SRI International

LOCAL SHADING ANALYSIS

Technical Note 272

November 1982

By: **Alex Paul Pentland**, Computer Scientist

Artificial Intelligence Center
Computer Science and Technology Division

SRI Project 1019

The final stages of the research reported herein was supported at SRI by Army Research Office Contract DAAG29-79-C-0216. A substantial portion of this research was accomplished at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research Contract N00014-80-C-0505, in part by National Science Foundation Grant 79-23110MCS.

SRI International®

333 Ravenswood Ave. • Menlo Park, CA 94025
(415) 326-6200 • TWX: 910-373-2046 • Telex: 334-486

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
NOV 1982		**00-11-1982 to 00-11-1982**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Local Shading Analysis	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **SRI International,333 Ravenswood Avenue,Menlo Park,CA,94025**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**		**40**	

ABSTRACT

Local analysis of image shading, in the absence of prior knowledge about the viewed scene, may be used to provide information about the scene. The following has been proved.

Every image point has the same image intensity and first and second derivatives as the image of an umbilical point (a point with equal principal curvatures) on a Lambertian surface; there is *exactly one* combination of surface orientation, curvature, (overhead) illumination direction and albedo times illumination intensity that will produce a particular set of image intensity and first and second derivatives. A solution for the unique combination of surface orientation, etc., at umbilical points is presented

This solution has been extended by using general position and regional constraints to obtain estimates of the following:

- Surface orientation at each image point
- Whether the surface is planar, singly or doubly curved at each point
- The mean illuminant direction within a region
- Whether a region is convex, concave, or is a saddle surface.

Algorithms to recover illuminant direction, identify discontinuities, and estimate surface orientation have been evaluated on both natural and synthesized images, and have been found to produce useful information about the scene.

I. Introduction

A spatially restricted analysis of an image is logically the first stage of any visual system. This initial stage of analysis is especially important because it determines what information will be available to the remainder of the visual system; if a rich description of the world can he computed locally, there is a smaller computational load placed on the remainder of the system. It is, therefore, important to acertain as much ahout the world as possible at this first stage of processing.

Biological visual systems conform to this principal. There is overwhelming evidence that they devote a large percentage of their neurons to an initial *local* analysis of the image. Thus, assessment of the limits and potential uses of a local analysis can be expected to provide insight into both machine and biological vision problems.

What information is available locally? When we examine a small neighhorhood around an image point, we often find only small changes in shading (changes in image intensity[1]). It is unusual to find a contour passing through an image point. Thus, if we are to learn about scene characteristics from local examination of an image, we must concern ourselves with shading[2] The main question posed in this paper will therefore be: What information can, or cannot, be recovered from an unfamiliar image through a local analysis of shading? In the following sections I shall first discuss the limitations that are inherent in any local analysis of image shading, and then show how information about the scene can be determined by means of additional constraints derived from general position and the distribution of data within homogeneous image regions. Proofs of the various propositions are presented in the appendix.

Previous work. Horn and his colleagues [2], [3] have analyzed the process of image formation and have developed several numerical integration schemes for using image intensity to solve for object shape. These shape-from-shading techniques, however, require considerable *a priori* knowledge of the scene, and they function by propagating constraint from houndary conditions (such as those provided by smooth occluding contours) over the surface whose shape is to be estimated. These techniques, therefore, cannot he applied to an unfamiliar, unanalyzed scene and do not perform the purely local analysis I wish to consider here.

Bruss [4] has addressed the question of whether shape can be derived from shading using a purely local analysis (again with a considerahle *a priori* knowledge of the scene assumed). She proved that no shape-from-shading technique can yield a unique solution without additional constraints which, in certain restricted cases (most importantly electron

[1]To avoid confusion, the term "image intensity" will be used throughout this document, rather than the technically more correct "image irradiance," both for the flux per unit area falling on the image plane and for the measured image irradiance. The two may be assumed to be numerically equal; and thus the distinction has little significance for the task at hand.

[2]For the purposes of this paper we shall restrict our attention to shading, because the problem of estimating shape from local texture information has already received much attention, e.g., [17] and [18].

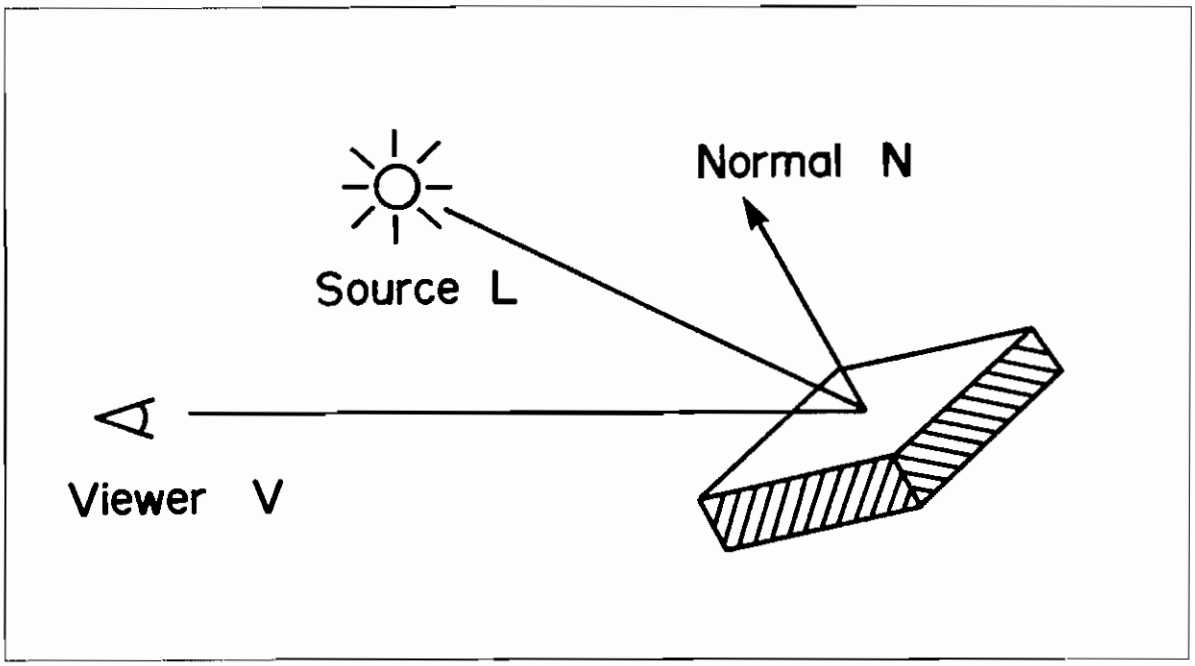

Figure 1. A simple model of image generation. **N** is the surface normal, **L** the illumination direction, **V** the viewer's direction. If λ is the flux emitted toward the surface, ρ the average reflectance of the surface, and we assume distant light source and a Lambertian reflectanee function for the surface, then the image intensity I is given by $I = \rho\lambda(\mathbf{N} \cdot \mathbf{L})$.

micrograpbs), may be provided by the bounding contour of the surface. Bruss, however, dealt mostly with the question of what *cannot* be obtained from an analysis of shading; tbe question of what *can* be accomplished with a local analysis of shading was not fully explored. It is this, consequently, that we discussed below.

A. Image Formation

Before we can make quantitative statements about the limitations or usefulness of a local analysis of shading, we must first develop a mathematical model of the image generation process. Figure 1 shows a simple model of image generation: a distant point-source illuminant at direction **L**, a patch of surface with surface normal **N**, and a viewer in direction[3] **V**. We will assume orthographic projection; note, however, because the model is purely local orthographic and perspective projection are identical except at points of discontinuity.

The surface normal **N**, the viewer's direction **V** and the illuminant direction **L** are unit vectors in Cartesian three-space. As they are unit vectors, two parameters suffice to specify them, the third being determined by the constraint that they have unit magnitude. Two parameters that are often chosen are the slant σ and the tilt τ. The tilt of a surface is the image-plane component of surface orientation and is equal to $\tan^{-1}(y_N/x_N)$, where x_N and y_N are the x and y components of the surface normal. The slant of the surface is the depth component of surface orientation and is equal to $\cos^{-1}(z_N)$, where z_N is the z

[3]All boldface variables (e.g., **N**, **L**, **p** etc.) represent three- dimensional vectors (x, y, z), all other variables are scalars. The (x, y) plane is taken parallel to the image plane, so that $\mathbf{V} = (0, 0, 1)$.

component of the surface normal.

The image intensity I is in general given by

$$I = \rho\lambda(\mathbf{N}\cdot\mathbf{L})R(\mathbf{N},\mathbf{L},\mathbf{V})(\mathbf{N}\cdot\mathbf{V})^{-1}$$

where ρ, the *albedo*, is the portion of incident light that is reflected, λ is the amount of light incident upon the surface and $R(\mathbf{N},\mathbf{L},\mathbf{V})$ is the reflectance function, which describes how much of the reflected light leaves in each direction. The amount of incident light reflected in the viewer's direction $\mathbf{V}$ is a function of the illuminant direction $\mathbf{L}$ and the surface normal $\mathbf{N}$. The term $(\mathbf{N}\cdot\mathbf{L})$ describes the amount of light incident upon the surface, while the term $(\mathbf{N}\cdot\mathbf{V})^{-1}$ describes the foreshortening that occurs during projection into the image[4] . A Lambertian reflectance function, an idealization of rough, matte surfaces, is defined as

$$R(\mathbf{N},\mathbf{L},\mathbf{V}) = \mathbf{N}\cdot\mathbf{V}$$

which is proportional to the reciprocal of the foreshortening caused by the projection term. Thus, for a Lambertian surface the reflectance function and the effect of projection cancel each other, and the equation for image intensity becomes

$$I = \rho\lambda(\mathbf{N}\cdot\mathbf{L}) \tag{1}$$

Thus, the assumption of a Lambertian reflectance function is equivalent to the assumption that the scattering of incident light is isotropic. We shall assume a Lambertian reflectance function.

Generality of the assumptions. The assumption of a distant point-source illuminant and a Lambertian reflectance function is not as restrictive as it might at first seem. We note, for instance, that for a Lambertian surface *any* constant distribution of illumination is equivalent to a single distant point-source illuminant; this follows from application of the mean value theorem. Because we are concerned only with local analysis, the requirement that the distribution of illumination be constant is almost trivially met[5] . Therefore, local inferences derived with this single-illuminant/ Lambertian model will generally be valid whenever the surface scatters incident light in an isotropic manner, regardless of the actual distribution of illumination.

B. The Derivatives Of Image Intensity

The image intensity I and the surface normal $\mathbf{N}$ are different at each point (x,y) in the image, and thus are perhaps better written $I(x,y)$ and $\mathbf{N}(x,y)$. However, when they are discussed at a particular point P, they will be written as simply I and $\mathbf{N}$. Similarly, we shall write dI and $d\mathbf{N}$ to designate the first derivative of image intensity and the surface normal, respectively, at a point P in the direction (dx,dy). The partials of I, $\mathbf{N}$, and other variables will be denoted by subscripts, i.e., $I_x = \partial I/\partial x$ and $\mathbf{N}_y = \partial\mathbf{N}/\partial y$.

[4]In other terminology, $\mathbf{N}\cdot\mathbf{L}$ is equal to the cosine of the incident angle, and $\mathbf{N}\cdot\mathbf{V}$ is equal to the cosine of the exitent angle.

[5]Only illumination near the "horizon" of the surface patch causes a problem; in this case there is some self-occlusion and thus somewhat different illumination at neighboring points.

If we are examining a small, homogeneous region of an image, it is reasonable to assume that the illumination and albedo of the surface change very little, and so we may treat $\mathbf{L}$, ρ, and λ as constants. If we also assume a Lambertian reflectance function, so that Equation (1) applies, then

$$dI = d(\rho\lambda(\mathbf{N}\cdot\mathbf{L})) = \rho\lambda(d\mathbf{N}\cdot\mathbf{L}) + \rho\lambda(\mathbf{N}\cdot d\mathbf{L}) = \rho\lambda(d\mathbf{N}\cdot\mathbf{L}) \tag{2}$$

The term $(\mathbf{N}\cdot d\mathbf{L})$ is zero because $\mathbf{L}$ was assumed constant. Similarly, the second derivative of image intensity is

$$d^2I = d(\rho\lambda(d\mathbf{N}\cdot\mathbf{L})) = \rho\lambda(d^2\mathbf{N}\cdot\mathbf{L}) + \rho\lambda(d\mathbf{N}\cdot d\mathbf{L}) = \rho\lambda(d^2\mathbf{N}\cdot\mathbf{L}) \tag{3}$$

Thus, the second derivative of image intensity depends upon the second derivative of the surface normal, just as the first derivative depended upon the first derivative of the surface normal.

II. Local Shading Information: Limitations And Potential

Before we can know what is possible to accomplish with local shading information, it is important to characterize what *cannot* be done. The following proposition describes the fundamental limitation which is inherent to any local analysis of image shading:

Proposition 1. The image of a Lambertian umbilical point (a point with equal principal curvatures) can produce any combination of image intensity I and derivatives I_x, I_y, I_{xx}, I_{yy} and I_{xy}.

This proposition says that when we view a point on a surface, regardless of what the actual surface curvatures are or what the actual surface reflectance function is, the resulting image point always looks like an umbilical point on a Lambertian surface. This proposition implies, therefore, that it is impossible for a local analysis of the image to determine unambiguously whether a surface is Lambertian and whether the principal curvatures are equal; there will always be the possibility that the observed point is an umbilical point on a Lambertian surface. We cannot resolve these ambiguities by resorting to higher derivatives of image intensity because, although more measurements are obtained by measuring the higher derivatives, each additional derivative brings in more unknowns than measurements.

We can see by the following argument that this proposition is likely to be true. Consider that at each point in an image we can measure the intensity, and its first and second derivatives to obtain six independent measurements, which are I, I_x, I_y, I_{xx}, I_{yy} and I_{xy}. To specify the image intensity of an umbilical point on a Lambertian surface[6] requires six independent parameters: τ the surface tilt[7] , σ the surface slant[8] R the radius of curvature, $(l_1, l_2, \sqrt{1-l_1^2-l_2^2})$ the illuminant direction, and $\rho\lambda$ the surface albedo times illuminant intensity. Solving for these six unknowns requires at least one measurement for

[6]The set of all possible images of umbilical points is obtained by considering surfaces of the form $z(x,y) = \sqrt{R^2-x^2-y^2}$ for particular values of $R > 0$, $R \geq x \geq -R$, $R \geq y \geq -R$.

[7]Tilt is the image-plane component of surface orientation.

[8]Slant is the depth-component of surface orientation.

each unknown; thus, the measurement of intensity, first and second derivatives can at most establish the six parameters required to specify a Lambertian umbilical point. No additional measurements are available to determine whether the surface curvatures are unequal or whether the reflectance function is Lambertian.

Proposition 1 leaves open the possibility that there may be a great many combinations of surface orientation, curvature, etc., corresponding to each combination of image intensity and its derivatives. If the equations for image intensity were linear, there would be exactly one combination of the six parameters that would correspond to the observed measurements. Although the equations are not linear, the following proposition shows that there are only two possible combinations of these factors that will yield a particular combination of image intensity aud derivatives.

Proposition 2. Given the image of an umbilical point on a Lambertian surface with image intensity I and derivatives I_x, I_y, I_{xx}, I_{yy} and I_{xy} there are two possible combinations of surface orientation, curvature, illuminant direction and surface albedo times illuminant intensity, one with the illuminant direction above the line of sight, the other exactly opposite surface tilt and illuminant tilt.

Thus, the ambiguity present in local image shading is not much greater than was evident from the first proposition; there is only the additional ambiguity that arises from a symmetry involving the illuminant direction and the tilt of the surface. This symmetry results in the interrelation of the direction of illumination and the convexity of the surface; if the illuminant direction is taken to have the opposite tilt (e.g., from above the line of sight to below it) the convexity of the surface will reverse. Therefore, one cannot determine the convexity of the surface unless something is known about the illuminant direction [1], [5].

Using propositions 1 and 2, we can produce an exhaustive characterization of the limitations of any local analysis of shading. A local analysis of shading cannot

- Determine the sign and magnitude of the surface curvatures[9] e.g., whether the surface is convex, concave or a saddle-shaped and whether or not the curvatures are unequal.
- Determine the surface reflectance function.
- Separate the surface albedo from illuminant intensity.

A. Solving For Image Formation Parameters At An Umbilical Point

The amount of information we can extract from a local analysis of shading (given that we are viewing the image of a Lambertian umbilical point) is surprising. For such image points we can solve for *every* parameter in the image formation process. This seems to indicate that there is approximately two degrees of freedom left undetermined by the local shading information when we view a point within a homogeneous region of an image: the ratio of the surface curvatures and the degree to which the surface is non-Lambertian.

The umbilical-point case is the most complex situation in which all of the image formation parameters may be recovered locally. The fact that there are only relatively

[9]With the additional constraint provided by general position, we can determine when the curvatures are zero. This is shown in the following section.

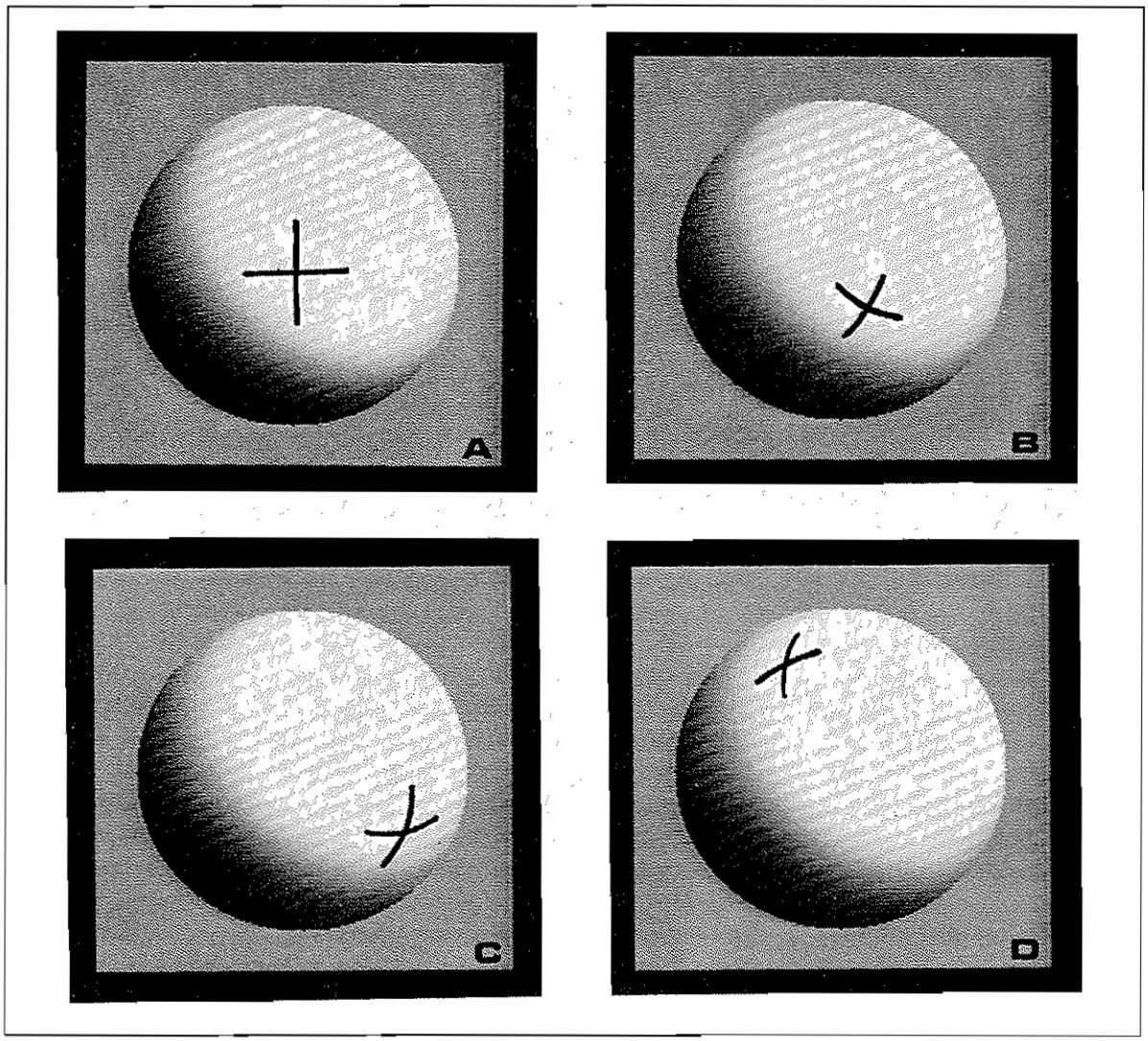

Figure 2. The manner in which image curvature "spreads" indicates the tilt of the surface. This may be understood by imagining that we could observe the lines of curvature on a surface directly. They would look just like the lines drawn in this figure. If we were looking straight down on the surface of a sphere, the lines of curvature would appear perpendicular, as in (a). As we tilted the surface to one side, the lines of curvature would appear progressively more spread, as in (b) and (c). Different directions of tilt cause spreading in different directions, as demonstrated in (d). The amount of spread depends on the slant of the surface.

few additional parameters required to obtain a reasonably general model suggests that the umbilical-point solution may provide us with a useful (albeit simplified) model of how the various portions of the image formation process evidence themselves in the image, and may also prove useful as a tool for analyzing image points. The umbilical-point solution for surface orientation, for example, is instructive to examine. How can surface orientation be determined from local shading information?

Imagine that we could observe the lines of curvature on a surface directly. They would look like the lines drawn in Figure 2. If we were looking straight down on the surface of a sphere, the lines of curvature would appear perpendicular, as in Figure 2 (a). As we tilted the surface to one side, the lines of curvature would appear progressively more spread, as in

Figures 2 (b) and (c). Different directions of tilt would cause spreading in different directions, as demonstrated in (d).

We cannot observe lines of curvature on the surface directly, of course, but we can observe the interaction of surface curvature with the illuminant in the second derivatives of image intensity. The second derivative of image intensity has three components: I_{xx} and I_{yy}, the "curvature" of image intensity along the x and y axes, respectively, and I_{xy}, the "spread" of those curvatures. Just as with the spread of the lines of curvature, the direction in which the spread term is greatest is also the direction of the surface tilt. The direction in which this spread is greatest also turns out to be the direction along which d^2I is greatest, and hence the following proposition:

Proposition 3. Given the image of an umbilical point on a Lambertian surface, the tilt of the surface τ is the image direction in which the second derivative of image intensity d^2I is greatest.

Thus, for Lambertian umbilical points the tilt may be determined from the second derivative of image intensity directly, without *a priori* knowledge. This leaves only the surface slant to be determined.

In Figure 2 the direction of the spread indicated the tilt of the surface. Similarly, the *amount* of the spread indicates the slant (depth) component of the surface orientation. Measuring the magnitude of this spread relative to the total curvature (as measured by the Laplacian) provides an indicator of the surface slant, as described in the following proposition.

Proposition 4. Given the image of an umbilical point of a Lambertian surface, the surface slant σ is given by

$$\sigma = \cos^{-1} \sqrt{\frac{k\nabla^2 I - (k^2+1)I_{xy}}{k\nabla^2 I + (k^2+1)I_{xy}}}$$

where $k = \tan^{-1} \tau$.

The slant and the tilt propositions together determine surface orientation exactly. Note that neither the slant nor the tilt estimate requires *any* knowledge of illuminant direction, surface albedo, curvature, or illuminant intensity.

B. Unlikely Umbilical-Point Solutions: Constraint From General Position

Although the umbilical-point solution will always provide us with an interpretation that is consistent with the local image data, it sometimes yields an interpretation of an image point that strikes us as *unlikely* because the umbilical-point solution requires an unlikely configuration of orientation, illumination, or viewer direction.

When we observe an image point for which the umbilical-point solution requires an unlikely configuration, we have available the additional constraint provided by general position to help us interpret the image data. This additional constraint allows us to reject the umbilical-point solution and infer that something special has occurred in the image formation process — something that may permit further analysis.

Zero second derivative. One such case arises when one or both of the second derivatives are zero. When one of the second derivatives is zero, the umbilical-point solution is a

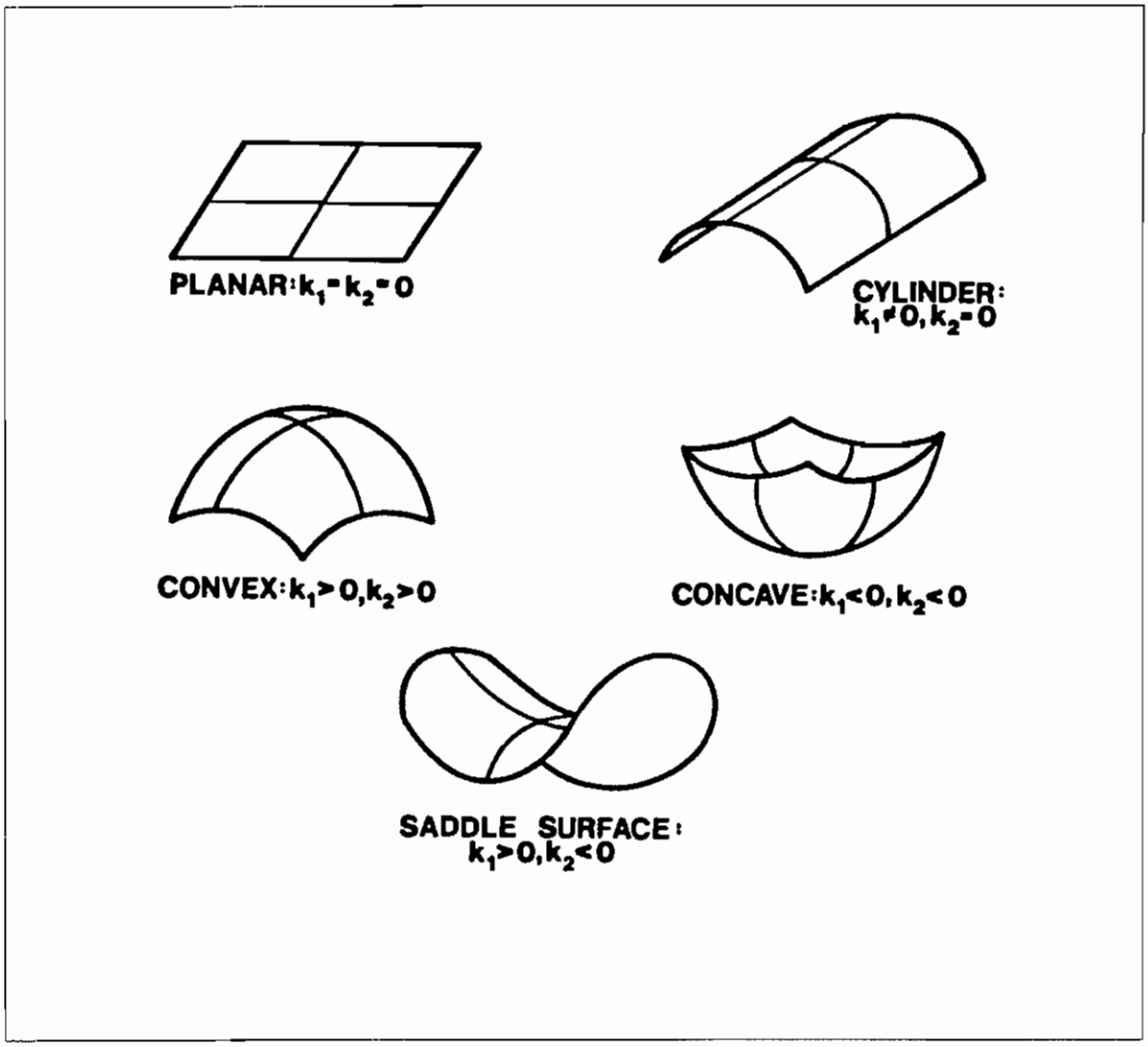

Figure 3. Surface types. Surfaces may be classified into five types: planar, cylindrical, convex, concave, or saddle surface. The classification of a surface depends on whether the two principal curvatures κ_1 and κ_2 are positive, negative, or zero.

surface patch whose orientation is exactly perpendicular to the line of sight. When both of the second derivatives are zero, the umbilical-point solution requires that the illuminant direction be exactly in tbe image plane. These interpretations of the image point are unlikely because precise alignment of surface orientation or of illuminant direction is necessary, i.e., this interpretation of the image point presupposes a violation of general position. The more likely inference when a zero second derivative is observed is that one of the surface curvatures is zero. The fact that this inference is valid (shown in the appendix) allows us to partially classify the surface type.

Surface points may be classified into five types: planar, cylindrical, convex, concave, or saddle surface. These five types are shown in figure 3. The classification of a surface depends on whether the two principal curvatures κ_1 and κ_2 are positive, negative, or zero:

plane	$\kappa_1 = 0$	$\kappa_2 = 0$
cylinder	$\kappa_1 \neq 0$	$\kappa_2 = 0$
convex	$\kappa_1 < 0$	$\kappa_2 < 0$
concave	$\kappa_1 > 0$	$\kappa_2 > 0$
saddle surface	$\kappa_1 > 0$	$\kappa_2 < 0$

One important step in identifying the type of surface is determining when the principal curvatures are zero, as this allows us to classify the surface as planar, cylindrical, or doubly curved.

When one of the principal curvatures is zero, the surface normal does not change as we travel along the surface in the direction of that principal curvature. Because the surface normal $\mathbf{N}$ does not change along that direction (let us specify the direction by (dx, dy)), we know that[10] $d\mathbf{N} = 0$ along (dx, dy). Since

$$dI = \rho\lambda d\mathbf{N} \cdot \mathbf{L} \tag{2}$$

we see that dI must also be zero along (dx, dy). Unfortunately, the reverse inference is not generally true, because dI is zero along *some* direction for every image point. Therefore, we cannot infer that $d\mathbf{N} = 0$ in direction (dx, dy) just because $dI = 0$ along that direction.

That problem does not occur when we observe $d^2I = 0$ along a direction (dx, dy). When the surface normal does not change along a direction (dx, dy), then $d^2\mathbf{N} = 0$ along (dx, dy). Since

$$d^2I = \rho\lambda d^2\mathbf{N} \cdot \mathbf{L} \tag{3}$$

we see that, when $d^2\mathbf{N} = 0$ along (dx, dy), then d^2I must also be zero along (dx, dy). For the second derivative, the reverse inference — that $d^2\mathbf{N} = 0$ because $d^2I = 0$ — is generally valid, for when we observe that $d^2I = 0$ along direction (dx, dy), we can conclude that either (1) $d^2\mathbf{N}$ is perpendicular to $\mathbf{L}$ or (2) that $d^2\mathbf{N} = 0$. As it is unlikely that $d^2\mathbf{N}$ is perpendicular to $\mathbf{L}$ for any distance, we may legitimately conclude that, if we observe that $d^2I = 0$ for some distance along direction (dx, dy), then $d^2\mathbf{N} = 0$. This implies that $d\mathbf{N}$ is constant along (dx, dy), and, if $d\mathbf{N}$ remains constant for some distance, we may use the constraint of general position to conclude that $\mathbf{N}$ is also constant (this is shown in the appendix).

We can now begin to classify the surface. If we observe that $d^2I = 0$ along a line in the image, then $\mathbf{N}$ does not change along that locus and we have a surface that is cylindrical along that line. If we observe that $d^2I = 0$ along a direction (dx, dy) throughout some region in the image, then the surface is a cylinder with an axis whose projection points in the (dx, dy) direction. Similarly, if we observe that $d^2I = 0$ along two orthogonal directions, then $\mathbf{N}$ does not change along either direction and thus the surface is planar. Finally, if $d^2I \neq 0$ in all directions, then the surface must be doubly curved, i.e., it is convex, concave, or a saddle surface. We can not distinguish among these alternatives on the basis of local shading information alone — a consequence of the previous propositions. Thus, we have the following proposition:

[10] Here the '0' in '$d\mathbf{N} = 0$' is the zero *vector*.

Proposition 5. The surface type at a point is partially determined by the number of directions in which $d^2I = 0$.

$d^2I = 0$ in no directions $\Rightarrow$ *convex/concave/saddle surface*
$d^2I = 0$ in one direction $\Rightarrow$ *cylinder*
$d^2I = 0$ in all directions $\Rightarrow$ *plane*

It is interesting to note that linear intensity gradients do not invalidate this classification scheme.

The detection of lines along which a surface is cylindrical is of considerable importance because it is only at such cylindrical lines that changes occur in the surface type (e.g., change from a convex to a saddle surface). As the surface changes from one type to another, the sign of at least one of the principal curvatures changes from positive to negative, or vice versa. In the course of a sign change the curvature is briefly zero, and so the surface is cylindrical along the locus where the surface changes type[11] Thus, lines along which $d^2I = 0$ are places where the surface is undergoing a change of type, and the set of such lines divides the surface into regions that are of the same surface type.

III. Generalization Of The Results: Regional Constraints

In real images, relatively few points are umbilical and relatively few surfaces are Lambertian. Therefore, we must find some additional constraints in order to obtain generally applicable formulas for surface orientation, illuminant direction, and so forth. Unfortunately, the thrust of the preceding propositions is that there is no point-wise local assumption that will generally be true; there will always be at least a two-parameter family of possible solutions.

One way we can obtain additional constraint is to expand our view: to consider *regions* rather than single points only. Once we allow discussion of regions, we find that there are many possibilities for obtaining a good estimate of the mean value of particular parameters within the region, by using inferences about the range or distribution of image data within the region. Having obtained an estimate for the mean value of a parameter, we can then solve for other parameters by *assuming* the already estimated value — i.e., by bootstrapping.

The mean value of a parameter within a region may be used either to comment upon the *average* properties of the region, or we may assume that the parameter is constant throughout the region and thus obtain point-by-point estimates. If we comment only about average properties, then the validity of our deductions depends solely on the accuracy of the initial estimate. If we desire point-by-point estimates, the validity of our inferences is also conditional upon the intraregional variance of the estimated parameter.

In the remainder of this paper I shall discuss results obtained by estimating the mean value of one parameter within a region and then using this estimate as an assumption to infer other properties of the scene. Examples of both regional and point-by-point inference will be presented. Estimates of average properties of a region, such as illumination direction and surface type (e.g., convex, concave, saddle, etc.), have been made by using the

[11] If the change takes place over an extended area, both curvatures will be zero and so the surface will be planar instead of cylindric.

maximum-likelihood estimate of change in surface normal. Point-by-point estimates of surface orientation have been made by using an estimate of the curvature within the region.

A. Finding The Illuminant Direction

Estimating the illuminant direction is difficult because image data are determined by both the surface normal and the illuminant direction. Since evidence relating to illuminant direction is confounded with the unknown direction of the surface normal, estimating the direction of illumination seems to require making some assumption about surface orientation or its derivatives.

One useful assumption is that change in surface orientation ($d\mathbf{N}$) is distributed isotropically within each image region. It is true that $d\mathbf{N}$ is isotropically distributed when considered over all scenes; furthermore, there is a large class of common image regions for which $d\mathbf{N}$ is isotropically distributed. This class of image regions includes all images of convex objects bounded entirely by a gradual occluding contour[12] , such as the image of a smooth pebble.

Given the assumption that changes in surface orientation are isotropically distributed, we can devise a procedure for estimating the illuminant direction $\mathbf{L}$ by looking for the regular biasing effect of the illuminant direction on $d\bar{I}$, the mean value of dI, along various image directions (dx, dy). The effect of the illuminant direction is to make $d\bar{I}$, vary according to

$$\begin{aligned} d\bar{I} &= \rho\lambda d\bar{\mathbf{N}} \cdot \mathbf{L} \\ &= \rho\lambda(d\bar{x}_N x_L + d\bar{y}_N y_L + d\bar{z}_N z_L) \end{aligned}$$

where $d\bar{\mathbf{N}} = (d\bar{x}_N, d\bar{y}_N, d\bar{z}_N)$ is the mean change in $d\mathbf{N}$ measured in image direction (dx, dy), and $\mathbf{L} = (x_L, y_L, z_L)$ is the illuminant direction. Under the assumption that change in surface normal is distributed isotropically within a region, then, along any *one* image direction (dx, dy) we find that $d\bar{x}_N$ is proportional to dx, the x-component of the image direction, that $d\bar{y}_N$ is proportional to dy, the y-component of the image direction, and that $d\bar{z}_N$ is zero. (see [5]) Therefore,

$$d\bar{I} = k(x_L dx + y_L dy) \tag{4}$$

where k is a constant determined by the albedo, illuminant strength, and the variance of the distribution of $d\mathbf{N}$ witbin the region.

Using (4), we can set up a linear regression that employs the mean of dI as measured along various image directions to obtain a maximum-likelihood estimate of the ratio of the unknowns x_L and y_L. This ratio is the tilt of the illuminant direction, which we shall use in identifying surface type. The constant k (and from this the values of x_L, y_L and z_L) can be estimated from the mean and variance of the distribution of dI along various image directions. In this procedure, most of the information about the illuminant direction comes from image points where there are large changes in image intensity, e.g., edges and specularities. This seems to agree with our introspective impression as to how we determine

[12]This may be proved by noting that the surface normals on such an object are perpendicular to $\mathbf{V}$ at the image boundary of such an object, and thus (given that the object is strictly convex) we may form a 1-1 onto map between the surface normals of the object and the Gaussian sphere, which has sum $d\mathbf{N}$ equal to zero.

the illuminant direction. Note that this estimation procedure establishes the tilt of the illuminant direction to within $\pm\pi/2$, leaving an ambiguity regarding illuminant position that is identical to the human perceptual ambiguity that obtains in the absence of cast shadow information.

Evaluation of the illuminant direction estimator. This illuminant estimation procedure has been compared with the answers given by fifteen human observers on a series of digitized pictures of natural objects, such as rocks and logs. The photographs of these objects were made in natural illumination, so that the imaged scenes do not have a Lambertian reflectance function or true point-source illumination. Digitized versions of the pictures were shown to the human subjects, so that both they and the computer procedure would receive exactly the same image information.

Figures 4 (a) and (b) show a comparison of human and computer estimates of illuminant direction. Previous experiments have documented that the fifteen subjects' mean estimates exhibit a standard error of ten degrees in this experimental condition. Thus, the human and computer estimates shown in Figure 4 concur to within experimental error.

Otber evidence concerning the equivalence of human and computer estimates comes from the variance of the two estimates. The illuminant direction estimator generates a confidence statistic for each image, along with its estimate. This confidence statistic is proportional to the variance of the estimate for that image (given the assumptions of the procedure). We can compare the variance of human estimates for a particular picture with the variance of the maximum-likelihood estimate (as predicted by the confidence statistic). This comparison is shown in Figure 4 (c). There is a correlation of 0.63 between the variance of the two sets of estimates, significant at the $p = 0.05$ level. The linear regression line relating the human and maximum-likelihood variance is shown as a dashed line; the coefficients of the regression are significantly different from zero at the $p = 0.01$ level. The significant relationship between the variance of the two estimation procedures (human and computer) shows that, when one of them finds enough information in the image to make a low-variance estimate, so does its counterpart.

One of the images employed is of particular importance, because it is an example of incorrect estimation by humans of the illuminant direction. When the image of the rock shown in Figure 4 (d) was presented to human subjects, they misestimated the illuminant direction by about 120 degrees (it is actually illuminated from top left, not top right as all but two of the fifteen subjects reported). The computer generated estimate, interestingly, agrees with the human ones — even though in both cases the estimates are objectively wrong. This image must violate the assumptions on which the human estimates of illuminant direction are based, because the human estimate is objectively wrong. The special significance of this case is that it also violates the assumptions of the computer estimation procedure in such a manner tbat it produces exactly the same estimate as the human subjects. This is strong evidence that the algorithm people employ to estimate illuminant direction is similar to the one described above.

B. Using The Illuminant Direction To Type The Surface

Once we have an estimate of **L** for a region, we can use this estimate as a basis for

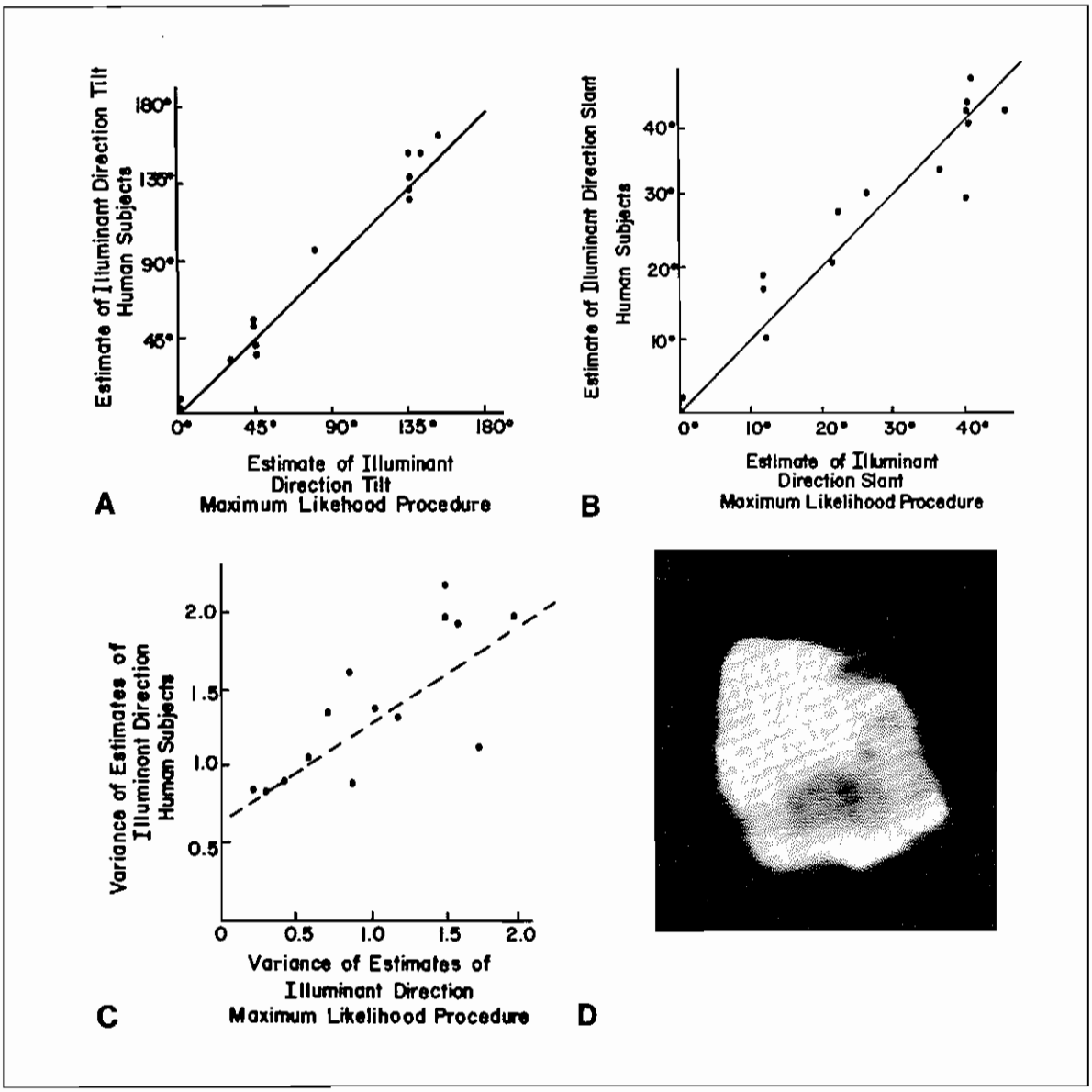

Figure 4. A comparison of human and computer estimates of illuminant direction in images of natural objects. Part (a) shows the comparison for the tilt component of the illuminant direction, and (b) shows the comparison for the slant component of the illuminant direction. Part (c) shows the relationship between the variance of human estimates of illuminant direction and the variance of the computer's estimate of illuminant direction. There is a correlation of 0.63 between the variances of the two sets of estimates, significant at the $p \leq 0.05$ level. The dashed line is the linear regression line relating the variance of the two estimation procedures. Part (d) is a picture of a rock for which both human estimates of illuminant direction and the maximum likelihood estimate agreed, but were objectively wrong. Actual illumination direction is top right, not top left as reported by all but two of the fifteen human subjects.

acquiring further information about the image formation process. One important use of **L** is to provide sufficient constraint to identify the surface as convex, concave or saddle — thus completing the typing of the surface.

Figure 5 contains an example of the "crater illusion." In this image, the shadow information is not prominent enough to determine the illuminant direction; consequently,

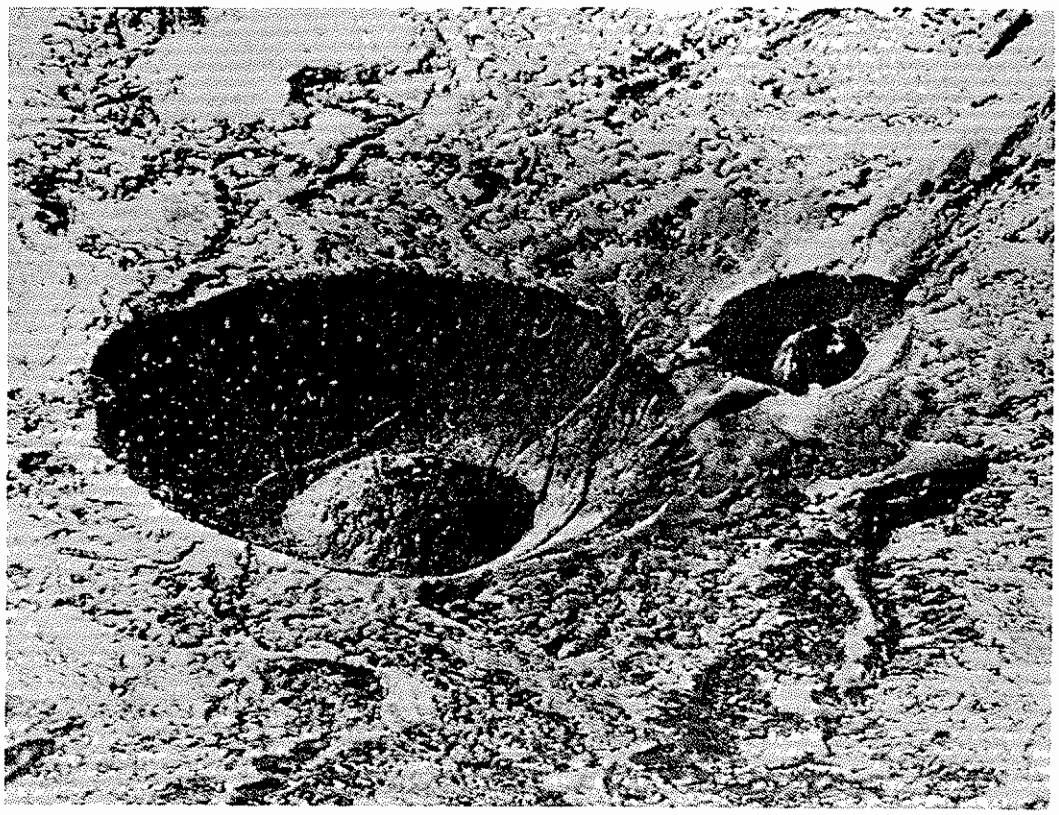

Figure 5. The Crater Illusion. Pictures of craters can look like bumps instead of depressions if we imagine the light source to be at the bottom of the picture rather than at the top; to see this, turn the figure upside down. This picture is of an ash cone in the Hawaiian Islands (courtesy of W. Richards).

no matter how the picture is turned the illumination is always (by default) seen to be coming from above[13] Thus, the direction of illumination (relative to the image) changes as the image is turned upside down. When the direction of illumination changes, the convexity of the imaged surface also changes — thereby demonstrating that people use the direction of illumination to determine the convexity of the surface.

How can information about the direction of illumination be used to determine the surface convexity? When we invert Figure 5, the sign of dI as we move toward the apparent illuminant reverses, because the perceived image-plane component of the illuminant direction shifts by $\pi/2$ radians. Let us consider what the sign of dI tells us about the convexity of the imaged surface.

Equation (2) shows that the change in image intensity dI is dependent upon $d\mathbf{N}$, the change in the surface normal:

$$dI = d(\rho\lambda(\mathbf{N}\cdot\mathbf{L})) = \rho\lambda(d\mathbf{N}\cdot\mathbf{L}) + \rho\lambda(\mathbf{N}\cdot d\mathbf{L}) = \rho\lambda(d\mathbf{N}\cdot\mathbf{L}) \qquad (2)$$

It turns out tbat $d\mathbf{N}$ is always perpendicular to $\mathbf{N}$, which can be sbown by observing that

$$2d\mathbf{N}\cdot\mathbf{N} = d(\mathbf{N}\cdot\mathbf{N}) = d(1) = 0$$

[13]This is an example of the above-mentioned $\pm\pi/2$ ambiguity in illuminant direction.

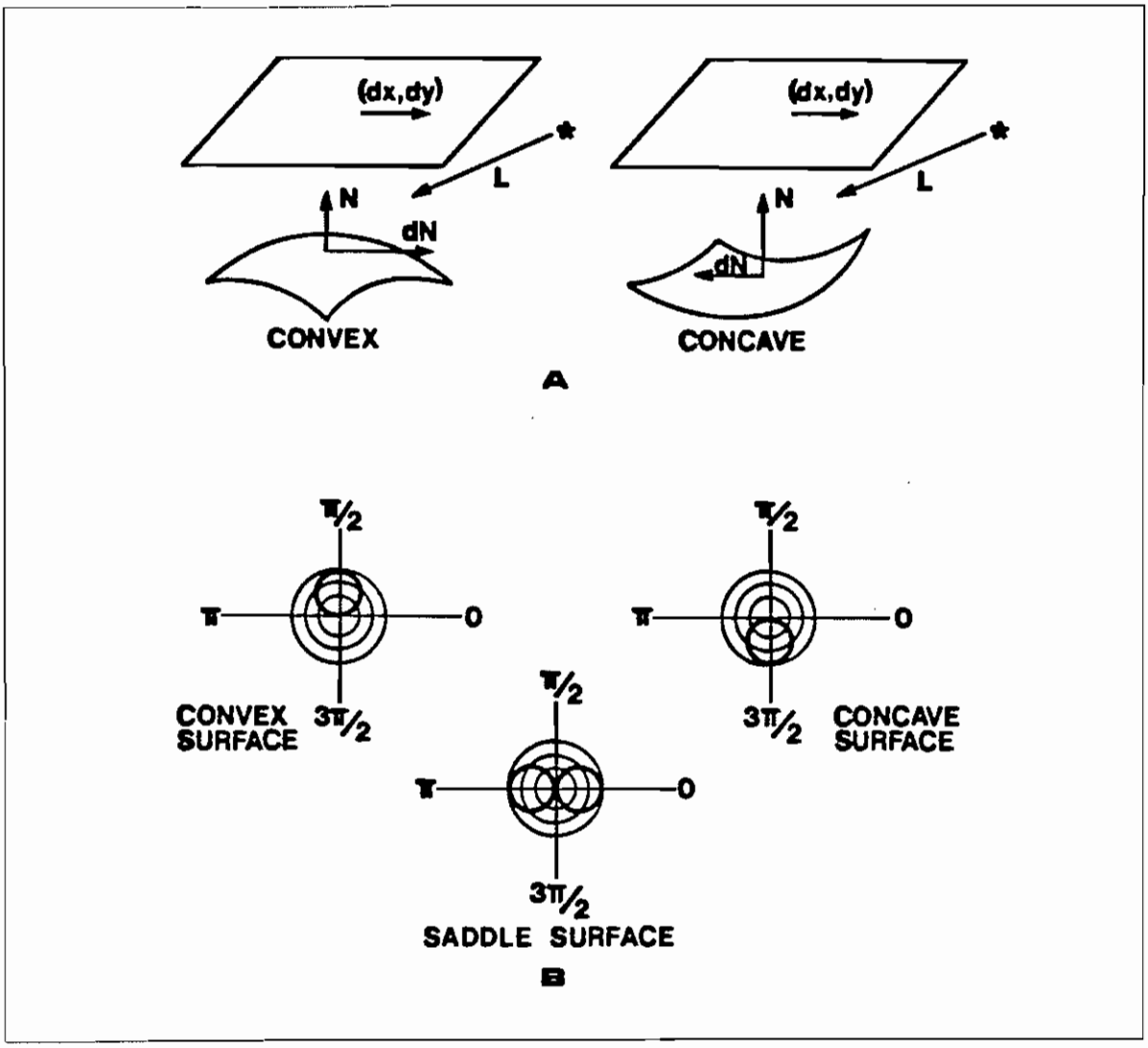

Figure 6. Estimation of surface type. (a) For a convex surface, $d\mathbf{N}$ measured along image direction (dx, dy) typically points in the direction $(dx, dy, 0)$, so that, if the image direction (dx, dy) is toward the light source, then $dI = \rho\lambda d\mathbf{N} \cdot \mathbf{L}$ is positive. For a concave surface, $d\mathbf{N}$ measured along image direction (dx, dy) typically points in the direction $(-dx, -dy, 0)$, so that $dI = \rho\lambda d\mathbf{N} \cdot \mathbf{L}$ is negative. Thus, the sign of dI in relation to the illuminant direction gives an estimate of the surface convexity along that direction. (b) The illuminant direction may be used to provide sufficient constraint to determine the qualitative type of surface. Each type of surface has a generic appearance, which may be characterized by the angle between τ_0, the direction in which $dI = 0$, and τ_L, the illuminant direction. The distribution of $\tau_0 - \tau_L$ is shown for each surface type, assuming that the change in surface normal is isotropically distributed, and taking $dI > 0$ to the right of τ_0. It can be seen that the appearance of the different types does not overlap much, so that a good identification of the surface type within the region may be made from this angle.

Whether dI is positive or negative along a particular direction depends upon whether $d\mathbf{N}$ points toward or away from the illuminant direction $\mathbf{L}$. This is illustrated in Figure 6 (a).

If we assume that change in surface orientation is isotropically distributed within an image region, then, for a convex surface, dI measured in the image direction (dx, dy) will typically be positive if (dx, dy) is toward $\mathbf{L}$. The sign of dI is positive because, for a convex

surface, $d\mathbf{N}$ measured along image direction (dx, dy) points on the average in the direction $(dx, dy, 0)$, so that $dI = \rho\lambda d\mathbf{N} \cdot \mathbf{L}$ is positive. In contrast, if the surface is concave dI will typically be negative because, for a concave surface, $d\mathbf{N}$ measured along image direction (dx, dy) points on the average in the direction $(-dx, -dy, 0)$, so that $dI = \rho\lambda d\mathbf{N} \cdot \mathbf{L}$ is negative. Thus, the sign of dI as we measure toward and away from the light source gives us an estimate of the convexity of the surface in that direction.

Unfortunately, $d\mathbf{N}$ measured along (dx, dy) does not usually point precisely along either the direction $(dx, dy, 0)$ or $(-dx, -dy, 0)$. Thus, even if we are given $\mathbf{L}$ and $\mathbf{N}$, there remain too many unknown factors to establish the surface type with certainty. Each surface type, however, does have a typical or generic appearance. Therefore, given the tilt of $\mathbf{L}$ and the assumption that change in surface orientation is distributed isotropically within a region, we can estimate the surface type by observing the sign of dI measured toward and away from the illuminant.

Sufficient information for estimating the surface type is provided by the angle between τ_0, the direction along which $dI = 0$, and τ_L, the tilt of the illuminant direction, as the sign of dI is positive on one side of τ_0, negative on the other. Thus, knowing τ_0 and τ_L enables us to estimate the surface type. Figure 6 (B) shows the probability distribution of τ_0 for each surface type given τ_L, the tilt of the illuminant direction[14] , and the assumption that surface orientation is isotropically distributed within the region. As can be seen by comparing the overlap between these probability distributions, the likelyhood of a correct identification is quite good.

Note that the ambiguity of $\pm\pi/2$ in the estimation of illuminant direction tilt leads to a global convexity/concavity ambiguity. Thus, just as with human perception, when a scene is sufficiently simple as to make $\mathbf{L}$ uncertain, the direction of illumination may be "switched" by $\pi/2$, which causes all the convexity/concavity determinations to change, as in Figure 5.

C. Estimation Of Surface Orientation

Although in real images relatively few points are umbilical and relatively few surfaces are Lambertian, the solution for surface tilt turns out to be fairly robust. The slant equation, however, depends critically on equal surface curvatures and on exact knowledge of the surface tilt. We must look further to find an estimator of surface slant that will be generally serviceable.

When a patch of surface is slanted away from the viewer, projection foreshortening occurs along the direction in which the surface tilts, causing an apparent increase in the surface curvature along that direction. This results in an increase in image curvature, i.e., the second derivative of the image intensities. Thus, for umbilical points (where the surface curvature is constant), the direction in which the second derivative of image intensity is greatest turns out to be the tilt of the surface. The slant of the surface can be measured by the amount of increase in image curvature.

The fact that increasing the surface slant results (all else being constant) in increased image curvature suggests that a measure of image curvature might be a good estimator of

[14]These distributions were determined by means of a Monte Carlo simulation.

slant. Image curvature, however, also depends on surface albedo, strength of illumination, surface curvature, and other factors. Still, if we investigate a homogeneous, uniformly lit region of a natural image we find that there is a good correlation between the values of a measure of image curvature such as the Laplacian ($\nabla^2 I$) and the surface slant. If the surface albedo or the illumination changes, however, there will be large changes in the Laplacian that have nothing to do with the surface slant — because the Laplacian values are directly dependent upon the surface albedo ρ and the illuminantion strength λ.

If we divide the Laplacian values by the image intensity values, we can remove the dependence on ρ and λ, thus eliminating two of the most important confounding factors (see Equations (2) and (3)). The division of $\nabla^2 I$ by I also introduces a factor that is dependent upon the illuminant direction; however, this dependency does not seem to affect performance seriously — especially in natural imagery where there is a large amount of diffuse and reflected light. Thus, the division of $\nabla^2 I$ by I yields a measure that depends primarily upon the surface curvature and surface slant. Thus, we are led to the following estimator of surface slant, which is analyzed in the appendix.

Proposition 6. Given the image of an umbilical point on a Lambertian surface and R, the radius of surface curvature, the following is an estimate of z_N, the z component of the surface normal, equal to the arccosine of the surface slant:

$$z_N \approx -R^{-1}\left(\left|\frac{\nabla^2 I}{I}\right| - R^{-2}\right)^{-\frac{1}{2}}$$

This estimate of surface slant turns out to be much more robust than the umbilical-point solution for surface slant, degrading slowly as the surface curvatures become progressively more unequal or as the reflectance function becomes non-Lambertian.

Estimation of R. To use this estimator, the constant R must be determined. A good estimate of the mean R within an image region can be made by applying the constraint that the resulting z_N must satisfy the inequality $0 \geq z_N \geq -1$ — i.e., visible surfaces must be facing the viewer. We can determine a likely value for R by using this constraint and the equation for z_N in light of the range of values of $\nabla^2 I/I$ within a region.

We can then *assume* that the estimated value of R holds throughout the region, and thus obtain an estimate of intraregional slant. If the variance of R is small we will obtain a good estimate of surface orientation. It can happen, however, that the value of R will vary considerably from point to point — unless we can place bounds on the range of R so that its variance is reduced to an acceptable level.

Using the values of $d^2 I$ to identify planar and near-planar regions (as discussed in the previously), we can place a bound on the minimum value of R. We can also place a bound on the largest value of R by blurring image of the region in which slant is to be estimated[15] Such blurring also has the effect of removing highlights, specularities, marks, textures, and the like, thus making the imaged surface more homogeneous and Lambertian. By bounding

[15]Noting that $I \otimes G = \rho\lambda(\mathbf{N}\cdot\mathbf{L}) \otimes G = \rho\lambda((\mathbf{N}\otimes G)\cdot\mathbf{L}) = \rho\lambda(\tilde{\mathbf{N}}\cdot\mathbf{L})$ where G is a two-dimensional Gaussian and $\otimes$ designates convolution, we see that a smoothed version of I may be considered the image of a surface with normal $\tilde{\mathbf{N}} = \mathbf{N}\otimes G = (N_x \otimes G, N_y \otimes G, N_z \otimes G)$, i.e., a smoothed version of the original surface.

R the variance can apparently be reduced to acceptable levels[16] , so that one may expect to obtain a useful estimate of surface shape within a homogeneous area.

Estimating tilt from the slant estimates. The umbilical-point solution for the surface tilt is the direction in which d^2I attains its maximum. To calculate the maximum, therefore, we require either the values of d^2I along many different orientations, or quite accurate values of[17] I_{xx}, I_{yy} and I_{xy}. This is a fair number of image convolutions; besides, biological visual systems seem to manage with only $\nabla^2 I$-like convolutions. It is, therefore, worth inquiring whether there exists a method of determining d^2I along a particular direction from the values of $\nabla^2 I$.

Let us examine the convolution filters required to calculate d^2I and $\nabla^2 I$, the differential quantities used to define the tilt and slant estimates, respectively. We can calculate the second derivative d^2I in the x direction by convolving the image with $d^2G(x,y,\sigma)/dx^2$, where $G(x,y,\sigma)$ is a two-dimensional Gaussian in the variables (x,y) with variance σ. Similarly, we can calculate the Laplacian $\nabla^2 I$ by convolving $\nabla^2 G(x,y,\sigma)$ with the image[18] These two filters are closely related: if we sum $d^2G(x,y,\sigma)/dx^2$ and its 90° rotation, $d^2G(x,y,\sigma)/dy^2$, we obtain $\nabla^2 G(x,y,\sigma)$. We can obtain an approximation to the second-derivative filter $d^2G(x,y,\sigma)/dx^2$ by using a weighted sum of several Laplacians along a straight line in the perpendicular y direction, e.g.,

$$d^2G(x,y,\sigma)/dx^2 \approx \sum_{\epsilon} G(\epsilon,\sigma)\nabla^2 G(x_0, y_0+\epsilon,\sigma)$$

where $G(\epsilon,\sigma)$ is a one-dimensional Gaussian, and $G(x_0,y_0,\sigma)$ designates a Gaussian centered about the point (x_0,y_0). In this manner we can obtain a close approximation to $d^2G(x,y,\sigma)/dx^2$ from $\nabla^2 G(x,y)$ filters (see [15], [16]). Applying this result we see that if we were to sum the quantity $\nabla^2 I/I$ (the input data for the slant estimator) along a straight line, we would obtain an approximation to d^2I/I. This approximation allows us to compute the direction of maximum d^2I from the slant estimation data without additional convolutions; we need only find the orientation along which the sum of $\nabla^2 I/I$ is a maximum.

In practice this approximation to d^2I results in slightly better performance than using the filter that corresponds exactly to d^2I. The difference arises primarily in low-slant regions where the slant estimator (and thus its gradient and this approximation) is more stable than the straightforward tilt estimator.

Evaluation of the surface orientation estimate with an analytic model. To ascertain how well the slant and tilt estimators might be expected to perform under ideal

[16]Although we can reduce the variance of R, we cannot remove systematic bias. Thus, for example, if our viewpoint and the surface shape such that the surface curvature varies inversely with the surface slant (e.g., a parabolic solid viewed point-on), we will obtain a poor estimate of slant. It is worth noting that people also perform poorly under such conditions. Luckily, such arrangements are unusual in natural scenes because surface slant depends on viewpoint, unlike surface curvature; thus, the two are rarely inversely related.

[17]Possession of these three values allows analytic solution for the direction of maximum d^2I; the solution is shown in the appendix.

[18]These convolutions may be regarded as calculating the exact values for a blurred version of the image; or, as mentioned earlier, the blurred image may be regarded as an exact image of a smoothed version of the original scene.

conditions, a computer program was written that used the analytic formulas to calculate the derivatives for images of a wide range of ellipsoidal solids. A Lambertian reflectance function was assumed and a wide range of illumination directions used. Three sets of solid shapes were utilized. The first set consisted only of spheres, so as to test the validity of the program. The second set consisted of ellipsoidal solids with a ratio of principal curvatures which ranged from 2:1 through 1:1 to 1:2. The third set consisted of ellipsoidal solids with a ratio of principal curvatures which ranged from 10:1 through 1:1 to 1:10. Thus, the third set of solids encompassed shapes ranging from almost completely cylindrical to spherical. Points were then sampled evenly from across the entire imaged surface and error statistics computed.

A summary of results for the tilt estimator over the three sets of solids, $\kappa_1 = \kappa_2$ (a sphere), $\kappa_1/\kappa_2 = 2{:}1$ (shapes between elongated eggs and spheres, i.e., common non-cylindrical shapes) and $\kappa_1/\kappa_2 = 10{:}1$ (shapes between cylinders and spheres, i.e., virtually all ratios of curvatures) is shown in Table 1. The direction of surface tilt was computed by using the approximation to the direction of maximum d^2I discussed in the appendix. All error figures are given in radians.

Table 1. Tilt estimator over all surface slants

Ratio Of Curvatures	Error: Bias, Variance	Correlation
$\kappa_1 = \kappa_2$	0.00, 0.052	0.891
$\kappa_1/\kappa_2 = 2{:}1$	0.00, 0.097	0.786
$\kappa_1/\kappa_2 = 10{:}1$	0.00, 0.114	0.742

Although the tilt estimator of Proposition 3 performs perfectly on spheres, the approximation used here shows some small errors. This loss of accuracy is offset by the greater stability that the approximation exhibits in low-slant regions. This table shows that as the range of curvatures increases the performance of the estimator degrades considerably. However, it is only in high-slant regions that errors in the tilt cause serious miscalculations in determining the surface shape; therefore, if the tilt estimator performs well in these regions the resulting shape estimate will still be accurate. Table 2 summarzes the tilt estimator's performance in the critical high-slant regions.

Table 2. Tilt estimator over surface slants greater than 30°

Ratio Of Curvatures	Error: Bias, Variance	Correlation
$\kappa_1 = \kappa_2$	0.00, 0.020	0.950
$\kappa_1/\kappa_2 = 2{:}1$	0.00, 0.066	0.835
$\kappa_1/\kappa_2 = 10{:}1$	0.00, 0.070	0.816

Table 2 describes the performance of the tilt estimator when the surface slant is greater than 30°. It can be seen that the tilt estimates remain quite reasonable for both the 2:1 and the 10:1 range of curvatures. Thus, the tilt estimator makes most of its errors in

low-slant regions, where such errors are relatively unimportant. Furthermore, if the slant estimator provides consistent estimates, we should be able to use it to distinguish between more reliable and less reliable tilt estimates.

Table 3. The slant estimator

Ratio Of Curvatures	Error: Bias, Variance	Correlation
$\kappa_1 = \kappa_2$	0.201, 0.012	0.924
$\kappa_1/\kappa_2 = 2{:}1$	0.182, 0.034	0.843
$\kappa_1/\kappa_2 = 10{:}1$	0.137, 0.101	0.674

Table 3 shows that the slant estimator, although biased, performs quite well on spheres. As the range of curvatures increases, the performance of the estimator degrades — but, even for the 10:1 range of curvatures, it is still good. For all these cases one estimate of R was used; therefore, in all except the case of spheres the estimated R is actually in error — in some instances by a factor of 10. This seems to indicate that the slant estimator is remarkably robust.

Table 4. The unbiased slant estimator

Ratio Of Curvatures	Error: Bias, Variance	Correlation
$\kappa_1 = \kappa_2$	0.00, 0.004	0.967
$\kappa_1/\kappa_2 = 2{:}1$	0.00, 0.006	0.948
$\kappa_1/\kappa_2 = 10{:}1$	0.00, 0.028	0.796

The bias of the slant estimator that appears in Table 3 also shows up in the equations discussed in the appendix; where it is also explained how the bias may be removed. When the slant estimator is made unbiased, its accuracy becomes even better, as shown in Table 4. It turns out that the performance of the slant estimator is approximately as good for regions of low slant as for those of higher slant. Therefore, the slant estimate produced by this estimator is useful in assessing the tilt estimator's reliability.

Evaluation on natural images. The surface orientation estimator (the "shape algorithm") has been tested on several natural images, and four such examples will be presented here. The shape algorithm produces estimates of the surface orientation; it was found, however, that displays of the estimated surface orientation do not allow an observer to evaluate the performance of the algorithm adequately. Therefore, for purposes of exhibiting the performance of the algorithm, the shape algorithm's estimates of surface orientation were integrated to produce a relief map of the surface. As these relief maps were found to give observers an adequate impression of the estimated surface shape, they constitute the output shown for the examples presented in this paper even though integration is not part of the shape algorithm *per se*.

Figure 7 (a) shows the image of a log, together with the relief map generated from the shape algorithm's estimates of surface orientation. Figure 7 (b) shows the image of a

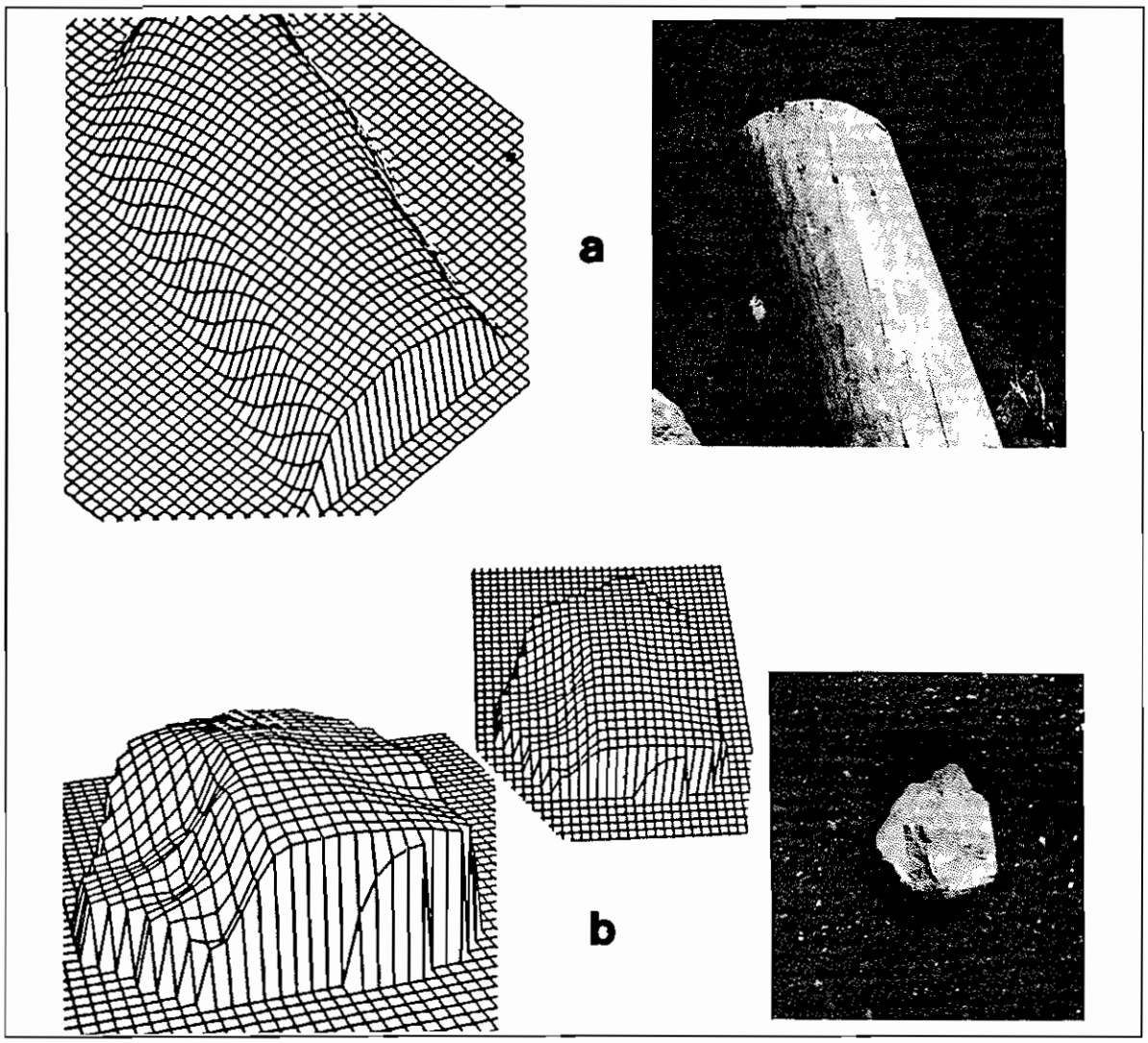

Figure 7. Evaluation on two natural images. (a) An image of a log and the relief map generated by the shape algorithm for that image. (b) An image of a rock and the relief map generated by the shape algorithm for that image.

rock, together with the relief map generated from the shape algorithm's estimates of surface orientation. The relief maps in Figure 7 (a) and 7 (b) correspond closely to the actual shapes of these two objects. The reader should compare his impression of shape from the images with the relief maps of Figure 7 (a) and 7 (b).

Figure 8 shows (a) the digitized picture of a small portion of a face (belonging to a woman named Lisa), and (b) a relief map of the surface slant estimated for that image (eye and eyebrow regions were masked out by hand). No relief map of the estimated surface shape is shown because the complexity of the shape made it difficult to integrate the slant and tilt estimates. In this slant-map representation regions with higher relief face toward the viewer, while lower relief regions face away from him. Note that many important details of the surface shape are apparent in this representation; for instance, the structure of the nose, the cheeks and the eyebrow ridges is plainly visible.

Figure 9 shows (a) the digitized image of Tuckerman's ravine (a skiing region on Mt.

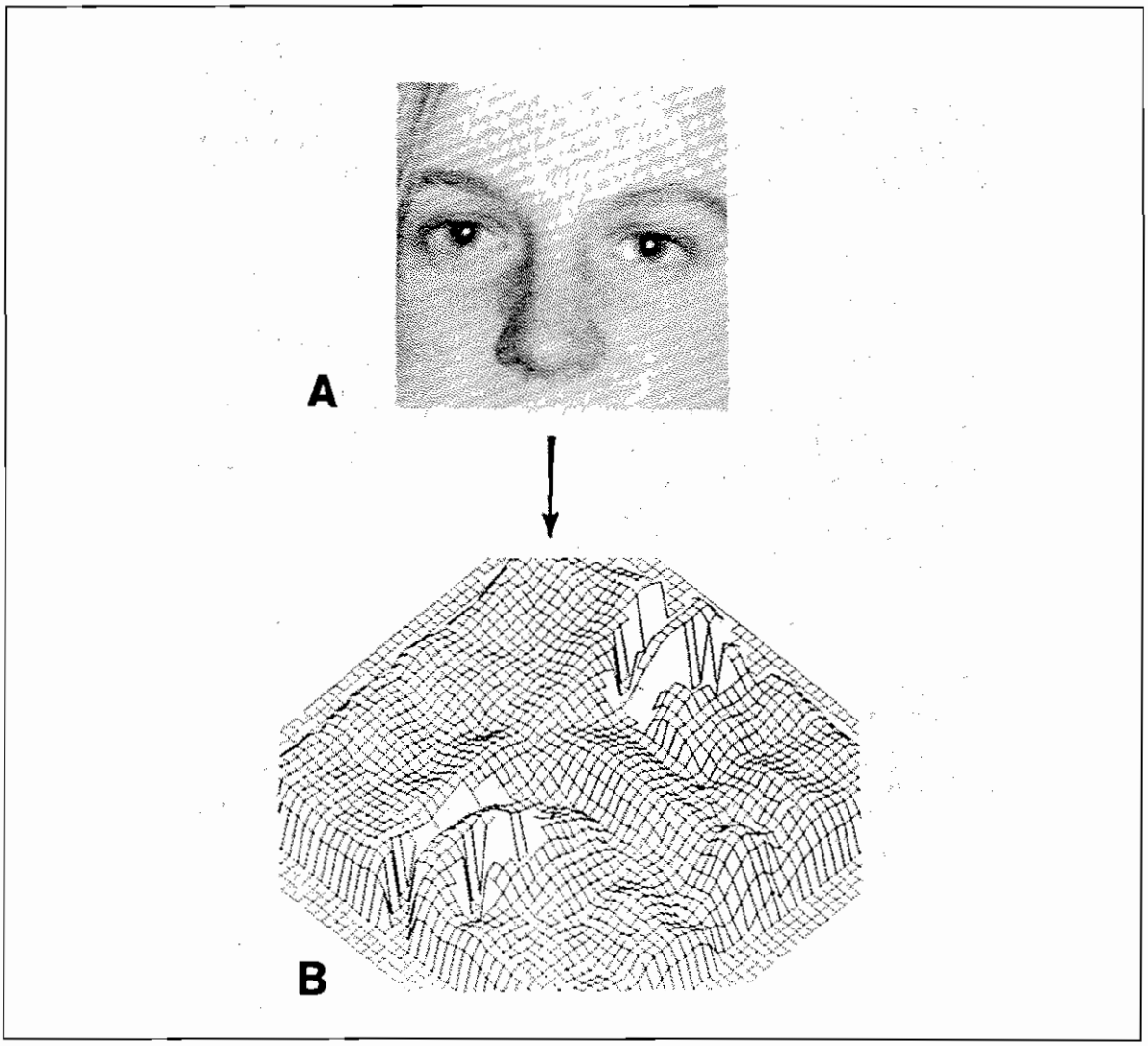

Figure 8. The Lisa Image. (a) The digitized image of Lisa and (b) a relief map showing the estimated surface slant in this image. High-relief areas in this representation correspond to regions facing the viewer, low-relief regions to regions slanting away from the viewer. Eye regions and eyebrow regions were masked out by hand. The integration of estimated slant and estimated tilt to show surface shape proved difficult because of the complexity of the surface. Note that many important details of the surface shape are apparent in this representation; for instance, the structure of the nose, cheeks and eyebrow ridges is plainly visible.

Washington, in New Hampshire), (b) a relief map showing a side view of the estimated surface shape, obtained by integrating the slant and tilt estimates. This relief map may be compared directly with a topographic map of the area; when we compare the estimated and actual shape, we find that the roll-off at the top of Figure 9 (b) and the steepness of the estimated surface are correct for this surface[19] ; this area of the ravine has a slope that averages 60°.

The comparison with a topographic map also shows that the relief of the lower right-hand portion of the image is somewhat underestimated. When people are asked to look at this image, however, it becomes clear that they also fail to perceive the shape of the ravine

[19]And people ski down this incline!

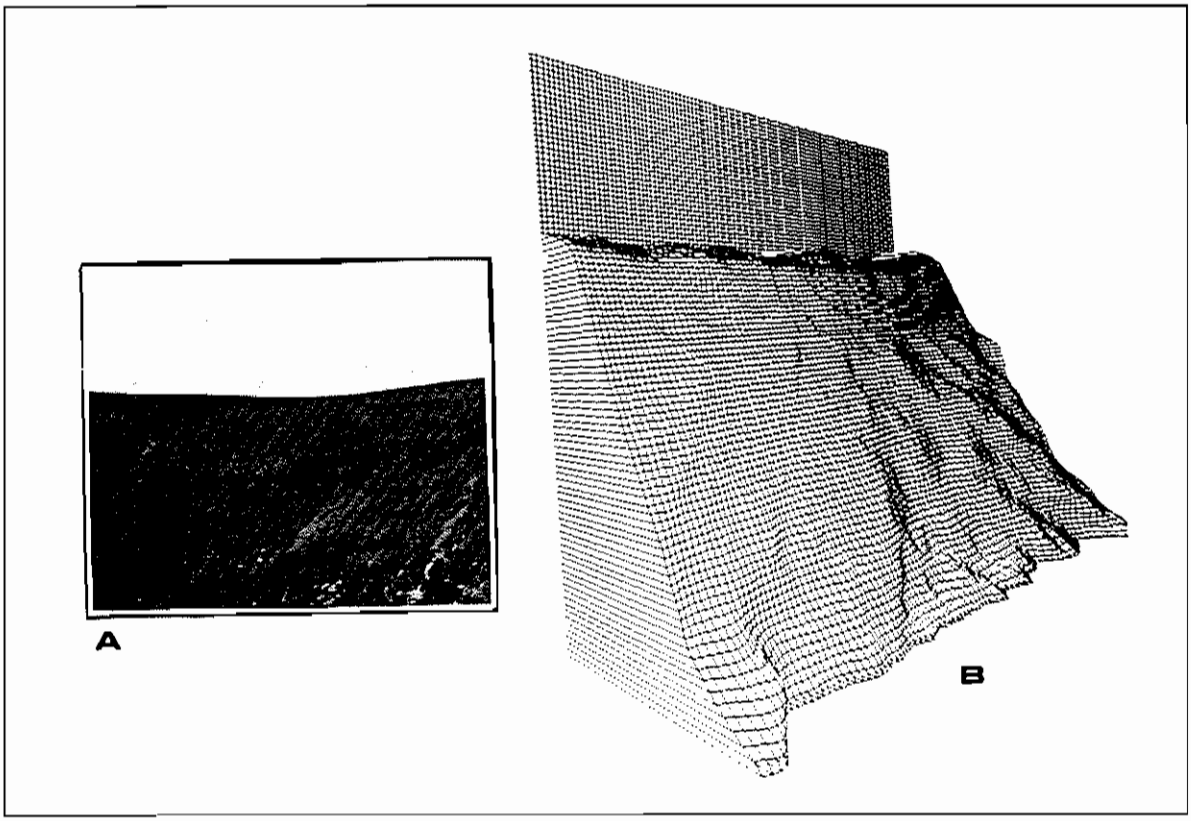

Figure 9. Tuckerman's Ravine. (a) The digitized image of Tuckerman's ravine and (b) a relief map showing a side view of the surface estimated for this image. Comparing a topographic map of the area with the estimated surface shape, we find that the roll-off at the top of (b) and the steepness of the estimated surface are correct for this surface. However, the relief of the lower right-hand portion of the image is somewhat underestimated. The underestimation of relief is similar to human perception of this image.

correctly: they also underestimate the relief of the lower right portion of this image[20] .

Evaluation On An Electron Microscope image. In addition to natural images, the electron microscope (EM) image shown in Figure 10 (a) was selected from the book *Magnifications* by D. Scharf [6]. People can use the shading information in EM images to perceive shape, as Figure 10 (a) confirms. This is surprising because these images have a reflectance function not found in natural scenes. This image, therefore, provides a critical test of the similarity between the human use of shading and this estimator of surface orientation.

Ikeuchi and Horn [3] measured the reflectance function for this image and found that the image intensities may be reasonably well described by

$$I = k(\mathbf{N} \cdot \mathbf{V})^{-1}$$

where k is approximately 0.8. If we carry out the required computations, we see that the tilt is still the direction along which d^2I is greatest, and that the z-component of the surface normal is approximately proportional to $\nabla^2 I/I$, as in normal images. We can thus expect to

[20]However, when people are able to view the original higher-resolution image or the entire image they perceive the surface correctly.

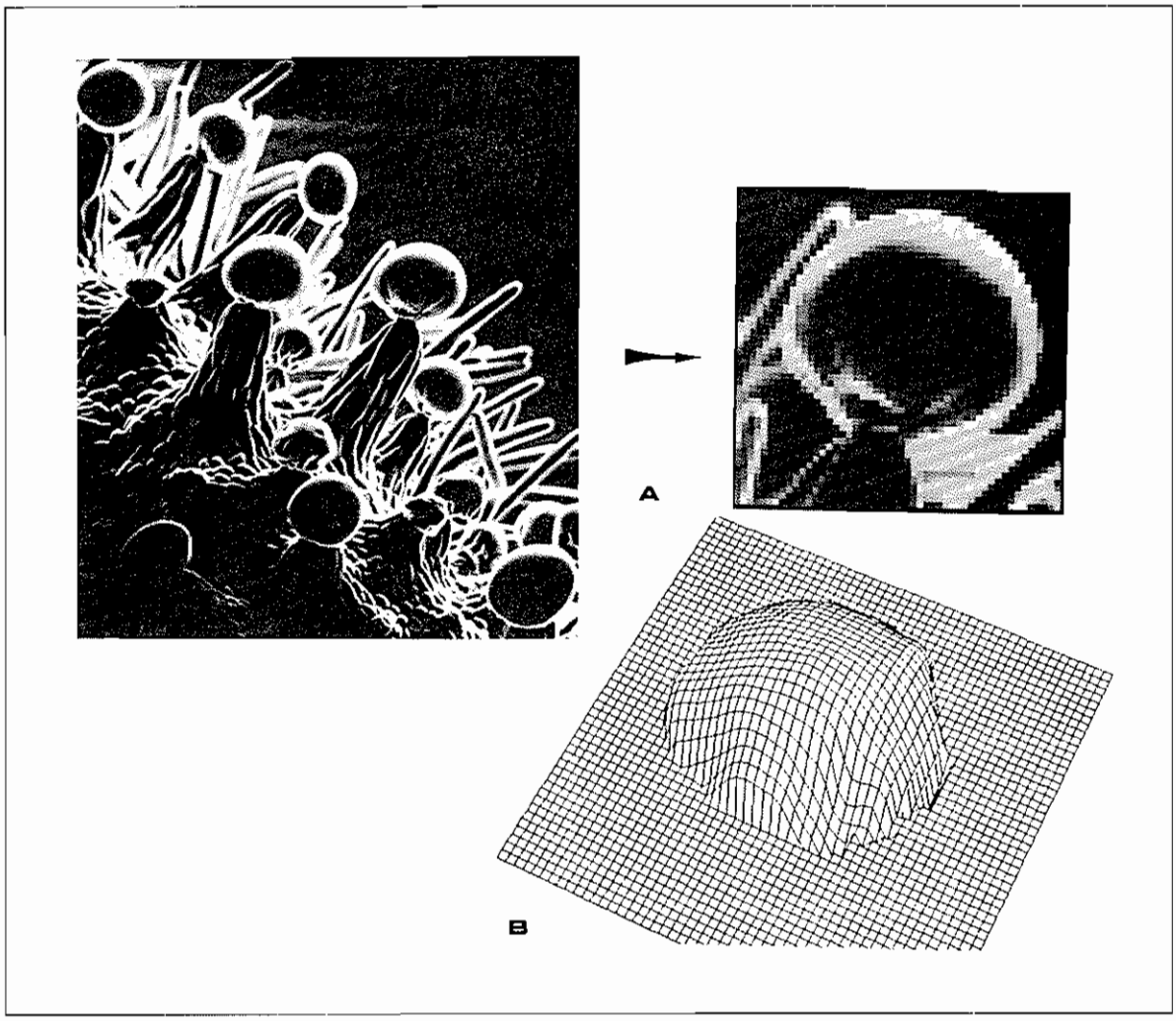

Figure 10. An electron microscope image. Part (a) is an electron microscope image of resin nodules on a flower of *cannabis sativa*, and the portion of the image for which shape was estimated. Part (b) shows a relief map showing a side view of the surface estimated for this image. The fact that both people and this algorithm can correctly use the shading information in electron microscope images may have important implications for understanding human vision.

obtain a reasonable shape estimate for EM images by using the same estimation technique developed for normal images.

Figure 10 (b) the portion of 10 (a) for which shape was estimated. Figure 10 (c) is a relief map showing a side view of the estimated surface shape, again obtained by integrating the slant and tilt estimates for this image. It can be seen that the estimated surface shape is quite accurate.

IV. Discussion

The preceding portions of this paper have shown that it is possible to obtain useful estimates of scene properties from natural images by using a local analysis of image shading. The analysis does *not* assume that any scene information is known beforehand; no knowledge

of scene characteristics or boundary conditions is used. Thus, the techniques are applicable to "raw" images. Because no *a priori* information is assumed, however, the recovery of information about the scene is necessarily imprecise.

One major problem which is inherent in obtaining shape from shading under general viewing conditions is that the measured image intensities are not equal to the image irradiance. Film, video camera, and other image transcription methods produce image intensity measurements which are (generally) non-linear transformations of the image irradiance. It is, therefore, surprising to note that humans have no problem in utilizing the shading in such transformed images even though the relationship between the transformed image and the original image irradiance is unknown. Thus, any shape-from-shading technique which will be as generally useful as the human capacity must function *despite* such transformations of the data.

The shape-from-shading techniques described in this paper are relatively unaffected by smooth, monotonic transformations of the image data — in marked contrast to previous methods of inferring shape from shading. This robustness is achived by dividing the Laplacian of the image intensities by the intensities themselves, thus removing the primary effects of any multiplicative terms in the image irradiance equation. The division also removes the effects of any linear scaling of the image intensity. Thus, division of the Laplacian by the intensity compensates for any transformation of the image irradiance which is locally approximately linear.

What is the use of a reasonably accurate, but certainly not infallible, local estimate of scene properties? Several potential applications spring to mind: to provide an initial "guess" for a more global shading analysis [20], to constrain stereo matching by providing a qualitative estimate of shape, or to help in the estimation of albedo. One other use that I have begun examining is classification of the type of imaged contours. This serves as a good demonstration of the potential usefulness of a reasonably accurate local estimate of surface orientation.

Once we are given the location of a contour we should be able to use our local estimate of surface orientation to acertain whether a contour is a smooth occluding contour (i.e., a contour formed by the surface curving smoothly out of sight, such as is found at the edge of an image of a sphere) by checking whether our estimates of surface orientation are appropriate for that contour. If the surface adjoining one side of a contour has a large slant and a tilt perpendicular to the contour, then it is likely a smooth occluding contour. On the other hand, if the estimated slant is small, or if the tilt is not perpendicular, then it is quite probably not a smooth occluding contour.

Figure 11 shows the results of applying this typing strategy to the contours extracted from two natural images. Part (a) of this figure shows the Moore sculpture image, and part (d) the Tuckerman's ravine image. Parts (b) and (e) depict the discontinuity contours found in these images. Parts (c) and (f) of this figure show the contours that were adjoined by regions whose estimated surface orientation was consistent with the contours' being a smooth occluding contour. When we compare the contours identified as smoothly occluding with the original images, we find that this criterion is quite apt in identifying the smooth occluding contours in these images.

Biological Vision Systems. The shape-from-local-shading theory presented in this paper

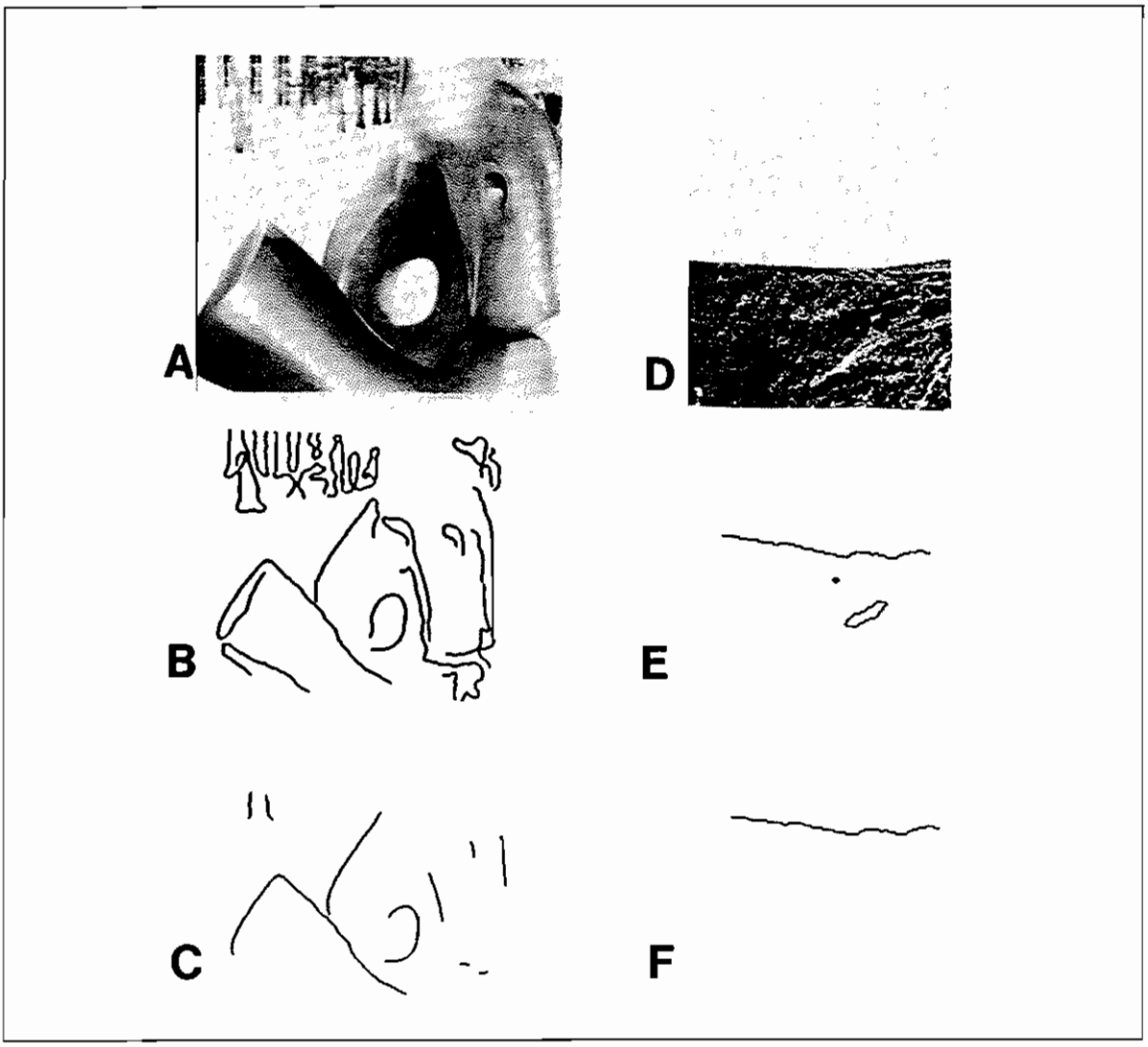

Figure 11. Results of typing the contours extracted from two natural images. Part (a) of this figure shows the Moore sculpture image, and part (d) shows the Tuckerman's ravine image. Parts (b) and (e) show the discontinuity contours found in these images. Parts (c) and (f) show the contours which were adjoined by regions whose estimated surface orientation was consistent with the contour being a smooth occluding contour. When we compare these contours to the original images, we find that this criterion does quite well at identifying the smooth occluding contours in these images.

seems to have considerable utility as a model of one aspect of the functioning of biological visual systems. It is known from studies of neurophysiology [8], [9] and human psychophysics [10], [11] that the retinal receptive fields in mammals have a center-surround organization that is well modeled by the filter $\nabla^2 G(x, y, \sigma)$. In addition, the responses of these retinal neurons are logarithmically scaled by the intensity, so that their response r to an image point $I(x_0, y_0)$ can reasonably be modeled by the following convolution and division:

$$r = \frac{\nabla^2 G(x, y, \sigma) \otimes I(x_0, y_0)}{I(x_0, y_0)} \tag{5}$$

The quantity r is the measurement needed by the slant estimator. Thus, it seems that the information required to estimate slant is present in the output of the mammalian retina.

It is also well established [12], [13], [14] that many (perhaps most) cortical neurons in the primary visual cortex have oriented receptive fields whose responses characteristics are closely modeled by the filter $d^2G(x,y,\sigma)/dx^2$. Moreover, there is very strong evidence [15], [16] that these cortical neurons are constructed by summing the center-surround receptive fields described by Equation (5), just as was done here. Thus, not only does it seem that the image data required by the tilt estimator is present in the mammalian primary visual cortex, but the information is apparently derived from the image by means of the same steps that have been employed here.

Acknowledgments. I would like to thank my fellow graduate students, including Dr. Andy Witkin, Donald Hoffman, Joseph Scheuhammer and Dr. Eric Grimson, who helped in the refinement and development of this work.

V. Appendix

In this section the previous five propositions will be proved. The strategy of proof will be to start with the six values for I, I_x, I_y, I_{xx}, I_{yy} and I_{xy} and then solve for surface orientation, curvature, illuminant direction and albedo times illuminant intensity under the assumption that the point is an umbilical point on a Lambertian surface. This solution will, at the same time, prove Propositions 1, 2, and 4. Some additional calculations using the results of this solution will then prove Propositions 3, 5 and 6.

A. Solution For Umbilical Points

Consider the surface of a sphere of radius R:

$$Z(x,y) = \sqrt{R^2 - x^2 - y^2}$$

This equation, with $R > 0$ $R \geq x \geq -R$ and $R \geq y \geq -R$, describes the set of all umbilical points. From this equation we see that $Z_x = -xT^{-1/2}$ and $Z_y = -yT^{-1/2}$ where $T = R^2 - x^2 - y^2$. Assume that the illuminant is unknown, so that we must consider all illumination directions $\mathbf{L} = (l_1, l_2, l_3)$. Then, if the surface is Lambertian, we have

$$I(x,y) = \rho\lambda\mathbf{N}\cdot\mathbf{L} = \frac{\rho\lambda(Z_x, Z_y, -1)\cdot(l_1,l_2,l_3)}{\sqrt{Z_x^2 + Z_y^2 + 1}} = \frac{\rho\lambda}{R}(-xl_1 - yl_2 - l_3T^{1/2}) \qquad (1)$$

where ρ is the surface albedo and λ is the illuminant intensity at the surface. The first and second derivatives are then

$$I_x = \frac{\rho\lambda}{R}(-l_1 + xl_3T^{-1/2}) \qquad (2)$$

$$I_y = \frac{\rho\lambda}{R}(-l_2 + yl_3T^{-1/2}) \qquad (3)$$

$$I_{xx} = \frac{\rho\lambda}{R}(l_3T^{-1/2} + x^2l_3T^{-3/2}) \qquad (4)$$

$$I_{yy} = \frac{\rho\lambda}{R}(l_3T^{-1/2} + y^2l_3T^{-3/2}) \qquad (5)$$

$$I_{xy} = \frac{\rho\lambda}{R}(xyl_3T^{-3/2}) \qquad (6)$$

Assume that the values of I, I_x, I_y, I_{xx}, I_{yy} and I_{xy} are known; we may now solve for surface orientation, curvature, iluminant direction and alhedo times illuminant intensity.

Solution

Using Equation (6) to solve for $\frac{\rho\lambda}{R}$ we obtain

$$\frac{\rho\lambda}{R} = \frac{I_{xy}T^{3/2}}{xyl_3} \qquad (7)$$

Using Equations (7) and (4) we obtain

$$I_{xx} = (l_3T^{-1/2} + x^2l_3T^{-3/2})\frac{\rho\lambda}{R} = (l_3T^{-1/2} + x^2l_3T^{-3/2})\frac{I_{xy}T^{3/2}}{xyl_3} = (\frac{T + x^2}{xy})I_{xy} \qquad (8)$$

and using Equation (7) with Equation (5) we obtain

$$I_{yy} = (l_3 T^{-1/2} + y^2 l_3 T^{-3/2})\frac{\rho\lambda}{R} = (l_3 T^{-1/2} + y^2 l_3 T^{-3/2})\frac{I_{xy} T^{3/2}}{xy l_3} = (\frac{T + y^2}{xy})I_{xy} \quad (9)$$

Using Equations (8) and (9) we see that

$$T = \frac{xy I_{xx}}{I_{xy}} - x^2 = \frac{xy I_{yy}}{I_{xy}} - y^2 \quad (10)$$

Using Equations (9) and (10) we see that

$$\begin{aligned} I_{yy} &= (\frac{T + y^2}{xy})I_{xy} \\ &= (\frac{\frac{xyI_{xx}}{I_{xy}} - x^2 + y^2}{xy})I_{xy} \\ &= I_{xx} - \frac{x}{y}I_{xy} + \frac{y}{x}I_{xy} \end{aligned} \quad (11)$$

Letting $k = \frac{y}{x}$ we see that Equation (11) is a quadratic in k, which we can solve to obtain

$$k = \frac{-(I_{xx} - I_{yy}) \pm \sqrt{(I_{xx} - I_{yy})^2 + 4I_{xy}^2}}{2I_{xy}} \quad (12)$$

From this we may obtain the surface tilt $\tau = \tan^{-1} k$ Thus the tilt of the sphere's surface may be determined without knowledge of the illuminant direction, the illuminant strength, the surface albedo or the surface curvature. To prove Proposition 3, which stated that the tilt of the surface is in the direction of maximum d^2I, it remains only to show that this solution for the surface tilt is the direction of maximum d^2I. The remainder of the proof will be presented in the following subsection.

Is this solution unique? Equation (12) yields two possible solutions:

$$k_1 = \frac{y_1}{x_1} \qquad k_2 = \frac{y_2}{x_2}$$

Note that we may also solve Equation (12) for k^{-1}; this yields

$$k^{-1} = \frac{-(I_{xx} - I_{yy}) \pm \sqrt{(I_{xx} - I_{yy})^2 + 4I_{xy}^2}}{-2I_{xy}} \quad (13)$$

Equation (13) also gives two solutions, k_3 and k_4,

$$k_3 = k_1^{-1} = \frac{x_1}{y_1} \qquad k_4 = k_2^{-1} = \frac{x_2}{y_2}$$

As the left-hand side of Equation (12) is the negative of the left-hand side of Equation (13) we find that either $k_3 = -k_1^{-1}$, which leads to a contradiction, or $k_3 = -k_2^{-1}$. Thus

$$k_1 = \frac{y_1}{x_1} = -k_2^{-1} = -\frac{x_2}{y_2}$$

This proves that the two solutions of Equation (12) are perpendicular. We shall see later that this allows us to discard one of the solutions, because it results in an illuminant direction that is behind the observed object.

Once the tilt τ is known, the only remaining component of surface orientation is σ, the slant of the surface, which is equal to the arccosine of the z component of the surface normal, z_N. Noting that

$$z_N = \frac{-1}{\sqrt{Z_x^2 + Z_y^2 + 1}} = \frac{-\sqrt{R^2 - x^2 - y^2}}{R}$$

we find that we can use Equation (10) to determine the surface slant. We may average the two expressions for T in Equation (10) to obtain

$$T = R^2 - x^2 - y^2 = \frac{xy(I_{xx} + I_{yy})}{2I_{xy}} - \frac{(x^2 + y^2)}{2} = \frac{xy\nabla^2 I}{2I_{xy}} - \frac{(x^2 + y^2)}{2} \tag{14}$$

If we add $x^2 + y^2$ to both sides of Equation (14) we then obtain

$$R^2 = \frac{xy\nabla^2 I}{2I_{xy}} + \frac{(x^2 + y^2)}{2} \tag{15}$$

Thus, from Equations (14) and (15):

$$\begin{aligned} z_N^2 &= \frac{R^2 - x^2 - y^2}{R^2} \\ &= \frac{\frac{xy\nabla^2 I}{2I_{xy}} - \frac{(x^2+y^2)}{2}}{\frac{xy\nabla^2 I}{2I_{xy}} + \frac{(x^2+y^2)}{2}} \\ &= \frac{xy\nabla^2 I - (x^2 + y^2)I_{xy}}{xy\nabla^2 I + (x^2 + y^2)I_{xy}} \\ &= \frac{k\nabla^2 I - (k^2 + 1)I_{xy}}{k\nabla^2 I + (k^2 + 1)I_{xy}} \end{aligned} \tag{16}$$

From z_N^2 we can obtain the surface slant $\sigma = \cos^{-1}\sqrt{z_N^2}$ This concludes the proof of Proposition 4. Note that there is only one solution for the surface slant, as $0 \geq z_N \geq -1$. With both the slant and tilt determined, we now know the surface orientation. Once again, this has been accomplished without prior knowledge of illuminant direction, illuminant strength, surface curvature, or surface albedo.

We may now proceed to solve for the remaining unknowns. Because we know the surface orientation, we can compute $\chi = x/R$ and $\gamma = y/R$.

$$\chi = \cos\tau\sin\sigma \qquad \gamma = \sin\tau\sin\sigma$$

The quantities χ and γ may be thought of as the x and y coordinates normalized to the unit sphere. Using χ and γ, we may define Γ, a unit sphere analogue to T

$$\Gamma = 1 - \chi^2 - \gamma^2 = \frac{T}{R^2}$$

Going back to Equation (7) we now know all of the terms on the left-hand side except l_3, thus, let us write

$$\frac{\rho\lambda l_3}{R} = \frac{I_{xy}T^{3/2}}{xy} = \frac{I_{xy}\Gamma^{3/2}R}{\chi\gamma}$$

and so

$$\rho\lambda l_3 = \frac{I_{xy}\Gamma^{3/2}R^2}{\chi\gamma} \tag{17}$$

We may then subsitute this equation into Equations (2) and (3) to give

$$\rho\lambda l_1 = \frac{I_{xy}\Gamma R^2}{\gamma} - I_x R \tag{18}$$

and

$$\rho\lambda l_2 = \frac{I_{xy}\Gamma R^2}{\chi} - I_y R \tag{19}$$

If we now convert Equation (1) to the variables χ, γ and Γ, we may substitute Equations (17), (18) and (19) into (1) to obtain a quadratic in R:

$$\begin{aligned} I &= \frac{\rho\lambda}{R}(-xl_1 - yl_2 - T^{1/2}l_3) \\ &= \rho\lambda(-\chi l_1 - \gamma l_2 - \Gamma^{1/2}l_3) \\ &= -\chi\left(\frac{I_{xy}\Gamma R^2}{\gamma} - I_x R\right) - \gamma\left(\frac{I_{xy}\Gamma R^2}{\chi} - I_y R\right) - \frac{I_{xy}\Gamma^2 R^2}{\chi\gamma} \\ &= (\chi I_x + \gamma I_y)R - \frac{I_{xy}\Gamma R^2}{\chi\gamma} \end{aligned} \tag{20}$$

We may solve this for R to obtain

$$R = \frac{\chi\gamma\left((\chi I_x + \gamma I_y) \pm \sqrt{(\chi I_x + \gamma I_y)^2 - \frac{4II_{xy}\Gamma R^2}{\chi\gamma}}\right)}{2I_{xy}\Gamma R^2} \tag{21}$$

Because **L** is a unit vector, we may now use Equations (17), (18) and (19) to determine $(\rho\lambda)^2$:

$$\begin{aligned} (\rho\lambda)^2 &= (\rho\lambda l_1)^2 + (\rho\lambda l_2)^2 + (\rho\lambda l_3)^2 \\ &= \left(\frac{I_{xy}\Gamma R^3}{\gamma} - I_x R\right)^2 + \left(\frac{I_{xy}\Gamma R^3}{\chi} - I_y R\right)^2 + \left(\frac{I_{xy}\Gamma^{3/2}R^2}{\chi\gamma}\right)^2 \end{aligned} \tag{22}$$

As $\rho\lambda \geq 0$, we may discard the negative root of $\sqrt{\rho\lambda}$, so that $\rho\lambda$ is uniquely determined. Using this value of $\rho\lambda$, we may now substitute into Equations (17), (18) and (19) to obtain l_1, l_2 and l_3.

Note that the signs of l_1, l_2 and l_3 depend on the signs of χ and γ. Thus, in solving Equation (12) one of the two solutions will make $\chi\gamma$ negative making l_3 negative — which corresponds to an illuminant behind the observed surface. Therefore, only one of the two solutions of (12) is physically possible.

Because Equation (12) gives only the solution for y/x, there are two possible pairs (x_1, y_1), (x_2, y_2) such that $k = y_1/x_1 = y_2/x_2$. These two solutions are each other's negative, i.e., $x_1 = -x_2$, $y_1 = -y_2$; consequently, the surface tilts for these two solutions are 180° apart so that one corresponds to a convex surface, the other to a concave surface. Because it is the signs of χ and γ that determine the signs of l_1 and l_2, choosing one of the (x, y) pairs results in an illuminant direction that is overhead (i.e., $\mathbf{L} \cdot (0, 1, 0) \geq 0$), while picking the other results in an illuminant direction that is below the viewing line. Thus, if we specify an illuminant direction which must be overhead and in front of the illuminated object, there is only one possible solution to Equation (12). The symmetry between the signs of x, y and l_1, l_2 is commonly familiar as the crater illusion, in which the convexity of the surface changes as the perceived illuminant direction shifts from overhead to below the viewing line.

We have now solved for each of the unknown quantities, and, by so doing have shown that to each set of measurements I, I_x, I_y, I_{xx}, I_{yy} and I_{xy} there corresponds exactly one combination of surface orientation, curvature, (overhead) illuminant direction, and factor $\rho\lambda$ for a Lambertian umbilical point. This concludes the proofs of Propositions 1 and 2.

B. Proposition 3

In the preceding subsection it was shown that

$$r = \tan^{-1} k = \tan^{-1} \frac{y}{x}$$

where

$$k = \frac{-(I_{xx} - I_{yy}) \pm \sqrt{(I_{xx} - I_{yy})^2 + 4I_{xy}^2}}{2I_{xy}} \tag{12}$$

It remains to show that this solution is equivalent to the proposition that the tilt is the image direction in which d^2I is greatest.

We know that, given I_{xx}, I_{yy} and I_{xy} we may obtain these quantitites in any other image plane coordinate system (x^*, y^*) that is a rotation of (x, y) by the angle ξ. First we note that

$$I_{x^*} = I_x \frac{dx}{dx^*} + I_y \frac{dy}{dx^*} \qquad I_{y^*} = I_x \frac{dx}{dy^*} + I_y \frac{dy}{dy^*}$$

The standard rotation transformation is

$$x^* = xc_\xi + ys_\xi$$
$$y^* = -xs_\xi + yc_\xi$$

where s_ξ and c_ξ are the sine and cosine of the angle ξ. The inverse of this rotation transformation is

$$x = x^* c_\xi - y^* s_\xi$$
$$y = x^* s_\xi + y^* c_\xi$$

Thus,

$$\frac{dx}{dx^*} = c_\xi \qquad \frac{dy}{dx^*} = s_\xi \qquad \frac{dx}{dy^*} = -s_\xi \qquad \frac{dy}{dy^*} = c_\xi$$

and so

$$I_x^* = I_x c_\xi + I_y s_\xi$$
$$I_y^* = -I_x s_\xi + I_y c_\xi$$

Similarly,

$$I_{x^*x^*} = \frac{d(I_{x^*})}{dx}\frac{dx}{dx^*} + \frac{d(I_{x^*})}{dy}\frac{dy}{dx^*}$$
$$I_{y^*y^*} = \frac{d(I_{y^*})}{dx}\frac{dx}{dy^*} + \frac{d(I_{y^*})}{dy}\frac{dy}{dy^*}$$
$$I_{x^*y^*} = \frac{d(I_{x^*})}{dx}\frac{dx}{dy^*} + \frac{d(I_{x^*})}{dy}\frac{dy}{dy^*}$$

resulting in

$$I_{x^*x^*} = I_{xx}c_\xi^2 + I_{yy}s_\xi^2 + 2I_{xy}s_\xi c_\xi$$
$$I_{y^*y^*} = I_{xx}s_\xi^2 + I_{yy}c_\xi^2 - 2I_{xy}s_\xi c_\xi$$
$$I_{x^*y^*} = -I_{xx}s_\xi c_\xi + I_{yy}s_\xi c_\xi + I_{xy}(c_\xi^2 - s_\xi^2)$$

To find the direction for which d^2I attains its maximum, we find the angle ξ for which $I_{x^*x^*}$ attains a maximum over all rotations of the image plane coordinate system. As $I_{x^*x^*}$ is equal to

$$I_{x^*x^*} = I_{xx}c_\xi^2 + I_{yy}s_\xi^2 + 2I_{xy}s_\xi c_\xi$$

the maximum of $I_{x^*x^*}$ occurs at

$$\begin{aligned} 0 &= \frac{d(I_{x^*x^*})}{d\xi} \\ &= (I_{yy} - I_{xx})2s_\xi c_\xi + 2I_{xy}(c_\xi^2 - s_\xi^2) \\ &= (I_{yy} - I_{xx})s_{2\xi} + 2I_{xy}c_{2\xi} \end{aligned}$$

which was obtained by using the relations $\sin 2\xi = 2\sin\xi\cos\xi$ and $\cos 2\xi = \cos^2\xi - \sin^2\xi$. Solving this for ξ we see that the angle ξ for which d^2I attains its maximum satisfies

$$\tan 2\xi = \frac{2I_{xy}}{I_{xx} - I_{yy}}$$

Using Equations (4), (5) and (6), we see that for a sphere

$$\tan 2\xi = \frac{2I_{xy}}{I_{xx} - I_{yy}} = \frac{2\frac{\rho\lambda}{R}(xyI_3T^{-3/2})}{\frac{\rho\lambda}{R}(x^2I_3T^{-3/2} - y^2I_3T^{-3/2})} = \frac{2xy}{x^2 - y^2}$$

However, noting that for a sphere $\tan\tau = y/x$ and that $\tan 2\tau = \frac{2\tan\tau}{1-\tan^2\tau}$ we have

$$\tan 2\tau = \frac{2\tan\tau}{1 - \tan^2\tau} = \frac{2xy}{x^2 - y^2}$$

Thus, $\xi = \tau$ and so Proposition 3 is proved.

C. Proposition 5

Let us assume that we have observed that $d^2I = d^2\mathbf{N}\cdot\mathbf{L} = 0$ in the direction (dx, dy), and that this situation continues for some distance along (dx, dy). In this situation, either (1) $d^2\mathbf{N}$ is perpendicular to $\mathbf{L}$ or (2) the magnitude of $d^2\mathbf{N}$ is zero. It is unlikely that $d^2\mathbf{N}$ is perpendicular to the illuminant over any distance; thus, if we see that $d^2I = 0$ it must be the case that the magnitude of $d^2\mathbf{N}$ is zero. Therefore, $d\mathbf{N}$ is some constant vector as we take a step along (dx, dy).

If the magnitude of $d\mathbf{N}$ along (dx, dy) is zero, then at least one of the surface curvatures is zero, i.e., the surface is cylindrical or planar. If the magnitude of $d\mathbf{N}$ is not zero, then, when we take a step in the direction (dx, dy) there is some change in surface orientation. In this case, either the amount of forshortening that occurs with each infinitesimal step along (dx, dy) will change or $\mathbf{N}$ will be constant, contrary to the assumption that $d\mathbf{N}$ was not zero. If there is change in the foreshortening and yet $d\mathbf{N}$ remains constant, the surface orientation relative to the viewer and the intrinsic surface curvature are in an exactly reciprocally relation, which is a violation of general position.

Thus, when $d^2I = 0$ for some distance along a direction (dx, dy), the surface is either cylindric or planar. If it is planar, $d^2I = 0$ in all directions. The converse — that if we also observe $d^2I = 0$ along the direction orthogonal to (dx, dy) the surface is then planar — is also true. From the previous reasoning, $d^2I = 0$ implies that $d\mathbf{N}$ is zero along that direction (if general position is assumed). If $d\mathbf{N}$ is zero in two directions, the surface is planar.

D. Proposition 6

Proposition 6 suggests that the following equation is a useful estimator of z_N, the z component of the surface normal, equal to the arccosine of the surface slant:

$$z_N = -c\left(\left|\frac{\nabla^2 I}{I}\right| - c^2\right)^{-1/2}$$

where c is a constant related to the surface curvature.

We may examine this estimator in the context of the calculations presented so far. First we use Equations (4) and (5) to find that

$$\nabla^2 I = I_{xx} + I_{yy} = \frac{\rho\lambda}{R}(2l_3T^{-1/2} + (x^2+y^2)l_3T^{-3/2})$$

Thus,

$$\frac{\nabla^2 I}{I} = \frac{\frac{\rho\lambda}{R}(2l_3T^{-1/2} + (x^2+y^2)l_3T^{-3/2})}{\frac{\rho\lambda}{R}(-xl_1 - yl_2 - l_3T^{1/2})} \tag{23}$$

If we assume that $-xl_1 - yl_2$ is zero (as is true on the average, although not neccessarily true for any one image point) Equation (23) becomes

$$\begin{aligned}\left|\frac{\nabla^2 I}{I}\right| &= 2T^{-1} + (x^2+y^2)T^{-2} \\ &= \frac{2R^2 - x^2 - y^2}{(R^2 - x^2 - y^2)^2} \\ &= \frac{R^2 + T}{T^2}\end{aligned} \tag{24}$$

To show how this estimator might work, let us first take $c = T^{-1/2} = (R^2 - x^2 - y^2)^{-1/2}$; we then find that, in fact,

$$\begin{aligned} -c(|\frac{\nabla^2 I}{I}| - c^2)^{-1/2} &= -T^{-1/2}\left(\frac{R^2+T}{T^2} - \frac{1}{T}\right)^{-1/2} \\ &= -\sqrt{\frac{T}{R^2}} \\ &= -\sqrt{\frac{R^2 - x^2 - y^2}{R^2}} = z_N \end{aligned}$$

The problem with actually using this choice of c is that it requires prior knowledge of the surface slant. We must find a constant for our choice of c. If we take $c = 1/R$, then we find that

$$-c(|\frac{\nabla^2 I}{I}| - c^2)^{-1/2} = -\sqrt{\frac{R^2 - x^2 - y^2}{R^2 - \epsilon}} \tag{25}$$

where

$$\epsilon = R^2\left(\frac{T}{R^2} - \frac{R^2}{T}\right) = R^2(z_N^2 - z_N^{-2}) \tag{26}$$

Because ϵ is a term in z_N^{-1} the estimated surface slant will be larger than the actual one, introducing a bias into our estimate. This bias can be removed, however, as follows. Let $s = c(|\frac{\nabla^2 I}{I}| - c^2)^{-1/2}$ be the estimated surface slant. Then, combining Equations (25) and (26), we have

$$s^2 = \frac{T}{R^2 - R^2(z_N^2 - z_N^{-2})}$$

Then

$$s^2(1 - z_N^2 + z_N^{-2}) = \frac{T}{R^2} = z_N^2$$

so that we obtain a quadratic in z_N^2

$$(1 + s^2)z_N^4 - s^2 z_N^2 - s^2 = 0 \tag{27}$$

Equation (27) can then be solved to obtain an unbiased estimate of z_N

$$z_N = -\sqrt{\frac{s^2 + \sqrt{4s^2 + 5s^4}}{2 + 2s^2}}$$

(There is only one solution as $0 \geq z_N \geq -1$.) This concludes the proof of Proposition 5.

This estimate of z_N, while unbiased, is not exact because it was necessary to assume that $-xl_1 - yl_2 = 0$. If we examine the conditions under which this factor causes significant error, we see that there will be large errors only when all of the following conditions occur simultaneously:

(1) The surface slant is relatively large

(2) The z-component of the illuminant direction, l_3, is small

(3) The surface faces closely toward or away from the illuminant.

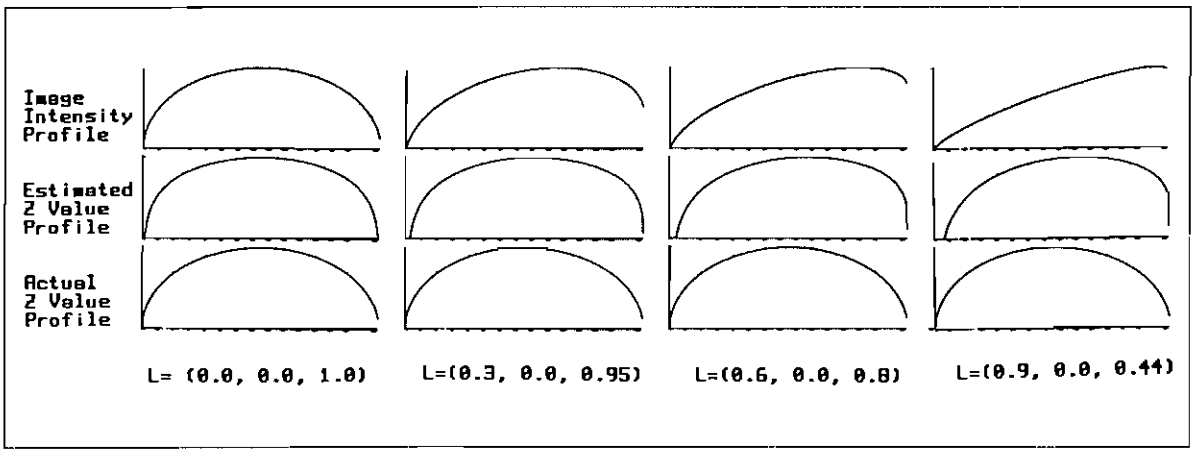

Figure 12. Bias due to illuminant direction. This figure shows the image intensity profile, the profile of the true surface shape (a sphere) and the profile of the reconstructed surface for four illumination conditions. In each case the profile is taken along the image line which goes through the center of the sphere and directly toward the illuminant. This is the direction along which the estimation errors are largest. The distributions of illumination are extended sources, such as would occur if the imaged sphere were placed on a desktop which was near a window. The leftmost distribution shown is centered directly behind the viewer at (0.0,0.0,1.0), the next (proceeding left to right) is centered at (0.3,0.0,0.954), the next at (0.6,0.0,0.8), and the rightmost at (0.9,0.0,0.436).

Figure 12 shows the bias due to illuminant direction which occurs during the estimation of surface orientation for the image of a sphere. This figure shows the image intensity profile, the profile of the true surface shape (a sphere) and the profile of the reconstructed surface for four illumination conditions. In each case the profile is taken along the image line which goes through the center of the sphere and directly toward the illuminant. This is the direction along which the estimation errors are largest.

The distributions of illumination are extended sources, such as would occur if the imaged sphere were placed on a desktop which was near a window. The leftmost distribution shown is centered directly behind the viewer at (0.0, 0.0, 1.0), the next (proceeding left to right) is centered at (0.3, 0.0, 0.954), the next at (0.6, 0.0, 0.8), and the rightmost at (0.9, 0.0, 0.436). Note that the rightmost distribution of illumination results in an almost linear gradient across the image. Thus, these examples approximately span the range of illumination directions found in natural scenes.

Comparing the true surface shape, shown across the bottom of Figure 12, to the reconstructed surface shape[21] shown across the middle of Figure 12, we see that the bias due to illuminant direction does not cause large errors

[21] As obtained by integrating the estimated surface orientation.

REFERENCES

[1] A. P. Pentland, "The Visual Inference Of Shape: Computation From Local Features," Ph.D. Thesis, Psychology Department, Massachusetts Institute of Technology (1982)

[2] B.K.P. Horn,"Shape From Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View", A.I. Technical Report 79, Project MAC, Massachusetts Institute of Technology, (1970)

[3] K. Ikeuchi and B.K.P. Horn,"Numerical Shape from Shading and Occluding Boundaries," Artificial Intelligence, Special Issue on Computer Vision, 15 (1981)

[4] A. Bruss, "Shape From Shading And Bounding Contour," Ph.D. Thesis, Dept. Elec. Engr.& Comp. Sci., Massachusetts Institute of Technology, (1981).

[5] A. Pentland "Finding The Illuminant Direction," Journal of the Optical Society of America, Vol 72, No 4 (1982)

[6] D. Scharf "Magnifications — Photography With The Scanning Electron Microscope," Schocken Brooks, New York, New York

[7] D. Marr and E. Hildreth, "Theory of edge detection," Massachusetts Institute of Technology AI Memo 518 (1979).

[8] S. R. Cajal, "Histologie du systeme nerveux de l'homme et des vertebrates," 2, Paris, Maloine (1911).

[9] S. W. Kuffler,"Discharge Patterns And Functional Organization Of The Mammalian Retina," J. Neurophys. 16 (1953), pp. 281-292.

[10] D. H. Kelly, "Spatial Frequency Selectivity In The Retina," Vision Research 15 (1975), pp. 665-672.

[11] D. H. Kelly, "Motion And Vision. II. Stabilized Spatio-Temporal Threshold Surface," Journal of the Optical Society of America 69, 10 (1979), pp. 1340-1349.

[12] D. H. Hubel and T. N. Wiesel, "The Ferrier Lecture: Functional Architecture Of Macque Monkey Visual Cortex," R. Soc. Lond. B. 198 (1977), pp. 1-59.

[13] P. H. Schiller, B. L. Findlay, and S. F. Volman, "Quantitative Studies Of Single Cell Properties In Monkey Striate Cortex I. Spatio-Temporal Organization Of Receptive Fields," J. Neurophys. 39 (1976), pp. 1288-1319.

[14] K. De Valois, R. L. De Valois, and E. W. Yund, "Responses Of Striate Cortex Cells To Grating And Checkerboard Patterns," J. Physiology 291 (1979) pp. 483-505.

[15] D. Rose, "Mechanisms Underlying The Receptive Field Properties Of Neurons In Cat Visual Cortex," Vision Research 19 (1979) pp. 533-544.

[16] J. G. Daugman, "Two-Dimensional Spectral Analysis Of Cortical Receptive Field Profiles," Vision Research 20 (1980) pp. 847-856

[17] A. P. Witkin, "Recovering Surface Shape and Orientation from Texture," Artificial Intelligence, 17, pp. 17-47 (1981).

[18] J. R. Kender, "Shape From Texture: An Aggregation Transform that Maps a Class of Textures Into Surface Orientation," Proceedings of the Sixth International Joint Conference on Artificial Intelligence, Tokyo, Japan (1979).

[19] A. P. Pentland, "Remote Sensing And Human Visual Processing," Proceedings of the Seventh Annual Conference on Remote Sensing of the Environment, San Jose, Costa Rica, (1980).

[20] C. M. Brown, D. H. Ballard, and O. A. Kimball, "Constraint Interaction In Shape-From-Shading Algorithms," Proceedings of the Image Understanding Workshop, Palo Alto, Ca., (1982).