LOCAL COMPUTATION OF SHAPE

Alex Paul Pentland
Artificial Intelligence Center
SRI International, 333 Ravenswood Ave, Menlo Park, CA. 94025

ABSTRACT

Local, parallel analysis of shading in an image may be used to infer surface shape without *a priori* knowledge about the viewed scene. The following has been proven: Given a reasonable assumption about surface smoothness, local analysis of the second derivative of image intensity allows the image-plane component of surface orientation to be determined exactly, and a maximum-likelihood estimate of the remaining depth component of orientation to be made. An algorithm has been developed to compute surface orientation, in parallel, for each image point without knowledge of scene characteristics. This algorithm has been evaluated on both natural and synthesized images, and produces a good estimate of shape.

I. Introduction

A spatially restricted analysis of a single image is logically the first stage of any visual system. This analysis might be relatively simple, e.g., measuring image intensity, or it might be quite complex. This first stage of analysis is especially important because it determines the information that is available to the remainder of the visual system, and, therefore, determines the requirements for the remainder of the system. If the first stage of analysis produces a rich description of the world, then the remainder of the visual system will be much simpler than if it had to deal directly with all of the ambiguities of the image. Determining as much about the world as is possible is therefore important at this first stage of processing. If a visual system can possibly calculate shape from local image information ir direct, bottom-up fashion then it should take full advantage that possibility.

When we examine a small neighborhood around a point an image, usually all we find in the neighborhood is small chan in shading (changes in image intensity). Finding a contour witl the neighborhood of a particular image point is unusual. Tl if we are to investigate what we may learn about surface sha from local examination of an image, we must concern oursel with shading. Thus this paper examines the question: how m surface shape be recovered from an unfamiliar image by usin local analysis of shading?

* Support for the final stages of this research was provided 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 Project Agency of the Department of Defence under Office of Naval Research contract N00014-75-C-0643 and in part by National Science Foundation Grant MCS77-07569.

II. Previous Work

Horn and his colleagues (e.g., [1], [2]) have developed seve numerical integration schemes for using image intensity to so for object shape given *a priori* knowledge of (1) the distributi of illumination, (2) the bidirectional reflectance function of t object's surface, and (3) the surface orientation along an init curve on the object's surface. These shape-from-shading te niques may be useful in situations for which there is suffici information known *a priori* about the image, such as in a f tory setting in which the illumination and the surface reflecta function is known beforehand.

However, because Horn's shape-from-shading theory sumes *a priori* knowledge of the scene, it fails to answer the qu tion of how to determine shape in an unfamiliar image. Furth none of these techniques satisfy our requirement of using o local image information; Horn's shape-from-shading techniq function by propagating constraint from boundary conditi (such as provided by smooth occluding contours) over the s face whose shape is to be estimated. Thus further analysis required.

III. The Estimation Of Surface Orientation

The problem of estimating surface shape from local inf mation is, essentially, the problem of determining the unkno surface normal, **N**, from image measurements of image intensity, I, or its derivatives. Solving for the unknowns in any system of equations requires having more measurements than unknowns. Because the surface normal **N** has two degrees of freedom, we require least two measurements at each image point to obtain a solution.

Therefore, we cannot determine the surface normal using image intensity alone, because it provides only one measurement per image point. More measurements per point can be obtained from the first, second, or higher derivatives of image intensity; however, while each additional derivative does give more measurements per point, it also brings an even larger number of unknowns into the equations. Therefore, we cannot solve for **N** simply by using derivatives to obtain more measurements per point; in order to solve for **N** locally, we must make some simplifying assumptions so that the number of unknowns is not larger that the number of measurements available to us.

A. The Tilt Of The Surface

When we observe a smooth surface, we obtain a strong impression of the *tilt* of the surface — that is, which direction the surface slants away from us. Because we have a strong impression of the image-plane component of surface orientation, it might be possible to compute the surface tilt directly. If we could determine the tilt of the surface exactly, then only one degree of freedom (slant, the depth-component of surface orientation) would be left undetermined in **N**.

How we might go about estimating the tilt? Imagine that we could observe the lines of curvature on a surface directly. These lines of curvature would look like the lines drawn in Figure 1. If we were looking straight down on a surface with no twist, the lines of curvature would appear perpendicular, as in Figure 1 (a). As we tilted the surface to one side, the lines of curvature would appear progressively more spread, as in Figure 1 (b) and (c). Different directions of tilt would cause spreading in different directions.

We can not 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, and I_{xy}, the "spread" of those curvatures. Just as with the spread of the lines of curvature, the direction in which this spread term is the greatest is also the direction of the surface tilt. The direction in which the spread is the greatest is also the direction along which d^2I is the greatest, and so the following proposition:

Proposition (Tilt of the Surface) Given an image of a smooth, homogeneous second-order surface with $I_{xx} \neq I_{yy}$, $I_{xy} \neq 0$, then the tilt of the surface is the image direction in which the second derivative of image intensity, d^2I, is greatest.

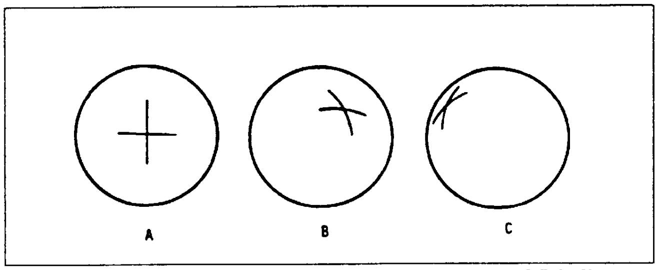

Figure 1 The manner in which image curvature "spreads" indicates the tilt of the surface. This may be understood by imagining that we could observe the lines of curvature on a surface directly. These lines of curvature would look just like the lines drawn in this figure. If we were looking straight down on a surface with no twist, the lines of curvature would appear perpendicular, as in (a). As we tilted the surface off 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.

Thus one of the two components of surface orientation, the tilt, may be determined from the second derivative of image intensity directly *without knowledge of the illuminant direction.* This leaves only the slant of the surface to be determined.

This proposition assumes that the surface we are observing is a second-order surface, that is, a surface that may be described by the Monge patch $\mathbf{p} = u\mathbf{e}_1 + v\mathbf{e}_2 + f(u,v)\mathbf{e}_3$, where $f(u,v)$ is of the form* $\alpha_1 u^2 + \alpha_2 v^2$. The approximation of an arbitrary surface by such a surface typically causes errors on the order of δ^3, where δ is the spacing between observed points on the surface. Thus if ten points are observed across a surface, the maximum error incurred by this approximation is on the order of $1/1000^{th}$ of the width of the surface. The largest errors occur for "twisting" surfaces, those with $\|\mathbf{p}_{uv}\| >> 0$. For such surfaces, the error can be on the order of $1/100^{th}$ of the width of the surface [3]. Thus the assumption that we are observing such a second-order surface does not introduce much error.

B. The Slant Of The Surface

Pentland [4] has proven that while the tilt of the surface may be exactly determined, the slant of the surface cannot be completely disentangled from the curvature of the surface. We may still make an unbiased estimate of the surface slant, as developed in the following propositions.

Proposition (Normalized Laplacian Of Image Intensity) Given an image of a smooth, homogeneous second-order surface, then

$$\frac{\nabla^2 I}{I} = -\kappa_n^2 z_N^{-2} - \kappa_m^2$$

where κ_n is the surface curvature along the surface tilt direction and κ_m is the surface curvature in the orthogonal direction, z_N is the z component of the surface

*Linear and constant terms in f may be accounted for by appropriately positioning the coordinate axes.

normal (which is equal to the arc cosie of the surface slant)

What this proposition shows is that $\nabla^2 I/I$ is a function *only* of the squared curvatures of the surface, κ_n^2 and κ_m^2, and the foreshortening. Note that the effects of the illuminant and the surface albedo do not appear in this quantity.

The range of relationships between the surface curvatures and the foreshortening for any particular observed value of $\nabla^2 I/I$ is limited. Therefore, if we were given that the magnitude of the surface curvature had a particular *a priori* distribution, say a uniform distribution, then for any observed value of $\nabla^2 I/I$, we could make a maximum likelihood estimate of the foreshortening. Because the foreshortening is proportional to z_N^{-2}, we then have an estimate of the slant of the surface, i.e., $\cos^{-1}(z_N)$. This leads to the following proposition.

Proposition (Estimation Of Slant) Assuming a uniform distribution of surface curvature the maximum-likelihood estimate of z_N (the z component of the surface normal, equal to the arccosine of the slant of the surface) is

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

where σ_κ^2 is the variance of the distribution of surface curvatures.

Pentland [4] and Bruss [5] demonstrate that at least one degree of freedom will remain undetermined by the shading information. The tilt proposition showed that one of the two parameters of surface orientation can be determined exactly, leaving only the slant undetermined. This proposition gives a maximum likelihood estimate of the slant, which by definition is the minimum-variance unbiased estimate. *Therefore, the slant and the tilt propositions together constitute the best estimate of surface that it is theoretically possible to make from local shading information.* Note that neither the slant estimate nor the tilt estimate require knowledge of the illuminant direction.

IV. Evaluation

On the basis of this theory, an algorithm for estimating the surface shape (the "shape algorithm") was implemented on one of M.I.T.'s LISP machines.*This algorithm has been tested on both synthetic and natural images. Using synthetic images allows the level of performance of the shape algorithm to be checked under ideal conditions, whereas the use of natural images allows the performance of the shape algorithm to be evaluated under the more varied, complex and noisy conditions found in natural scenes. Some examples of the application of this shape algorithm on both synthetic and natural scenes are presented here.

*The algorithm is a straightforward implementation of the calculations in the slant and tilt propositions above.

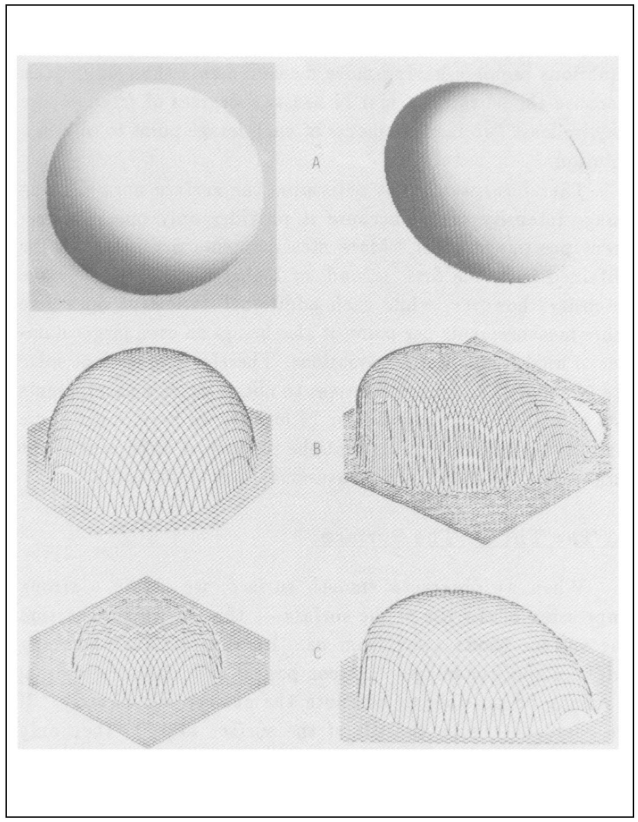

Figure 2 Evaluation on synthetic images. (A) Artificial images of a sphere and an ellipsoidal shape used to test the surface shape estimation algorithm, (B) side view of relief maps showing the true shape. (C) Relief maps showing the estimate of surface shape made by the shape algorithm for the sphere and ellipsoidal shape images. Compare these relief maps to those of 4 (B).

A. Synthetic Images

The shape algorithm was tested using the syntheticly generated images which appear in Figure 2 (A). Figure 2 (B) shows the side view of a relief map for this surface, showing the surface shape which was used to generate this image.

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 adequately evaluate the performance of the algorithm. Therefore, for purposes of displaying the performance of the algorithm, the shape algorithms' estimates of surface orientation were integrated to produce a relief map of the surface. These relief maps were found to give observers an adequate impression of the estimated surface shape, and so they are the output shown for the examples presented in this paper even though integration is not part of the shape algorithm *per se*.

The relief map which results from integrating the shape algorithm's estimate of surface orientation is shown in Figure 2

(C). Comparing the relief map which results from integrating the estimate of surface orientation to the original "true" relief map in Figure 2 (B) shows that the shape algorithm can attain a very high level of accuracy. For a 200 × 200 pixel image of a sphere, using convolutions with 21 × 21 pixel masks to calculate d^2I and $\nabla^2 I$, the correct shape was recovered to within 0.01 %. It is important to remember that these shape were calculated in parallel from purely local image measures, without any *a priori* knowledge of scene characteristics.

B. Natural Images

The shape algorithm has also been tested on several natural images, and two such examples will be presented here. Figure 3 (A) shows the digitized image of a log, together with the relief map generated from the shape algorithms estimates of surface orientation. Figure 3 (B) shows the digitized image of a rock, together with the relief map generated from the shape algorithms estimates of surface orientation. The relief maps in Figure 3 (A) and 3 (B) correspond as closely as can be determined to the actual shapes of these two objects. The reader should compare his impression of shape from the digitized images with the relief maps of Figure 3 (A) and 3 (B).

The shape algorithm has also been successfully employed on other natural images, and on electron microscope images. In each case the estimates of surface shape produced correspond closely to the actual surface shape. These examples demonstrate that the parallel, local computation described in this paper, which does not require any *a priori* knowledge of the scene, is sufficient to obtain a useful estimate of surface shape in natural images.

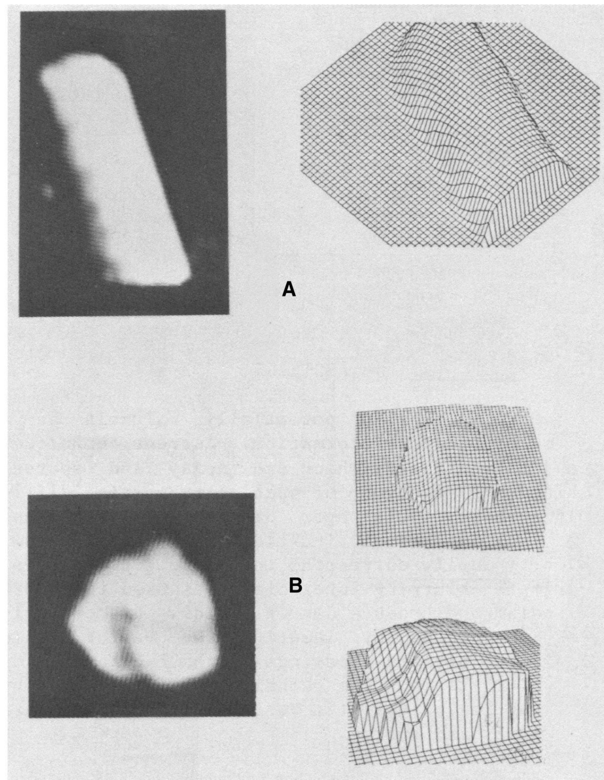

Figure 3 Evaluation on natural images. (A) The digitized log image and the relief map generated by the shape algorithm for that image. (B) The digitized rock image and the relief map generated by the shape algorithm for that image.

REFERENCES

[1] Horn, B.K.P., "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)

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

[3] Issacson, E., and Keller, H. B., "Analysis Of Numerical Methods," (John Wiley and Sons, New York, New York, 1966)

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

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