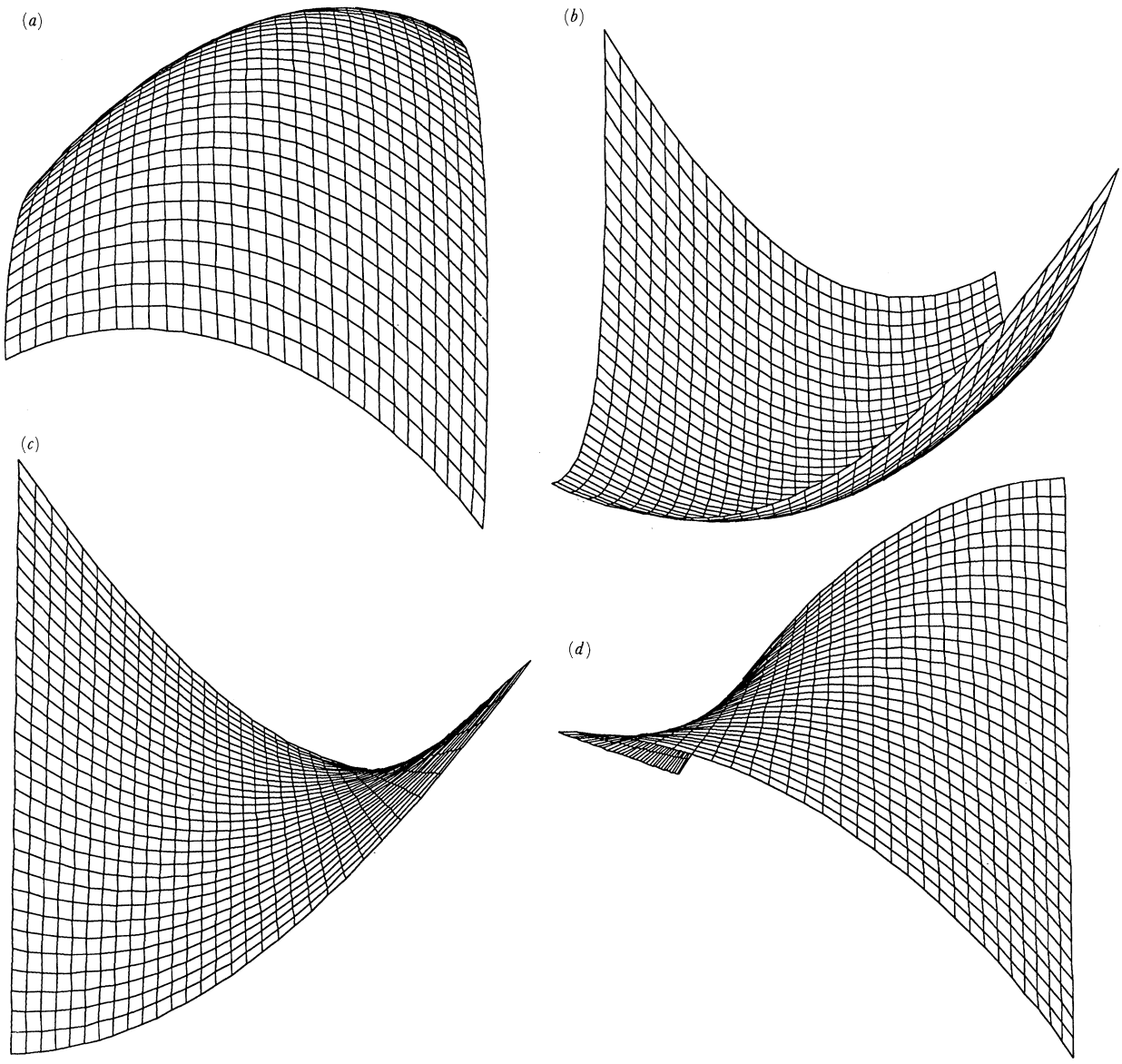


Figure 7. Observation of a specularity, generated by a spherical source of known size, determines local surface curvature up to a four-fold ambiguity. The possible surface interpretations of a particular specularity are shown, one convex, one concave and two hyperbolic.

specularity, only partially constrains the three parameters of surface curvature at a point, leaving one degree of freedom. If the visual system were also using monocular cues they could, in principle, be combined with specular stereo cues to determine surface curvature fully at any surface point, even on an unsymmetrical surface.

(ix) The results described above for specular stereo also have counterparts for specular motion. Thus moving specularities might be expected similarly to act as curvature cues.

(x) Under extended motion of the viewer it is possible, in principle, to compute the position of a light source from the image path of a specular reflection (Zisserman *et al.* 1989). For extended motions (of the right kind) it is therefore predicted that specular motion should serve as a quantitative curvature cue.

7. PSYCHOPHYSICAL EXPERIMENTS

This section describes two experiments aimed to test whether the human visual system exploits constraints introduced in the previous sections. For the first experiment an adjustment task was devised in which the subject interactively changed both epipolar and circumpolar disparities of a specularity, while keeping the disparity of the underlying surface constant. For the second experiment an ambiguous non-disparate image of surface was used in which the relative disparity of a specularity was fixed at one of two values corresponding to a convex or concave surface. The model predicts the influence of relative disparity on perceived curvature by the following constraint:

No convex surface can generate a convergent (−) relative epipolar disparity; a concave surface does not generate a divergent (+) one, except under the conditions mentioned above which are unusual in practice.

(*a*) *Stimulus generation*

Images of glossy, textured, curved surfaces are generated with a solid modelling software package

(S-Geometry, Symbolics Inc.) and displayed on a high-resolution colour monitor (Mitsubishi UC-6912, short persistence phosphor). To view these surfaces in 3D, stereoscopically disparate images were interlaced (even lines for the left-hand image and odd-lines for the right-hand image) with a frame-rate of 30 Hz. These images were viewed through shutter glasses (Stereo-Optic Systems, Inc.), which were triggered by the interlace signal to present only the appropriate images to the left and right eyes. The images were shown with a true resolution of 254 grey levels by using a 10-bit digital-to-analog converter. To get a sufficient reference for the relative disparity of a specularity a textured surface was used instead of a homogeneous surface. Note that Bülthoff & Mallot (1988) showed that an edge-token on a shaded surface can have a pulling effect reducing the depth between token and surface. This is not the case if the surface has enough features, in other words is textured. A bi-level texture of sufficient density and contrast (90%) to furnish strong cues for curvature from edge-based stereo was generated as a level-crossing of a Brownian fractal noise function ($1/f$ spatial frequency spectrum, 1st and 2nd harmonic removed) using the Symbolics Fractal Software. The texture (256×256 pixels) was smoothly mapped onto the curved surface patches using (UV-spherical MIP mapping (Williams 1983). Gloss was simulated with the Phong (1975) shading model by using the following equation:

$$I = 0.15 + 0.45(0.5 \cos\theta + 1.5(\cos\phi)^{40}),$$

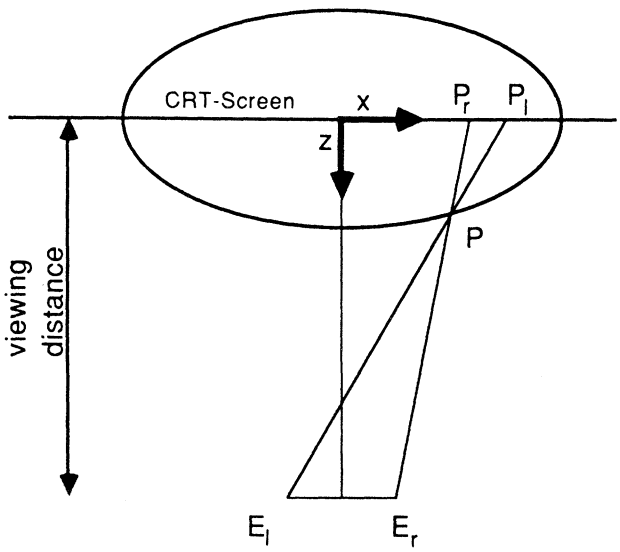


Figure 8. Imaging geometry. Projection onto the x–z plane. Viewing distance is 114 cm. Nodal points of left and right eyes are e_L, e_R respectively, separated by 6.5 cm. A point p is imaged onto the screen at p_l for the view from the left eye and at p_r for the view from the right eye.

where θ is the angle of incidence (between surface normal and light source direction) and ϕ the angle of emittance (between reflected ray and viewing direction).

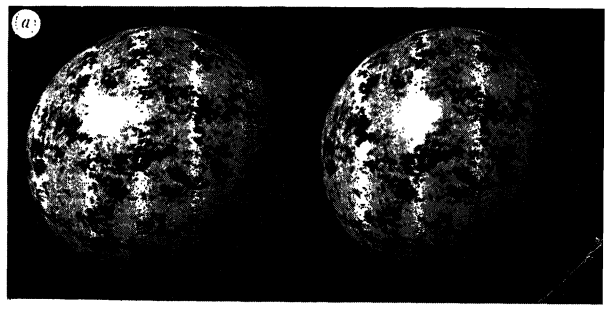
The imaging geometry for generating the disparate images is shown in figure 8. It differs from the usual camera geometry in that the image is constructed on a screen that is not perpendicular to the optical axis of the eyes. Note that the imaging geometry, and therefore the image itself, does not depend on the fixation point as long as the nodal points of the two eyes remain fixed at the positions E_l and E_r, respectively. Images were computed for a viewing distance of 114 cm and an interpupillary separation of 6.5 cm. The viewing angle of the 400×400 pixel image with a diameter of 10.9 cm on the screen was 5.5 deg. Panum's fusional area of ± 10 arc min (Ardity 1986) corresponds to an interval from 4.3 cm to 15.2 cm in front of the screen if the fixation point is 10 cm in front of the screen. Images were viewed on the CRT-screen under reduced room light conditions which allowed the sides of the monitor to be seen. The dimensions and viewing geometry for the 3 objects are given in table 1. Object orientation and light-source direction are given in a coordinate system with x–y axes in the image plane and z axis towards the viewer. The light source was always infinitely far away (parallel illumination). The specular exponent in the Phong shading model was set to 40 for the sphere and 10 for the ellipsoids. Concave versions of the two ellipsoids were also used.

(*b*) *Experiment 1: judgement of surface material quality*

Simulation of surface gloss causes a specularity to appear superimposed on the texture, as in figure 9*a*. Previous studies have shown (Bülthoff & Mallot 1988) that edge-based stereo cues largely override cues such as monocular or disparate shading. We might therefore expect also that specularity cues should be overridden; that is precisely what happens. When the specular relative disparities are veridical, specularity behind the surface, the whole surface appears glossy as in figure 9*a* (not just in the vicinity of the specularity (Beck 1972)). However, when epipolar relative disparity is non-veridical the surface ceases to look glossy. For example, if the specularity is in front of the surface with excessively convergent (−) epipolar relative disparity, surface quality is reported to be matte and opaque, with a puff of cloud in front of the surface. The cloud patch is not 'seen' as a specularity and therefore there is no reason why the surface should look glossy. For excessively divergent (+) relative disparity subjects usually report that the surface looks transparent, with a source of light behind it (like a frosted glass light bulb).

Table 1. *Object geometry*

surface	principal axes/cm	orientation	light source
sphere	10.0	1 1 1	−1 1 1
horizontal ellipsoid	8.0 8.0 12.0	−1 1 1	0 0 1
oblique ellipsoid	4.5 4.5 9.0	−1 1 0.25	0 0 1



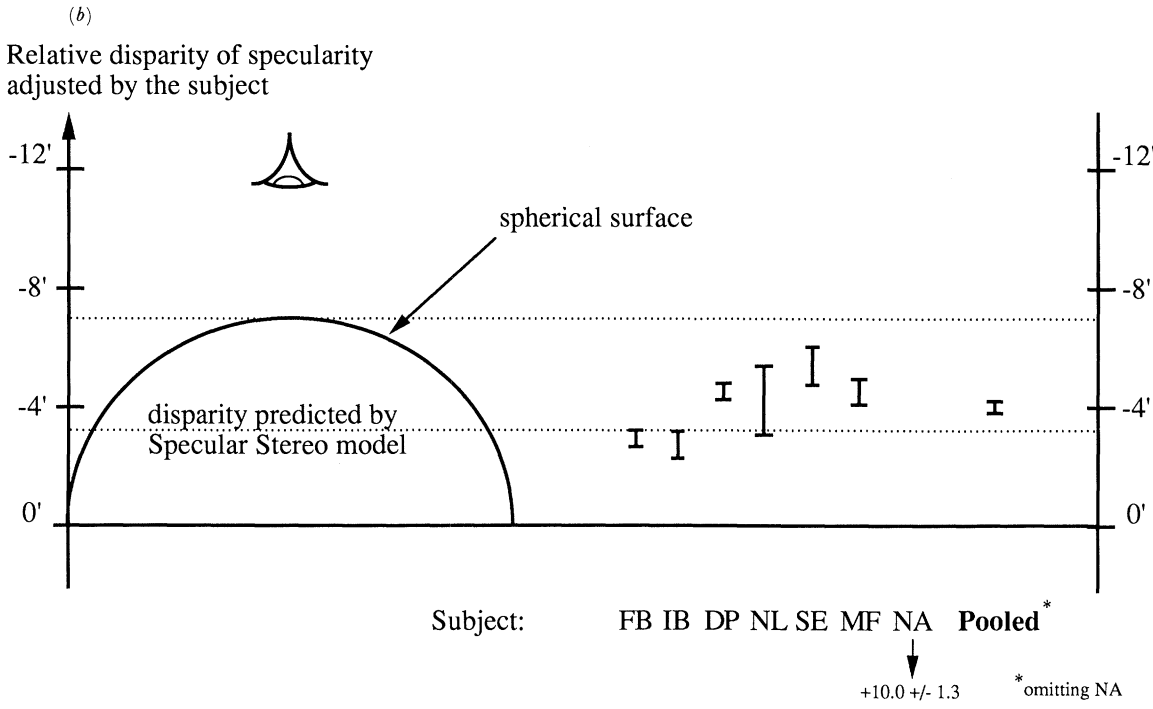


Figure 9. The perception of surface properties can change by moving a specular highlight relative to the surface. The surface of the sphere (*a*) (uncrossed view) looks metallic because the highlight is in the correct position behind the surface. If the highlight is in front of the surface the surface looks duller, and not at all mirror-like (metallic). The human visual system seems to exploit the laws of reflection in the three-dimensional interpretation of two-dimensional images. In the psychophysical adjustment task with the sphere, most subjects place the specular highlight close to the predicted apparent depth (or, equivalently, relative epipolar disparity) (*b*).

Again an incorrect position (relative disparity) of the specularity discounts the bright patch as a specularity and the visual system finds a different interpretation for the way in which the patch was generated. The interpretation of surface property changes from opaque to transparent. When the relative disparity is zero the simulated specularity looks like a powdery patch on the surface and the surface does not look glossy. Note however that in non-stereo images (like any photograph) surfaces can look glossy even even with zero relative disparity. In this case a cue conflict does not really exist because all surfaces are flat and relative disparity does not have any meaning in such images.

In an informal two-alternative forced choice (2AFC) experiment, 11 out of 12 naive observers who were asked which of two presented surfaces was 'polished', chose one shown in figure 9*a*, on which the specularity has divergent relative epipolar disparity, rather than one with convergent relative epipolar disparity. This is in agreement with the prediction of the model.

We wished to test whether the visual system is sensitive not just qualitatively but also quantitatively to stereoscopic relative disparity. We devised an adjustment task in which naive subjects were asked to obtain the appearance of highest surface gloss. They repeatedly pressed buttons which (unknown to the naive subjects) caused the relative disparity of a specularity to vary. They were simply told was that pressing the two buttons would make the surface appear more or less shiny. Either circumpolar disparity was held constant (at the value determined by the ray-optic model) while epipolar disparity was varied or vice versa. Steps in specular disparity for each button press were sufficiently small (2 pixels or about 1.5 arc min) that most of the subjects did not perceive the specularity to be moving in depth. Five test surfaces were used in the adjustment task, a convex sphere, two convex ellipsoids and their reversed, concave counterparts.

Results for the convex sphere (figure 9*b*) show that, on average, subjects adjusted values were not significantly different from veridical for epipolar disparities

Table 2. *Epipolar disparity for convex surfaces (arc min)*

surface	subject	surface disparity	predicted disparity	adjusted disparity	n
horizontal ellipsoid	FB	−12.4	−9.1	−4.60±0.67	20
	IB			−0.41±0.69	20
	DP			−11.25±0.43	40
oblique ellipsoid	FB	−5.8	0.0	−2.87±0.56	20
	IB			−2.59±0.46	20
	DP			−5.09±0.35	40

Table 3. *Epipolar disparity for concave surfaces (arc min)*

surface	subject	surface disparity	predicted disparity	adjusted disparity	n
horizontal ellipsoid	FB	12.4	9.1	12.03±0.64	20
	IB			15.81±0.24	20
	DP			12.75±0.42	40
oblique ellipsoid	MF	5.8	0.0	5.19±0.88	19
	SE			5.54±0.60	20

Table 4. *Circumpolar disparity for convex surfaces (arc min)*

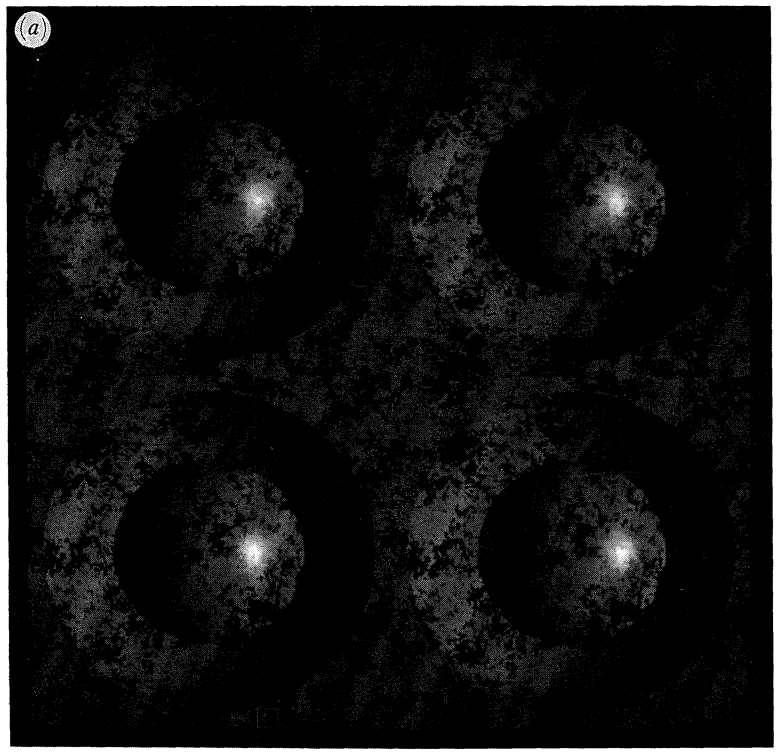
surface	subject	predicted disparity	adjusted disparity	n
sphere	IB	0.5	0.94±0.95	20
	FB		2.29±0.60	19
	DP		0.06±0.54	40
	Pooled		0.82±0.40	79
horizontal ellipsoid	IB	0.7	5.92±1.15	20
	FB		2.79±0.94	20
	DP		0.76±0.34	40
oblique ellipsoid	SE	−3.3	2.51±0.49	20
	MF		−0.29±0.99	20
	AB		−2.26±0.29	20
	HB		−1.60±0.46	20
	IB		−1.52±0.52	40
	FB		−1.44±0.85	20
	DP		−1.97±0.49	20
	Pooled		−1.65±0.24	140

Table 5. *Circumpolar disparity for concave surfaces (arc min)*

surface	subject	predicted disparity	adjusted disparity	n
horizontal ellipsoid	IB	−0.7	10.97±1.29	20
	FB		0.86±0.64	20
	DP		2.18±0.44	40

($p < 0.001$, $F = 2$). It is difficult to get significant circumpolar (vertical) disparity effects for this surface because the veridical circumpolar disparity is close to zero (0.5 arc min). Four naive subjects adjusted the circumpolar disparity close to zero (table 4) but it is conceivable that there is anyway some regression towards zero. We therefore tested a situation in which the correct circumpolar disparity of a specularity was quite different from zero. This is the case for the oblique oriented ellipsoid shown in figure 6. The data under 'oblique ellipsoid' in table 4 show that all five naive subjects and the two authors made adjustments whose signs were as predicted by the model. This is particularly interesting because non-zero circumpolar disparity is more or less solely associated with specular reflections. The visual system apparently has some dedicated competence for analysis of specularities. However, the adjustments are not in quantitative agreement with the model (at a 95% significance level). All adjusted values were biased towards zero disparity. This again suggests regression towards zero, perhaps because most of the time, for non-specular features, circumpolar disparities are close to zero.

Similar results are obtained also for another convex ellipsoid with horizontal orientation of its major axis. Again, the sign of the epipolar relative disparity is



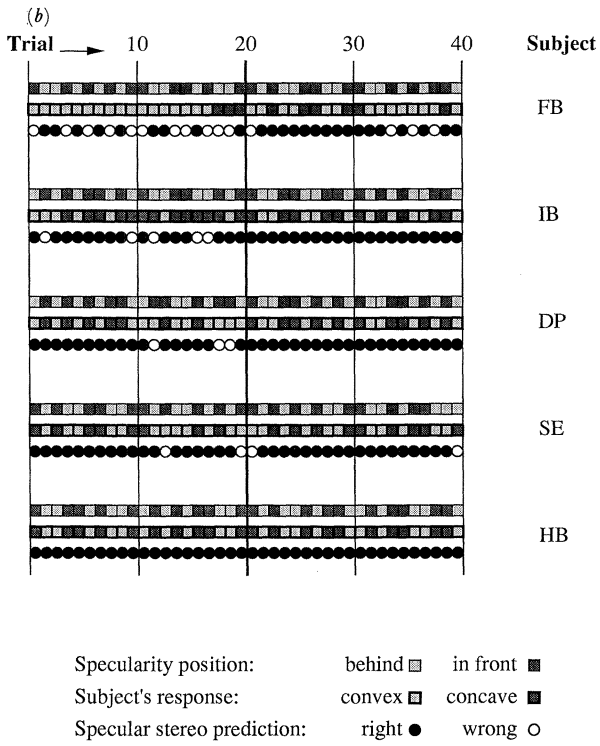


Figure 10. The perception of surface curvature can change with the position of a specular highlight. To determine whether judgements of surface shape are influenced by specular relative disparity we used a stereo image whose three-dimensional interpretation can flip between two states (convex–concave). If a highlight is added to the image the 3D interpretation of the inner part of the surface is biased somewhat towards convex (left image in (*a*)) but still reversible. A stereo pair (*a*) was made with zero disparity, with a specularity superimposed. The specularity could have either convergent or divergent disparity, flipping randomly between the two, with 5- or 10-second exposures separated by a random-dot masking frame. Subjects make a two-alternative forced choice (2AFC) between convex and concave. After a short training period (20 exposures) they make choices which conform to the predictions of the model (*b*). Note that the effect is difficult to reproduce in print because of the limited dynamic range.

always correct (table 2). This corresponds to robust discrimination between convex and concave surfaces as predicted earlier (prediction 5).

Poor agreement is obtained in the case of the concave ellipsoid (tables 3, 5), even the signs of the relative epipolar circumpolar disparities, after adjustment, are inconsistent. Subjects reported that, for this surface, the adjustment task was difficult to perform. The results for all tested surfaces are summarized in figure 9 and tables 2, 3, 4 and 5.

The conclusion of this experiment is that the human visual system models the physics of specular reflection well enough to predict relative disparity effects. Agreement with predictions is good qualitatively (sign is preserved), and there a degree of quantitative agreement, but apparently with some regression towards zero. In particular, in the case of a convex sphere for which we can associate epipolar disparity with depth, the visual system 'expects', correctly, that a specularity lies behind not on a surface (figure 9*b*).

(*c*) *Experiment 2: judgement of surface geometry (curvature)*

The second experiment is complementary to the first. Can the visual system accommodate to variations in specular relative disparities by changing its hypothesis about surface curvature, rather than its hypothesis of glossiness?

We devised the stimulus of figure 10*a*, a stereo, textured variant of an ambiguous (reversible) shaded surface (Ramachandran 1988). The texture elements all have zero disparity, consistent with a fronto-parallel surface. None the less, monocular shading or texture cues are not entirely overriden (Bülthoff & Mallot 1988), so that subjects can usually see both convex (like a ball in a saucer) and concave (like a dog-bowl) interpretations. A superimposed specularity, with either convergent (−) or divergent (+) relative disparity strongly influences the interpretation. As the specular stereo model predicts, convergent (−) relative disparity biasses the subjects' interpretation away from convex. Similarly, divergent(+) relative disparity biasses interpretations away from concave. The prediction was tested by a two-alternative forced choice (2AFC) between a convex or a concave surface interpretation, when subjects were presented with simulated specularities, divergent or convergent (±5 arc min), in random sequence. The 5-, 10- or 15-s exposures were separated by random-dot masking frames. Note that subjects were asked whether the surface appeared convex or concave, not whether the specularity was behind or in front of the surface. Before recording responses subjects could familiarize themselves with the task with a short run of 20 exposures but no feedback as to the correctness of the responses was given at any time.

The effect develops gradually with repeated exposures. Time sequences (figure 10*b*) show, that whilst initially subjects may be locked into one or other interpretation, after around 20 exposures they reliably pick the interpretation that is consistent with the sign of epipolar relative disparity. Note that the change in position of the specularity is contrary to that of the surface, when the specularity is furthest away (divergent epipolar relative disparity) the surface is predicted to be convex, so its centre is nearer to the viewer and vice-versa. Any explanation in terms of a pulling effect exerted by the specularity on the surface is thereby excluded.

8. CONCLUSIONS

In conclusion, let us return to the predictions made earlier and consider which have been addressed by these experiments and what remains to be investigated. Prediction 1 is substantially addressed by the adjustment experiment. Subjects are able, at least for certain surfaces, to detect discrepancies in epipolar and circumpolar specular disparity which they correct by adjustment. When presented informally with a surface in which the specular epipolar disparity was physically inconsistent they reported a loss of the glossy appearance. Even for those convex surfaces for which adjustments of epipolar disparity of specularities accorded less accurately with the model, subjects did not place the specularity on the surface, but behind it, in fulfilment of prediction 3. For concave surfaces behaviour did not accord with the model at all, for reasons which are not altogether clear. For prediction 4, that for reasons of computational robustness the specular stereo cue could overwhelm non-specular shape cues we have only circumstantial evidence. In experiment 2, the stereoscopic cue is actually for a flat surface whilst there is monocular evidence for curvature. In that case the specular stereo cue is able to dominate. Other experiments should be tried to test the prediction more directly, for example, by using textures of varying density. Prediction 5 concerned the robustness of convex–concave discrimination from the sign of specular relative epipolar disparity. This is strikingly demonstrated by experiment 2, in which the discrimination succeeds despite the presence of prominent edge-based stereo cues and strong cue-conflict.

Remaining predictions relate to motion analogues of the stereoscopic effects presented here, to the interaction of monocular and stereoscopic cues for specularity and to stronger effects that could result from extended motions. It would be surprising, given the similarity of the models, if similar results could not be obtained for motion; it remains to be seen whether they can. The other cases are less clear-cut. Does the visual system exploit the extra geometric information that is inherent in extended motions? Does it utilize the monocular shape of the specularity as a further cue to shape? How is that integrated with the specular stereo cue?

We gratefully acknowledge the support of a grant from the Office of Naval Research, Cognitive and Neural Sciences Division (H.H.B.) and the SERC, EEC, the Fellowship of Engineering, Exeter College, Oxford and the University of Oxford (A.B.). We are grateful for discussions with R. Cipolla, J. P. Frisby, H. C. Longuet-Higgins, T. Poggio, S. Ullman and A. Zisserman, and for the assistance of G. Brelstaff in preparation of figures.

REFERENCES

Ardity, A. 1986 Binocular vision. In *Sensory process and perception*, vol. 1 (Handbook of perception and human performance) (ed. K. R. Boff, L. Kaufman & J. P. Thomas). New York: Wiley.

Beck, J. 1972 *Surface color perception.* New York: Cornell University Press.

Blake, A. 1985 Specular stereo. *Proc. 9th IJCAI Conf.* pp. 973–976.

Blake, A. & Brelstaff, G. J. 1988 Geometry from specularities. In *Proc. Int. Conf. Computer Vision*, pp. 394–403. IEEE, Washington, DC.

Blake, A. & Bülthoff, H. H. 1990 Does the brain know the physics of specular reflection? *Nature, Lond.* **343**, 165–168.

Blake, A. & Cipolla, R. 1990 Robust estimation of surface curvature from deformation of apparent contours. *Proc. 1st European Conf. Computer Vision*, Nice, France. (In the press.)

Blake, A. & Brelstaff, G. J. 1988 Detecting specular reflections using Lambertian constraints. In *Proc. Int. Conf. Computer Vision*, pp. 297–301. Washington DC: IEEE.

Bülthoff, H. H. & Mallot, H. A. 1988 Integration of depth modules: stereo and shading. *J. Opt. Soc. Am.* **5**, 1749–1758.

Bruce, J. W. & Giblin, P. J. 1984 *Curves and singularities.* Cambridge University Press, Cambridge, UK.

Bülthoff, H. H., Little, J. & Poggio, T. 1989 A parallel algorithm for real-time computation of optical flow. *Nature, London* **337**, 549–553.

do Carmo, M. P. 1976 *Differential geometry of curves and surfaces.* New Jersey: Prentice-Hall.

Hildreth, E. C. 1984 Computations underlying the measurement of visual motion. *Art. Int. J.*, **23**, 309–354.

Julesz, B. 1971 *Foundations of cyclopean perception.* London: The University of Chicago Press Ltd.

Koenderink, J. J. & van Doorn, A. J. 1976 Geometry of binocular vision and a model for stereopsis. *Biol. Cybernetics* **21**, 29–35.

Koenderink, J. J. & van Doorn, A. J. 1980 Photometric invariants related to solid shape. *Optica Acta*, **27**, 7, 981–996.

Longuet-Higgins, H. C. & Prazdny, K. 1980 The perception of a moving retinal image. *Proc. R. Soc. Lond.* B **208**, 385–397.

Longuet-Higgins, H. C. 1982 The role of the vertical dimension in stereoscopic vision. *Perception* **11**, 377–386.

Marr, D. 1982 *Vision.* San Francisco: W. H. Freeman & Co.

Marr, D. & Poggio, T. 1979 A computational theory of human stereo vision. *Proc. R. Soc. London.* B **204**, 301–328.

Maybank, S. J. 1988 The angular velocity associated with the optical flow-field arising from motion through a rigid environment. *Proc. R. Soc. Lond.* A **401**, 317–326.

Mayhew, J. E. W. & Frisby, J. P. 1981 Towards a computational and psychophysical theory of stereopsis. *AI Journal* **17**, 349–385.

Mayhew, J. E. W. & Longuet-Higgins, H. C. 1982 A computational model of binocular depth perception. *Nature, Lond.* **297**, 376–379.

Phong, B. T. 1975 Illumination for computer generated pictures. *Comm. ACM*, **18**, 311–317.

Ramachandran, V. S. 1988 Perception of shape from shading. *Nature, Lond.* **331**, 163–166.

Rogers, B. J. & Cagnello, R. 1989 Disparity curvature and the perception of three-dimensional surfaces. *Nature, Lond.* **339**, 135–137.

Thorpe, J. A. 1979 *Elementary topics in differential geometry.* New York: Springer-Verlag.

Yuille, A. L. & Grzywacz, N. M. 1988 A computational theory of coherent motion perception. *Nature, Lond.* **333**, 71–74.

Williams, L. 1983 Pyramidal Parametrics. *Proc SIGGRAPH Conf.*, **17**, 1–11.

Zisserman, A., Giblin, P. & Blake, A. 1989 The information available to a moving observer from specularities. *Image and Vision Computing*, **7**, 38–42.

Received 15 October 1990; accepted 30 October 1990

APPENDIX

(*a*) *Derivation of the specular motion equation*

In this appendix it is shown how the equations describing viewing geometry and surface geometry combine to produce the specular motion equation (5).

Image and source positions. The distance from the viewer at $\boldsymbol{v}(t)$ to the point of reflection $\boldsymbol{r}(t)$ (figure 4) is $\lambda(t)$:

$$\lambda(t) = |\boldsymbol{r}(t) - \boldsymbol{v}(t)|. \tag{21}$$

Similarly the distance from the source at $\boldsymbol{S}$ to the point of reflection is $\mu(t)$:

$$\mu(t) = |\boldsymbol{r}(t) - \boldsymbol{S}|. \tag{22}$$

Image motion. Viewer position $\boldsymbol{v}$, point of reflection $\boldsymbol{r}$ and the image (on the unit image sphere) $\boldsymbol{T}$ of the specularity are related by

$$\boldsymbol{r} = \boldsymbol{v} + \lambda \boldsymbol{T}. \tag{23}$$

Differentiating and projecting into the plane orthogonal to $\boldsymbol{T}$ yields the following equation for image motion:

$$\lambda \boldsymbol{T}_t = \pi_{\mathrm{T}}(\boldsymbol{r}_t - \boldsymbol{v}_t). \tag{24}$$

Relative image motion. The fixed surface feature, whose image is $\boldsymbol{T}^*$, is stationary and assumed to be instantaneously coincident with the point of specular reflection so that, in place of (24), we have:

$$\lambda \boldsymbol{T}_t^* = -\pi_{\mathrm{T}} \boldsymbol{v}_t$$

and, subtracting this from (24), we obtain an equation for the relative image motion $\boldsymbol{\Delta}_t = \boldsymbol{T}_t - \boldsymbol{T}_t^*$:

$$\lambda \boldsymbol{\Delta}_t = \pi_{\mathrm{T}} \boldsymbol{r}_t.$$

Projecting orthogonally to $\boldsymbol{n}$ gives $\boldsymbol{r}_t$ in terms of $\boldsymbol{\Delta}_t$:

$$\boldsymbol{r}_t = \lambda \mathscr{P} \boldsymbol{\Delta}_t. \tag{25}$$

Incident light ray. Having dealt with the reflected ray, the incident ray obeys a similar relation, with a stationary source in place of the viewer, and with vector $\mu \boldsymbol{L}$ ($|\boldsymbol{L}| = 1$) from source to point of reflection:

$$\mu \boldsymbol{L}_t = \pi_{\mathrm{L}} \boldsymbol{r}_t. \tag{26}$$

Specular reflection. The geometry of specular reflection is expressed as a relation between $\boldsymbol{T}$, $\boldsymbol{L}$ and the surface normal $\boldsymbol{n}$ ($|\boldsymbol{n}| = 1$):

$$\nu \boldsymbol{n} = \boldsymbol{T} + \boldsymbol{L}, \tag{27}$$

where ν is a normalising constant. Differentiating and projecting with π_{n}:

$$\boldsymbol{n}_t = \frac{\pi_{\mathrm{n}}(\boldsymbol{T}_t + \boldsymbol{L}_t)}{2\boldsymbol{T} \cdot \boldsymbol{n}}. \tag{28}$$

Surface curvature. Finally surface curvature is represented by the Weingarten map W which relates $\boldsymbol{n}_t, \boldsymbol{r}_t$:

$$\boldsymbol{n}_t = W(\boldsymbol{r}_t). \tag{29}$$

Substituting in (29) for $\boldsymbol{n}_t$ from (28) and then for $\boldsymbol{T}_t, \boldsymbol{L}_t$ from (24) and (26), we obtain

$$W(\boldsymbol{r}_t) = \tfrac{1}{2}\pi_n\left(\frac{\pi_T(\boldsymbol{r}_t - \boldsymbol{v}_t)}{\lambda} + \frac{\pi_L \boldsymbol{r}_t}{\mu}\right).$$

Now, using (27) and the tangency condition $\boldsymbol{r}_t \cdot \boldsymbol{n} = 0$, it follows that

$$\pi_n \pi_T \boldsymbol{r}_t = \pi_n \pi_L \boldsymbol{r}_t,$$

so now

$$W(\boldsymbol{R}_t) = \tfrac{1}{2}\Pi\left(\frac{\boldsymbol{r}_t - \boldsymbol{v}_t}{\lambda} + \frac{\boldsymbol{r}_t}{\mu}\right),$$

where Π is as defined in (6). Finally, substituting for $\boldsymbol{r}_t$ from (25) yields the specular motion equation (5).

(*b*) *Convex–concave discrimination: proof of result*

A compact proof follows of the rule for discrimination of convexity–concavity. Given an observation of the sign of epipolar relative disparity

$$\boldsymbol{\Delta}' \cdot (\pi_T \boldsymbol{v}_t),$$

we wish either to disprove that the Weingarten map W is positive definite (convex surface) or to disprove that it is negative definite (concave surface). Now from (18)

$$\begin{aligned}\text{sign}\,(\boldsymbol{\Delta}' \cdot \pi_T \boldsymbol{v}_t) &= \text{sign}\,(\pi_T \boldsymbol{v}_t \cdot (W^{-1}\pi_n \pi_T \boldsymbol{v}_t)),\\ &= \text{sign}\,((\pi_n \pi_T \boldsymbol{v}_t) \cdot (W^{-1}\pi_n \pi_T \boldsymbol{v}_t)),\end{aligned}$$

since W is restricted, both in domain and range, to the tangent plane on the surface. Finally, writing

$$\boldsymbol{z} = \pi_n \pi_T \boldsymbol{v}_t,$$

we have

$$\text{sign}\,(\boldsymbol{\Delta}' \cdot (\pi_T \boldsymbol{v}_t)) = \boldsymbol{z}^T W^{-1} \boldsymbol{z},$$

which is positive only if W^{-1} and hence W is not negative definite, so that the surface is either convex or hyperbolic. The proof for concave–hyperbolic curvature is similar.

(*c*) *Degeneracy on the caustic*

Here is the proof that the full model (5) is degenerate, whether or not the Weingarten map is of full rank, when the viewer hits a caustic surface. The model equation (5) becomes singular when $\Pi \boldsymbol{v}_t = 0$ for some $\boldsymbol{r}_t \neq 0$. Now $\boldsymbol{v}$ is on the caustic when it is on a reflected ray,

$$\boldsymbol{v} = \boldsymbol{r} + \lambda \boldsymbol{T},$$

and when also, for some $\boldsymbol{r}_t \neq 0$, when also

$$v_t = 0,$$

(Bruce & Giblin 1984). In that case, certainly

$$\Pi \boldsymbol{v}_t = 0,$$

so that the model equation is singular.