Pattern Recognition Vol. 18, No. 5, pp. 333–347, 1985.
Printed in Great Britain.

0031-3203/85 $3.00 + .00
Pergamon Press Ltd.

SHAPE ESTIMATION FROM TOPOGRAPHIC PRIMAL SKETCH

TING-CHUEN PONG
Computer Science Department, University of Minnesota, Minneapolis, MN 55455, U.S.A.

LINDA G. SHAPIRO and ROBERT M. HARALICK*
Machine Vision International, 325 E. Eisenhower, Ann Arbor, MI 48104, U.S.A.

(*Received* 2 *October* 1984; *received for publication* 21 *March* 1985)

Abstract—In this paper, a technique for estimating object shape from topographic primal sketch is investigated. Given a gray tone image of a three-dimensional object, a topographic labeling of the image indicates the peaks and pits, ridges and valleys, and flats and hillsides of the underlying, continuous, gray tone surface. The patterns of these topographic labels capture information about the original three-dimensional object in the scene and about the illumination. In order to determine if estimation of three-dimensional shape from a topographic labeling is feasible, we have both analytically and experimentally determined the topographic labelings for images of some mathematically generated surfaces with varied directions of illumination. Our results indicate that such patterns do exist and will be useful in determining three-dimensional shape from two-dimensional images. A scheme for partial classification of three-dimensional object surface is proposed. Preliminary results are illustrated.

Topographic primal sketch Computer vision Three-dimensional shape Surface classification
Topographic patterns

1. INTRODUCTION

Consider an image of a three-dimensional object illuminated by an arbitrary light source and viewed from an arbitrary position. Although ambiguities are possible, frequently the human viewer can estimate (a) the three-dimensional shape of the object, (b) the camera position, and (c) the location of the light source. The original 'shape-from-shading' techniques[1] solve systems of differential equations to derive three-dimensional shape from gray tone intensity variations and operate under a limiting set of restrictions. In addition to low level shading cues, we believe that the human viewer also recognizes *patterns* in the image that give cues leading to estimation of the shape of the object.

Extracting patterns from the original gray tone image is, in most nontrivial cases, an impossible task. In fact, it is for this reason that syntactic pattern recognition systems have had to first extract descriptions consisting of primitives, their properties, and their interrelationships from the image and then to parse these descriptions according to the rules of a grammar. Instead of trying to recognize patterns at the gray-tone intensity level, we propose to work at the topographic labeling level.

To obtain a topographic labeling, a gray tone image may be viewed as a three-dimensional surface whose height at each row–column position corresponds to the intensity value of the image at that position. While the image is a discrete matrix of values, the underlying surface is continuous. Each point of the surface may be labeled as part of a peak, pit, ridge, valley, saddle, hillside, or flat area. Hillsides can be further broken down into the subcategories inflection point, convex hill, concave hill, saddle hill, and slope. In Haralick *et al.*[2] these categories are defined mathematically and the topographic classification of image pixels is described.

Our goal is to use patterns expressed in terms of ridges and valleys, peaks and pits, flats and hillsides to estimate three-dimensional shape. In Section 2 of this paper, results employing two methods for determining such topographic patterns from gray tone intensity images of simple surfaces are described. Section 3 discusses how surface orientations can be estimated from topographic structures. Section 4 describes an object surface classification scheme based on the topographic structures extracted from the image. In the remainder of this section, the imaging geometry and the illumination model to be used in our discussion are defined. The topographic primal sketch is briefly summarized, and some related literature is discussed.

1.1. *Imaging geometry*

The relationship between scene coordinates and image coordinates is illustrated in Fig. 1. We assume that the camera lens is at the origin and that the z-axis

* To whom correspondence should be addressed.

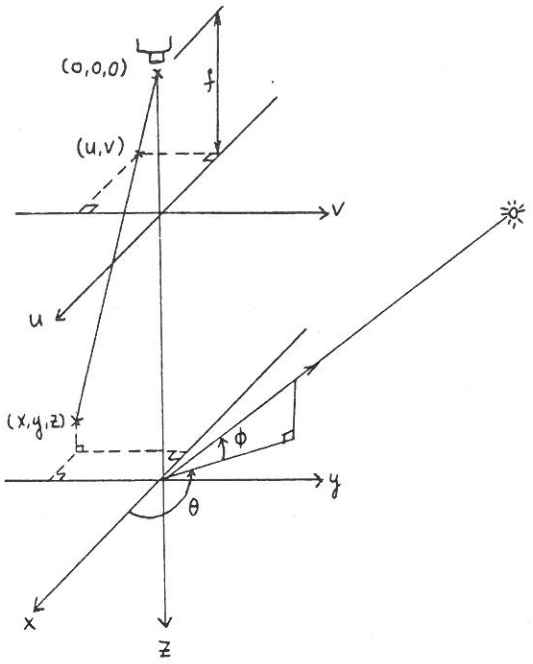

Fig. 1. Relationship between scene coordinates and image coordinates.

is directed towards the image plane which is in front of the lens. The image plane is placed at a distance f, the focal length of the lens, in front of the origin so that the image is oriented in the same way as the scene. As seen from Fig. 1, the following relations hold for perspective projection:

$$u = \frac{fx}{z}, \quad v = \frac{fy}{z}.$$

In our discussion, the perspective projection is approximated by an orthographic projection. This approximation is good when the size of the objects being imaged is small compared to the viewing distance. In this case, appropriate coordinate systems can be chosen such that the following relations hold:

$$u = x, \quad v = y.$$

1.2. *Illumination model*

In the following discussion, we will use a simple illumination model that assumes a distant point light source and a Lambertian reflectance model. A Lambertian surface scatters light equally in all directions. The brightness of a Lambertian surface illuminated by a distant point light source is given by:

$$I = I_0 N \cdot L \tag{1}$$

where I_0 is a constant depending on the surface albedo and the intensity of the light source, N is the unit surface normal vector, and L is the unit vector of the illumination direction.

The unit vector which points in the direction of the light source can be specified by the two angles shown in Fig. 1. The first is the azimuth (θ) which is the angle between the x-axis and the projection of the vector onto the x–y plane, while the second is the angle of elevation (ϕ) of the light source. If we represent this unit vector by $[a, b, c]$, then

$$\begin{aligned} a &= \cos\theta \cos\phi, \\ b &= \sin\theta \cos\phi, \\ c &= -\sin\phi. \end{aligned} \tag{2}$$

In our discussion we will consider only positive values of ϕ. Therefore, c is always less than zero.

If the height of the object surface above the x–y plane is expressed as a function of x and y,

$$z = S(x, y),$$

then the surface normal is given by the vector:

$$N = [S_x, S_y, -1]/(1 + S_x^2 + S_y^2)^{1/2}$$

where S_x and S_y denote first partials of S with respect to x and y, respectively. By carrying out the dot product in Equation (1), it follows that

$$I = I_0 \frac{aS_x + bS_y - c}{(1 + S_x^2 + S_y^2)^{1/2}}.$$

1.3. *Extracting topographic structures*

The topographical primal sketch can be used to represent the underlying intensity surface of a digital image. A complete mathematical treatment of the topographic primal sketch is given in Haralick *et al.*[2] We will summarize here the concepts necessary to understand the remainder of this paper. A digital image may be interpreted as a sampling and quantizing of a real valued function f. While the image is a discrete matrix of values, the underlying surface is continuous. Since the underlying surface is continuous, we can work with such well-defined concepts as its gradient magnitude and its first and second directional derivatives.

The topographic labeling scheme is based on the estimation of the values of the gradients and the directional derivatives of the surface. In order to obtain these values, we need first to assume some kind of parametric form for the underlying function f. If we assume that the neighborhood around each pixel is suitably fitted by the bivariate cubic:

$$\begin{aligned} f(x, y) = {} & k_1 + k_2 x + k_3 y + k_4 x^2 + k_5 xy \\ & + k_6 y^2 + k_7 x^3 + k_8 x^2 y + k_9 xy^2 + k_{10} y^3, \end{aligned}$$

then the parameters k_1–k_{10} can be estimated by a least squares fit to each pixel in the neighborhood. Once these parameters are estimated, the gradient vector (∇f) is given by $(\partial f/\partial x, \partial f/\partial y)$ and its magnitude ($\| \nabla f \|$) is

$$[(\partial f/\partial x)^2 + (\partial f/\partial y)^2]^{1/2}.$$

The first and second directional derivatives may be calculated by forming the Hessian matrix

Table 1. Mathematical properties of topographic structures

$\|\nabla f\|$	X_1	X_2	$\nabla f \cdot w_1$	$\nabla f \cdot w_2$	Label
0	−	−	0	0	Peak
0	−	0	0	0	Ridge
0	−	+	0	0	Saddle
0	0	0	0	0	Flat
0	+	−	0	0	Saddle
0	+	0	0	0	Ravine
0	+	+	0	0	Pit
+	−	−	−, +	−, +	Hillside
+	−	*	0	*	Ridge
+	*	−	*	0	Ridge
+	−	0	−, +	*	Hillside
+	−	+	−, +	−, +	Hillside
+	0	0	*	*	Hillside
+	+	−	−, +	−, +	Hillside
+	+	0	−, +	*	Hillside
+	+	*	0	*	Ravine
+	*	+	*	0	Ravine
+	+	+	−, +	−, +	Hillside
+	*	*	0	0	Impossible

Where:

∇f = gradient vector of the gray tone intensity function f;
$\|\nabla f\|$ = gradient magnitude;
H = Hessian matrix of f,
w_1 = direction in which the second directional derivative has greatest magnitude (first eigenvector of H);
w_2 = direction orthogonal to w_1 (second eigenvector of H);
X_1 = value of the second directional derivative in the direction w_1 (first eigenvalue of H);
X_2 = value of the second directional derivative in the direction w_2 (second eigenvalue of H);
$\nabla f \cdot w_1$ = value of the first directional derivative in the direction of w_1;
$\nabla f \cdot w_2$ = value of the first directional derivative in the direction of w_2.

$$H = \begin{bmatrix} \partial^2 f/\partial x^2 & \partial^2 f/\partial x\,\partial y \\ \partial^2 f/\partial x\,\partial y & \partial^2 f/\partial y^2 \end{bmatrix}.$$

The gradient magnitude and the directional derivatives obtained from the Hessian are used in determining the topographic labeling of the surface. The mathematical properties of the topographic structures defined in Haralick *et al.*[2] are summarized in Table 1. Each entry in the table is either 0, *, + or −. The 0 means an entry is not significantly different from zero, * means an entry does not matter, and +(−) means an entry is significantly different from zero on the positive (negative) side. If the values of the gradient magnitude, eigenvalues and eigenvectors of the Hessian at a given point of the surface satisfy the constraints defined by one row of Table 1, then that point is classified by the label at the end of that row. It is important to note that any combination of $\|\nabla f\|$, X_1, X_2, w_1, and w_2 corresponds to one and only one row of Table 1.

1.4. *Related literature*

Many 'shape-from-...' approaches have been proposed for determining three-dimensional shape from two-dimensional images. Among these approaches, stereopsis[3, 4] and motion[5, 6] are currently the most studied subjects. Four other important sources of shape information are shading, texture, shadow and contour.

Smooth intensity variation (or shading) is an important clue for determining surface orientation. The shape-from-shading idea was first formulated by Horn.[1] Since then, a great deal of work has been done in this area.[7–9] If we assume a uniformly textured surface, surface orientations may be inferred from the way the coarseness of the image texture changes across the image. Shape-from-texture is another area of recent research.[10, 11] When shadows are located in an image, the shapes of the shadows can be used to determine three-dimensional information about the objects in the scene. Shadow analysis can be referred to as the process of locating shadow regions, finding correspondences between shadow casting objects and shadow regions, and deducing three-dimensional information about the objects involved in the shadow formation process. Theoretical work on shadow analysis can be found in Shafer and Kanade.[12] Shape-from-shadow methods have been found to be useful for estimating heights of objects in aerial images.[13] Three-dimensional surface shape can also be inferred from the two-dimensional shapes of edges or curves in an image. Shape-from-contour methods[14, 15] have been found to be effective in determining the shape of a visible surface.

The shape modules use various sources of information in the images to infer intrinsic scene characteristics. Since intrinsic scene characteristics of the visible surface in an image are captured as observed intensities in the image, an important goal of early vision is to extract a rich symbolic representation of the gray tone intensity changes in the image. Among the many representations proposed, the following are the most important ones: (1) the primal sketch[16] or zero-crossing edges[17] of Marr, (2) the intrinsic images of Barrow and Tenenbaum,[18] (3) the relational trees of Ehrich and Foith,[19] and (4) the topographic primal sketch of Haralick *et al.*[2]

The idea of the facet model[20, 21] is used in the construction of the topographic primal sketch. The facet model assumes that image intensity values are noisy sampled observations of an underlying intensity surface. Thus, any interpretation made on the basis of a neighborhood of pixel values should be understood through the analysis of its underlying intensity surface. In the past, the facet model has been proven to be useful in edge and region analysis,[20, 22, 23] noise removal,[24] corner detection,[25] optical flow,[26] shape-from-shading and stereo matching.[27] These successes suggest that there is indeed something fundamental about the approach. We believe that a unified approach based on the facet model and the topographic primal sketch can be developed and used to solve the computer vision problem.

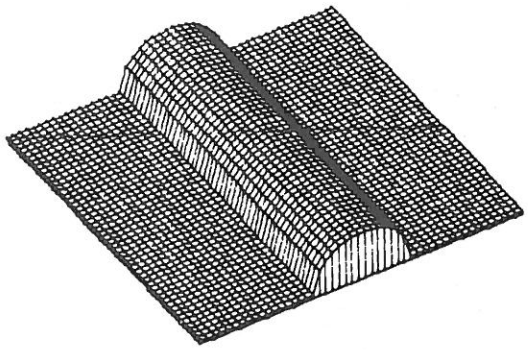

Fig. 2. The cylindrical object used in our experiments.

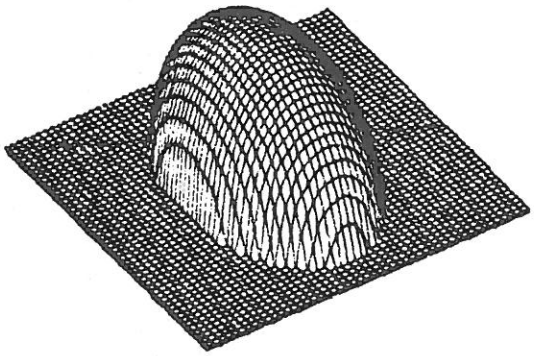

Fig. 4. The ellipsoid used in our experiments.

2. SHAPE FROM TOPOGRAPHIC PATTERNS

There are two possible methods for determining the pattern of topographic labels that will appear, given a particular three-dimensional shape category, a particular reflectance model, a particular light source, and a particular viewpoint. The first method is to work the problem analytically, obtaining exact equations for the illuminated surface. At each point the gradient, eigenvectors, and eigenvalues can be computed in order to determine precisely which sets of points have the various topographic labels. The second method is to work the problem experimentally, using software to generate digital images of illuminated three-dimensional surfaces, to fit these image with either polynomials, splines, or discrete cosines, and to assign topographic labels to each pixel. The first method has the advantage of exactness and the disadvantage of becoming extremely difficult for all but the simplest surfaces. The second method has the advantage of being applicable to a wide variety of surfaces and illuminating conditions and the disadvantage of yielding some inaccurate results due to possible errors in fitting the gray tone image. We have begun to experiment with both methods, starting with very simple surfaces, the Lambertian reflectance model, and point light sources. We have worked with four simple surfaces: (1) the top half of a cylinder, (2) the upper hemisphere of a sphere, (3) the top half of an ellipsoid, and (4) the upper half of a hyperboloid. Figures 2–5 illustrate the four three-dimensional surfaces.

2.1. *Method 1: the experimental approach*

The process for topographic classification can be done in one pass through the image. At each pixel of the image, the following four steps, which are discussed in more detail in Haralick *et al.*,[2] need to be performed.

1. Calculate the least-squares fitting coefficients of a two-dimension cubic polynomial in an $n \times n$ neighborhood around the pixel.
2. Use the coefficients calculated in step 1 to find the gradient, the gradient magnitude, and the eigenvalues and eigenvectors of the Hessian at the center of the pixel's neighborhood.
3. Search in the direction of the eigenvectors calculated in step 2 for a zero-crossing of the first directional derivative within the pixel's area.
4. Recompute the gradient, gradient magnitude, and

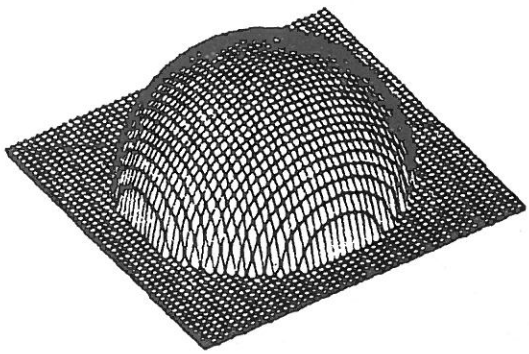

Fig. 3. The spherical object used in our experiments.

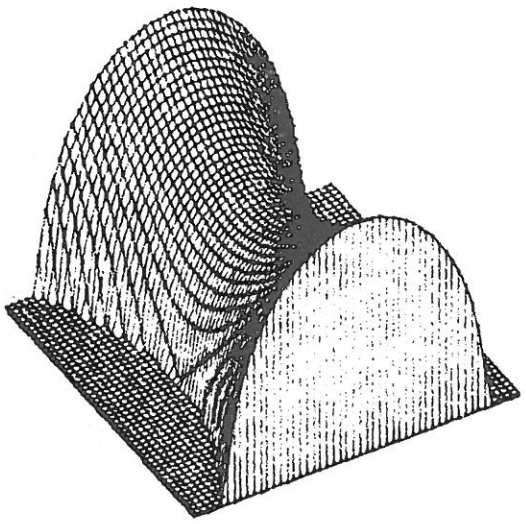

Fig. 5. The hyperboloid used in our experiments.

values of second directional derivative extrema at each zero crossing. Then classify the pixel based on Table 1.

2.2. *Method 2: the analytical approach*

2.2.1. *Topographic labels on the cylinder.* Consider a cylindrical surface given by:

$$S(x, y) = d - (r^2 - y^2)^{1/2} \quad \text{for} \quad -r \leqslant y \leqslant r \quad (3)$$

where d is the distance of the x–y plane from the camera down the z-axis and r is the radius of the cylinder. This surface, in which the axis of the cylinder lies along the x-axis, was chosen to simplify calculations. Notice that since only the top half of the cylinder is considered, the sign of the square root in equation (3) is taken as positive. By differentiating S with respect to x and y, we obtain

$$S_x = 0, \quad S_y = y(r^2 - y^2)^{-1/2}.$$

It follows from equation (2) that the intensity of the cylinder illuminated from direction (a, b, c) is given by:

$$I(x, y) = I_0(by - c(r^2 - y^2)^{1/2})r \quad (4)$$

After some simplifications, the first and second partials of I are found to be:

$$I_x = I_{xx} = I_{xy} = I_{yx} = 0,$$
$$I_y = I_0(b + cy(r^2 - y^2)^{-1/2})/r,$$
$$I_{yy} = I_0 cr(r^2 - y^2)^{-3/2},$$

where the subscripted I's denote partial differentiation with respect to the subscript(s).

Since I_x is equal to zero, the gradient magnitude ($\| \nabla f \|$) is equal to the absolute value of I_y. Therefore, $\| \nabla f \| = 0$ when

$$I_y = I_0(b + cy(r^2 - y^2)^{-1/2})/r = 0$$

which implies

$$b(r^2 - y^2)^{1/2} = -cy. \quad (5)$$

Upon solving equation (5), we obtain

$$y^2 = r^2b^2/(b^2 + c^2).$$

Because c is always negative, the sign of y is taken to be the same as that of b in order for equation (5) to be satisfied.

To determine the second directional derivative extrema values and the first directional derivatives taken in the directions which extremize second directional derivatives, we form the Hessian:

$$H = \begin{bmatrix} 0 & 0 \\ 0 & I_0 cr(r^2 - y^2)^{-3/2} \end{bmatrix}.$$

The eigenvalues of the Hessian are obtained as:

$$X_1 = I_0 cr(r^2 - y^2)^{-3/2}, \quad (6)$$
$$X_2 = 0; \quad (7)$$

their associated eigenvectors are:

$$w_1 = (0, 1),$$
$$w_2 = (1, 0).$$

Recall that c is always negative, therefore, X_1 is always negative for $-r < y < r$. By taking the dot product of the gradient with the eigenvectors, we obtain:

$$\nabla I \cdot w_1 = I_y = I_0(b + cy(r^2 - y^2)^{-1/2})/r,$$
$$\nabla I \cdot w_2 = 0.$$

To determine the topographic labels, we need to consider two cases: (1) zero gradient magnitude and (2) positive gradient magnitude.

Case 1: zero gradient magnitude. If we let $y_0 = rb(b^2 + c^2)^{-1/2}$, it follows from equation (5) that $\| \nabla I \| = 0$ when $y = y_0$. By equations (6–7), the second directional derivative extrema values at $y = y_0$ are

$$X_1 = I_0 cr(r^2 - y_0^2)^{-3/2} \quad (8)$$
$$X_2 = 0. \quad (9)$$

Since X_1 is always less than zero, it follows directly from Table 1 that a ridge is located at $y = y_0$.

Case 2: positive gradient magnitude. If the gradient magnitude ($\| \nabla I \|$) is taken to be positive, then the value of the first directional derivative in the direction of $w_1(\nabla I \cdot w_1)$ is always non-zero because $\nabla I \cdot w_1 = I_y$ and $\| \nabla f \| = |I_y|$. In this case, since X_1 is always negative and X_2 is always zero, it follows from row 11 of Table 1 that hillsides are located at those places where the gradient magnitudes are positive.

2.2.2. *Topographic labels on the sphere.* In the case of the sphere, the equation of a spherical surface with radius r is given by:

$$S(x, y) = d - (r^2 - x^2 - y^2)^{1/2} \quad \text{for} \quad -r \leqslant x \leqslant r \quad \text{and} \quad -r \leqslant y \leqslant r \quad (10)$$

Its intensity illuminated from direction $[a, b, c]$ is given by:

$$I(x, y) = I_0 * [ax + by - c(r^2 - x^2 - y^2)^{1/2}]/r \quad (11)$$

After some simplifications, the first and second partials of I are found to be:

$$I_x = I_0[a + cx(r^2 - x^2 - y^2)^{-1/2}]/r,$$
$$I_y = I_0[b + cy(r^2 - x^2 - y^2)^{-1/2}]/r,$$
$$I_{xx} = I_0 c(r^2 - y^2)(r^2 - x^2 - y^2)^{-3/2}/r,$$
$$I_{xy} = I_{yx} = I_0 cxy(r^2 - x^2 - y^2)^{-3/2}/r,$$
$$I_{yy} = I_0 c(r^2 - x^2)(r^2 - x^2 - y^2)^{-3/2}/r.$$

The gradient magnitude ($\| \nabla I \|$) is given by:

$$\| \nabla I \| = (I_x^2 + I_y^2)^{1/2},$$

which is zero when $a(r^2 - x^2 - y^2)^{1/2} + cx = 0$ and $b(r^2 - x^2 - y^2)^{1/2} + cy = 0$ are satisfied simultaneously. By squaring and invoking the constraint $a^2 + b^2 + c^2$

$= 0$ on the unit vector $[a, b, c]$, the solution to the simultaneous equations is found to be:

$$x = ra, \quad y = rb.$$

The Hessian for the intensity surface of the illuminated sphere is given by:

$$H = \frac{I_0 c}{r(r^2 - x^2 - y^2)} * \begin{bmatrix} r^2 - y^2 & xy \\ xy & r^2 - x^2 \end{bmatrix}.$$

Its eigenvalues are found to be:

$$X_1 = I_0 cr(r^2 - x^2 - y^2)^{-3/2},$$
$$X_2 = I_0 c(r^2 - x^2 - y^2)^{-1/2}/r.$$

Notice that both eigenvalues are always less than zero since c is always less than zero. The eigenvector corresponding to X_1 is given by:

$$w_1 = [x(x^2 + y^2)^{-1/2}, y(x^2 + y^2)^{-1/2}]$$

and the eigenvector corresponding to X_2 is given by:

$$w_2 = [-y(x^2 + y^2)^{-1/2}, x(x^2 + y^2)^{-1/2}].$$

The dot product of the gradient with w_1 is

$$\nabla I \cdot w_1 = \frac{I_0}{r(x^2 + y^2)^{1/2}} \left[x\left(a + \frac{cx}{(r^2 - x^2 - y^2)^{1/2}}\right) + y\left(b + \frac{cy}{(r^2 - x^2 - y^2)^{1/2}}\right)\right] \quad (12)$$

and the dot product of the gradient with w_2 is

$$\nabla I \cdot w_2 = I_0(-ay + bx)(x^2 + y^2)^{-1/2}/r. \quad (13)$$

We determine the topographic labels by considering two cases.

Case 1: zero gradient magnitude. The gradient magnitude is equal to zero when $(x, y) = (ra, rb)$. Since both eigenvalues are less than zero on the illuminated sphere, it follows directly from Table 1 that a peak is located at $(x, y) = (ra, rb)$.

Case 2: positive gradient magnitude. In the case when the gradient magnitude is given to be positive, since both eigenvalues are known to be negative, it follows from Table 1 that there is a ridge at those locations where either $\nabla f \cdot w_1 = 0$ *or* $\nabla f \cdot w_2 = 0$ is satisfied. We obtain from equations (12) and (13) that

$$\nabla I \cdot w_1 = 0 \text{ when } (ax + by)(r^2 - x^2 - y^2)^{1/2} + c(x^2 + y^2) = 0,$$

$$\nabla I \cdot w_2 = 0 \text{ when } -ay + bx = 0.$$

Table 1 also says that hillsides appear at places where both $\nabla f \cdot w_1$ and $\nabla f \cdot w_2$ are non-zero.

2.2.3. *Other conic surfaces.* We have also attempted to derive analytical results for the following conics surfaces:

(1) ellipsoid

$$S(x, y) = d - (l^2 - m^2x^2 - n^2y^2)^{1/2};$$

(2) hyperboloid of one sheet

$$S(x, y) = d - (l^2 - m^2x^2 + n^2y^2)^{1/2};$$

(3) hyperboloid of two sheets

$$S(x, y) = d - (m^2x^2 - n^2y^2 - l^2)^{1/2};$$

(4) paraboloid

$$S(x, y) = d - (nx^2 + my^2).$$

Unfortunately, no simple closed form solution has been obtained for the first and second directional derivatives of these surfaces. Some of the analytical equations of these surfaces are given in Pong.[26] Results of the topographic patterns on these surfaces are described in Section 5.5. Results show that patterns also emerged for these surfaces.

3. ESTIMATION OF SURFACE ORIENTATION

In this section, we will show that the topographic labels along with their quantitative measurements bear a strong relationship to the surface orientation of the three-dimensional object in the scene. Consider a spherical surface as described in Section 2.2.2. The unnormalized surface orientation of such a surface can be represented in the gradient space by the vector $[p, q, -1]$, where

$$p = \frac{x}{(r^2 - x^2 - y^2)^{1/2}}, \quad q = \frac{y}{(r^2 - x^2 - y^2)^{1/2}}.$$

An alternative way of specifying surface orientation is the tilt and slant representation. Tilt specifies the orientation of the projection of the surface normal onto the image plane. Slant is the angle between the surface normal and the viewing direction. The tilt and slant representation and the gradient space representation can be related by the following formulas:

$$\tan\theta = q/p \quad \text{and}$$
$$\tan\phi = (p^2 + q^2)^{1/2} \quad \text{or}$$
$$\cos\phi = (1 + p^2 + q^2)^{-1/2}.$$

In the case of a sphere,

$$\tan\theta = y/x,$$
$$\cos\phi = (r^2 - x^2 - y^2)/r^2.$$

To see how the surface orientation of a spherical Lambertian surface can be derived from the topographic analysis of the image intensity surface, we need first to complete the analytical results of Section 2.2.2 by considering the lower half of the sphere. The equation of the lower hemisphere of a sphere whose center is at $(0, 0, d)$ is given by:

$$S(x, y) = d + (r^2 - x^2 - y^2)^{1/2} \quad \text{for} \quad -r \leqslant x \leqslant r \quad \text{and} \quad -r \leqslant y \leqslant r.$$

Differentiating the above equation with respect to x and y, we obtain

$$p = \frac{-x}{(r^2 - x^2 - y^2)^{1/2}}, \quad q = \frac{-y}{(r^2 - x^2 - y^2)^{1/2}}.$$

Notice the sign difference between the surface orientations of the upper and lower hemispheres.

As for the upper hemisphere, after some simplification, we obtain for the lower hemisphere a similar set of expressions for I and its partials,

$$I = I_0[-ax - by - c(r^2 - x^2 - y^2)^{1/2}]/r,$$
$$I_x = I_0[-a + cx(r^2 - x^2 - y^2)^{-1/2}]/r,$$
$$I_y = I_0[-b + cy(r^2 - x^2 - y^2)^{-1/2}]/r,$$
$$I_{xx} = I_0c(r^2 - y^2)(r^2 - x^2 - y^2)^{-3/2}/r,$$
$$I_{xy} = I_{yx} = I_0cxy(r^2 - x^2 - y^2)^{-3/2}/r,$$
$$I_{yy} = I_0c(r^2 - x^2)(r^2 - x^2 - y^2)^{-3/2}/r.$$

Notice that the second partials of I are the same for both halves of the sphere. Since the second partials make up the Hessian, it follows that the eigenvalues and eigenvectors for the two hemispheres are also identical. Recall that the eigenvalues and eigenvectors are given by

$$X_1 = I_0cr(r^2 - x^2 - y^2)^{-3/2},$$
$$X_2 = I_0c(r^2 - x^2 - y^2)^{-1/2}/r,$$
$$w_1 = [x(x^2 + y^2)^{-1/2}, y(x^2 + y^2)^{-1/2}],$$
$$w_2 = [-y(x^2 + y^2)^{-1/2}, x(x^2 + y^2)^{-1/2}].$$

If we take the ratio of the smaller over the larger eigenvalue, we obtain

$$\frac{X_2}{X_1} = \frac{r^2 - x^2 - y^2}{r^2}.$$

This ratio is the square root of the cosine of the surface slant. Note that the signs of both X_1 and X_2 depend only on the sign of c which is the negative of the sine of the angle of elevation of the light source. Therefore, the ratio is always positive and its square root is always justifiable. Furthermore, the ratio is always less than or equal to one since X_2 is the smaller eigenvalue. Thus, we can obtain surface slant by taking the arc-cosine of the square root of X_2/X_1. The resulting angle is determined uniquely because the surface slant for a visible surface always lies between 0 and $\pi/2$.

The remaining component to be determined for the unit surface normal is the surface tilt. By considering w_1, the eigenvector corresponding to the larger eigenvalue, we can obtain the direction θ in which the second directional derivative of I is extremized. That is,

$$\tan\theta = \frac{y}{(x^2 + y^2)^{1/2}} \cdot \frac{(x^2 + y^2)^{1/2}}{x} = \frac{y}{x},$$

which is identical to the tangent of the surface tilt. Thus we have $t = \theta$ or $t = \theta + \pi$. Unfortunately, there are two possible solutions. This is expected because each solution corresponds to one half of the sphere. This shows the ambiguity in local analysis of image shading.

Recently, Pentland[9] proposed a method that estimates surface orientation locally at each image point. This method is based on the claim that the image of a point on a spherical Lambertian surface can produce any combination of image intensity I and derivatives I_x, I_y, I_{xx}, I_{yy} and I_{xy}. The implication of this claim is that it is impossible to determine, based on local analysis of shading, whether or not an image point is resulted from a point on a spherical Lambertian surface. In what follows, we will show that Pentland's claim is incorrect and we will then suggest a scheme for partial classification of three-dimensional object surfaces.

Without any assumption about the location of the light source, we found from the above analytical results that the topographic labels on the underlying intensity surface resulting from a spherical Lambertian surface can only be peak, pit, ridge, valley, convex or concave hillside. This is because any combinations of I_{xx}, I_{yy} and I_{xy} resulting from a spherical Lambertian surface can produce only either a semi-positive or semi-negative definite Hessian. Therefore, not all combinations of I_{xx}, I_{yy} and I_{xy} are possible.

Furthermore, if we approximate three-dimensional surfaces locally by spherical surfaces, it is expected that the radii of the approximating spheres for points on a spherical surface are constant. Recall that the radius, r, and the eigenvalues, X_1 and X_2, of a spherical Lambertian surface are related by the following expressions:

$$X_1 = \frac{I_0cr}{(r^2 - x^2 - y^2)^{3/2}}, \quad X_2 = \frac{I_0c}{r(r^2 - x^2 - y^2)^{1/2}}.$$

We obtain from the expression of X_1

$$(r^2 - x^2 - y^2)^{1/2} = \frac{I_0c}{X_2r}.$$

We then have from the expression of X_2

$$X_1 = I_0cr\left(\frac{I_0c}{X_2r}\right)^{-3},$$

or

$$r^4 = \frac{X_1I_0^2c^2}{X_2^3}.$$

Since I_0 and c are fixed, we conclude that X_1/X_2^3 is constant for a spherical Lambertian surface. Therefore, an image point can be determined as resulting from a point on a spherical Lambertian surface only if it is labeled as a peak, pit, ridge, valley, convex hillside or concave hillside, and the radii of the approximating spheres at pixels within the neighborhood around that point are similar enough. We thus showed that Pentland's claim is incorrect. What this suggests is that we should estimate surface tilt and surface slant locally from the eigenvalues and eigenvectors of the Hessian of the underlying intensity surface only if the underlying intensity surface is compatible with that of a spherical Lambertian surface.

It can be observed from the expressions of the

eigenvalues that a pit, valley and convex hillside classification of the intensity surface of a spherical Lambertian surface corresponds to a positive c. This implies a light source below the object surface. Although this is physically possible, such illuminating condition can usually be ignored when solving practical problems. We thus further assume that a spherical Lambertian surface can only result in peak ridge and concave hillside classifications.

4. CLASSIFICATION OF OBJECT SURFACES

We will propose here a scheme for partial classification of three-dimensional object surfaces. The basic goal of this classification scheme is to group together pixels which are likely to come from the same surface patch. We will limit our consideration to five types of object surfaces. They are planar, developable*, spherical, elliptical, and hyperbolic surfaces.

Based on the above discussions, it is evident that topographic labels together with the signs and magnitudes of their second directional derivatives bear a strong relationship to the nature of the three-dimensional object surface in the scene. This evidence leads to the assumption that maximally connected sets of pixels having the same topographic label belong to the same surface patches. To extract these connected sets of topographic structures, the feature extraction process described in Pong[27] is employed. The resulting structures of this feature extraction process are arcs, regions and topographic labels. The desired topographic structures are then determined by applying a connected components algorithm[28] to the topographic labels within each region segment.

The assembled topographic structures may be divided into three categories: (1) areal structures which consist of convex hillsides, concave hillsides, saddle hillsides, flat surfaces, and sloped surfaces; (2) arc structures which include edges, ridges, and valleys; (3) point structures which include peaks, pits, and saddle points. In what follows, we suggest hypotheses that can be made about the three-dimensional objects based on the analytical results that we have derived and the results of the experiments that we have performed. We believe that three-dimensional object shape can be inferred by feeding this knowledge into a hypothesis based reasoning system such as that of Mulgaonkar.[29]

4.1. *Areal structures*

Flat. A flat is a simple surface with zero gradient and no curvature. That is, the gray level intensity is constant in a connected flat structure. Since the surface normal vectors within a planar surface are constant, we can be almost certain that pixels belonging to a connected flat structure come from the same planar surface. Although this may not hold for shadow areas, we can usually separate shadow areas by identifying flat structures with relatively low intensity averages.

Hillsides. We first hypothesize that a concave/convex hillside assembly is part of a spherical, elliptical or developable surface, and a saddle hillside assembly is part of a hyperbolic surface. Our first hypothesis is driven by the analytical and experimental results of the cylindrical, spherical, elliptical and hyperbolic surfaces that we have considered in Section 2.2.

We further postulate that a concave/convex hillside assembly is part of a developable surface if it is adjacent to a straight and horizontal ridge. In particular, it is part of a cylindrical surface if it is concave and the second directional derivative of the hillside in the direction of the ridge is zero.

As a result of Section 3, a hillside assembly can be considered as part of a spherical surface if it is concave and the radii of the approximating spheres within the hillside assembly are similar enough. We have not been able to derive a complete classification scheme for all the areal structures. Nonetheless, since the assembled regions are likely to come from the same surface patches, they are good starting regions for the shape-from-shading method described in Pong.[27]

4.2. *Line and point structures*

While edges are considered to be good indications of surface discontinuities, peaks and ridges are found to be significant structures in the images of the conic surfaces that we have considered. The following observations are gathered from the topographic structures of the conic surface.

1. The ridge arcs obtained from the images of the sphere are found to be symmetrical around the peaks.
2. Ridges for the images of the ellipsoid and the hyperboloid are found to be symmetrical around the peaks only if the projection of the light vectors are parallel to one of the axes of these conic surfaces.
3. Straight ridge lines are found in the images of the cylinder. The gray tone intensities along the ridge lines are found to be constant.
4. While the ridges around the peaks found in the image of the ellipsoid curve away from the light sources, those of the hyperboloid curve towards the light sources.
5. The peaks located in the images of the conic surfaces correspond to locations where the surface normals are pointing towards the light sources.

A pictorial illustration of the above observations is given in Fig. 6.

In this section, we have listed a set of criteria that can be used to infer shapes of three-dimensional objects. These criteria are supported by the experimental results to be described in Section 5. We believe that an integrated knowledge based vision system can be benefited by effectively using the knowledge about

* A developable surface is one which may be unrolled onto a plane without distortion.

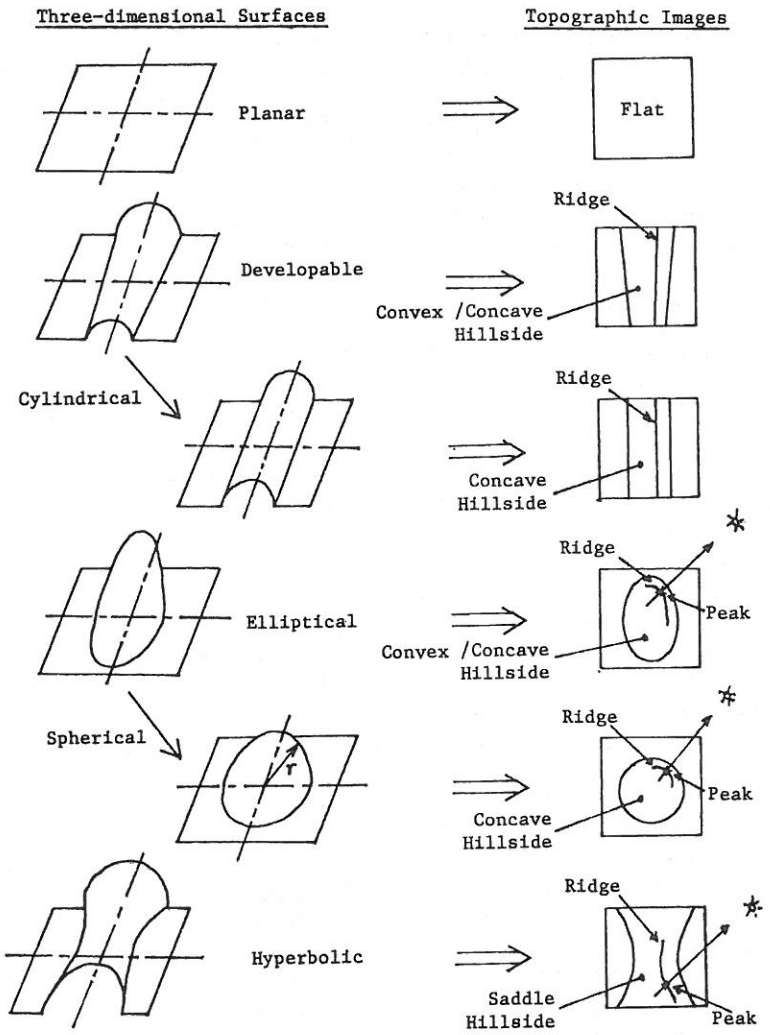

Fig. 6. An illustration of the topographic patterns.

these topographic structures.

5. RESULTS

In Sections 5.1–5.3, we show the analytical and experimental results of the topographic patterns on the cylinder and sphere of Fig. 2 and 3. Three illumination conditions are considered for each surface: (1) the light direction is (0, 0, −1) which means directly above the center of the surface; (2) the light direction is $(0, \sqrt{3}/2, -1/2)$ which translates to azimuth 0 degrees and elevation 30 degrees; (3) the light direction is $(1/2, 1/2, -1/\sqrt{2})$ which translates to azimuth 45 degrees and elevation 45 degrees. The illuminated surfaces of the cylinder and the sphere are shown in Fig. 7 and Fig. 8, respectively.

5.1. *Analytical results for the cylinder*

When the light direction is from azimuth 0 degrees, elevation 90 degrees, analytical results in Section 2.2.1 indicate a ridge parallel to the axis of the cylinder and running along the center of the top half, as shown in Fig. 9a. When the light direction is from azimuth 0 degrees and elevation 30 degrees, the ridge appears as in Fig. 9b. When the light direction is from azimuth 45 degrees and elevation 45 degrees, the ridge appears as in Fig. 9c. In all three cases, the remaining points of the cylinder are hillsides.

5.2. *Analytical results for the sphere*

When the light source is directly above the center of the sphere, the gradient magnitude is zero at (0, 0), therefore, a peak is located at the center of the sphere. The gradient magnitude is positive and the first directional derivative in the direction w_2 is zero at the remaining points of the sphere. It follows from our analytical results that ridges locate at these points.*

* This result is due to the analysis used in Haralick *et al.*[2] and may not be intuitively obvious to all readers.

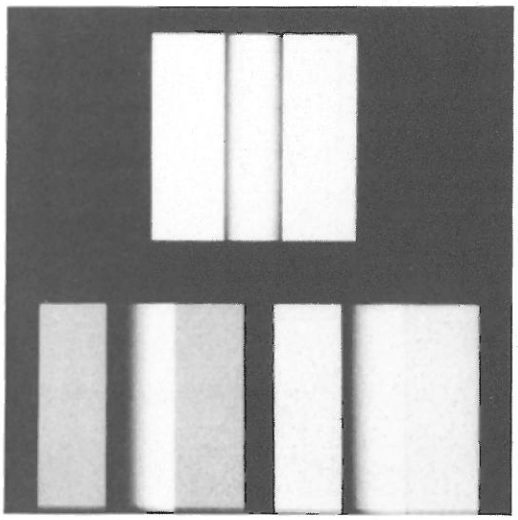

Fig. 7. Shaded images of the cylinder of Fig. 2.

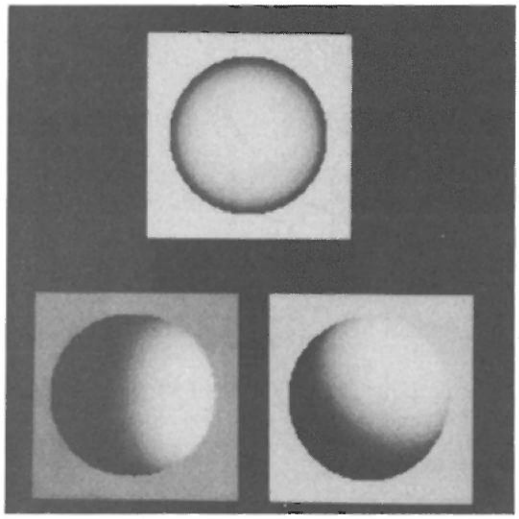

Fig. 8. Shaded images of the sphere of Fig. 3.

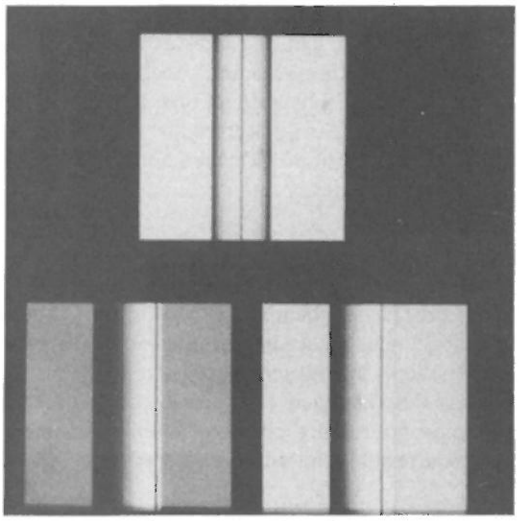

Fig. 9. The analytically derived topographic labeling of the cylinder.

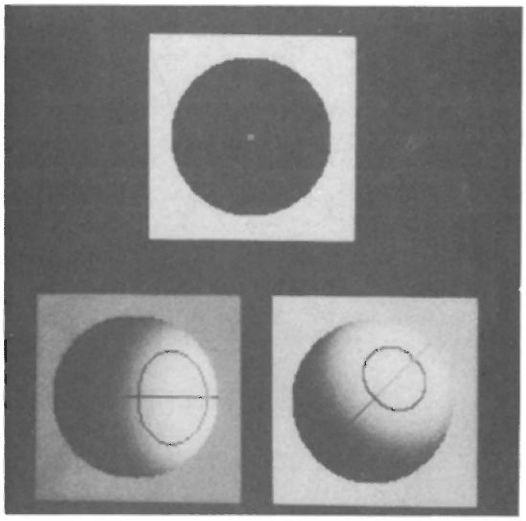

Fig. 10. The analytically derived topographic labeling of the sphere.

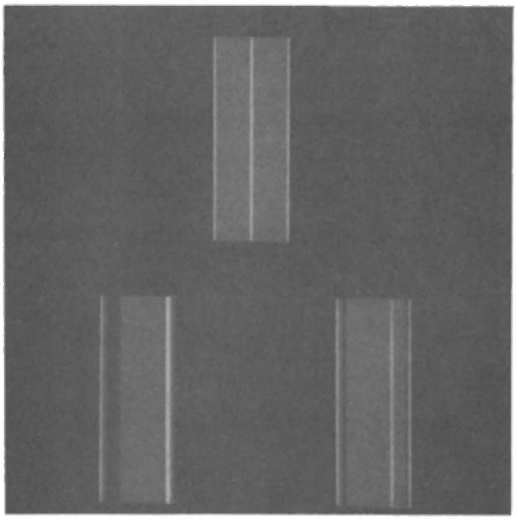

Fig. 11. The experimental results for the cylinder.

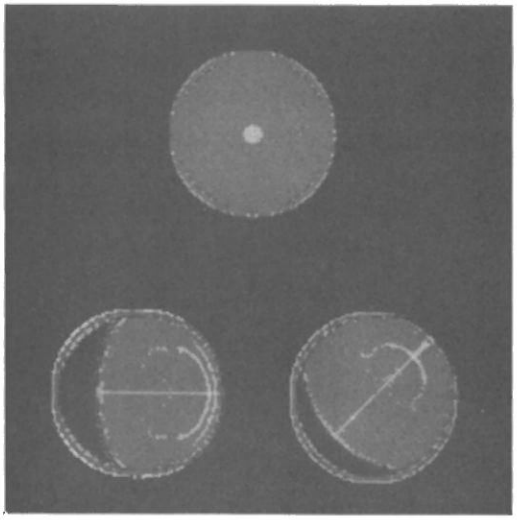

Fig. 12. The experimental results for the sphere.

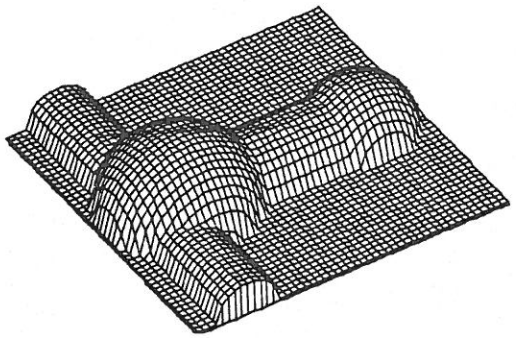

Fig. 13. A synthetic three-dimensional object.

When the light direction is $(0, \sqrt{3}/2, -1/2)$, a peak is found at $(0, \sqrt{3}r/2)$. At the remaining points,

$$\nabla I \cdot w_1 = 0 \quad \text{when} \quad (x^2 + y^2) = \sqrt{3}/2y(r^2 - x^2 - y^2)^{1/2},$$

$$\nabla I \cdot w_2 = 0 \quad \text{when} \quad x = 0.$$

Therefore, there are ridges when either one of the above two equations is satisfied and hillsides otherwise.

Similarly, a peak is found at $(r/2, r/2)$ when the light direction is $(1/2, 1/2, -1/\sqrt{2})$. Ridges can be located at places where either

$$\sqrt{2}(x^2 + y^2) = (x + y)(r^2 - x^2 - y^2)^{1/2}$$

or $\quad x = y$ is satisfied.

At the remaining points, hillsides are the correct categories. Figure 10 shows the topographic labels for the illuminated spheres.

5.3. *Experimental results*

Experimentally, we are working in the GIPSY (General Image Processing System) environment. There currently exist GIPSY commands to construct three-dimensional surfaces, to produce images of these surfaces from various viewpoints and light directions, to fit these images with either cubic polynomials, splines, or discrete cosines, and to calculate the topographic labelings. Figures 11 and 12 show experimental results for the cylinder and sphere using cubic polynomial surface fitting. Experimental results show very good correspondence with the analytical results, except for the sphere when the light direction is $(0, 0, -1)$. In this case, when points are labeled ridge and have neighboring points in a direction orthogonal to the gradient that are also labeled ridge, our software reclassifies these ridge continuums as hillsides. A detailed discussion of ridge and valley continua can be found in Haralick *et al.*[2]

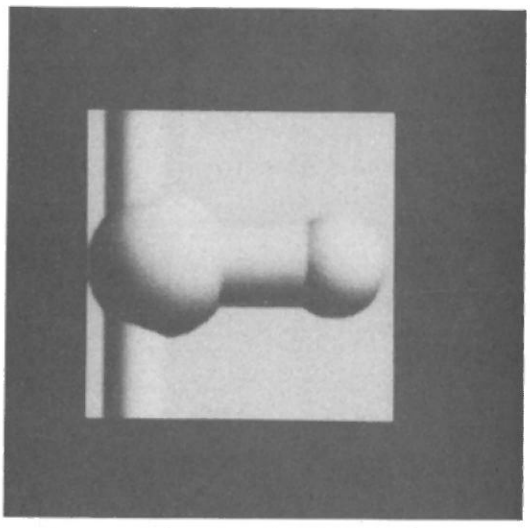

Fig. 14. Gray tone image of the synthetic object illuminated from azimuth 45 degrees and elevation 45 degrees.

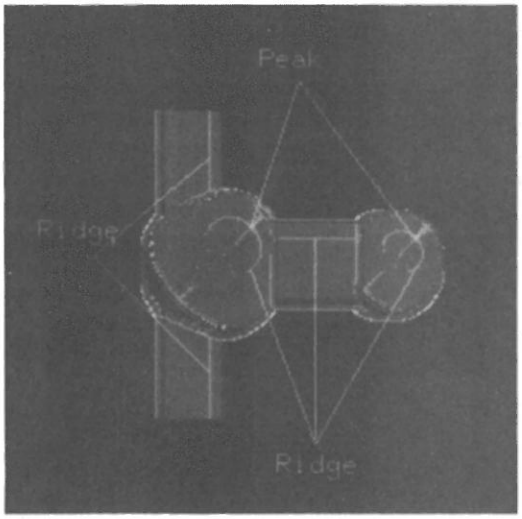

Fig. 15. The topographic labeling of the synthetic object image.

In addition to the images of the two simple surfaces, a synthetic image of a more complex surface was also used in testing. The surface of Fig. 13 is composed of cylindrical and spherical surface patches. Figure 14 shows the image of the surface when illuminated from azimuth 45 degrees and elevation 45 degrees. Figure 15 illustrates the topographic labels that resulted from the experimental method. Most of the resulting topographic labels are located at places where they are predicted by the analytical method.

Our results show that the most informative features found in the images of the cylinder and sphere are ridges and peaks. While the ridges found in the cylinder images are intuitive, the ellipse-like ridges found in the sphere images are unexpected. Although most ridge points found in the sphere images are weak ridges, experimental results show that these ridges are detectable. These ellipse-like ridges will be a definite clue to 3-dimensional surface identifications. Once the shape of the 3-dimensional surface is hypothesized as cylindrical or spherical, information such as the

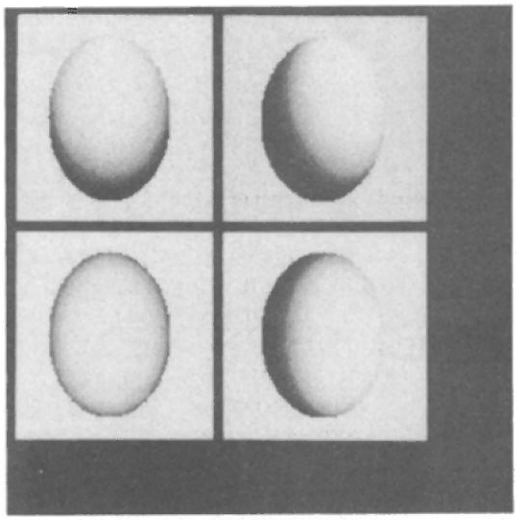

Fig. 16. Shaded images of the ellipsoid of Fig. 4.

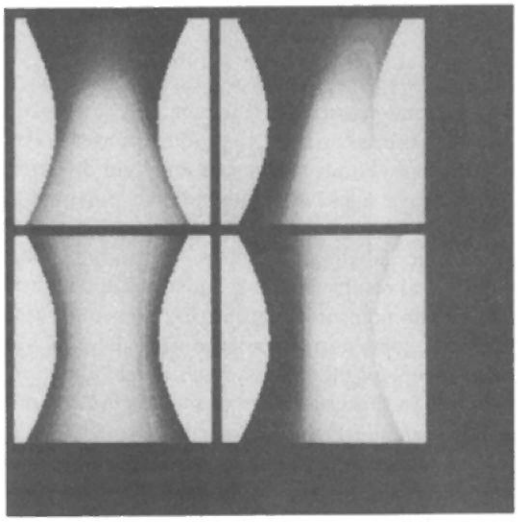

Fig. 17. Shaded images of the hyperboloid of Fig. 5.

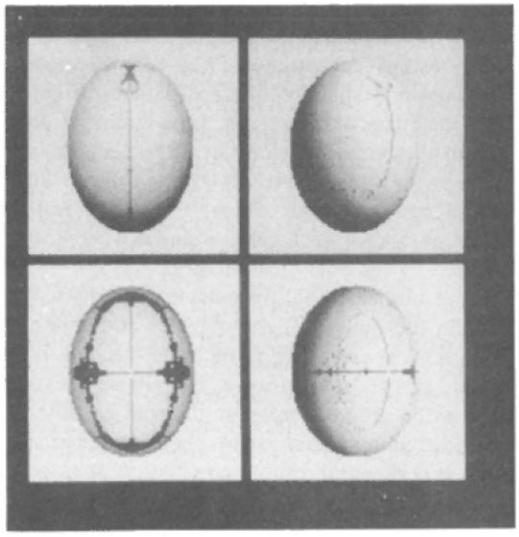

Fig. 18. The analytically derived topographic labelings of the images of the ellipsoid.

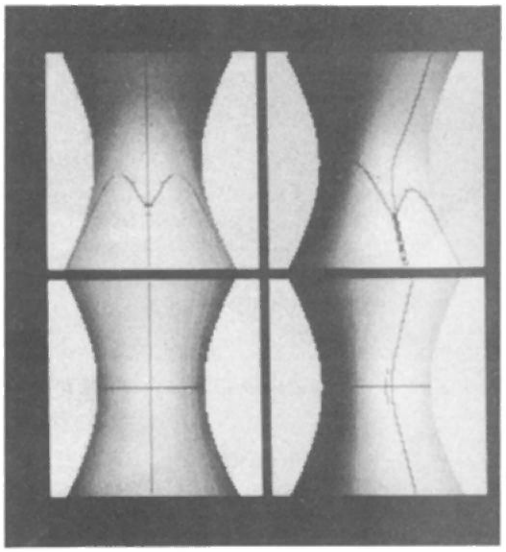

Fig. 19. The analytically derived topographic labelings of the images of the hyperboloid.

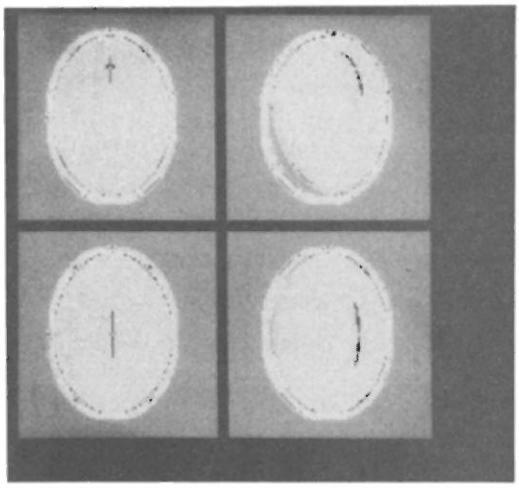

Fig. 20. The experimental results of the images of the ellipsoid.

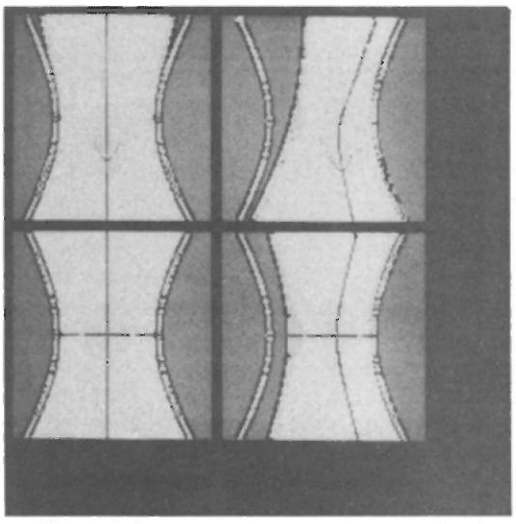

Fig. 21. The experimental results of the images of the hyperboloid.

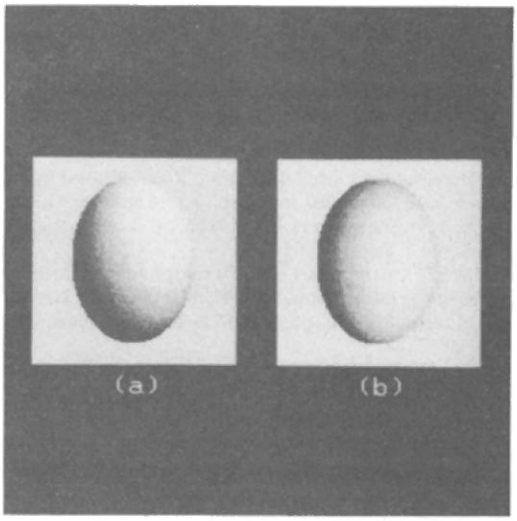

Fig. 22. Noisy images of an ellipsoid.

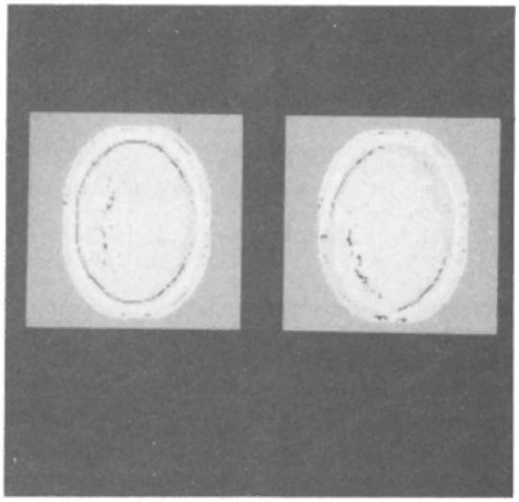

Fig. 23. The experimentally derived topographic labelings of the images of Fig. 23.

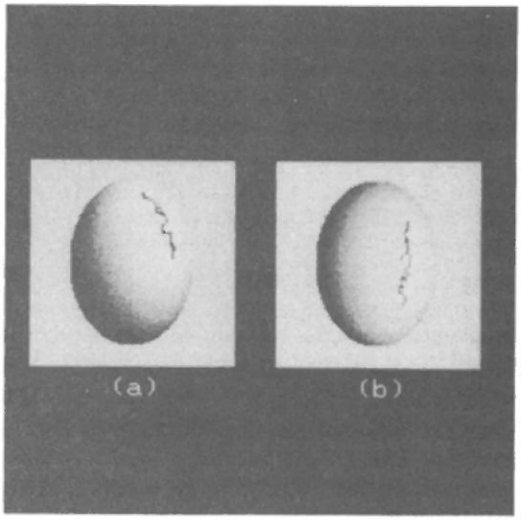

Fig. 24. Results of the ridge linking process.

direction of the light source and the cylinder/sphere radius may also be estimated by examining the topographic labels.

5.4. *Results for the other conic surfaces*

Experiments have been performed on the ellipsoid and hyperboloid of Figs. 4 and 5. Shaded images of these conic surfaces are shown in Figs. 16 and 17. Figures 18 and 19 illustrate the topographic labels that resulted from the analytical equations given.[27] Figures 20 and 21 show the topographic labels that resulted from the facet surface fitting model. Results show that the dominant structures in the images of the ellipsoid and hyperboloid are, respectively, concave and saddle hillsides. It was also found that only portions of the ridges that appear in Figs. 18 and 19 were detected by the experimental method. Careful examinations showed that the ridges detected by the experimental method were dominant ridges which gave high curvatures (or eigenvalues), and the undetected ridges were weak ridges which gave very small curvatures. It was also observed that the intersection of the two ridge lines produced a peak.

To examine the performance of the facet fitting topographic labeling scheme on noise images, random noise of mean = 0 and standard deviation = 100 was added to the shaded images of an ellipsoid. The gray level intensities of the shaded images range from 0 to 1000. The noisy images are shown in Fig. 22. Figure 23 shows the topographic labels that resulted from the facet surface fitting model. It was found that the ridge labels detected were fragmented into small pieces. In order to group the fragmented ridges into more meaningful structures, a modification of the arc segment extraction procedure as described in Pong[27] was used to link the fragmented ridges. The modified procedure differs from the original procedure in that linking is done based on ridge orientations instead of edge orientations. In particular, ridge orientation is defined to be the direction perpendicular to the eigen direction which gives the first directional derivative zero-crossing. Results of the ridge linking process are shown in Fig. 24. It was also found that the detected ridges became unrecognizable when more noise was added to these images.

6. SUMMARY

Both the analytical and experimental results so far indicate that there are definite patterns emerging that can help indicate the shape of the original three-dimensional surface and the direction of the light source. For both methods, we anticipate a great deal more work. We would like to carry out the analytic approach for several more simple surfaces, perhaps using a symbol manipulator to help solve the equations. For the experimental methods, we need to work on getting results as close as possible to the analytic results. We will then perform a large series of

experiments with various surfaces, combinations of surfaces forming objects, viewpoints, reflectance functions and lighting conditions. Only then will we be able to complete the work of analyzing the patterns of topographic labels produced and predicting three-dimensional shape from these patterns.

7. REFERENCES

1. B. K. P. Horn, Obtaining shape from shading information, *The Psychology of Computer Vision*, P. H. Winston, ed., pp. 115–155. McGraw-Hill, New York (1975).
2. R. M. Haralick, L. T. Watson and T. J. Laffey, The topographic primal sketch, *Int. J. Robotics Res.* **2,** 50–72 (1983).
3. D. Marr and T. Poggio, A theory of human stereo disparity, *Science* **194,** 283–287 (1976).
4. W. Grimson, *From Images to Surfaces.* MIT Press, Cambridge, MA (1981).
5. S. Ullman, *The Interpretation of Visual Motion.* MIT Press, Cambridge, MA (1979).
6. B. K. P. Horn, Determining optical flow, *Artif. Intell.* **17,** 185–203 (1981).
7. K. Ikeuchi and B. K. P. Horn, Numerical shape from shading and occluding boundaries, *Artif. Intell.* **17,** 141–184 (1981).
8. R. J. Woodham, Photometric method for determining surface orientation from multiple images, **19,** 139–144 (1980). (Journal title not supplied by authors.)
9. A. Pentland, Local shading analysis, SRI International Technical Note 272 (1982).
10. A. P. Witkin, Recovering surface shape and orientation from texture, *Artif. Intell.* **17,** 17–45 (1981).
11. J. R. Kender, A computational paradigm for deriving local surface orientation from local textural properties, *Proc. IEEE Workshop on Computer Vision,* Rindge, NH, pp. 143–152 (1982).
12. S. A. Shafer and T. Kanade, Using shadows in finding surface orientations, *Comput. Vision Graphics Image Process.* **22,** 145–176 (1983).
13. A. Huertas, Using shadows in the interpretation of aerial images, University of Southern California ISG Report 104, R. Nevatia, ed., pp. 55–84 (1983).
14. K. A. Stevens, The visual interpretation of surface contours, *Artif. Intell.* **17,** 47–73 (1981).
15. M. Brady and A. Yuille, An extremum principle for shape from contour, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-6,** 288–301 (1984).
16. D. Marr, Early processing of visual information, *Phil. Trans. R. Soc. B* **275,** 483–534 (1976).
17. D. Marr and E. C. Hildreth, Theory of edge detection, *Proc. R. Soc. B* **207,** 187–217 (1980).
18. H. Barrow and J. M. Tenenbaum, Recovery of intrinsic scene characteristics from images, *Computer Vision Systems,* A. Hanson and E. Risemen, pp. 3–26. Academic Press, New York (1978).
19. R. W. Ehrich and J. P. Foith, Topology and semantics of intensity arrays, *Computer Vision Systems,* A. Hanson and E. Risemen, eds, pp. 111–128. Academic Press, New York.
20. R. M. Haralick, Edge and region analysis for digital image data, *Comput. Graphics Image Process.* **12,** 60–73 (1980).
21. R. M. Haralick and L. T. Watson, Facet model for image data, *Comput. Graphics Image Process.* **15,** 113–129 (1981).
22. R. M. Haralick, Digital step edges from zero crossing of second directional derivatives, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-6,** 58–68 (1984).
23. R. M. Haralick and Y. Yasuoka, Peak noise removal by a facet model, *Pattern Recognition* **16,** 23–29 (1983).
24. T. C. Pong, L. G. Shapiro, L. T. Watson and R. M. Haralick, Experiments in segmentation using a facet model region grower, *Comput. Vision Graphics Image Process.* **25,** 1–23 (1984).
25. O. A. Zuniga and R. M. Haralick, Corner detection using the facet model, *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* Washington, DC pp. 30–37 (1983).
26. R. M. Haralick and J. S. Lee, The facet approach to optic flow, *Proceedings of the Image Understanding Workshop,* Washington, DC (1983).
27. T. C. Pong, Determining intrinsic scene characteristics from images, Virginia Polytechnic Institute and State University, Ph.D. Dissertation (1984).
28. R. Lumia, A new connected components algorithm for virtual memory computers, *Proc. IEEE Conference on Pattern Recognition and Image Processing,* pp. 560–565 (1982).
29. P. G. Mulgaonkar, Analyzing perspective line drawings using hypothesis based reasoning, Virginia Polytechnic and State University, Ph.D. Dissertation (1984).

About the Author—TING-CHUEN PONG was born in Hong Kong on 27 June 1957. He received a B.S. in Physics and Mathematics from the University of Wisconsin, Eau Claire, in 1978, and M.S. and Ph.D. degrees from Virginia Polytechnic Institute and State University, in 1981 and 1984, respectively, both in Computer Science. He is presently an Assistant Professor of Computer Science at the University of Minnesota, Minneapolis. His research interests include computer vision, image processing, artificial intelligence and computer graphics.

Dr Pong is a member of the IEEE Computer Society and the Association for Computing Machinery.

About the Author—LINDA G. SHAPIRO was born in Chicago, Illinois, in 1949. She received a B.S. degree in Mathematics from the University of Illinois, Urbana-Champaign, in 1970, and M.S. and Ph.D. degrees in Computer Science from the University of Iowa, Iowa City, in 1972 and 1974, respectively.

She was an Assistant Professor of Computer Science at Kansas State University, Manhattan, from 1974 to 1978, an Assistant Professor of Computer Science at Virginia Polytechnic Institute and State University, Blacksburg, from 1979 to 1981, and an Associate Professor at Virginia Polytechnic Institute from 1981 to 1984. She is currently Director of Intelligent Systems at Machine Vision International in Ann Arbor, Michigan. Her research interests include computer vision, pattern recognition, intelligent spatial information systems, computer graphics, and data structures. She has completed an undergraduate textbook on data structures with R. Baron.

Dr Shapiro is currently editor of *Computer Vision, Graphics, and Image Processing.* She is a senior member of the IEEE Computer Society, and a member of the Association for Computing Machinery, the Pattern Recognition Society and the American Association for Artificial Intelligence.

About the Author—ROBERT M. HARALICK received a Ph.D. degree from the University of Kansas, Lawrence,

in 1969, where he served as Professor until 1979. From 1979 to 1985 he was a Professor in the Department of Electrical Engineering and Computer Science and Director of the Spatial Data Analysis Laboratory, Virginia Polytechnic Institute and State University, Blacksburg. He is currently Vice President of Research at Machine Vision International. Ann Arbor, MI. He has done research in pattern recognition, multi-image processing, remote sensing, texture analysis, image data compression, clustering, artificial intelligence and general system theory.

Dr Haralick is an Associate Editor of the *IEEE Transactions on Systems, Man, and Cybernetics, Pattern Recognition* and *Computer Graphics and Image Processing.* He is on the Editorial Board of the *IEEE Transactions on Pattern Analysis and Machine Intelligence* and is the Chairman of the IEEE Computer Society Technical Committee on Pattern Analysis and Machine Intelligence. He is the Computer Vision and Image Processing Editor of *Communications of the ACM.* He is also a member of Association for Computing Machinery, the Pattern Recognition Society and the Society for General System Research.