

Ambiguity of shading and stereo contour

Andrew Blake¹

Department of Computer Science,
University of Edinburgh,
King's Buildings, Mayfield Rd,
Edinburgh EH9 3JZ,
Scotland.

ABSTRACT

Imagine a smooth surface patch, bounded by a closed contour which is observed stereoscopically. The interior of the patch presents an interpretation problem because of its lack of matchable features. Shading inside the patch is a potentially rich source of information. Its exploitation is facilitated by the stereoscopically matched contour, which supplies boundary conditions on surface shape.

Interpretation of the shape of the patch from shading and the stereo contour, is subject to certain ambiguities. The well known Necker-like inversion ambiguity of the surface, with corresponding reflection of the source, is generally capable of being resolved by the stereo contour. But there are three other forms of ambiguity which can occur. The first relates to source position; it may be determined entirely, or else restricted to a set of positions. The second ambiguity is of the contour strip - the contour labelled with depth and surface orientation - which, for general source position, can undergo just a single, global inversion. The third is of the surface patch itself, given source and contour strip: it is generally determined, but bistable ambiguity may be induced by the presence of two or more points of maximal intensity.

1. Introduction

Stereoscopic vision [29,30] retrieves richest shape information when it is directed at a densely textured surface. The depth map is irregularly sampled but can be converted to a dense, regularly sampled form by interpolation or approximation, known as surface reconstruction [20,21]. At this point it is similar to the dense depth map obtained from an optical rangefinder. It may already be useful, for example for collision avoidance, in this form. A mobile viewer can compute the intersection, if any, of a proposed path with the visible surface. Alternatively, for matching parts of stored object models to the observed surface, further descriptive processes may be needed. Edge detection can be used to delineate occluding and connect boundaries in the visible surface [10,33]. Moreover reconstruction and edge detection can be combined in a single process that is viewpoint invariant and has superior accuracy in edge localisation [9]. Edges may themselves be subject to segmentation and description [2,8]; shape descriptors may be computed also for surface patches [11]. All this information might be organised into a symbolic relational description in the form of a graph [1,14], suitable for matching with appropriately expressed object models.

Over texture, reconstruction is unambiguous because texture elements and hence stereo depth values are denser than the "scale of interest" - the characteristic scale used in reconstruction. When visible features are sparsely distributed however, as when viewing smooth surfaces of uniform albedo, the only photometrically invariant features are those generated by connect edges, and occluding and extremal contours. The viewer can recover these contours as space curves but, as far as stereo processing is concerned, surface shape between contours is grossly underdetermined. There is however another potential source of information - the shading distribution between contours.

Grimson's [20] original reconstruction scheme was based on a minimal use of shading information - broadly that smoothness of shape suggests smoothness of the underlying surface. Turning this notion rigorously into a reconstruction scheme that captures just the right constraints on surface shape proves difficult however [4]. The advantage of Grimson's scheme is in making minimal use of shading, requiring no specific information about scene illumination. It is therefore robust and relatively tolerant of poor intensity data. The disadvantage is that, in some measure, the baby has been thrown out with the bathwater - there is insufficient information to make unambiguous reconstructions (figure 1).

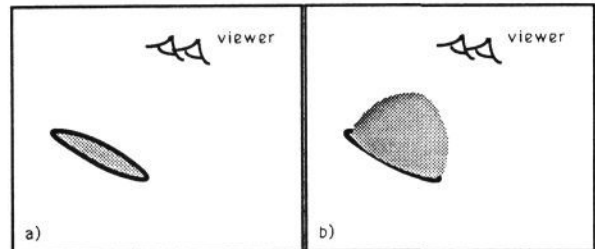


Fig 1. A circular contour bounding a smooth shaded surface patch is viewed stereoscopically. There are many possible patches consistent with such a contour, including a disc (a) or an egg shape (b). Grimson's reconstruction algorithm opts for (a), but in reality it has no grounds for excluding (b). Appropriate processing of the patch's shading could resolve this problem.

Algorithms for recovery of "shape from shading" have been investigated extensively [24,26,34]. Most effort has been directed at the case of a shaded patch with an extremal boundary, a sphere for example, viewed under orthogonal projection. However Ikeuchi [27] investigated a variant of the problem which is the one we address here: viewing a smooth, shaded patch, bounded by a stereoscopically viewed contour (figure 2). Stereo processing establishes the "contour generator" as a space curve. The precise photometric properties of the scene illumination are used to derive the "contour strip" - 3D position and surface orientation along the space curve. The contour strip provides boundary conditions for shape from shading by relaxation. A drawback of this method is the requirement for a precise characterisation of illumination. But the reward is the ability to derive precise surface shape.

This paper considers just what sort of information is in principle available from stereo contour and shading. To what degree is inferred shape ambiguous? Is it necessary to be furnished in advanced with precise photometric specifications of illumination and illuminated surface? In order to address these questions, some assumptions must of course be made, for example about imaging geometry and the photometric properties of the surfaces.

Projection: parallel projection is assumed, as a reasonable approximation to perspective projection when viewing distance is large compared with the size of the surface patch under consideration.

Reflectance map: It is assumed that locally, over a particular surface patch, image intensity is describe by the irradiance equation $E(x) = R(p)$, where $x = (x, y)$ is position in the image, $p = (p, q)$ is surface orientation and R is the "reflectance map" [25].

¹The Author's current address is: Department of Engineering Science, University of Oxford, Parks Rd, Oxford.

Surface properties: surfaces are assumed to be smooth over the area of interest. This is a reasonable assumption, for example, in the feature-free interior of a stereo-matched contour. Surface reflectance is assumed to be a combination of lambertian [25] and specular, as is common in computer graphics. For a lambertian surface under a source characterised by a vector l (in direction $l/|l|$ with albedo/source-strength product $|l|$) the reflectance map is

$$R(p) = l \cdot n, \text{ when } l \cdot n > 0, \quad R(p) = 0 \text{ otherwise,} \quad (1)$$

where n is the surface normal vector corresponding to orientation p :

$$n = (p, 1) / \sqrt{1 + |p|^2}$$

and

$$p = (n_x/n_z, n_y/n_z).$$

It is assumed that specularities (highlights) can be detected [12] and excised. Note that there is never any need to know absolute surface albedo; only the product of albedo and source strength affects the form of $R(p)$.

Illumination: The basic model is that there is one point source and a constant level of ambient illumination so that

$$E(x) = R(p) = R_{min} + l \cdot n \text{ when } l \cdot n > 0, \\ = R_{min} \text{ otherwise.} \quad (2)$$

The brightest possible intensity generated by the reflectance map is $R_{max} = R_{min} + |l|$. Even with multiple sources l_i , reflectance maps add to give an effective source $l = \sum l_i$ (provided none of the sources are shadowed).

Shading analysis is assumed to be monocular. Stereoscopic analysis of shading - observation of small differences of shading at corresponding points in the 2 eyes' views - is unusable with lambertian surfaces and unstable with specular ones [22]. Stereoscopic area correlation [3] is effective only at short range [7], whereas shape from shading with stereo would be effective at medium range; only shading cues can be useful at long range.

Under the above assumptions, this paper presents two results that provide at least partial answers to the question of ambiguity. The first (section 2) applies when illumination parameters (direction, strength, ambient) are known. It has two parts: one relates to the ambiguity of a "contour strip" - inferred surface orientation along a closed, stereo contour, which has a bistable ambiguity. The other result shows that, once the contour strip has been established, the surface inside the contour is uniquely determined, except in some special cases. The second result (section 3) concerns estimation of illumination parameters. Suppose albedo-source-strength product and ambient illumination (R_{min} and R_{max}) have been obtained by observation of shadows and "maximal" points. It is shown that the **irradiance** and **orientation** constraints define a set of directions on the gaussian sphere, inside which the source direction must lie (figure 2). An initial source position inside that feasible set could be adopted, and refined iteratively by a procedure such as that of Brooks and Horn [13] or Brown et al. [15]. Quite often, the source direction is located exactly, at one of the "double-degenerate" points on the boundary of the feasible set.

2. Uniqueness

2.1 Surface reversal ambiguity

It is well known that for given shading, source and a consistent interpretation of the underlying surface $Z = u(x)$, there is a reversed version of the surface $Z = A - u(x)$ where A is a constant, that is consistent with the same shading and a reflected source $(-l_x, -l_y, l_z)$ (viewing along the z -axis). This is easily seen by inspection of the irradiance equation (2), replacing p by

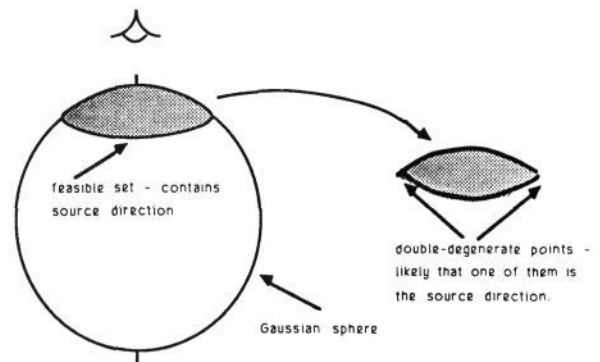


Fig 2. The irradiance and orientation constraints generate a feasible set on the gaussian sphere in which the source direction must lie. An argument based on the assumption of general source position shows that the source is very likely to be at one of the "double-degenerate" points.

$-p$. Of course, a fixed source is assumed here, in which case such a reversal could not occur. But even if the source were allowed to move, a reversal should not, in general, occur when there is also a contour C , visible in stereo. This is because depths along the fixed contour $u(x)$, x on C , cannot reverse (unless $Z = u(x) = \text{const}$, x on C). However, in the human visual system a similar reversal, the Necker ambiguity, can be sufficiently strong to persist even in the presence of such conflicting stereo information [19].

Maximal points of the image intensity, where $E(x) = R_{max}$, are important in reversal ambiguity. Maximal points always appear as local maxima of $E(x)$, although a local maximum might alternatively be a parabolic point of the surface [28]. The normal at a maximal point, from (2), must be $n = l/|l|$ - it points towards the source. Hence a contour of maximal points lies in a plane whose normal points towards the source. When source and viewer are in the same position then, it is plain from the previous paragraph, reversal can occur (figure 3).

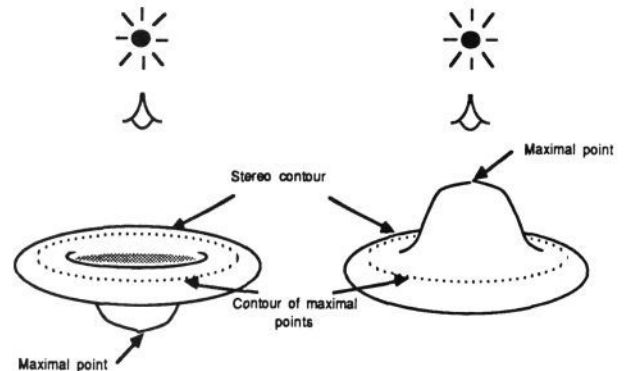


Fig 3. The surface on the left can reverse to give the surface on the right. This ambiguity is enabled by the presence of maximal points, regardless of the form of the stereo contour.

When the source is elsewhere, the reversal cannot occur in the same way, but the proof of the uniqueness theorem (stated later) may fail, suggesting a possible ambiguity. Unfortunately, we have not found any example of ambiguity under those circumstances.

2.2 Establishing the contour strip

The contour generator is recovered by stereo processing as a space-curve $(x, u(x))$, for $x(r)$ on C , which will usually be assumed to be closed, parametrised by arc-length r and enclosing a shaded, edge-free patch. The curve tangent is $t(r)$ and the surface normal at $x(r)$ is $n(r)$. Now the curve tangent must lie in the surface, and this constrains n :

$$n \cdot t = 0. \quad (3)$$

And if the image intensity just inside the contour at $\kappa(r)$ is $E(\kappa)$, then the angle θ between l and n is given by the irradiance equation (2):

$$\cos\theta = (E - R_{min}) / (R_{max} - R_{min}). \quad (4)$$

These 2 constraints can be drawn on the gaussian sphere, as in figure 4, with (3) represented by a great circle, and (4) by another circle that forms the base of a cone of semi-angle θ . The intersection of the circles produces 2, 1, or 0 solutions for the direction n . If there are no solutions (figure 4c), that suggests that original assumptions have not been met in some way. The prime possibility is that source parameters are wrong; this forms the basis of the "irradiance constraint", as used later to constrain source position. The normal situation is to have 2 solutions for n (figure 4a). In between, there is a degenerate case (figure 4b) in which the two solutions coincide; we call a point on the contour at which this occurs a **degenerate point**. It plays an important role in the ambiguity of the contour strip.

2.3 Degenerate points and ambiguity of the contour strip

The procedure just described for computing the surface normal n is applied to obtain $n(r)$ at all points on the contour. Generally there will be a twofold ambiguity locally, at each contour point, generated by the 2 solutions on the gaussian sphere. But that does not mean that $n(r)$ can flip independently for each r . The assumed smoothness of the underlying surface means that $n(r)$ must be a continuous function of r . The normal $n(r)$ can only flip as r passes through a degenerate point on the contour; this is because the 2 solutions for $n(r)$ coincide at such a point, so that the flip can occur there without causing a discontinuity in $n(r)$. It looks rather as if the contour is divided into segments by its degenerate points, and that flipping of $n(r)$ can occur independently in each segment (figure 5).

In principle, the contour strip can have an independent bistable ambiguity in each of the intervals between degenerate points. In practice, it is now shown, this occurs but rarely. For general source position, it transpires that there is just a single, global, bistable ambiguity for the whole strip. Only for very special source positions can independent ambiguities happen. To see why this should be, it necessary to look more closely at degenerate points. Careful inspection of figure 4b shows that a degenerate point is a position r along the contour at which $l, n(r), t(r)$ all lie in the same plane. To put it another way, a degenerate point is a zero of the function $P(r)$, defined as:

$$P(r) = l \cdot (n(r) \times t(r)). \quad (5)$$

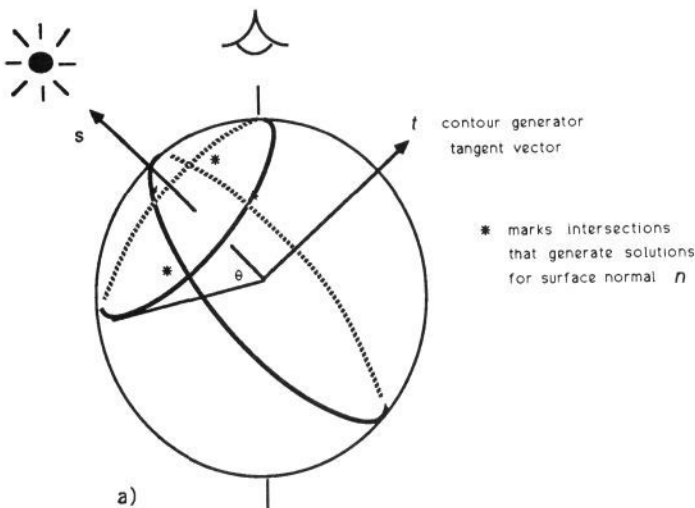


Fig 4. Irradiance equation at a contour point.

Constraints on the surface normal, at some point along the contour, generate two solutions (a), or no solutions (c) for the surface normal n . In the intermediate case (b) circles on the gaussian sphere representing the constraints touch to produce one solution only. The 2 solutions in (a) have coincided. This is called a degenerate point and plays an important role in the ambiguity of the contour strip.

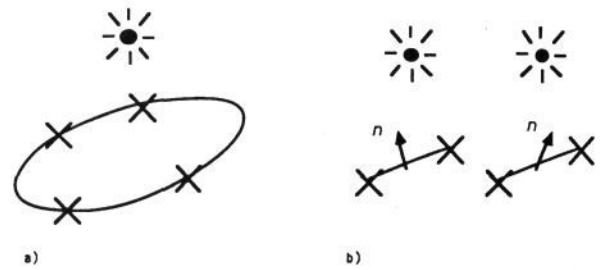


Fig 5. Role of degenerate points in ambiguity of the contour strip.

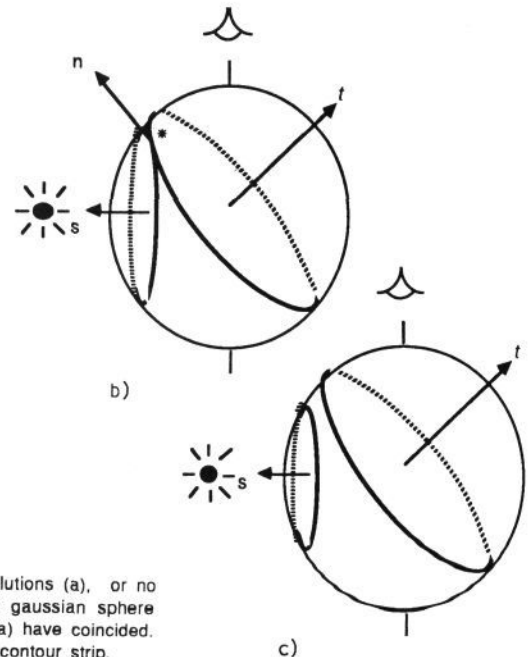
Solving for the normal, several degenerate points may be found on the contour, as marked in (a) by crosses. The contour strip can flip independently in each of the intervals between degenerate points (b).

The sign of $P(r)$ says on which side of the plane containing the contour generator tangent t , and the surface normal n , the source l lies; each sign of $P(r)$ corresponds to one of the two solutions in figure 4a. To be stable, with respect to perturbations in l , a zero of $P(r)$ must of course be a zero-crossing. So for general source position source l passes through the plane of t and n . This means that the contour strip does flip at a degenerate point. Flipping is not independent in each interval between degenerate points because the sign of $P(r)$ has to change from one interval to the next. So there is just one global reversal ambiguity for the entire contour strip. This is true whether the contour is closed or not. Furthermore, on a closed contour, over which r is periodic, there must be an even number of zeros of $P(r)$, and therefore an even number of degenerate points.

2.4 Ambiguity of the surface patch

The contour strip forms a boundary (assumed closed for the time being) enclosing a shaded patch in which the irradiance equation $E(\kappa) = R(\rho)$ holds. Position and gradient $\kappa(r)$ and $\rho(r)$ around the contour constitute boundary conditions for the irradiance equation which is a first order partial differential equation (FOPDE). The pressing question is whether, given contour strip, shading (E) and illumination (R) there is enough information to determine the shape $u(\kappa)$ of the surface. In other words, does the FOPDE have a uniquely determined solution?

The appropriate tool to investigate this question is the "characteristic strip" [18]. It converts the FOPDE which has 2 spatial variables $\kappa = (x, y)$ into a family of ordinary differential equations, in a single path parameter s . From a starting position $\kappa(0)$, at which gradient $\rho(0)$ is known, gradient $\rho(s)$ is



given by the ODE along a path $\kappa(s)$. The path with its associated gradient information is a characteristic strip. They were first investigated for solving shape from shading problems by Horn [24], who found them to be computationally unstable. Errors due to noise accumulate along the strip. Bruss [16] however found that they were very helpful for solving the sort of uniqueness problem addressed here. She successfully demonstrated uniqueness for shape from shading with an extremal contour and source at the eye. (Note that any such closed extremal contour must be lit by a source at the eye; if the source were elsewhere some of the contour must be self-shadowed, in which case the shape of that area of surface in shadow is entirely undetermined - non-unique. So Bruss's problem only makes sense for a source at the eye, unlike shape from stereo and shading.)

Before considering the ambiguity question, is there necessarily even a single consistent solution of the FOPDE, for the shape of the patch? The answer [6] is: not necessarily. It is perfectly possible to find examples, under any specified illumination conditions, of a contour strip enclosing a shaded patch which no surface could possibly have generated. Of course, under our assumptions about lambertian reflectance, illumination etc. such inconsistent data could never occur (although noise will introduce minor inconsistencies in practice). So if there is an inconsistency this must be because the assumptions did not hold. The prime suspect must be illumination. Imagine, as an armchair experiment, trying many different light sources until one is found that produces a consistent solution for the surface. That suggests that there is spare information in the system available for inference of source position. And it is precisely this hope that is exploited in section 3.

Uniqueness Theorem Given a closed, smooth contour strip $\kappa(r), p(r)$, bounding a patch containing a smooth surface $u(x)$, shaded with image intensity $E(x)$, with reflectance map $R(p)$ as above, and with at most one "maximal" point (where $\partial R/\partial p=0$), then the surface $u(x)$ is uniquely determined.

Some comments on the conditions are in order. The smoothness assumptions preclude extremal boundaries. The reflectance map assumption means that no part of the surface is self-shadowed with respect to the source; self-shadows can be observed in the image, and if one is present, its boundary can be used as part of a contour around remaining non-self-shadowed areas. Maximal points are detectable in the image as places where $E=R_{max}$ (and $\partial E/\partial x=0$).

Proof

A proof of a restricted version of the theorem (source at viewer) appears in [6]. The full proof is given in [7].

There are several ways in which the preconditions for uniqueness can be extended.

1. The above argument works not only for a lambertian reflectance map, but for any reflectance map which in some frame (e.g. a source frame) has the form $R(p)=f(|p|)$, for some function f . This does not cover specular reflection except when the source is at the viewer. However, if a specularly is detected [12] and excised, leaving only lambertian surface, the above result will still hold, at least for that portion of the patch that is accessed by characteristic strips emanating from the contour.

2. The one maximal point allowed in the above conditions, could equally well be a connected maximal area. All points in such an area have the same altitude u in the source frame, so no characteristic strip can visit the patch twice.
3. The conditions assumed a known contour strip. It was established earlier, that the contour strip has a bistable ambiguity. This suggests that there might be 2 solutions. It is reasonable to conjecture however, given the spare information in contour strip, illumination and shading (section 2.5 on existence), that a solution surface will not exist for both bistable states.
4. If the contour is not closed, then the uniqueness result applies to the patch of the visible surface that is accessible to characteristic strips emanating from the contour. But there is no obvious way of determining this patch without fully solving the shape from shading problem.
5. Additional maximal points or areas, it can be shown [7], either preserve uniqueness or introduce bistable modes of inversion. An example of inversion ambiguity induced by maximal points was given, in the special case of a source at the viewer, in figure 4. Other examples are hard to come by. A reasonable conjecture seems to be that, for fixed illumination, even with several maximal points (with non-vanishing 2nd derivatives) in the patch, inversion ambiguity could occur only for special alignments of source/viewer/surface, and is therefore negligible.

3. Inferring illumination

The inference of shape from stereo contour and shading requires an advance specification of the illumination field. Of course, illumination can be determined if appropriate assumptions are made - that the surface is isotropically textured [31] for example. And prior information about surface orientation on a given patch can be used to compute the illumination field (figure 6): for example, surface orientation can be inferred directly at a tangent discontinuity along a contour, on a smooth surface; surface orientation is known throughout a surface of revolution (suggested by symmetrically opposed contours). Measured intensity at 4 points whose surface normals are not all coplanar, is enough to deduce the 4 parameters of the lambertian reflectance map (including ambient illumination). Moreover, certain fortuitous observations may constrain illumination, such as the conjunction of a self-shadow boundary and a specularly, which imply a source located almost along the line of sight.

The discussion of existence in section 2.4 strongly suggested that some information about illumination should be available from the very stereo contour and shading themselves. That is because surface shape is generally overconstrained by the combined information contained in illumination, stereo contour and shading. Given just stereo contour and shading, therefore, most illumination fields would be inconsistent, leaving perhaps one or just a few consistent ones.

Although characteristic strips have provided motivation, they seem not to be a tractable technology for estimating the illumination field. But constraints on illumination can be found, by other means. First, it is necessary to estimate R_{min} and R_{max} , by appropriate observation of self-shadows and maximal points. Then two constraints can be obtained, in the form of "feasible sets" on the gaussian sphere, within which the source vector l must lie. The "irradiance" and "orientation" constraints prove to

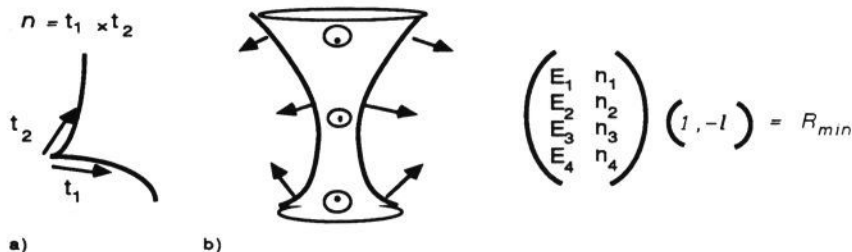


Fig 6. A priori knowledge of surface normals allows illumination to be inferred. Surface normal is known at a tangent discontinuity in a contour that lies on a smooth surface (a). And surface orientation is known throughout a surface of revolution (b). If intensity is measured at 4 points with non-coplanar normals, then the reflectance map can be obtained by solving 4 simultaneous equations, as shown (or using pseudo-inverse with more than 4 points).

be mutually complementary. The irradiance constraint most powerfully constrains the source where surface orientation n is close to the source vector l . The orientation constraint is strongest in exactly the opposite case: where the surface is almost self-shadowed (n is almost orthogonal to l).

3.1 Determining source strength and ambient illumination

The irradiance and orientation constraints apply only to the direction $l/|l|$ of the source vector; it is assumed that R_{min} and $|l|$ or equivalently R_{min} and R_{max} are known. They can be determined by observation of maximal points and self-shadows, in the following way. A self-shadow is characterised by the sort of intensity profile shown in figure 7; the intensity at a self-shadow is precisely R_{min} . So whenever part of the stereo contour is a self-shadow, R_{min} will be known. As for R_{max} it is, of course, the intensity at a maximal point. The problem is in observing a maximal point: a local maximum of intensity may be a maximal point but it might also be a parabolic point on the surface. Compensation for imperfect knowledge, arising from defects in detection of or absence of shadows and maximal points, can be made by observing nearby patches. The ratio $A = R_{max}/R_{min}$ is independent of surface albedo; it is a property of the illumination field only and can be expected to vary slowly across the image. Thus, if one of R_{max}, R_{min} has been measured on the patch, the other can be computed using A from a nearby patch. A further compensation is that, even when self-shadow is visible in the patch, the darkest point of the patch gives an upper bound for R_{min} , which will mean that the irradiance constraint can still be used. If there is no maximal point either, but A has been obtained from a nearby patch, then AR_{min} is an upper bound for R_{max} . Again, this is sufficient to enable irradiance constraint to be used.

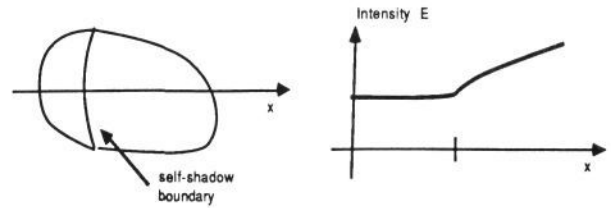


Fig 7. A self-shadow boundary has a recognisable intensity profile

3.2 The irradiance constraint

The irradiance constraint represents the fact that, at each contour point, the situation in figure 5c in which there is no solution for surface orientation, is forbidden to occur. Several steps are needed to get from figure 5c, to a feasible set on the gaussian sphere in which the direction of the source vector l must lie (figure 8).

The shaded band (constraining the direction of the source vector l) in figure 11c is bounded by two circles C_1 and C_2 , which are described algebraically by

$$l \cdot t(r) = \pm \sin \theta \quad (6)$$

The band represents the constraint on source position due to one point $x(r)$ on the contour. To obtain the combined constraint for all points of the contour, the band itself must be swept as r varies around the contour, and a set constructed containing points which lie in the band for all r . In other words, the set is the intersection of the whole family of bands.

The extent of the bands is further restricted by the visibility constraint: for all r on the contour, the surface normal $n(r)$ should lie in the visible half of the gaussian sphere:

$$n \cdot v > 0.$$

This restricts each band (figure 8c) as it continues round to the invisible half of the gaussian sphere, to lie within the circles

$$l \cdot (v \times t) < \cos \theta \text{ and } -l \cdot (v \times t) < \cos \theta.$$

A weaker form of the visibility constraint, which is very easy to compute, may be derived entirely from the brightest point on the contour, for which θ in (4) is smallest, say $\theta = \theta_{min}$. The

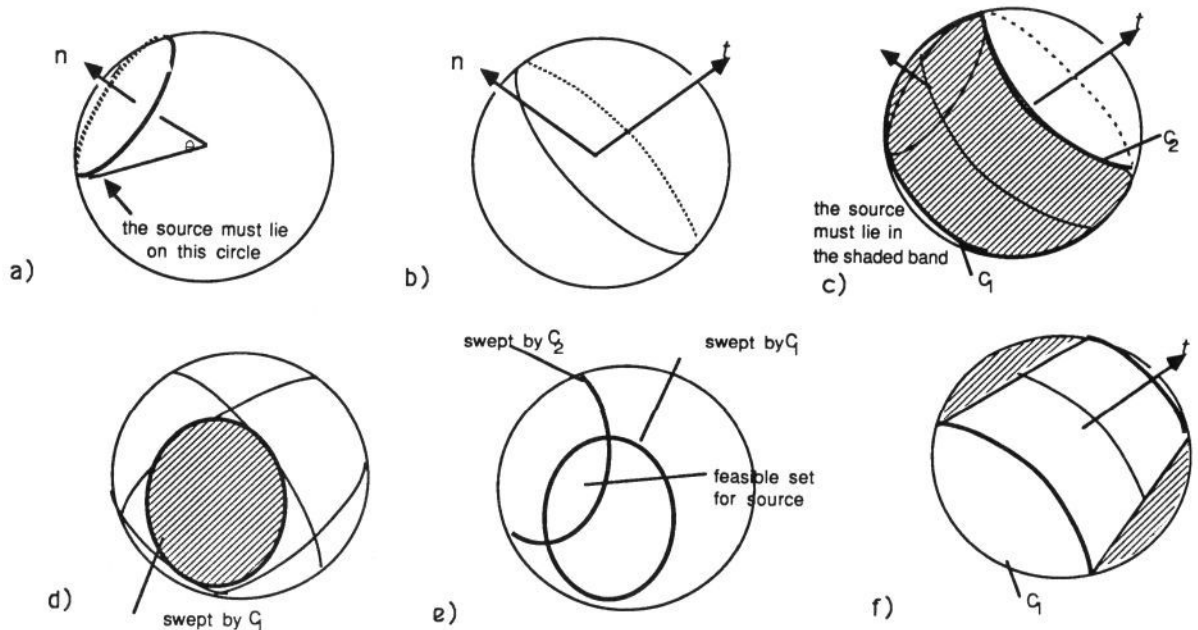


Fig 8. Construction for the irradiance constraint, on the gaussian sphere.

For a given surface normal vector, and known intensity, the source lies on a certain circle (a). But the normal at a contour point lies anywhere on a great circle orthogonal to the contour tangent (b) so the circle in (a) sweeps out a band (c), bounded by curves C_1 and C_2 , in which the source must lie. Now, combining information from all points on the contour, the source must lie in the intersection of all the bands. This is formed by sweeping (d) both of the bounding curves to form two sets which are intersected (e) to give, at last, the feasible set for the source. Only the visible half of the gaussian sphere is shown. In (f) the band in (c) is shown as it appears on the invisible hemisphere; it is restricted to the shaded regions, by the visibility constraint.

constraint on the source vector l , is that

$$l \cdot v > -\sin\theta_{min}. \quad (7)$$

Envelope construction

The construction of figure 8 can be simplified under certain circumstances. The brute force way to compute the feasible set (figure 8e) is to construct bands (figure 8c) for each (sampled) point on the contour and paint them into an array representation of the gaussian sphere. The painting operation must include intersection with the current contents of the array so that, when all bands have been painted, the array contains just points that lie in every band. A much cheaper and more elegant approach is to construct, algebraically, the envelopes of curves C_1 and C_2 , as

r sweeps around the contour. This is done by differentiating (6) with respect to r to obtain

$$l \cdot (dt/dr) = \pm \cos\theta \, d\theta/dr$$

or, using (4),

$$-\kappa(r)l \cdot N(r) = \pm \cot\theta \left((dE(x(r))/dr) / (R_{max} - R_{min}) \right) \quad (8)$$

where $\kappa(r)$ is the curvature of the contour generator at $x(r)$, and $N(r)$ is its normal vector (the contour generator's, as distinct from the surface normal $n(r)$). The combination (6) and (8), solved simultaneously, specify the envelopes for C_1 and C_2

(taking \pm to be $+$ and $-$ respectively, in both (6) and (8)). The envelopes form set boundaries (the heavy curve in figure 8d), and the sets are then intersected to form the feasible set (figure 8e). Computational complexity is much improved, involving only a few algebraic operations for each point on the (sampled) contour; the brute force method required an entire band (figure 8c) to be painted, for each contour point.

The envelope construction method is appropriate only for a closed contour (so that the envelopes are closed). It can be shown that a sufficient condition for the envelopes to be valid is that each pass the following test: there must exist a plane such that, when the envelope is projected from the centre of the gaussian sphere, onto the plane, its projected image is convex. For example, gradient space is constructed by just such a projection; in practice it would be convenient to inspect the envelopes for convexity in the viewer's gradient space. Even if the envelope is not convex, then its convex hull certainly contains the feasible set - although that may be as expensive to compute as the brute-force painting method.

One more point is worth making. The construction of the feasible set under the irradiance constraint does not depend crucially on knowing exact values for R_{min} and R_{max} . Upper bounds on these values are good enough. This is because, in (4), θ varies monotonically with both R_{min} and R_{max} . An upper bound on R_{min} is readily obtainable, even in the absence of a self-shadow (which gives an exact value), simply by observing the darkest intensity occurring in the patch. And if $A = R_{min}/R_{max}$ is known (section 3.1) then that also gives an upper bound on R_{max} . Because of the monotonicity, these upper bounds for θ at each contour point mean a looser constraint (a larger band) in figure 8c. The feasible set obtained therefore still contains the source, and so is still valid.

3.3 Double-degenerate points may give exact source direction

The boundary of the feasible set under the irradiance constraint is important. It can be shown [7] that, under an assumption of general source position, points on the gaussian sphere at which the envelopes of C_1 and C_2 cross - "double-degenerate points" as marked on figure 9 - are particularly likely directions for the source vector.

Provided the source vector l does not lie inside the cone (figure 10b), the source position must be given by one of the double-degenerate points. This is clearly least likely to work when the stereo contour is close to being extremal, and most likely when the contour lies on an "almost" planar surface.

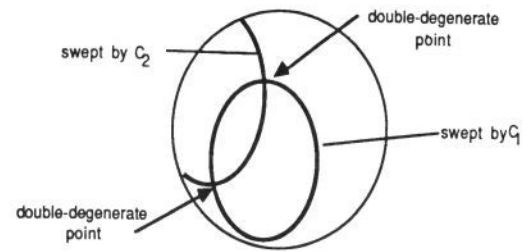


Fig 9. Double-degenerate points are likely source directions

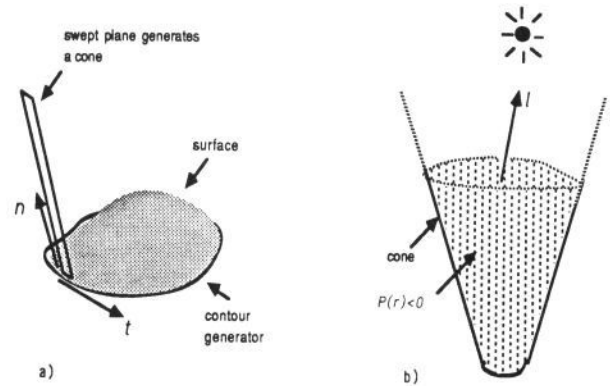


Fig 10. Cone construction for degenerate points.

The plane containing tangent and normal is swept around the contour (a). If the source vector lies inside the cone there are no degenerate points (b).

A reasonable strategy to capitalise on this would be as follows. First construct the feasible set for the irradiance constraint. Then test the double-degenerate points (if any), perhaps by using them as source positions in iterative search by shading [27] or possibly a qualitative variant based on [17]. If any leads to a low error residual in the surface fit, it is a consistent source and a consistent surface has been found. (In principle, it is also possible to run a purely local test on the double-degenerate points; this is because characteristic strips run parallel to the contour at a degenerate point, so that the existence check of figure 8 is entirely local.) Otherwise the source lies strictly inside the feasible set boundary, and its position is further constrained by the orientation constraint as described in the next section.

3.4 The orientation constraint

Suppose, following application of the irradiance constraint, it is found that the source direction indeed lies strictly inside its feasible set. This means that there are no degenerate points on the contour, so that throughout the contour $P(r) \neq 0$ - either $P(r) > 0$ or $P(r) < 0$. The source must retain a fixed orientation, staying on one or other side of the plane containing $n(r)$ and $t(r)$. Combining this fact with the requirement that n be visible ($n \cdot v > 0$) further restricts the bands of figure 8c, as shown in figure 11. For a single contour point, the band of allowed source directions (figure 8c) has a disc eaten out of it as in figure 11a (when $P(r) > 0$) or figure 11b (when $P(r) < 0$). One of the half-discs (C_3 or C_4) forms part of the new boundary of the band. Each of C_3 and C_4 is swept out (as r varies around the contour) to form a pair of swept curves which are simply copies of one another, rotated through 180° about the viewer direction v . They bound areas on the gaussian sphere whose union is the feasible set under the orientation constraint (figure 11c). This feasible set is then intersected with the feasible set under the irradiance constraint, to arrive at a final feasible set.

The equation for C_3 is

$$l \cdot n(r) = \cos\theta \text{ with } P(r) = l \cdot (n(r) \times t(r)) > 0, \quad (9)$$

where $n(r) = v \times t(r)$. As with the irradiance constraint, it may be possible, when the contour is closed, to construct the swept curves as envelopes by differentiating (9) to give:

$$l \cdot (v \times dt/dr) = -\sin\theta \, d\theta/dr$$

or

$$\kappa l \cdot (v \times N(r)) = -(dE/dr) / (R_{max} - R_{min}) \quad (10)$$

which is then solved, simultaneously with (9), to obtain the direction of l .

Finally, the earlier claim that the irradiance and orientation constraints are complementary is easily justified. From figure 8a, it is apparent that the irradiance constraint is strongest when θ is small, that is when image intensity at the contour is greatest, and the surface normal n is nearly aligned with the source vector l . Whereas (figure 11a) the orientation constraint is strongest in just the opposite situation, when θ is large and the contour is nearly extremal.

3.5 Results

The computation of feasible sets has been simulated for the surface and viewing geometry shown in figure 12. A dome shaped surface, with a visible contour strip around its base is viewed along its axis of symmetry, with a source off axis. The cone construction of figure 10 can be performed to show that there will be degenerate points on the contour whenever $\alpha > \beta$. Results given in figure 13, illustrate one condition with degenerate points (figure 13a,b) and one without (figure 13c-f). The displays show the visible half of the gaussian sphere.

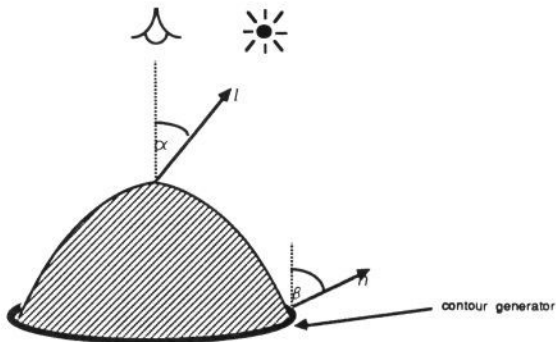


Fig 12. Surface used in simulation of computation of feasible sets.

Whereas figure 13 illustrates the computer plots of analytical formulae for the situation of figure 12, a program has also been written to construct feasible sets from stereo images. Computer generated images for the surface of figure 12 (with $\alpha=40\%$, $\beta=20\%$) are displayed in figure 14a. Noise with standard deviation of 5% of dynamic range has been added. Results are shown in figure 14b and should be compared with figure 13b. The method was as follows:

- edge detection to find the contour
- smooth with a discontinuity-preserving weak membrane [5]
- track inside the contour (known to be fronto-parallel)
- construct all C_1 and C_2 curves as in (6) (straight lines on projection shown) and their envelopes.

lines on projection shown) and their envelopes.

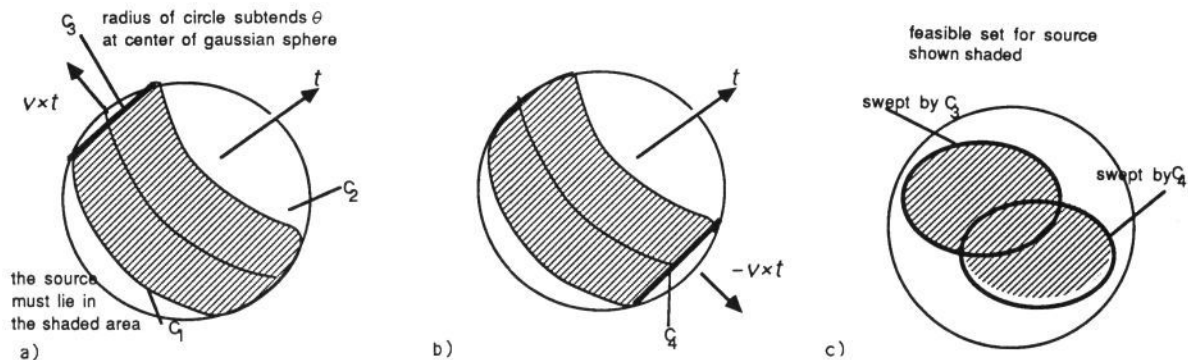


Fig 11. Construction on the gaussian sphere for the orientation constraint

The band of allowed source positions, for a given contour point (fig 8c), is further restricted to give either (a) or (b), depending on which side of the plane of curve tangent and surface normal the source vector falls. The new pieces of boundary are C_3 (a) or C_4 (b). Moving along the contour, they sweep out curves which are copies of one another, rotated through 180 degrees (c) and the union of the areas they enclose is the feasible set under the orientation constraint.

Acknowledgements

The author is very grateful for comments from Dr A. Gilbert, Dr A. Zisserman, Professor H. Longuet-Higgins, Dr G. Scott and G. Brelstaff. Thanks also to G. Brelstaff for his invaluable programming expertise, and to IBM UK Scientific centre for use of the WINSOM solid modeller. The generous support is acknowledged of the SERC (grant GR/D 1439.6), the Royal Society of London for its IBM research fellowship, and the University of Edinburgh for provision of facilities.

References

1. Ambler, A.P., Barrow, H.G., Brown, C.M., Burstall, R.M. and Poplestone, R.J. (1975). A versatile system for computer controlled assembly. *Artificial Intelligence*, 6, 129 - 156.
2. Asada, H. and Brady, M., *The Curvature Primal Sketch*, MIT AI Memo 758, February 1984.
3. Baker, H.H. (1981). Depth from edge and intensity based stereo. *Proc. IJCAI, 1981*, 583-588.
4. Blake, A. (1985). Fresh views on old news. *Journal of Mathematical Psychology*, 29, 3, 360-365.
5. Blake, A. and Zisserman, A. (1986). Weak continuity constraints in computer vision. Report CSR-197-86, Dept. Computer Science, University of Edinburgh, Edinburgh, Scotland.
6. Blake, A., Zisserman, A. and Knowles, G. (1985). Surface descriptions from stereo and shading. *Image and Vision Computing*, 3, 4, 183-191.
7. Blake, A. (1986). On the geometrical information obtainable from the simultaneous observation of stereo contour and shading. Report CSR-205-86, Dept. Computer Science, University of Edinburgh, Edinburgh, Scotland.
8. Blake, A., Zisserman, A. and Papoulias, A.V. (1986). Weak continuity constraints generate uniform scale-space descriptions of plane curves. *Proc. ECAI, Brighton, England*.
9. Blake, A. and Zisserman, A. (1986). Invariant surface reconstruction using weak continuity constraints. *Proc. CVPR Miami*, 62-67.
10. Bolles, R.C., Horaud, P. and Hannah, M. (1983). 3DPO: a three-dimensional part orientation system. *Proc. 8th IJCAI*, 1116-1120.
11. Brady, J.M., Ponce, J.M., Yuille, A. and Asada, H. (1985). Describing surfaces. *Robotics Research*, Hanafusa, H. and Hirochika, I. eds., MIT Press, Cambridge, USA, 5-16.
12. Brelstaff, G.J. and Blake, A. (1986). Detecting specularities: an approach based on lightness computation. Working paper, Dept. Computer Science, University of Edinburgh, Edinburgh, Scotland.
13. Brooks, M.J. and Horn, B.K.P. (1985). Shape and Source from Shading. *Proc. IJCAI 1985*, 932-936.
14. Brooks, R.A. (1981). Symbolic reasoning among 3-D models and 2-D images. *AI Journal*, 17, 285-348.
15. Brown, C.M., Ballard, D.H. and Kimball, O.A. (1982). Constraint interaction in shape-from-shading algorithms. *Proc. DARPA Image understanding workshop*, Palo Alto, CA.
16. Bruss, A.R. (1983). Is what you see what you get? *Proc. IJCAI conf. 1983*, 1053-1056.

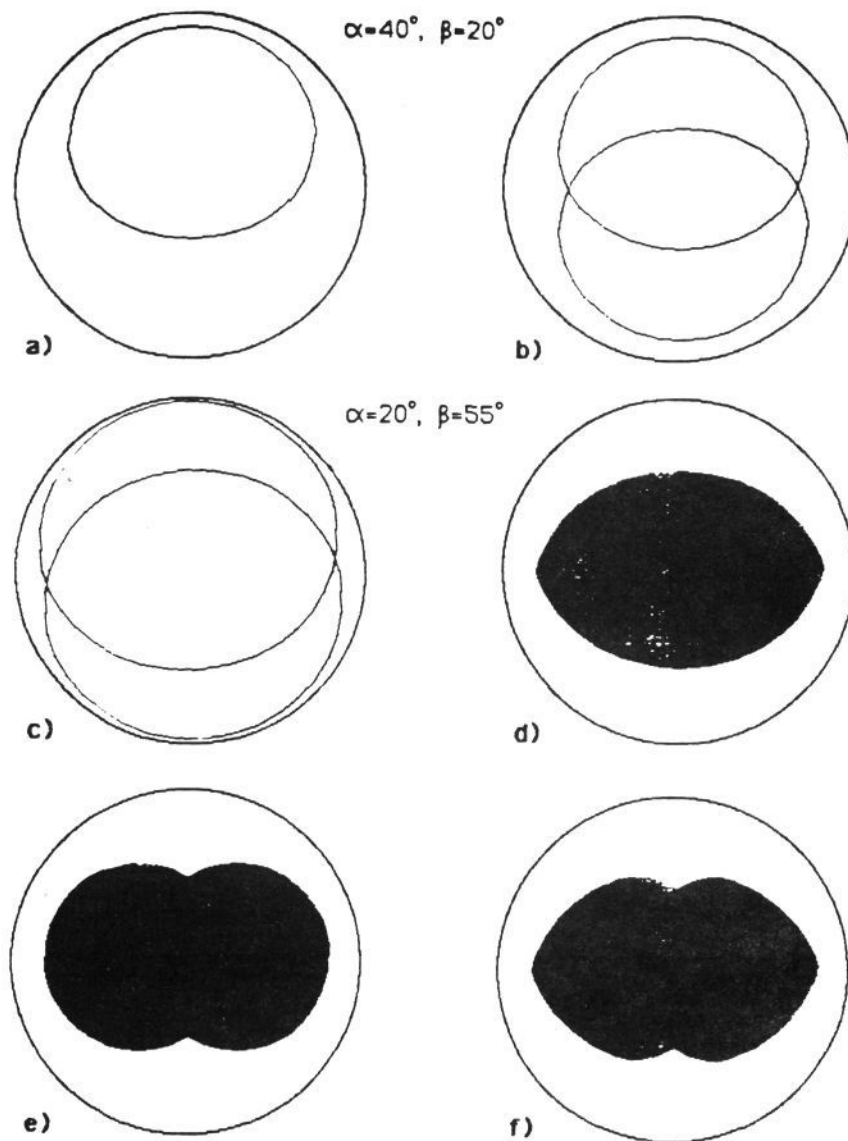


Figure 13. Results of feasible set computation with surface in figure 12. One irradiance constraint envelope (a). Both envelopes intersect (b) at double-degenerate points, one of which is the true source position. In the next example, neither of the double-degenerate points is the true source position (c). So feasible sets for the irradiance constraint (d) and the orientation constraint (e) are intersected to give a final feasible set (f).

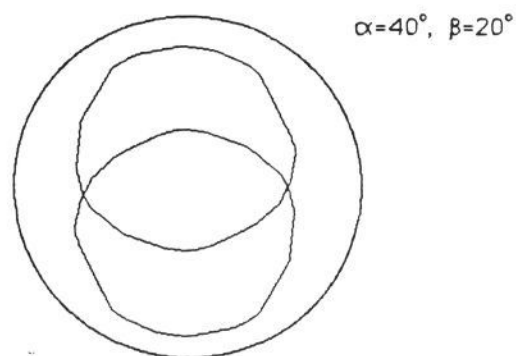


Figure 14. Results of feasible set computation with computer generated images.

Stereo images of the object in figure 12 are generated by a solid body modeller, and noise is added. Irradiance constraint envelopes are computed from the images. Compare them with figure 13b - it is clear that the computation is quite robust to noise.

17. Cernuschi-Frias, B., Bolle, R.M. and Cooper, D.B. (1983). A new conceptually attractive and computationally effective approach to shape from shading. *Proc. IJCAI 1983*, 966-968.
18. Garabedian, P. (1964). *Partial differential equations*. Wiley, New York.
19. Gregory, R.L.
20. Grimson, W.E.L. (1981). *From images to surfaces*. MIT Press, Cambridge, USA.
21. Grimson, W.E.L. (1983). An implementation of a computational theory of visual surface interpolation. *Computer Vision Graphics and Image Processing*, *22*, 39-69.
22. Grimson, W.E.L. (1984). Binocular shading and visual surface reconstruction. *Computer Vision Graphics and Image Processing*, *23*, 19-43.
23. Hochstadt, H. (1964). *Differential equations*. Holt, Rinehart and Wilson, New York.
24. Horn, B.K.P. (1975). Obtaining shape from shading information. In Winston, P.H. (ed.) *Psychology of computer vision*, McGraw-Hill, New York, 115-155.
25. Horn, B.K.P. and Sjoberg, R.W. (1979). Calculating the reflectance map. *Applied Optics*, *18*, 11, 1770-1779.
26. Ikeuchi, K. and Horn, B.K.P. (1981), Numerical shape from shading and occluding boundaries, *Computer Vision*, ed. Brady, J.M., 141-184.
27. Ikeuchi, K. (1983). Constructing a depth map from images. *AI Lab. Memo. 744*. MIT, Cambridge, USA.
28. Koenderinck, J.J. and van Doorn, A.J. (1980). Photometric invariants related to solid shape. *Optica Acta*, *27*, 7, 981-996.
29. Marr, D. and Poggio, T. (1979). A computational theory of human stereo vision. *Proc. R. Soc Lond. B*, *204*, 301-328.
30. Mayhew, J.E.W and Frisby, J.P. (1981). Towards a computational and psychophysical theory of stereopsis. *AI Journal*, *17*, 349-385.
31. Pentland, A.P. (1982). Finding the illuminant direction. *J. Opt. Soc. Am.*, *72*, 4.
32. Purdue University (1970). *Thermophysical properties of matter*, vols 7-9, Plenum, New York.
33. Smith and Kanade (1985). Autonomous scene description with range imagery. *CVGIP*, *31*, 322-334.
34. Woodham, R.J. (1977). A cooperative algorithm for determining surface orientation from a single view. *Proc IJCAI 1977*, 635-641.

