MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY
and
CENTER FOR BIOLOGICAL INFORMATION PROCESSING
WHITAKER COLLEGE

A.I. Memo No. 1264
C.B.I.P Memo No. 60

October 1990

The shape of shading

Daphna Weinshall

Abstract

This paper discusses the relationship between the shape of the shading, the surface whose depth at each point equals the brightness in the image, and the shape of the original surface. I suggest the shading as an initial local approximation to shape, and discuss the scope of this approximation and what it may be good for. In particular, qualitative surface features, such as the sign of the Gaussian curvature, can be computed in some cases directly from the shading. Finally, a method to compute the direction of the illuminant (assuming a single point light source) from shading on occluding contours is shown.

This report describes research done at the Massachusetts Institute of Technology within the Artificial Intelligence Laboratory and the Center for Biological Information Processing in the Department of Brain and Cognitive Sciences and Whitaker College. The Center's research is sponsored by a grant from the Office of Naval Research (ONR), Cognitive and Neural Sciences Division and by the National Science Foundation grant IRI-8719394. The Artificial Intelligence Laboratory's research is sponsored by the Advanced Research Projects Agency of the Department of Defense under Army contract DACA76-85-C-0010 and in part by ONR contract N00014-85-K-0124. Dr. Weinshall is supported by an M.I.T. Postdoctoral Fellowship from MIT.

1 Introduction

The main goal of this work is to explore approximate shape from shading representations that are easy to compute. Such representations are useful when shape from shading is viewed as complementary to other processes, e.g. shape from occluding contours, and shading is used to interpolate the surface between the contours. For some purposes, such as satellite image analysis, it may be necessary to obtain the exact depth map of a surface from shading. But for a real-time intelligent agent it may be sufficient to get from each image only crude information quickly, using an algorithm that is sufficiently correct often enough. This crude information may be sufficient for the computation of many relevant surface features. Errors in surface classification can be corrected using other cues (such as occluding contours), images taken from different viewpoints, active exploration, etc.

The problem of inferring shape from shading seems to be one of the most difficult in low-level vision. Even with the simplest (Lambertian) shading model, assuming constant albedo and a single point light source, significant ambiguity remains. For example, there exist concave, convex and saddle-like surfaces that appear the same from certain viewpoints[1].

All shape from shading algorithms impose some constraints on the reflectance function, such as Lambertian reflectance. Given some a priori knowledge about the objects in the image (e.g., the depth along occluding contours or boundaries), the exact shape from shading problem becomes solvable, though still computationally difficult ([Woo80], [IH81], [Hor86]). These exact shape from shading algorithms are computationally expensive. In addition, since they solve sets of nonlinear differential equations by propagating boundary or initial conditions, errors introduced by the simplifying assumptions, in addition to noise in the input, are accumulated in the integration.

These difficulties stimulated the exploration of possible local shape from shading techniques. There have been a few attempts to obtain shape from local shading analysis. This analysis is restricted to cases where the depth function is spherical ([Pen86], [LR89]) or for surfaces where the angle between the light source and the normal is sufficiently large [Pen88]. Others have studied the behaviors of isophotes (lines of equal brightness) and their relationship to geometrical invariants of surfaces ([KvD80], [Yui89], [Bra83]).

In this paper I follow this line of research by exploring approximations to shape from local shading analysis, and the computation of additional ge-

[1] e.g., under orthographic projection and when the light source is behind the viewer, the surfaces $z = x^2 + y^2$, $z = x^2 - y^2$, and $z = -x^2 - y^2$ all appear the same.

ometrical invariants of surfaces from isophotes. I assume similar restrictive assumptions on the reflectance function as had been used before, trying to minimize the computation and avoid the propagation of errors by the use of local analysis. This paper is organized as follows:

Section 2 explores the shading itself as one candidate for local shape estimation. The question becomes: when is the shape of the shading a good approximation for the surface? If a linear transformation between the shading and the surface shape is allowed, there exist families of surfaces for which equality holds. One such family is the locally-spherical surfaces, the only surfaces for which local shape from shading can be computed precisely [Pen86]. My results show that in the case of such surfaces, and others, the shading is the shape, very little computation is required. For additional families of surfaces, the individual isophotes, or lines of equal brightness, are identical to the contour lines on the surface. The shading approximation to shape can be regarded as the approximation of surfaces by members of these families.

The shading approximation to shape may be useful for various limited purposes. Some examples are the following:

- If the task requires a a simple transformation, e.g. the prediction of the image of the surface illuminated from a different direction, the approximation may be sufficient for many surfaces and small changes in light source direction. In that case, the transformation is very simple: the rotation of the shading surface by the amount of rotation of the light source.

- If a full-blown iterative shape from shading algorithm is to be performed, the shading approximation may give a better initial guess of the surface.

- The shading approximation gives relative depth of points on the surface in an unknown coordinate system, which may be sufficient for the computation of bumps and other surface features. This is further discussed in the next paragraph.

In section 3 I discuss the information that can be obtained on the surface directly from the shading, making use of the shading approximation discussed in section 2 and other cues. In particular, the computation of the Gaussian curvature of the surface is discussed. The computation of such qualitative surface features is the ultimate goal of the depth reconstruction for object recognition and representation. Their estimation from the shading is easier, and in some cases (as I will discuss) may be more accurate than their estimation from the output of a full-blown shape from shading algorithm.

Section 4 deals with the computation of the direction of the light source (see also [BH89], [Pen86], [SK83]). If the surface is illuminated by a single point light source, the shading approximation is obtained in the coordinate system of the light source. If the exact depth map of the surface is needed rather than surface features, it may be necessary to compute the light source direction. This is done directly from the shading on occluding contours in smoothly receding objects in the image.

2 The shape of shading

2.1 The contour and isophote maps:

A contour map, a map of the lines of equal depth on the surface, is an alternative surface representation to a depth map. The contour and depth maps are equivalent when each contour is assigned a depth value. However, the contour map can loose the precise depth assignment and still hold some meaningful information on the surface. Thus this representation degrades more gracefully than a depth map. If the depth of the contours is given up to an unknown scaling factor, the contour map is equivalent to a depth map given with an unknown scaling factor. The contour map is more ambiguous when only a monotonic function of the depth on the contour lines is given. In the most ambiguous representation, the depth of the contour lines is not given at all.

In this section I study the use of the shading itself to approximate shape. I define good approximation to be when the contour map of the surface and the contour map of the 3D shading surface (to be defined shortly) can be derived from each other with a monotonic function. More specifically, the question addressed here is when the isophotes (lines of equal brightness on the surface) are also the contour lines (lines of equal depth on the surface), and what is the function that relates the intensity on the isophotes to the depth on the contours. The cases of interest are when this function is linear, when this function is monotonic, and when no such function exists. In the last case the isophote map is the least useful, for some surfaces it is misleading. I will show how local stability analysis can detect these "bad" cases. I will characterize the family of surfaces that belong to the first two categories, and discuss how the isophote approximation deteriorates in the last case.

2.2 The 3D shading function:

Assume a Lambertian surface $z(x, y)$ and fixed albedo. (Ambient illumination can also exist.) The reflectance function of a Lambertian surface depends

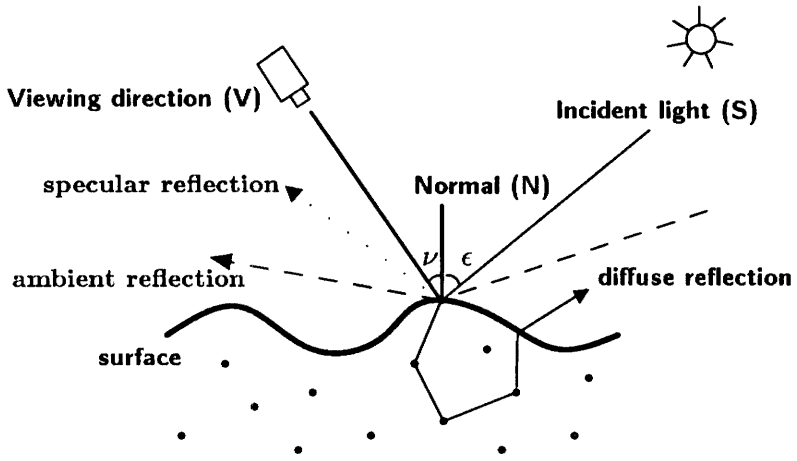


Figure 1: The reflectance function of a surface at a point is composed of three basic components: diffusive, specular, and ambient.

only on the angle between the light source direction and the surface normal, ϵ in figure 1. Assume that the brightness of this surface at a point with normal $\vec{N}$ is:

$$I = \alpha \vec{R} \cdot \vec{N} + \beta \ , \tag{1}$$

for some constants α, β and a fixed direction $\vec{R}$.

Examples of light source distributions for which (1) is accurate are the following: a single distant point light source, in which case $\vec{R}$ points at the light source; or a hemispherical sky of uniform intensity, in which case $\vec{R}$ is the vertical direction relative to the earth. In all the examples in this section, a uniform sky will be assumed.

Since the surface is assumed to be Lambertian, the position of the viewer does not affect the shading at each point on the surface. Therefore it is possible to define shading on the surface in 3D, and only later compute the projection onto the image plane. In the following discussion I will consider this 3D shading function, which assigns a shading value to every point on the surface in 3D. Henceforth in this section, isophote map will refer to this 3D shading map. The discussion of the projection to 2D is postponed to the next section.

In order to analyze the isophote map, which does not depend on the viewer, it is more convenient to choose the coordinate system where the depth Z is parallel to $\vec{R}$, and the $X - Y$ plane is perpendicular to $\vec{R}$. This coordinate

system will be denoted $\Re$. This selection is different from the usual viewer-centered coordinate system.

With this selection of coordinate system, $\vec{R}$ in (1) is $(0,0,1)$. The shading at point $(x,y,z(x,y))$, with normal $(z_x, z_y, -1)$, is:

$$I = \frac{1}{\sqrt{1+z_x^2+z_y^2}} .$$

Take a region where the intensity changes monotonically. The isophotes are contour lines, and their values change in the same (or inverse) direction, if there exists a monotonic function $\Phi(z)$ such that:

$$\Phi(z) = \frac{1}{\sqrt{1+z_x^2+z_y^2}} .$$

$\Phi(z)$ exists whenever $z_x^2 + z_y^2$ is some monotonic function of z. Thus the problem can be rephrased as follows: the shading depends monotonically on the depth when there exist a monotonic function $h(z)$ such that

$$z_x^2 + z_y^2 = h(z) . \tag{2}$$

There are (at least) two solutions to this differential equation: Radially symmetric surfaces and unidirectional surfaces. The isophotes of the first are concentric circles (figure 2), the isophotes of the second are straight lines (figure 5). For a given surface patch, if there exists a function from these families that describes the studied surface to a sufficient accuracy, the isophote approximation gives a good estimate.

I: Radially symmetric surfaces

Let $z(x,y)$ be radially symmetric in the coordinate system $\Re$, namely:

$$z(x,y) = z(r) \quad \text{for} \quad r = \sqrt{x^2+y^2} .$$

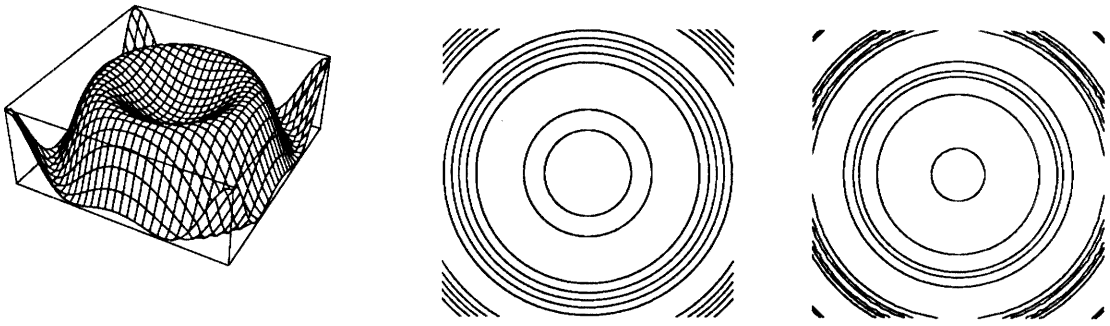
Note that this surface, which is radially symmetric in $\Re$, is not necessarily radially symmetric in other coordinate systems.

To see why shading depends only on depth for these surfaces, note first (as can be readily verified) that:

$$z_x^2 + z_y^2 = z_r^2 .$$

Since $I(x,y) = I(r)$ is monotonic, z_r must be of the same sign and nonzero in the region. Thus there exists an inverse function $r(z)$. z_r is therefore a function of z only.

Figure 2 shows the contour map and isophote map of a volcano-like radially symmetric surface facing upwards. Note that the shape of the isophotes is identical to the shape of the contour lines, both concentric circles. This is the isophote shape for all radially symmetric surfaces. The spacing between the isophotes and the contour lines, which reflects the scaling of the maps, is different between the two maps.



a) b) c)

Figure 2: a) An oblique view of a certain surface. b) The contour map of the surface in (a) illuminated by a uniform hemispherical sky, c) the isophote map.

Toruses and spheres: brightness is a linear function of depth

I will now characterize the sub-family of radially symmetric surfaces whose brightness gives the depth up to a constant scaling factor. These are surfaces $z(r)$ for which the following equation has a solution:

$$I = az + b \tag{3}$$

for two constants a and b. The solution is (as can be readily verified):

$$z(r) = \sqrt{\lambda^2 - (r-\mu)^2} + \nu$$

for constants λ, μ, ν (a torus when $\mu > 0$, a sphere when $\mu = 0$). Note that a sphere is radially symmetric in any coordinate system, not only $\Re$.

Thus in a region that can be approximated by a patch on a sphere or a patch on a torus facing $\vec{R}$, the brightness itself gives almost all the information we can hope to get about z from a single image. Figure 3 shows the contour and isophote maps for a torus.

a) b)

Figure 3: a) The contour map of a torus illuminated by a uniform hemispherical sky, b) the isophote map of the torus.

II: Unidirectional surfaces

A unidirectional surface is a surface that changes as a function of only one direction in the image, $z(x, y) = z(ax + by + c)$. The shading depends only on depth for these surfaces, with the following particular solution of (2):

$$x + cy = \int \sqrt{\frac{1 + c^2}{h(z)}} dz - d \tag{4}$$

for two constants c, d. In this solution, which can be readily verified, z varies with a single direction in the image plane ($x + cy + d$). Figure 4 gives an example of a unidirectional surface.

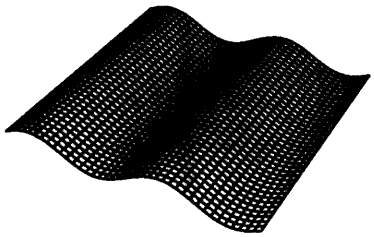


Figure 4: An oblique view of a unidirectional surface.

Cylinders: brightness is a linear function of depth

The sub-family of unidirectional surfaces whose brightness gives a linear transformation of the depth is:

$$z(x,y) = \sqrt{\lambda^2 - \alpha((x + cy + d) - \mu)^2} + \nu$$

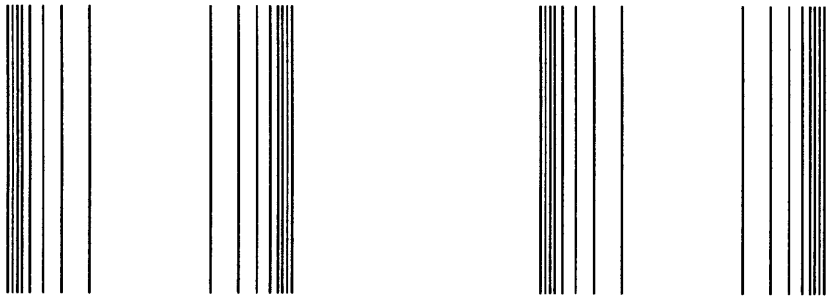
for constants $\lambda, \alpha, \mu, \nu$. This surface is a cylinder, as illustrated in figure 5.

III: General unidirectional surfaces

General Unidirectional surfaces are surfaces that can be described as $\tilde{z}(\tilde{x}, \tilde{y}) = \tilde{z}(\tilde{x} + a\tilde{y} + b)$ in some Cartesian coordinate system $\tilde{X}, \tilde{Y}, \tilde{Z}$ different from $\Re$. The isophote map of such surfaces is still composed of the same straight lines on the surface as characterized above. Thus the isophote map is a monotonic function of the depth $\tilde{Z}$.

IV: Other surfaces

For many surfaces $z_x^2 + z_y^2$ is not a function of z only. In many cases, the approximation of the surface by isophotes degrades gracefully as the following



a) b)

Figure 5: a) The contour map of a cylinder illuminated by a uniform hemispherical sky, b) the isophote map of the cylinder.

examples show. This issue is elaborated on in section 3, where the information in contour lines is discussed.

Ellipses:

Take an ellipse $z = c\sqrt{R^2 - (\frac{x}{a})^2 - (\frac{y}{b})^2}$ such that $a \neq b$ and the ellipse is not radially symmetric. The brightness is:

$$\begin{aligned} I &= \frac{z}{c\sqrt{z^2 + \frac{x^2c^4}{a^4} + \frac{y^2c^4}{b^4}}} \\ &= \frac{z}{c\sqrt{z^2 + \frac{c^2}{a^2}(c^2R^2 - z^2) + \frac{c^4y^2}{b^2}(\frac{1}{b^2} - \frac{1}{a^2})}} . \end{aligned}$$

The only dependence on x, y is in the last term $\frac{c^4y^2}{b^2}(\frac{1}{b^2} - \frac{1}{a^2})$. Thus the isophote approximation deteriorates continuously as a gets further away from b.

Figure 6 shows an example of an ellipse. As can be seen, the isophotes approximation only amplifies the effect of $\frac{a}{b}$, the measure of anisotropy between the x and y directions. Ellipses behave "well" since they are an intermediate case between a sphere and a cylinder (in a certain parametric representation).

a) b)

Figure 6: a) The contour map of an ellipse, with $a = 1.5$, $b = 2$, $c = 1.25$, $R = 1$, illuminated by a uniform hemispherical sky, b) the isophote map of the ellipse.

General radially symmetric surfaces:

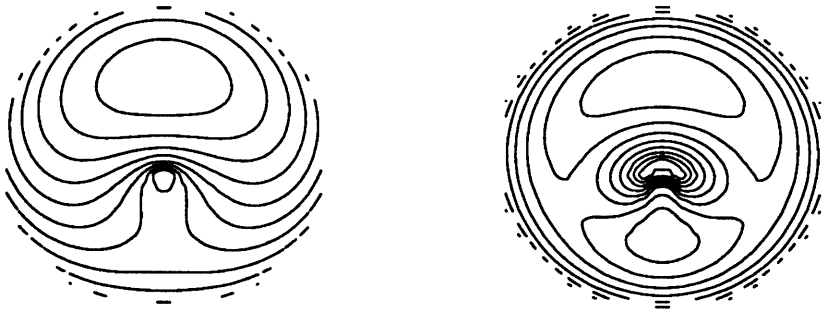
Figure 7 shows a rotated torus that does not face upwards (therefore it is not radially symmetric in $\Re$), its isophote and contour maps. In many regions the isophote map captures important aspects of the contour map (this is discussed in more detail in section 3), with distortions at increasing levels of severeness.

"bad" surfaces:

Finally, for some surfaces the isophote approximation is quite wrong, as shown in figure 8. This is an example of a hyperbolic saddle-like surface. In the next section I will discuss how such "bad" cases can be detected. It is not the case, though, that the isophote approximation fails for all hyperbolic regions, as the torus example shows.

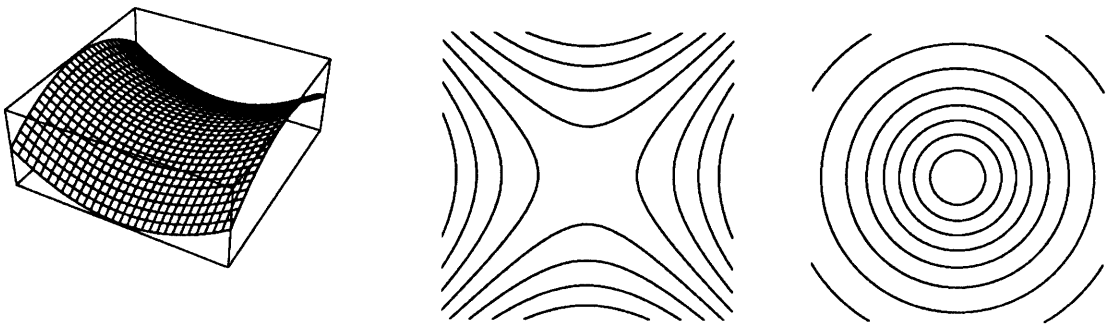
2.3 Projection to the image plane and stability analysis of the shading approximation

When the viewing direction is identical to $\vec{R}$ (the vertical direction or the light source direction), the projection of the isophote map onto the image plane is identical to the isophote maps shown in the figures in the previous section.



a) b)

Figure 7: a) The contour map of a rotated torus illuminated by a uniform hemispherical sky, b) the isophote map.



a) b) c)

Figure 8: a) An oblique view of a hyperbola. b) The contour map of the hyperbola illuminated by a uniform hemispherical sky, c) the isophote map.

Typically, however, the isophotes are foreshortened. In the following I assume orthographic projection.

Consider first the case where isophotes are contour lines. Here the isohpotes on the surface are planar curves lying on parallel planes $Z = \text{const}$ in some unknown Cartesian coordinate system. These parallel planar curves, are projected obliquely to the image plane, and all of them are uniformly foreshortened. This foreshortening is responsible for contraction along one axis of the planar curve, as is illustrated in appendix A. Foreshortening preserves many curve features: straight lines are projected to straight lines, elliptical close contours to elliptical close contours, and inflection points to inflection points. Moreover, this projection is stable in the sense that as the camera (or the light source) moves, the curves uniformly get more or less contracted along some axis.

The case is different for isophotes that are not contour lines. Here the isophotes are not planar curves, and therefore their projection to the image plane is not foreshortening. The further away the isophotes are from parallel planar curves, the less like foreshortening their projection to the image plane looks. This gives a heuristic to detect the "bad" cases where the isophotes are not a good approximation to shape. By moving the camera, if the isophotes do not change by a uniform contraction and expansion along a single axis, if features such as inflection points appear and disappear, than the shading cannot be used to approximate shape. As an example, figure 9 shows the projection of a particular isophote on the hyperbola of figure 8. For comparison, the projection of the corresponding isophote on an elliptic surface, where the isophotes undergo expansion/contraction along one axis only, is also given.

2.4 An approximate local shape from shading algorithm:

The isophote map approximates the contour map in the coordinate system $\Re$, which is often different from the viewer-centered coordinate system. Thus typically the map of contour lines is foreshortened when projected into the image, as discussed above. This foreshortened map can be used to infer properties of the surface directly, as can be imagined from figure 10 (and will be discussed in the next section).

The contour map is approximated in regions where the brightness changes monotonically. Based on the above analysis, an algorithm to construct a shape representation is the following:

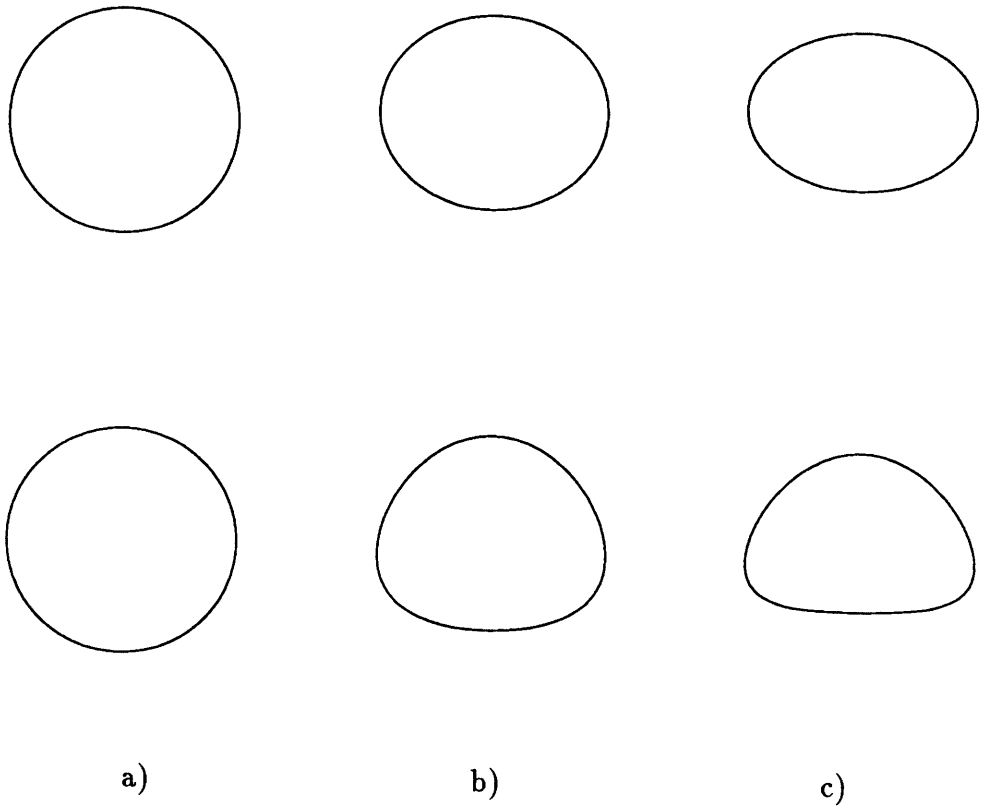


Figure 9: The projection onto the image of an isophote on the surface $z = -(x^2 + y^2)$ (elliptic, above) and $z = x^2 - y^2$ (hyperbolic, below, see also figure 8), illuminated by a uniform hemispherical sky; a) viewing angle is vertical, similar to $\vec{R}$, b) viewing angle with vertical is 30^o, c) viewing angle with vertical is 45^o.

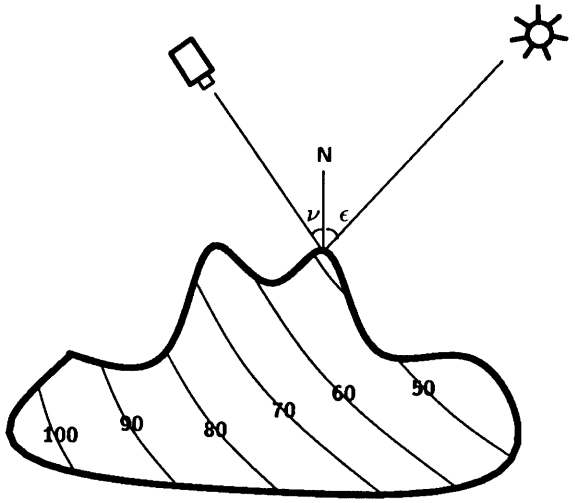


Figure 10: A Lambertian surface illuminated by a single point light source. In this case, the coordinate system in which the shading is the depth is the coordinate system of the light source.

1. Find curves or points of brightness extrema;
2. compute a contour map in each region between these curves using the isophote map;
3. assign direction of depth change (whether the depth increases with increasing brightness or decreases);
4. Compute the relevant surface features.

For complex surfaces, step 3 becomes nontrivial algorithmically.

In a minimalistic implementation of this algorithm, the depth on curves or points of brightness extrema that have been found in step 1 is either computed using other cues, such as motion or occluding contours, or set at an arbitrary constant, and the shading is then used to interpolate the depth between these features.

3 Geometrical properties of surfaces

In the previous section it has been shown that the contour map in some unknown (for the moment) coordinate system can be computed (or approxi-

mated) directly from the shading for many surfaces. In this section I discuss some geometrical information on the surface that can be obtained from this map without depth recovery and transformation of coordinate system. I will not discuss here relative depth information, though the existence of ordering of image points (in any coordinate system) can be used to locate bumps and other important surface features. The ability of humans to judge shape information without knowledge of the light source direction is demonstrated and discussed by Mingolla and Todd in [MT86].

Surfaces of objects can sometimes be concisely described as a collection of simpler parts, each of which described by a few parameters (e.g. generalized cylinders [Bin87]; see also [KvD79] and [Ett88]). Classifying regions according to the sign of their Gaussian curvature, namely as elliptic (convex/concave), planar, cylindrical, or hyperbolic, provides one important intrinsic surface feature (see also [BJ86], [VMA86], [Wei88], [Nal88], [BZK89] and [Wei89]). With this classification of parts as areas of the same sign of Gaussian curvature, part boundaries within an object are located on parabolic lines. The parts produced by this segmentation are often qualitatively similar to the parts produced by the generalized cylinders based scheme. I will now discuss cues in the contour map to the Gaussian curvature of the surface.

3.1 Geometrical properties of surfaces near global shading maxima:

Consider a Lambertian surface with reflectance (1). If the distribution of surface normals span a significant portion of the Gaussian sphere, most likely the global maxima of the shading, assuming it is not on the boundary of the surface, will be obtained in points where the surface normal is parallel to R. Choose such a point P. In this case the coordinate system $\Re$, where Z parallels $\vec{R}$ and P defines the origin, is a very natural one to use. It is possible to choose directions X and Y corresponding to the directions of principal curvature on the surface at P. With this selection, for $|k_1| \leq |k_2|$ the principal curvatures in P, the surface is:

$$z(x,y) = \frac{k_1 x^2 + k_2 y^2}{2} + R, \tag{5}$$

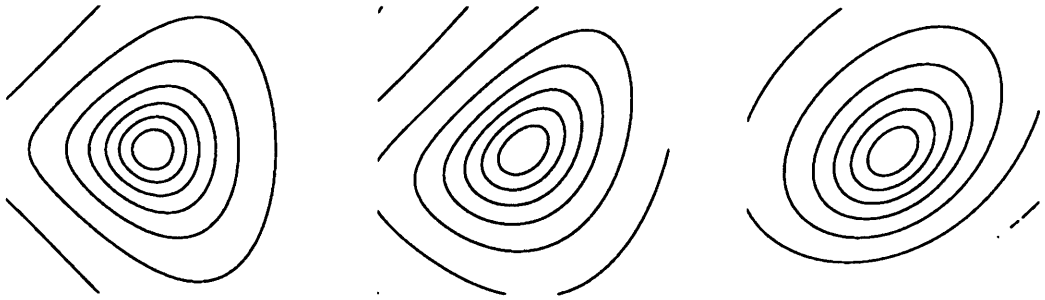
where

$$\lim_{(x,y)\to(0,0)} \frac{R}{x^2+y^2} = 0.$$

If P is parabolic, namely $k_1 = 0$, the isophotes near P are contour lines. Since locally the contour lines are straight lines, they are always projected to locally straight lines on the image.

If $|k_1| = |k_2|$, the isophotes are circles on the $X - Y$ plane (figure 9a). If P is elliptic ($k_1 = k_2$), the isophotes near P are again contour lines and define circles on the $X - Y$ plane. As argued above, they are projected to foreshortened circles in the image, as shown in the upper row of figure 9b,c. If P is hyperbolic ($k_1 = -k_2$), the projection of the isophotes looks very different, as shown in the lower row of figure 9b,c.

As $|k_1|$ gets further away from $|k_2|$, the distinction between the hyperbolic and elliptic cases becomes less sharp. The projected isophotes still differ in the same way, as is shown in figure 11. Figure 11b, depicting a hyperbolic surface where $|k_1| = \frac{2}{3}|k_2|$, maintains some of the triangular shape of figure 11a. Figure 11c, depicting an elliptic surface where $|k_1| = \frac{2}{3}|k_2|$, looks more like foreshortened ellipses.

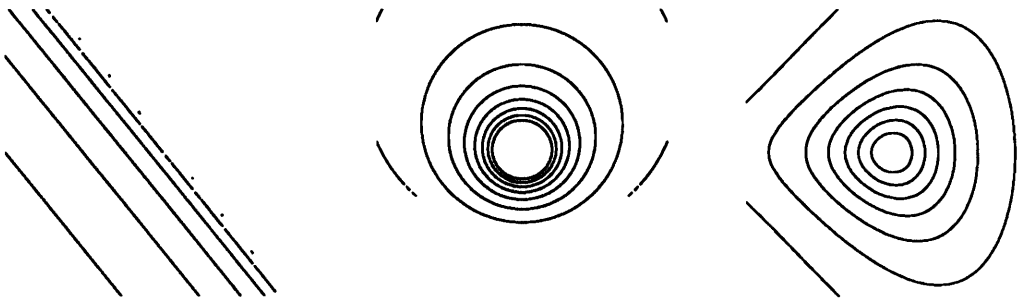


a) b) c)

Figure 11: The isophote map, projected onto the image, near a shading maximum. The viewing coordinate system is obtained by a rotation of 45° around the Z axis, followed by a rotation of 10° around the X. *a)* A generic hyperbolic surface $z = \frac{x^2-y^2}{2}$; *b)* a generic hyperbolic surface $z = \frac{x^2-\frac{2}{3}y^2}{2}$; *c)* a generic elliptic surface $z = -\frac{x^2+\frac{2}{3}y^2}{2}$.

This analysis suggests a method for the characterization of the Gaussian curvature of a surface near global shading maxima (that result from the diffuse, rather than the specular, component of the illumination):

- if the isophote projections are straight lines locally, the surface is parabolic (figure 12a);
- if the isophote projections are foreshortened circles, the surface is elliptic (figure 12b);
- if the isophote projections are concentric triangles, the surface is hyperbolic (figure 12c);



a) b) c)

Figure 12: Classification of surfaces near global shading maxima: *a)* parabolic, *b)* elliptic, *c* hyperbolic.

For this method to work, the viewing direction cannot parallel $\vec{R}$, it should be oblique. Also, $k_1 = 0$ or $k_1 = k_2$ is assumed. The classification becomes harder as k_1 gets further away from k_2. I should note that shading extrema often cling to parabolic points [KvD80]. The above method applies only to global maxima. Finally, this method applies in a region around P where R in (5) is small enough.

3.2 Geometrical properties of surfaces near inflection points:

3.2.1 Contour lines:

Let P be a point on the contour line $z = \text{const}$. Assume that the $X - Y$ plane is not tangent to the surface at P[2]. There exists a direction (which I will call Y without loss of generality) such that $z_y \neq 0$. In the neighborhood of P the contour line is some function $y(x)$, such that:

$$\begin{aligned} y'(x) &= -\frac{z_x}{z_y} \\ y''(x) &= -\frac{1}{z_y^3}(z_y^2 z_{xx} - 2z_x z_y z_{xy} + z_x^2 z_{yy}) \ . \end{aligned}$$

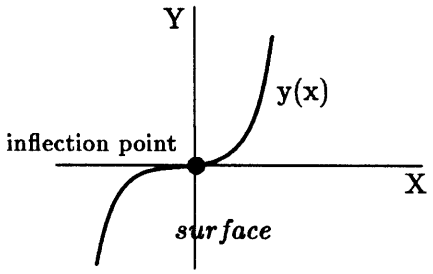


Figure 13: Inflection points in the contour lines can be seen from almost any viewpoint. They will correspond to inflections in the isophotes for surfaces whose isophotes are their contour lines.

In an inflection point $y''(x) = 0$, namely

$$z_y^2 z_{xx} - 2z_x z_y z_{xy} + z_x^2 z_{yy} = 0.$$

This equation has a solution if $z_{xy}^2 - z_{xx} z_{yy} \geq 0$, namely, the Gaussian curvature of the surface is not positive. In other words, the surface around an inflection point on a contour line is not elliptic, it is parabolic or hyperbolic.

Let me choose the $X-$ and $Y-$ directions such that $z_y \neq 0$ and $z_x = 0$ (figure 13). Then

$$y'(x) = 0, \quad y''(x) = -\frac{z_{xx}}{z_y}, \quad y'''(x) = -\frac{z_{xxx}}{z_y} \ . \tag{6}$$

[2]Inside a region where the intensity changes monotonically, it cannot be that $z_x = z_y = 0$. Thus the $X - Y$ plane is not the tangent plane to the surface at an internal point P.

P is an inflection point if and only if $y''(x) = 0$ and $y'''(x) \neq 0$, namely, $z_{xx} = 0$ and $Z_{xxx} \neq 0$. Thus z_{xx} changes sign at P whereas z_{yy} almost always does not. We can therefore conclude that any inflection point is one of the following:

- A hyperbolic point (less common);
- A parabolic point dividing an elliptic region from a hyperbolic region, and X is a principal direction (an example is given in figure 14).

Once again, the second case, where the inflection point is parabolic, can be identified by stability analysis. An inflection point in a hyperbolic region will move in all directions when the camera, or the light source, is moved a little around its location. An inflection point on a parabolic line will move on the parabolic line or disappear (possibly into the hyperbolic region).

3.2.2 Isophotes are contour lines:

Let $(z(x,y)$ be the actual surface and $I(x,y) = f(z)$ the shading surface for some monotonic function f (i.e. $f'(z) > 0$). The Gaussian curvatures K of both surfaces are related as follows:

$$K_I = [f'(z)^2 \; K_z \; + \; f''(z)f'(z)(z_y^2 z_{xx} - 2z_x z_y z_{xy} + z_x^2 z_{yy})]\frac{(1 + I_x^2 + I_y^2)^2}{(1 + z_x^2 + z_y^2)^2} \; . \quad (7)$$

We can now conclude the following:

- From (7) it follows that near an inflection point, the sign of the Gaussian curvature of the shading surface $I(x,y)$ is identical to the sign of the Gaussian curvature of the surface $z(x,y)$. The shading can be used to determine whether the surface is locally parabolic or hyperbolic.
- From (7) it also follows that the sign of the Gaussian curvature of the shading surface $I(x,y)$ is identical to the sign of the Gaussian curvature of the surface $z(x,y)$ whenever $f''(z) = 0$, namely, when the shading is a linear function of the depth (cylinders and spheres).
- An inflection on an isophote is either hyperbolic or a parabolic point dividing an elliptic region from a hyperbolic one.

  In the reverse direction: on a parabolic point on an isophote, the X direction as define in (6) is a principal direction [Yui89], namely, $Z_{yy} = 0$ or $Z_{xx} = 0$. If $Z_{xx} = 0$, P is an inflection point.

Finally, if $\vec{R}$ is not parallel to the image plane, the inflection point on the contour line (which is a planar curve) is projected to an inflection point on the projection of the contour line on the image. It can therefore be detected regardless of whether $\vec{R}$ is known.

3.2.3 An example:

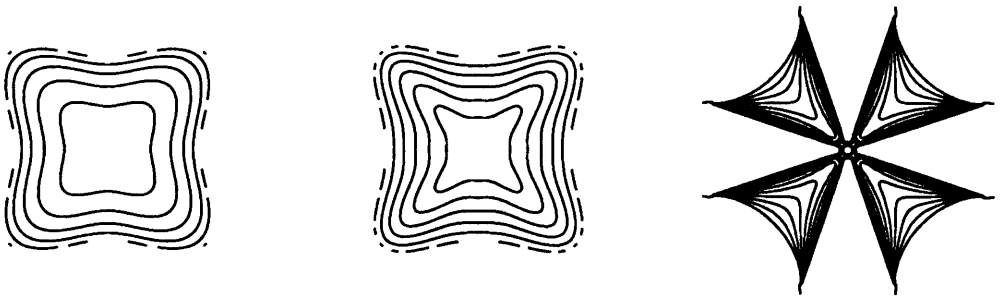
Assume a close contour line. Since $z_y \neq 0$, any sign change of the curvature of the contour line (as determines by $y''(x)$) can only happen with sign changes of z_{xx} (the direction X changes continuously with P). A convex segment curves towards the direction of increase in depth on the surface, a concave segment curves away from it, and both curve towards the inside of the close contour line when the Gaussian curvature does not change sign on the curve. (Thus close elliptic contour lines often resemble ellipses, see figure 6). Since an inflection point is either hyperbolic or it is separating a hyperbolic region from an elliptic one, contour line segments that have negative curvature relative to the inside of the contour line tend to be hyperbolic, and segments with positive curvature tend to be elliptic.

Figure 14 illustrates the implication of these results. Figure 14d in particular shows that the parabolic lines intersect the contour lines at an inflection point, the regions of the contour line with positive curvature relative to its inside are elliptic, and the regions with negative curvature relative to its inside are hyperbolic. Figure 14e shows that these qualitative relationship hold approximately for an isophote, as well as a contour line, though for this surface the isophotes and contour lines are not identical.

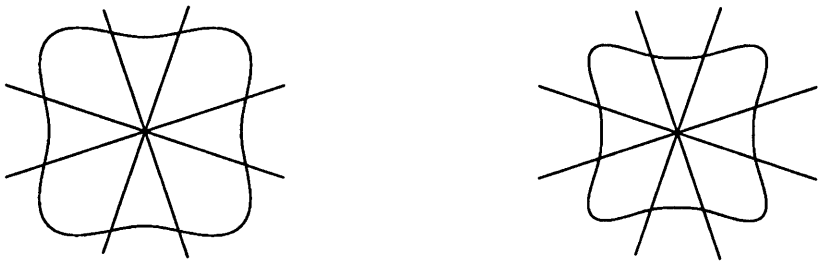
3.3 Geometrical properties of surfaces from isophotes:

For some surfaces the analysis of inflection points is not very useful, since the isophotes do not approximate the contour lines well, or there are no inflection points in regions of monotonic change of intensity. In such cases it is still possible to learn some shape properties from the isophotes. This section extends the work of Koenderink & van Doorn [KvD80] and Yuille [Yui89].

Yuille [Yui89] has shown that at a parabolic point the isophote points along the line of curvature (or a principal direction) at the point. He concluded that "typically the parabolic lines give rise to ridges, or valleys, in the image intensity". I will argue that in many interesting cases the isophotes are perpendicular, or almost perpendicular to the parabolic line. In particular, in these cases, the parabolic line is perpendicular to any ridge or valley in the



a) b) c)



d) e)

Figure 14: For the surface shown in figure 18a: a) the contour map of the surface, b) the isophote map, c) the filled-in regions mark elliptic regions, d) the intersection of the parabolic lines and a contour line, e) the intersection of the parabolic lines and an isophote. The contour line in d) and the isophote in e) were chosen randomly and are unrelated.

image intensity. These cases may prove more typical, and suggest a possible heuristic for the computation of parabolic lines directly from isophotes.

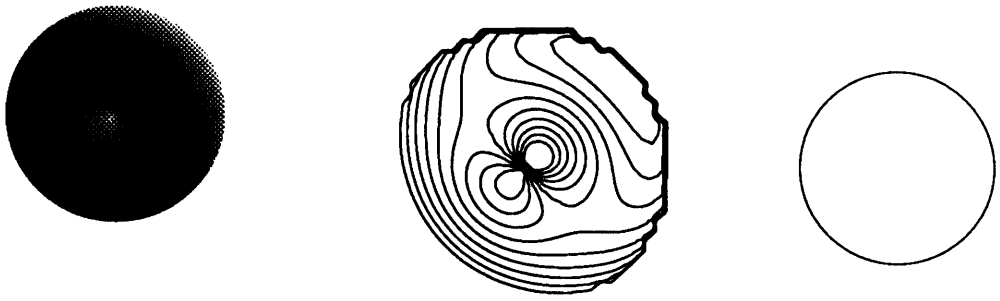


Figure 15: A torus illuminated from direction $(1,1,1)$, its isophote map, and its parabolic line.

Consider the family of surfaces of revolution (generalized cylinders where the main planar axis is a straight line). For this family it is known [DoC76] that the parallels (lines parallel to the generator curve) and meridians (lines defined by a given point on the generator curve as it sweeps around) are lines of curvature, and that parabolic lines are meridians. Thus, for surfaces of revolution, the isophotes are either parallel to the parabolic line or perpendicular to it. The first case leads to parabolic lines being ridges or valleys in the image intensity as discussed by Yuille, but it is the rarer case of the two. An example of a torus is given in figure 15 (here the parabolic line is a ridge when the light source is behind the viewer). More often the tangents to the isophotes are perpendicular to the parabolic line and the isophotes have a local extremum or an inflection point. More specifically, the isophotes bend at the parabolic line (figure 16a and 16b) if the parabolic line separates a hyperbolic region from an elliptic one. If the parabolic line lies between two elliptic regions, the isophote has only an inflection point (figure 16c). Proof is given in appendix B.

This result gives a heuristic for the detection of parabolic lines separating hyperbolic and elliptic regions: when all the isophotes bend and the line

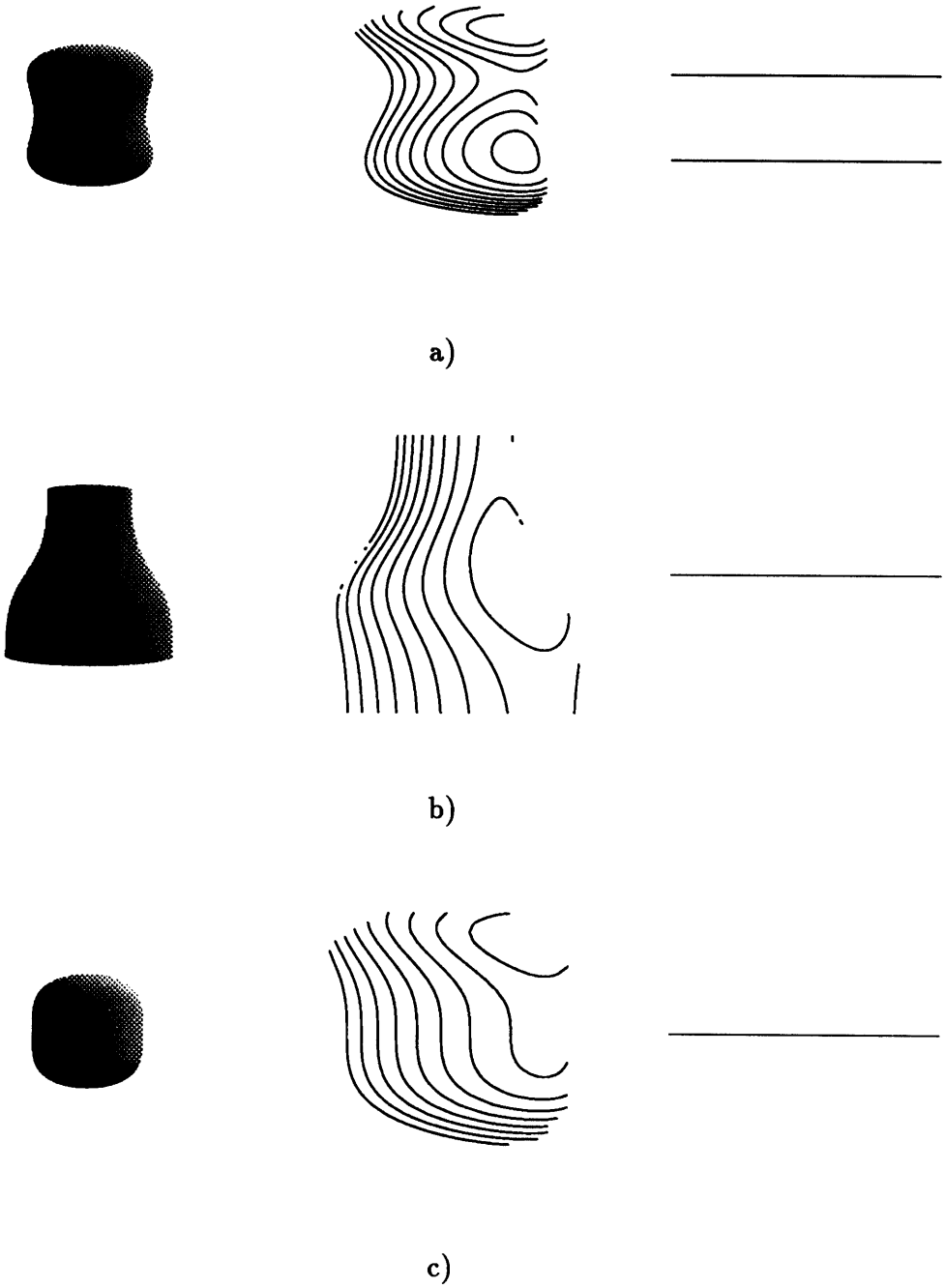


Figure 16: Each row, from left to right, shows a surface of revolution illuminated from direction $(1,1,1)$, its isophote map, and its parabolic lines. For illustration purposes, the object shown on the left of each row is actually illuminated from $(1,1,0.2)$.

through the extrema of the bend is roughly perpendicular to the isophote, this line is a good candidate for a parabolic line. It works well for surfaces of revolution, as figure 16 shows, though distortions occur when the orthographic projection of a straight angle is not straight.

4 Direction of illumination from occluding contours

Assume a single light source at a large distance from the surface. The source's position is defined by two angles (figure 17): tilt – the angle between the projection of the light direction on the image plane (denoted the $X-Y$ plane) and the X-axis, and slant – the angle between the light direction and the Z-axis (the viewing direction). In the following I discuss the computation of these two angles from shading on occluding contours and self-shadow edges. This computation is mostly based on a general shading model with Lambertian, ambient and specular components.

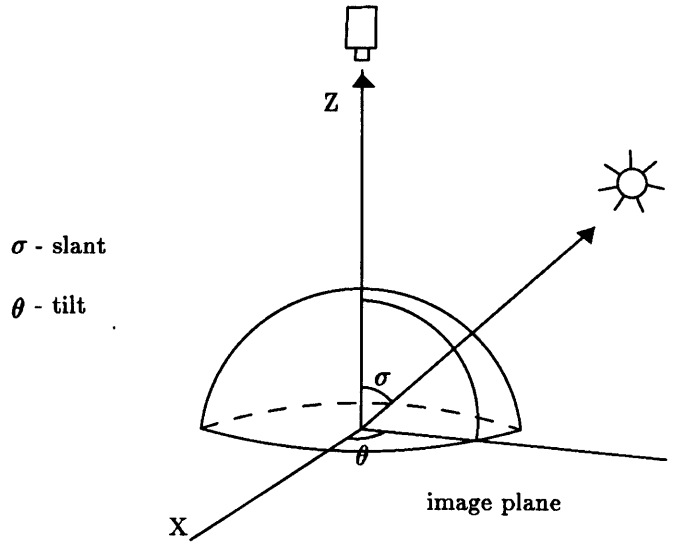


Figure 17: The tile and slant of a point light source.

4.1 Tilt computation

The following computation assumes a single point light source and constant albedo. If more than one light source exists, it should then be used with the self–shadow edges of each light source separately, assuming that the edges are separable.

Tilt from shading on occluding contours:

Assume an occluding contour where the normal to the surface is perpendicular to the viewing direction (or in the $X - Y$ plane).

Proposition 1 *The direction of the tilt of the illuminant is parallel or perpendicular to the occluding contour at points of shading extrema on the contour. This holds for points of extremum where the occluding contour does not have a singularity (such as a cusp).*

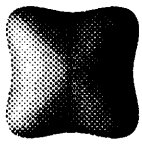
1. *At any point on the occluding contour where the minimum of shading occurs (must be the value of the ambient illumination, at the beginning of a self-shadow line), the angle of the tangent to the occluding contour (and the self-shadow line) is the angle of tilt of the illuminant.*

2. *At any point on the occluding contour where the maximum of shading occurs, the angle of the tangent to the occluding contour is perpendicular (in the image plane) to the angle of tilt of the illuminant.*

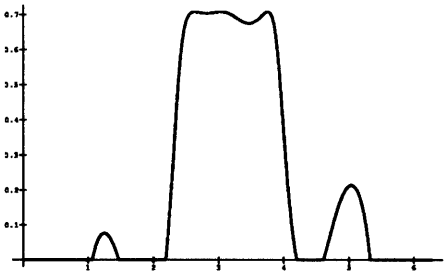
A proof is given in part 1 of appendix C. This proposition identifies a group of points on the occluding contour (at least 3, if all the occluding contour is visible) where the tangent to the occluding contour gives the angle of tilt of the illuminant. The tilt computation is therefore quite robust, since only one of these points should be visible (not occluded). Figure 18 shows an example of the occluding contour for a complex surface and the tangent to the occluding contour at the extrema of shading on the contour.

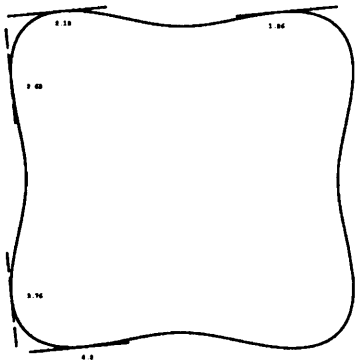
Tilt from self-shadow edges:

If both edges of self-shadow and cast-shadow are given, matching them can give the direction of illumination as well. As pointed out in [SK83], the angle of a line between a feature on a self-shadow line and its match on the cast-shadow line is the angle of tilt.



a)





b) c)

Figure 18: a) A bird's eye view of a surface illuminated from direction $(-1, -0.1, 1)$. b) The shading on the occluding contour plotted as a function of polar angle (in radians) around the center of the object. c) The occluding contour of the surface with its tangent plotted at 3 points of shading minima (solid line) and 2 points of shading maxima (dashed line). Each point is identified by its polar angle (as plotted in b).

4.2 Slant computation

Slant from shading on occluding contours:

The occluding contour in the neighborhood of a point of global maximum of shading gives the direction of tilt of the illuminant. If the albedo of the surface and the intensity of the illuminant are known, the shading at such a point, which does not have to be unique, gives the angle of slant of the illuminant. If the albedo and light intensity are not known, the ratio between the global maximum of shading on the occluding contour to the global maximum of shading on the surface gives the angle of slant according to the following formula:

$$\cos\sigma = \frac{\max\limits_{\partial\Omega} I}{\max\limits_{\Omega} I} \tag{8}$$

for brightness I, slant σ, surface of an object Ω, and the occluding contour $\partial\Omega$.

This formula assumes a Lambertian surface (proof is given in part 2 of appendix C). It may be used by humans to judge the slant of the light source, as discussed by Reichel & Todd [TR89]. In general, $\min\limits_{\partial\Omega} I$ should be subtracted from both numerator and denominator to eliminate the component due to ambient illumination. If the surface is also specular, the denominator becomes $\max\limits_{\Omega\prime} I$ where $\Omega\prime$ is the surface area not including the regions of specular reflection (the use of $\Omega\prime$ may lead to an underestimation).

4.3 An example

Figure 19 shows an example, an image of a gourd, for which the direction of illumination has been computed from the shading on the occluding contour. First, regions of maximal and minimal intensity on the occluding contour have been identified (figure 19c,d). The self-shadow edge in figure 19b does not actually intersect the occluding contour: the intersection was computed by locating the points on the occluding contour whose intensity was the closest to the mean intensity on the self-shadow edge (which was almost constant). The tangent to each of the edge segments in figure 19c,d was computed, giving an estimate of -104^o to the tilt direction of the illuminant. The angle of slant was computed from (8) to be 75^o.

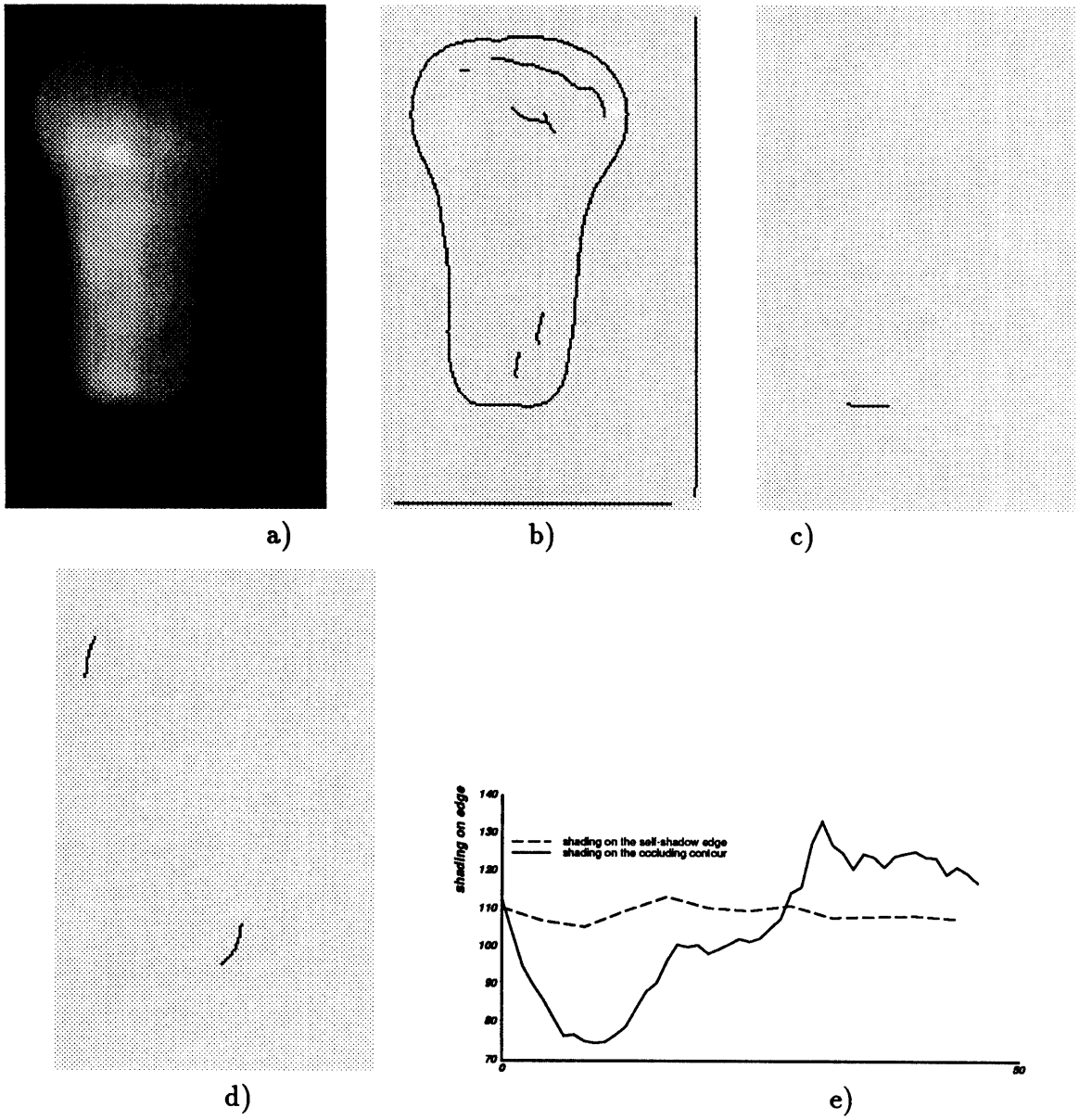


Figure 19: a) A gourd image; b) the edges of the image computed using Canny's algorithm (1986) [Can86]; c) a piece of the occluding contour where the intensity is maximal; d) two pieces of the occluding contour where the intensity is minimal (the beginning of a self-shadow edge); e) the (smoothed) shading profiles of the occluding contour and the self-shadow edge. The scale and absolute values of the units on the abscissa are unrelated for the two edges. The points of interest on the occluding contour are the maximum (giving the edge element in c) and the intersection of the contour line with the self-shadow edge (giving the edge elements in d).

5 Discussion

The approximation to shape by shading, discussed in section 2, is in some sense a generalization of Pentland's analysis of local shape from shading [Pen86]. Pentland showed that only spherical points can be precisely recovered with local shading analysis. The analysis in section 2 shows that local shading can be used to give the shape of a richer family of surfaces.

If the surface is assumed to be locally spherical, as in [Pen86], the depth is uniquely defined by the shading. The direction of the single light source that illuminates the surface can be computed from the foreshortening of the isohpotes. Once the light source direction is determined, the linear transformation relating the depth and the shading is uniquely defined by the second derivative of the intensity perpendicular to the isophotes. (The constant additive term in the linear transformation is not computable with orthographic projection.) This leaves no free parameters that need to be computed. Thus in the spherical case, Pentland's method gives the same results as the shading approximation.

The shading seems to be a better local approximation to shape than precise calculations since it gives the correct surface for a richer family of surfaces, not only spherical. It also gives an integrable consistent solution all over the surface when the approximation is very local, e.g., when each surface patch is approximated by a different spherical function. Finally, and most importantly, the shading approximation is computationally free, only the two parameters of the light source direction should be computed to obtain a vierer-centered depth map.

The shading approximation, which is the least sensitive to errors in the simplifying assumptions on the reflectance function and noise in the brightness data, may be the best initial local estimate to the surface. The estimate could then be improved by global shape from shading methods. This approximation is useful for application domains where some immediate shape approximation is needed instantly and where there do not exist the computation resources for exact shape from shading recovery. In general, this approximation cannot be used in isolation, it should be complemented by information from other images or cues (such as occluding contours, stereo, etc) to detect surfaces for which the isophotes do not approximate the contour lines well. Local stability analysis can also detect such surfaces, as discussed in section 2.3.

6 Summary

Assuming a particular Lambertian reflectance function (1) with a possible ambient component, orthographic projection, constant albedo and no mutual illumination, and using a particular coordinate system $\Re$, I have shown that the brightness at each point is a linear transformation of the depth of the surface at that point for spheres, toruses, and cylinders (the last two should be aligned with the $Z-$axis). For larger families of surfaces, such as radially symmetric surfaces and unidirectional surfaces, the depth is some monotonic function of the intensity in regions where the intensity changes monotonically. Thus using the map of isophotes to approximate the contour map of the surface gives the correct result for surfaces in these families and some approximate result for other surfaces. Using the shape of the shading to approximate the shape of the surface is useful, therefore, in many interesting cases.

The isophotes have been used to approximate the contour map of a surface in the coordinate system $\Re$. Features of this map, such as inflection points on the contours and the sign of the contour's curvature relative to its inside (assuming it is close), can be computed without any knowledge on $\Re$. These features have been shown to give strong clues to the sign of the Gaussian curvature, or the actual Gaussian curvature, of the surface. Heuristics for the computation of parabolic lines from isophotes were also described. Parabolic lines are useful since they enclose hyperbolic regions, and hyperbolic regions are often the most "natural" places to segment a surface into parts.

In the last part of this work I showed a method to compute the light source direction. The tangent to the occluding contour in points of shading extrema on the occluding contour gave the tilt of the light source. The value of maximal shading on the occluding contour gave the slant of the light source.

Appendix A

When the isophotes are contour lines, they are planar 3D curves that are projected onto the image plane. This projection is described in figure 20.

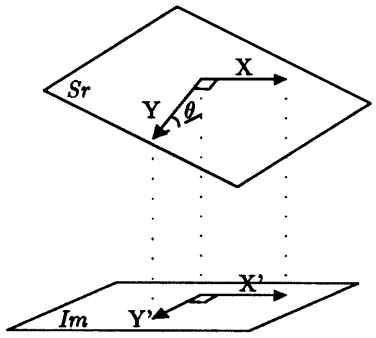


Figure 20: Illustration of foreshortening.

Let Sr be the plane on which the isophote lie, and Im the image plane. Select directions X, Y in plane Sr so that X is parallel to the image plane Im, and Y perpendicular to X. These directions are projected to perpendicular directions X', Y' in Im. Let θ be the angle between axes Y and Y'. (X is parallel to X' by definition.) A curve on Sr of the form $y = f(x)$ is projected to a curve $y' = \cos\theta f(x')$. Thus the projected curve is foreshortened, namely, it undergoes uniform scaling direction Y.

Appendix B

Let me use the following coordinate system to describe a surface in the neighborhood of a point P. Let the tangent plane at P be the $X - Y$ plane, with the $X-$ and $Y-$ directions corresponding to the two principal directions at P. Let Z be the direction of the normal to the surface at P. Let the origin be at P. Let the surface be a function $z(x, y)$.

With this selection, the depth function z and the first two derivatives z_x and z_y at the origin P are 0. z_{xx} and z_{yy} are the two principal curvatures λ_1 and λ_2, and $z_{xy} = 0$.

For $\Xi = \frac{1}{\sqrt{1+z_x^2+z_y^2}}$, an isophote curve on the surface is defined by

$$(-z_x\Xi, -z_y\Xi, \Xi) \cdot (s_1, s_2, s_3) = const$$

for a distant light source at (s_1, s_2, s_3). The tangent to the isophote lies in the $X - Y$ plane by definition, and is defined by:

$$\begin{aligned}(\tilde{x}, \tilde{y}) &= (s_1(z_{xy}\Xi + z_x\frac{\partial\Xi}{\partial y}) + s_2(z_{yy}\Xi + z_y\frac{\partial\Xi}{\partial y}) - s_3\frac{\partial\Xi}{\partial y}, \\ &\quad -(s_1(z_{xx}\Xi + z_x\frac{\partial\Xi}{\partial x}) + s_2(z_{yx}\Xi + z_y\frac{\partial\Xi}{\partial x}) - s_3\frac{\partial\Xi}{\partial x})) \\ &= (s_2\lambda_2, -s_1\lambda_1)\end{aligned}$$

namely,

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix} \cdot \begin{pmatrix} s_2 \\ -s_1 \end{pmatrix} . \tag{9}$$

From (9) it follows that at a parabolic point, where $\lambda_2 = 0$, the isophote points along the direction Y. Let the parabolic line point along the X direction (for a surface of revolution it must point along either X or Y). It follows that in the neighborhood of P the Y component of the isophote changes sign when the principal curvature λ_1 changes its sign at P (i.e. when the parabolic point is a transition between elliptic and hyperbolic regions) and the isophote bends. If λ_1 does not change sign (in the transition between two elliptic regions, the isophote has only an inflection point. Note that when $\lambda_1 = \lambda_2$, the isophotes are perpendicular to the projection of the light source on the $X - Y$ plane (which is not the image plane). On a self-shadow edge $s_3 = 0$, in which case the light source direction itself is perpendicular to the edge.

Appendix C

PART 1:

Proposition 1 *The direction of the tilt of the illuminant is parallel or perpendicular to the occluding contour at points of shading extrema on the contour. This holds for points of extremum where the occluding contour does not have a singularity (such as cusp).*

1. *At any point on the occluding contour where the minimum of shading occurs (must be the value of the ambient illumination, at the beginning of a self-shadow line), the angle of the tangent to the occluding contour (and the self-shadow line) is the angle of tilt of the illuminant.*

2. *At any point on the occluding contour where the maximum of shading occurs, the angle of the tangent to the occluding contour is perpendicular (in the image plane) to the angle of tilt of the illuminant.*

Proof:

1. Take a point P where a self-shadow edge begins on the occluding contour. The normal to the surface at P, N, lies in the image plane (by the definition of an occluding contour). Therefore the tangent plane to the surface at P projects to a straight line L on the image plane. Both the tangent to the occluding contour and the light source direction lie in this plane (since P is on a self-shadow line), thus both project to L in the image. The tangent to the projection of the occluding contour on the image plane is the projection of its tangent on the image plane, namely L, the projection of the light source on the image plane.

   Note: using the same reasoning, L is also the tangent to the projection of the self-shadow line at P.

2. Among all the directions of normals in the image N, the maximum of $N \cdot S$ is obtained when N is the projection of S on the image plane. Thus for Lambertian reflectance, at a point on the occluding contour where the highest brightness is obtained the normal to the surface is parallel to the projection of the light source, therefore the tangent to the occluding contour there is perpendicular to the projection of the light source.

PART 2:

Given a Lambertian surface:

$$\cos\sigma = \frac{\max_{\partial\Omega} I}{\max_{\Omega} I}$$

for brightness I, slant σ, surface Ω of an object, and the occluding contour $\partial\Omega$.

Proof:

The shading of a Lambertian surface with normal N is $\rho\lambda N \cdot S$ for albedo ρ and light intensity λ. $\max_{\Omega} I$ on a convex surface is therefore $\rho\lambda$. In the absence of occlusion, it is visible unless $\max_{\partial\Omega} I = \max_{\Omega} I$. From the proposition above $\max_{\partial\Omega} I$ is obtained when N is parallel to the projection of S on the image plane, namely, $N \cdot S = \cos\sigma$, which finishes the proof.

ACKNOWLEDGEMENTS: I'm very grateful to W. Richards, A. Zisserman, T. Poggio, S. Ullman, E. Hildreth and A. Shashua for many helpful suggestions regarding the manuscript. I also thank T. Breuel for helping me using his image processing package.

References

[BH89] M. J. Brooks and B. K. P. Horn. Shape and source from shading. In B. K. P. Horn and M. J. Brooks, editors, *Shape from Shading*, pages 53–68. MIT Press, Cambridge, MA, 1989.

[Bin87] T. O. Binford. Generalized cylinders representation. In S. C. Shapiro, editor, *Encyclopedia of Artificial Intelligence*, pages 321–323. Wiley, New York, 1987.

[BJ86] P. J. Besl and R. C. Jain. Invariant surface characteristics for 3D object recognition in range images. *Computer Vision, Graphics, and Image Processing*, 33:33–80, 1986.

[Bra83] M. J. Brady. Criteria for representations of shape. In J. Beck, B. Hope, and A. Rosenfeld, editors, *Human and machine vision*, pages 39–84. Academic Press, New York, 1983.

[BZK89] A. Blake, A. Zisserman, and G. Knowles. Surface descriptions from stereo and shading. In B. K. P. Horn and M. J. Brooks, editors, *Shape from shading*, pages 29–52. The MIT press, Cambridge, MA, 1989.

[Can86] J. F. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:679–698, 1986.

[DoC76] M. P. DoCarmo. *Differential geometry of curves and surfaces.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1976.

[Ett88] G. J. Ettinger. Large hierarchical object recognition using libraries of parametrized model sub-parts. In *Proceedings of the 1st International Conference on Computer Vision*, pages 32–41, December 1988.

[Hor86] B. K. P. Horn. *Robot vision.* MIT Press, Cambridge, Mass., 1986.

[IH81] K. Ikeuchi and B. K. P. Horn. Numerical shape from shading and occluding boundaries. *Artificial Intelligence*, 15:141–184, 1981.

[KvD79] J. J. Koenderink and A. J. van Doorn. The internal representation of solid shape with respect to vision. *Biological Cybernetics*, 32:211–217, 1979.

[KvD80] J. J. Koenderink and A. J. van Doorn. Photometric invariants related to solid shape. *Optica Acta*, 27(7):981–996, 1980.

[LR89] C-H. Lee and A. Rosenfeld. Improved methods of estimating shape from shading using the light source coordinate system. In B. K. P. Horn and M. J. Brooks, editors, *Shape from Shading*, pages 323–348. MIT Press, Cambridge, MA, 1989.

[MT86] E. Mingolla and J. T. Todd. Perception of solid shape from shading. *Biological Cybernetics*, 53:137–151, 1986.

[Nal88] V. S. Nalwa. Representing oriented peicewise c2 surfaces. In *Proceedings of the 1st International Conference on Computer Vision*, pages 40–51, December 1988.

[Pen86] A. P. Pentland. Local shading analysis. In A. P. Pentland, editor, *From pixels to predicates*, pages 40–77. ablex, New Jersey, 1986.

[Pen88] A. Pentland. Shape information from shading: a theory about human perception. In *Proceedings of the 2nd International Conference on Computer Vision*, pages 404–413, Tarpon Springs, FL, 1988. IEEE, Washington, DC.

[SK83] S. A. Shafer and T. Kanade. Using shadows in finding surface orientation. *Computer Vision, Graphics, and Image Processing*, 22:145–176, 1983.

[TR89] J. T. Todd and F. D. Reichel. Ordinal structure in the visual perception and cognition of smoothly curved surfaces. *Psychological Review*, 96(4):643–657, 1989.

[VMA86] B. C. Vemuri, A. Mitiche, and J. K. Aggrawal. Curvature-based representation of objects from range data. *image and vision computing*, 4:107–114, 1986.

[Wei88] I. Weiss. Projective invariants of shapes. In *Proceedings of the 1st International Conference on Computer Vision*, pages 291–297, December 1988.

[Wei89] D. Weinshall. Direct computation of 3D shape and motion invariants. A.I. Memo No. 1131, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, May 1989.

[Woo80] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19:139–144, 1980.

[Yui89] A. L. Yuille. Zero crossings on lines of curvature. *Computer Vision, Graphics, and Image Processing*, 45:68–87, 1989.