COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING 44, 139–154 (1988)

On Shape from Shading

ALFRED M. BRUCKSTEIN

Faculty of Electrical Engineering, Technion, IIT, 32000, Haifa, Israel

Received July 21, 1986; revised May 6, 1988

We present a new method for the recovery of a bivariate function $H(x, y)$ that describes a "nice," almost everywhere differentiable height profile, from shading information. The given shading data is assumed to be a result of diffuse, Lambertian reflection of light from the surface. This implies that, if the scene is uniformly illuminated from above, the shading yields information on the cosine of the angle between the vertical and the surface normal at each point. Given the shading information in the plane, the shape from shading problem is to determine all height profiles consistent with the data, and some boundary conditions, such as points of known height and surface orientation, or height profiles along continuous curves in the image plane. The new shape-from-shading method that we discuss is based on a recursive way of determining equal-height or level contours of the surface starting at a given level curve.

1. INTRODUCTION

Shape-from-shading problems have a long and interesting history. The first researchers to address the problem of determining shape from shading information were apparently those concerned with the photometric analysis of the lunar topography (see [5] and the references therein). Several approaches to the analysis of three-dimensional scenes via their images or shading maps were investigated by researchers interested in computer vision. Many stressed the importance of a huge prior-information data-base in scene analysis, and adopted AI-type approaches [11, 17, 20]. Others dealt with the recovery of depth from stereo images [10], from multiple images taken under different illumination conditions [18], or from sequences of images resulting from camera motion [13]. It is clear that the shading information plays, along with stereo vision and motion clues, an important role in the depth perception process. The theoretical question of how much depth information can be obtained from a single view of a scene from shading alone, thus arises quite naturally. Algorithms for determining the shape, or depth-profile from a single image produced by various shading rules were considered by B. K. P. Horn and his coworkers in the 1970's [5, 6, 18]. More recent work on shape determination from a single image concentrated on the importance of singular points, surface models and occluding boundaries, in providing initialization for an algorithm proposed by Horn that recovers height on a data-directed curve in the (x, y) plane called a characteristic strip [5]. Some iterative, relaxation-type techniques were also invented, relying on surface smoothness constraints [1, 7, 8, 16].

In this paper we discuss the basic shape-from-shading problem, stressing the importance of considering the behavior of *equal-height contours*. We show that when one such contour is available we can devise a simple algorithm that reconstructs all the equal-height curves of the surface of interest in a well-defined region and also clearly displays the inherent ambiguities of the given problem. This algorithm is easily derived and, in contrast to the classical characteristic-strip expansion method,

0734-189X/88 $3.00

does not use the derivatives of the shading data. This is a result of a natural way of exploiting lateral constraints in the parallel propagation of the recovery algorithm. We show how the algorithm works on some basic examples and discuss the ambiguities and possible ways to exploit topological constraints on the behavior of nice surfaces to help the shape recovery process in ambiguous situations.

2. THE SHAPE-FROM-SHADING PROBLEM AND SOME BASIC CONCEPTS

Suppose we are given a continuous function of two variables (x, y), $H(x, y)$, describing a surface in three dimensions,

$$z = H(x, y). \tag{2.1}$$

The shaded image of $H(x, y)$ is defined as a light intensity map, $A(x, y)$, so that the value $A(x, y)$ depends on the surface properties, its orientation at (x, y), and the illumination. The shape from shading problem that we address is to recover the function $H(x, y)$ over a region $\mathbb{D}$, from the image $A(x, y)$ over that region and some possible further information, e.g., the values of $H(x, y)$ over some continuous curve in $\mathbb{D}$.

The function $A(x, y)$ is defined via a shading rule. It is customary to define the shading rule via a so-called *reflectivity function*, that characterizes the surface properties and provides an explicit connection between $A(x, y)$ and the surface orientation. A thorough discussion of surface reflectivity properties and various types of reflectivity functions can be found in [6]. In the case of a surface with so-called Lambertian diffuse reflection properties and uniform illumination $A(x, y)$ is simply the cosine of the angle $\alpha(x, y)$, between the surface normal at (x, y) and the direction from which the light falls on the scene. For simplicity we shall always assume that the illumination is uniform and falls on the surface vertically from above, i.e., from $z = +\infty$.

Define the directional derivatives of the height profile $H(x, y)$ along the x and y directions as

$$\begin{aligned} p(x, y) &= \frac{\partial}{\partial x} H(x, y) \\ q(z, y) &= \frac{\partial}{\partial y} H(x, y). \end{aligned} \tag{2.2}$$

The surface normal at (x, y) is clearly perpendicular to the plane determined by the vectors $[1, 0, p]$ and $[0, 1, q]$, therefore it is along the direction of their vector product $[-p, -q, 1]$. The normal vector at (x, y) is thus

$$N(x, y) = \frac{1}{\sqrt{1 + p^2 + q^2}} [-p, -q, 1] \tag{2.3}$$

and the cosine of the angle between $N(x, y)$ and the vertical direction [0 0 1] is

$$\cos\alpha(x, y) = \frac{1}{\sqrt{1 + p^2 + q^2}}. \tag{2.4}$$

In the Lambertian case with light falling perpendicularly from above we therefore have the shading rule

$$A(x, y) = \frac{1}{\sqrt{1 + p^2 + q^2}} = R_L(p, q). \tag{2.5}$$

Note that the reflectivity function R is defined on the (p, q) plane—called the "gradient space." A general (not necessarily Lambertian) shading rule is defined via

$$A(x, y) = R(p(x, y), q(x, y)), \tag{2.6}$$

where $R(\cdot,\cdot)$ is a given function. Equation (2.6) is a nonlinear partial differential equation that has to be satisfied by the surface $H(x, y)$. Therefore solving the shape-from-shading problem amounts to solving a nonlinear partial differential equation, and a set of boundary conditions are necessary.

3. FROM SHADING INFORMATION TO SHAPE

Given the image $A(x, y)$ it is, in general, impossible to unambiguously recover the height profile $H(x, y)$. As an immediate example of ambiguity simply consider the function $-H(x, y)$, which, under a Lambertian shading rule, maps into the same image as $H(x, y)$. Some futher information on the function $H(x, y)$ is therefore needed. This is usually given as some smoothness constraint on the surface defined by $z = H(x, y)$ (for example, C^1 or C^k continuity) and exact or approximate values of $H(x, y)$ at either a discrete set of points $\{(x_i, y_i)\}$, together with the corresponding surface orientations $\{(p_i, q_i)\}$, or on a continuous curve on the (x, y)-plane (boundary conditions). The given boundary conditions and smoothness assumptions are not always enough to remove ambiguities, and it is in fact very difficult to determine, in general situations, sufficient conditions for a unique solution surface [2].

3.1 The Characteristic-Strip Expansion Method

Assume that $z = H(x, y)$ is a smooth surface, i.e., the partials $p(x, y)$ and $q(x, y)$ defined in (2.2), and also second derivatives exist everywhere. Consider now that at some point (x_o, y_o) in the plane we know that height $H(x_o, y_o)$ together with the surface orientation $\{p(x_o, y_o), q(x_o, y_o)\}$. Horn observed that in this case one can determine the height profile and the surface orientation along a well-defined curve in the (x, y)-plane called a characteristic. This name is imported from the theory of partial differential equations. Horn's method is in fact a particular case of a general procedure for solving Cauchy-type boundary value problems associated to nonlinear partial differential equations (see, e.g., [9, 14]). The characteristic curve is entirely determined by propagating a coupled set of differential equations, driven by the shading data $A(x, y) = R(p(x, y), q(x, y))$, from the starting point (x_o, y_o), with initial conditions $\{H(x_o, y_o), p(x_o, y_o), q(x_o, y_o)\}$. Suppose we take a step in the (x, y)-plane away from (x_o, y_o) so that

$$(x, y) = (x_o, y_o) + (\Delta x, \Delta y). \tag{3.1}$$

Then we have for the change in height

$$\Delta H(x, y) = p(x_o, y_o)\Delta x + q(x_o, y_o)\Delta y \tag{3.2}$$

and also

$$\begin{aligned} \Delta p(x, y) &= p_x(x_o, y_o)\Delta x + p_y(x_o, y_o)\Delta y \\ \Delta q(x, y) &= q_x(x_o, y_o)\Delta x + q_y(x_o, y_o)\Delta y. \end{aligned} \tag{3.3}$$

Note that $p_y = q_x$ by the smoothness constraint. If, by using the shading information $A(x, y)$, we find a direction $(\Delta x, \Delta y)$ so that, at the new point, both $H(x, y)$ and (p, q) can be determined, then we have a way of computing the height profile on a data-determined curve in the plane. This idea is exactly what underlies the characteristic strip expansion method. Using the chain rule for differentiation we obtain from (2.6) that

$$\begin{aligned} A_x &= \left[\frac{\partial}{\partial x} A(x, y)\right] = R_p(p, q)p_x + R_q(p, q)q_x \\ A_y &= \left[\frac{\partial}{\partial y} A(x, y)\right] = R_p(p, q)p_y + R_q(p, q)q_y. \end{aligned} \tag{3.4}$$

Now simply note that if

$$(\Delta x, \Delta y) = (R_p \Delta s, R_q \Delta s) \tag{3.5}$$

for some small Δs, we shall have

$$\begin{aligned} \Delta H &= (p \cdot R_p + qR_q)\Delta s \\ \Delta p &= A_x \cdot \Delta s \\ \Delta q &= A_y \cdot \Delta s. \end{aligned} \tag{3.6}$$

This is indeed quite remarkable since (3.6) enables us to propagate for both height and orientation along the curve recursively determined via (3.5). The data-directed curve defined by (3.5) is called a "characteristic strip" and, as we shall see later, in the case of a rotationally symmetric reflectance map, i.e., when $R(p, q) = f(p^2 + q^2)$, the characteristics are lines of steepest ascent on the surface $H(x, y)$.

The result presented above is the basis of Horn's classical shape-from-shading method. He proposed to look at the brightness map $A(x, y)$ and start the height recovery around singular points—where $A(x, y)$ attains the maximum value of 1, i.e., where $p = q = 0$ (under the Lambertian shading rule). From these points, one would propagate the characteristics outwards and in parallel, and one can also use certain neighborhood rules in propagation—such as not allowing crossover of adjacent strips and interpolating new characteristic strips when neighboring strips separate too far. Note that when $p = q = 0$ we have a start-up problem for the algorithm, since (3.5) will not move us out from the singular points. To solve this problem we need to add further assumptions about the behavior of $H(x, y)$ about each such starting point (i.e., to classify singular points as a local maxima or minima; of course, problems arise at saddle points). Implicitly we have to assume the knowledge of the initial slopes (p, q) on a small loop around the singularities [5]. In later work it was stressed that occluding boundaries and other visual clues have a crucial importance in providing just such initial conditions for the start-up of characteristic strip expansions along several curves in parallel. Ikeuchi and Horn have also analyzed a relaxation-type iterative algorithm using second-derivative surface smoothness constraints, together with shading data and occluding boundary information to recover a height profile that minimizes a smoothness measuring cost function. This algorithm uses a special representation of the surface-orientation profile, via the so-called "stereographic projection." It works on a grid of pixels and iteratively assigns the orientations in the stereographic projection plane so as to meet the shading requirements while minimizing orientation changes over neighboring pixels [8]. Recently, Horn and Brooks [7], provided a systematic analysis of variational approaches to the shape-from-shading problem, leading to iterative solutions on pixel grids. Their analysis yields improved algorithms, generalizing the methods previously proposed by Strat, Brooks, and Smith [1, 7, 15]. In the sequel we return to the classical continuous problem and discuss a new shape-from-shading algorithm that uses the conceptual framework of the characteristic strip expansion method but stresses the importance of recursively determining the equal-height contours of the profile $H(x, y)$.

3.2 Shape-from-Shading via Equal-Height Contours

It is clear that Horn's characteristic strip expansion method enables one to use various types of prior information—provided it can readily be translated into height and orientation data at a set of points. As a result many ideas for using this method under various circumstances were advanced, see, e.g., [2, 5, 19]. Others attempted to exploit different ways of formulating surface continuity and smoothness constraints in order to arrive at practical shape-from-shading algorithms that work on pixel-grids directly [1, 7, 8, 15, 16]. It seems, however, that none of the above-quoted works emphasized the potential of using as data and then trying to determine the *equal-height contours* of the profile $z = H(x, y)$. An equal-height contour or a *level curve* is a continuous curve in the (x, y)-plane on which the function $H(x, y)$ is constant. If $\{x(\theta), y(\theta)\}\theta \in \Theta$ is the parametric representation of the curve we have

$$\frac{d}{d\theta}H(x(\theta), y(\theta)) = 0. \tag{3.7}$$

One might argue that such a contour contains a lot of information and is scarcely available to us. This is true; however, many of the previously proposed algorithms for recovering shape-from-shading require very similar types of prior information (like the surface orientations, or the height profile on some curve in the image plane, or occluding boundaries). Also in some practical situations one is quite naturally able to determine equal height contours, or portions of it. As an example, the shores of a lake in a landscape readily provide a closed equal-height contour; as is the case when an island raises from the sea (any shoreline is an equal height curve). Furthermore, in robot vision systems one might be able to provide illumination which actively delineates one or more equal height contours.

In the sequel the assumption will be that we are given an equal height contour which is almost everywhere differentiable. By definition, along such a curve we have zero height gradient, which yields

$$dH = p\,dx + q\,dy = 0. \tag{3.8}$$

Therefore along the given contour $\{x(\theta), y(\theta)\}$ we have determined a relation between the two directional derivatives of the surface p and q. For almost all θ's we have, rewriting (3.8),

$$p(x(\theta), y(\theta))\frac{d}{d\theta}x(\theta) = -q(x(\theta), y(\theta))\frac{d}{d\theta}y(\theta). \tag{3.9}$$

Together with the shading information at that point $A(x(\theta), y(\theta))$, this relation determines p and q, up to an inherent sign ambiguity. Indeed the Lambertian shading rule

$$A(x, y) = R(p, q) = \frac{1}{\sqrt{1 + p^2 + q^2}} \tag{3.10}$$

yields $p^2 + q^2 = (1 - A^2)^{1/2}/A$, which together with $px' + qy' = 0$, provides

$$\begin{aligned} p &= \pm\frac{y'}{\left(x'^2 + y'^2\right)^{1/2}}\frac{\left(1 - A^2\right)^{1/2}}{A} \\ q &= \mp\frac{x'}{\left(x'^2 + y'^2\right)^{1/2}}\frac{\left(1 - A^2\right)^{1/2}}{A} \qquad \text{at } [x(\theta), y(\theta)]. \end{aligned} \tag{3.11}$$

We get two pairs of solutions, corresponding to a certain (p, q) vector and its negative counterpart. This is expected, since at each point on the equal height curve the same grey level would be produced by the shading rule if the tangent plane had the direction of maximum ascent given either by ϕ or by $\phi + \pi$. Note also that we could determine the (p, q) pairs up to a similar ambiguity along *any continuous path* on which the height profile is known a priori. In case of equal heights curves, the direction of the data contour determines the direction of the maximal surface ascent/descent. Suppose we know that the height profile is a mountain raising from the sea. This immediately settles the direction of the steepest ascent as the vector pointing towards the inner region defined by the equal height contour of the shorelines. Using this information we may determine an equal height contour situated a bit above the sea level and so on, we can recursively climb and reconstruct the height profile—provided no "problems" occur. Problems arise, as we shall see, if the mountain is not a nice and unimodal profile, and we further discuss these issues after a description of the basic profile reconstruction algorithm.

From a Level Curve to the Ones Nearby

Assume that $(x(\theta), y(\theta)]\theta \in [0,1]$ is a closed curve and that, as θ goes from 0 to 1 we trace the curve in the counterclockwise direction (see Fig. 1).

A tangent vector at θ is simply given by $[x'(\theta), y'(\theta)]$; the unit normal to it pointing inside the curve will be

$$\mathbf{n}_\theta = \frac{1}{\left[x'(\theta)^2 + y'(\theta)^2\right]^{1/2}}\left[-y'(\theta), x'(\theta)\right]. \tag{3.12}$$

From (3.9) it is clear that in the direction n_θ, we have to go a certain distance d_θ, in order to *climb* a given amount ΔH. If ΔH is small, this distance is quite accurately determined by the shading data alone, since, in the Lambertian example, $A(x(\theta), y(\theta))$ yields the cosine of the angle between the surface normal at $(x(\theta), y(\theta))$ and the vertical direction. As the direction of the maximal ascent is known to be (3.12), we have (see Fig. 1) from geometrical considerations only, that

$$d_\theta = \Delta H \frac{A}{\sqrt{1 - A^2}}. \tag{3.13}$$

Note that the same result would be obtained by writing (3.8), and substituting (3.11)

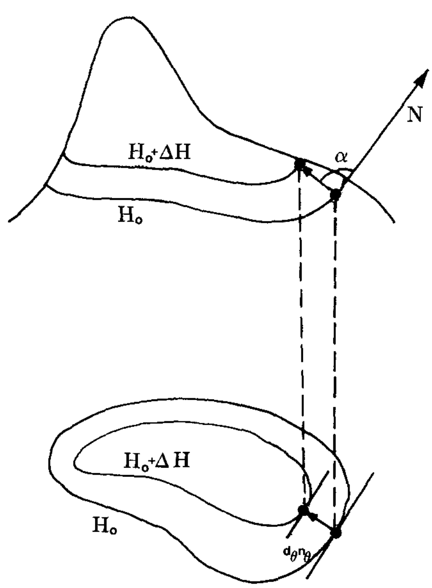

FIG. 1. Description of the level-contour climbing process on a simple unimodal height profile.

into it. Then

$$\Delta H = p\,dx + q dy = \frac{\sqrt{1 - A^2}}{A} \frac{1}{\sqrt{x'(\theta)^2 + y'(\theta)^2}} [-y'(\theta)\,dx + x'(\theta) dy]$$

and using the requirement that $[dx, dy] = d_\theta \mathbf{n}_\theta$, obtain

$$\Delta H = \frac{\sqrt{1 - A^2}}{A} d_\theta$$

which again yields (3.13). This second derivation is also more general than the first, geometric argument, and would also work in the case of height data provided along any differentiable curve in the image plane (not necessarily an equal-height curve). We shall briefly return to this point later.

Therefore, given a closed equal-height contour assumed to be at a reference level H_o, a closed contour situated at the level $H_o + \Delta H$ is determined via

$$\begin{aligned} [x(\theta, \Delta H), y(\theta, \Delta H)] &= [x(\theta), y(\theta)] + d_\theta \mathbf{n}_\theta \\ &= [x(\theta), y(\theta)] + \frac{1}{\sqrt{x'(\theta)^2 + y'(\theta)^2}} \\ &\quad \times \frac{\Delta H \cdot A_\theta}{\sqrt{1 - A_\theta^2}} [-y'(\theta), x'(\theta)]. \end{aligned} \tag{3.14}$$

Our derivation leads to a system of first-order nonlinear partial differential equations for the functions $x(\theta, h)$ and $y(\theta, h)$ representing "doubly parametrized" equal-height curves in the (x, y) plane. Indeed, if $[x(\theta, h), y(\theta, h)]$ is defined as a contour corresponding to $H = h$, (3.14) is equivalent to the set of partial differential equations

$$\begin{aligned} \frac{\partial}{\partial h}\begin{bmatrix} x(\theta, h) \\ y(\theta, h) \end{bmatrix} &= \frac{A(x(\theta, h), y(\theta, h))}{[1 - A^2(x(\theta, h), y(\theta, h))]^{1/2}} \\ &\quad \times \frac{1}{\left[\left[\frac{\partial}{\partial\theta}x(\theta, h)\right]^2 + \left[\frac{\partial}{\partial\theta}y(\theta, h)\right]^2\right]^{1/2}} \begin{bmatrix} -\frac{\partial}{\partial\theta}y(\theta, h) \\ \frac{\partial}{\partial\theta}x(\theta, h) \end{bmatrix} \end{aligned} \tag{3.15}$$

with initial conditions $[x(\theta, 0), y(\theta, 0)] = [x(\theta), y(\theta)]$.

Note that (3.15) is a *nonlinear initial value problem* that has to be integrated to obtain the equal-height curves of the profile that yield the shading $A(x, y)$. It is implicit in our derivations that the surface is smooth enough to provide almost everywhere differentiable equal-height contours at all heights h. It is also assumed that those contours are "well-behaved" as, for example, in the case of a unimodal

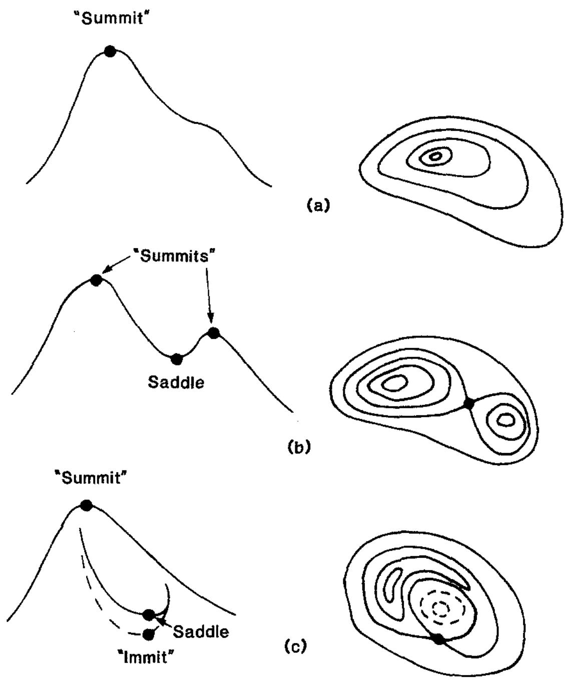

FIG. 2. The simplest possible types of level profiles.

$H(x, y)$ over the region of interest (say the interior of the first equal-height contour), when they are nicely nested "generalized" rings (see Fig. 1).

It is clear that the recursions (3.14) and their differential counterparts (3.15) are valid generally, provided we are given information on which side of the original equal height curve the surface increases. The data $[x(\theta, y(\theta)]$ can be any curve that is differentiable and if it is not a closed contour we will get, using (3.14), the reconstruction of a well-defined slice of the surface $z = H(x, y)$. If we do start with a closed contour and at some level we obtain a self-intersecting (i.e., not "well behaved") equal-height curve, this means that we encountered a saddle area which separates peaks, or peaks and dips, in $H(x, y)$. In this case the contour should be separated into nonintersecting parts and the algorithm may be continued with the separated closed parts as initial equal-height curves. An equal-height curve may also approach a saddle point from one side only, and there it will become necessary to continue a partial reconstruction (see Fig. 2 for types of level contour profiles that may be encountered).

A thorough discussion of what can happen to the equal-height contour profile of a smooth surface, based on topological constraints can be found in two classical papers by Cayley [3] and Maxwell [12]. In a modern interpretation, see [4]; Maxwell proves the so-called mountaineers' theorem, stating that if a surface has isolated simple singularities, i.e., "summits" (local maxima), "immits" (local minima), or saddle points, then within an equal-height contour we must have that

$$\#\text{of summits} + \#\text{of immits} - \#\text{of saddle points} = 1. \qquad (3.16)$$

This shows that the profiles shown in Fig. 2 are indeed the simplest cases, modulo a height inversion, since they correspond to one and two local extrema, respectively.

The practical implementation of the algorithm will, of course, be based on (3.14), the $[x(\theta), y(\theta)]$ curve being given (perhaps in a suitably chain-coded way) on a finite grid of θ values. Then we can use several methods to estimate the derivatives $x'(\theta)$ and $y'(\theta)$ that appear in the recursion formula. Also we can leave open the choice of the steps in height (ΔH) taken so as to enable the use of various adaptive schemes (that are particularly useful if the approach to a saddle area is detected). We also note that in practice, $A(x, y)$ is known only on a grid of pixels, thus we have to somehow interpolate for the values needed in (3.14) which are points situated on equal-height curves.

It becomes clear by looking at the system of Eqs. (3.15) that trouble arises when we approach singular point where $A(x, y) = 1$, indicating $p = q = 0$. A singular point can be either a local extremum or a saddle point of some sort. At an isolated singular point we can define an "unsafe" neighborhood and when an equal-height curve enters such a neighborhood, we disregard that portion of it but continue to propagate the algorithm from the remaining contour. Some portions of the (x, y) plane will, of course, remain uncovered using this method. The "singular" areas/curves in the plane for which $A(x, y) = 1$ provide boundaries of possible flips in the directions of maximal ascent and a practical shape-from-shading process should keep track of these and, based on natural constraints on the behavior of equal-height contours, choose the direction assignments which yield consistent final reconstructions. If a priori we know that the surface is unimodal then no such problems arise, the solution being unique up to height reversal (see also Bruss [2]). Note also that we can live with nondifferentiability at a finite set of points along each equal-height contour and reconstruct the profile by matching the slices corresponding to differentiable portions.

Level Curves and Characteristic Strips

The differential equations governing the evolution of equal height contours are related to the evolution of characteristic strips. The characteristic-strip expansion method, in the case when $R(p, q) = R(p^2 + q^2)$ (e.g., for Lambertian surfaces) determines a path of maximal ascent/descent on the height profile $H(x, y)$. This is so since we have

$$[dx, dy] = [R_p, R_q]\,ds = -\frac{1}{(1 + p^2 + q^2)^{3/2}}[p, q]\,ds; \tag{3.17}$$

therefore the vector $[dx, dy]$ points to the direction of maximal ascent. We also have that

$$dh = [pR_p + qR_q]\,ds = -\frac{p^2 + q^2}{(1 + p^2 + q^2)^{3/2}}\,ds. \tag{3.18}$$

In terms of the differential climb dh we can express the direction of the characteristic strip expansion as

$$[dx, dy] = \frac{1}{p^2 + q^2}[p, q]dh. \tag{3.19}$$

This yields

$$\frac{d}{dh}\begin{bmatrix} x \\ y \end{bmatrix} = \frac{\sqrt{p^2 + q^2}}{p^2 + q^2}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \frac{1}{\sqrt{p^2 + q^2}}\begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \tag{3.20}$$

where $[\alpha, \beta]$ is the unit vector in the direction of the maximal ascent. But we have $A = 1/\sqrt{1 + p^2 + q^2}$, thus $p^2 + q^2 = (1 - A^2)/A^2$ and therefore we have

$$\frac{d}{dh}\begin{bmatrix} x \\ y \end{bmatrix} = \frac{A}{\sqrt{1 - A^2}}\begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \tag{3.21}$$

where the vector $[\alpha, \beta]$, may be determined, if a θ-parameterized equal height curve is available, as

$$\frac{1}{\sqrt{x'(\theta, h)^2 + y'(\theta, h)^2}}\frac{\partial}{\partial\theta}\begin{bmatrix} -y(\theta, h) \\ x(\theta, h) \end{bmatrix}.$$

Therefore, if we ask for the evolution of level contours, parameterized as $[x(\theta, h), y(\theta, h)]$ we again obtain (3.15),

$$\frac{\partial}{\partial h}\begin{bmatrix} x(\theta, h) \\ y(\theta, h) \end{bmatrix} = \frac{a}{\sqrt{1 - A^2}}\frac{1}{\sqrt{x'(\theta, h)^2 + y'(\theta, h)^2}}\frac{\partial}{\partial\theta}\begin{bmatrix} -y(\theta, h) \\ x(\theta, h) \end{bmatrix}.$$

This shows that the curves defined by $[x(\theta_o, h), y(\theta_o, h)$ for a given fixed θ_o are identical to the characteristic strips. However, note that in the characteristic strip expansion method the parameterization by s is not simply related to the height h. Although Horn noted that scale changes are possible to give various natural interpretations to the s-parameter like arc length, etc.; the possibility of using the height h as a parameter equivalent to s was never mentioned. More importantly, the reconstruction of the characteristic strips uses the derivatives of the data field $A(x, y)$ instead of the local direction of the level curves. The use of level contour directions is in fact a very direct exploitation of lateral information if such information is available. In [5] Horn proposes another way to use lateral information as obtained via parallel propagation of many characteristic strips emanating from the neighborhood of a singular point, assumed to be a local maximum or minimum. He defines "rings of equal arc length" s about the singular point and uses those rings to interpolate for new characteristic strips when those propagated so far separate too much. The level curve approach is a more natural implementation of this idea. What we are doing is applying a special control on the "speed of expansion" of the characteristic strips to ensure that their "wavefront" is an equal height curve, and use this wavefront to determine the directions of the local maximal ascent/descent on the surface. This replaces the need to differentiate the picture intensity information $A(x, y)$ with natural "laterality" constraints. Differentiating a noisy intensity map would clearly lead to very bad effects in the propagation of characteristic strips; therefore the above-described procedure has the potential to significantly improve the numerical properties of the surface reconstruction process.

3.3 Remarks on Extensions of the Shape-from-Shading Algorithm

Recovering General Parametrized Curves on the Surface

The approach outlined above is clearly not limited to level contour initial conditions and level curve propagation. If we parameterize a general curve on the surface by $\{x(\theta, s), y(\theta, s), z(\theta, s)\}$ where $s = 0$ corresponds to the given data curve we could also propagate the set of equations

$$\frac{\partial}{\partial s}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} R_p \\ R_q \\ pR_p + qR_q \end{bmatrix}, \tag{3.22}$$

together with the nonlinear solutions for (p, q) of the equations

$$\begin{aligned} \frac{\partial}{\partial \theta} z(\theta, s) &= p\frac{d}{d\theta}x + q\frac{d}{d\theta}y \\ A(\theta, s) = R(p, q) &= \frac{1}{\sqrt{1 + p^2 + q^2}}. \end{aligned} \tag{3.23}$$

This, of course, is much more tedious and does not have as natural an interpretation as using level contours.

Shape-from-Shading under Perspective Viewing

If we assume that the image $A^P(u, v)$ is the shaded image, when viewing a Lambertian surface under the perspective transformation, the image plane coordinates (u, v) of a point $(x, y, z(x, y))$ are related to the true spatial coordinates (x, y) as follows:

$$u = x/z(x, y) \qquad \text{and} \qquad v = y/z(x, y). \tag{3.24}$$

In the above equations it was, of course, assumed that the function $z(x, y)$ measures depth from the viewing plane and the focal distance is unity. Given a parametrized curve in the image plane, $(u(\theta), v(\theta))$, that is known to be an equal-height contour, we have by (3.24) that the corresponding contour in the (x, y)-coordinates is a scaled version of the image–plane curve. The scaling factor is $1/z$.

The direction of the maximal ascent/descent is also immediately obtained in both image and true coordinates, being perpendicular to the tangent to the equal-height curve. The slope information is, as before, contained in the image intensity $A(u(\theta), v(\theta))$, and from it we can determine an equal-height curve nearby. The reconstruction algorithm will proceed to recursively determine equal-height contours, exactly as before. Assuming a known initial height, the scaling factors for each level curve is also recursively determined. This process replaces the strip-expansion algorithm of Horn [5] that requires the propagation of five differential equations in the case of perspective projection.

4. CONCLUDING REMARKS

A new shape-from-shading algorithm was discussed. The numerical properties of our shape-from-shading procedure are currently under investigation. Some results

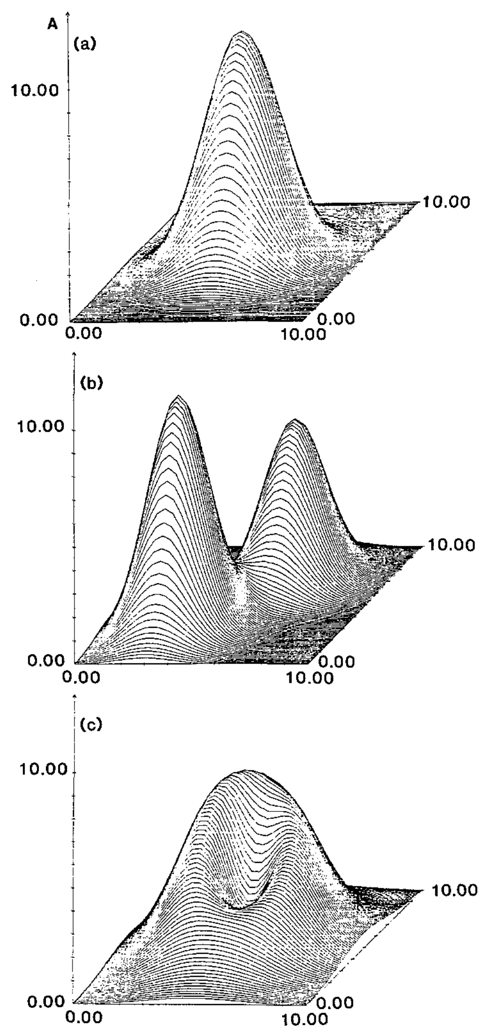

FIG. 3. Algorithm testing on the three types of basic level profiles: A. the level profiles; B. the level-climbing algorithm (with adaptive ΔH); C. comparison of level-climbing with characteristic strip expansion.

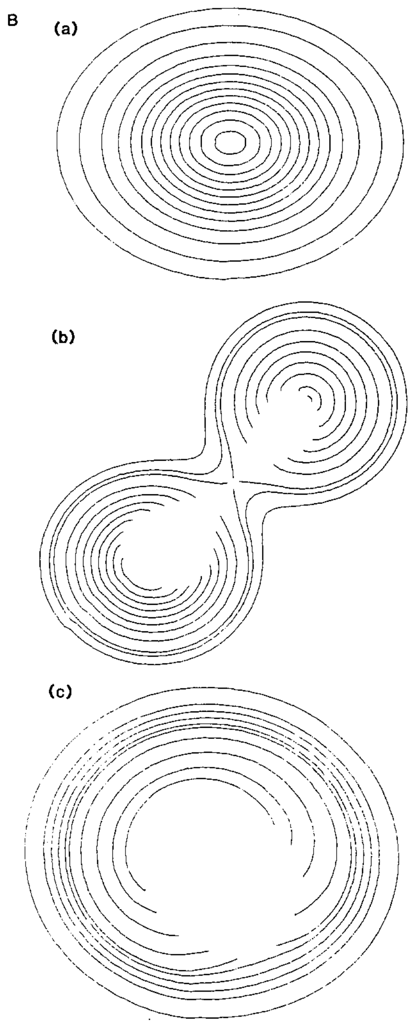

FIG. 3.—*Continued.*

are shown in Fig. 3. Figure 3B shows the behavior of the equal-height contour propagation algorithm on three typical examples: a gaussian profile, a sum of two displaced gaussians, and a surface resulting from subtracting a noncentral gaussian from a centrally located one (Fig. 3A). Figure 3C compares the characteristic strips propagated with identical data to the equal-height curves resulting from our algorithm. As we can see the behavior of both algorithms is good, the third profile leading to a significant uncovered region, as expected, due to the lateral encounter of an equal-height contour with a saddle point. From the preliminary results obtained by running these algorithms on this set of examples, with precise image irradiance information, we can conclude that their behavior is very good on unimodal or simple shapes having few saddle points located between local maxima or minima. However, more sophisticated backtracking or more side-information is needed when many singular points exist in the regions of interest. In this context, a

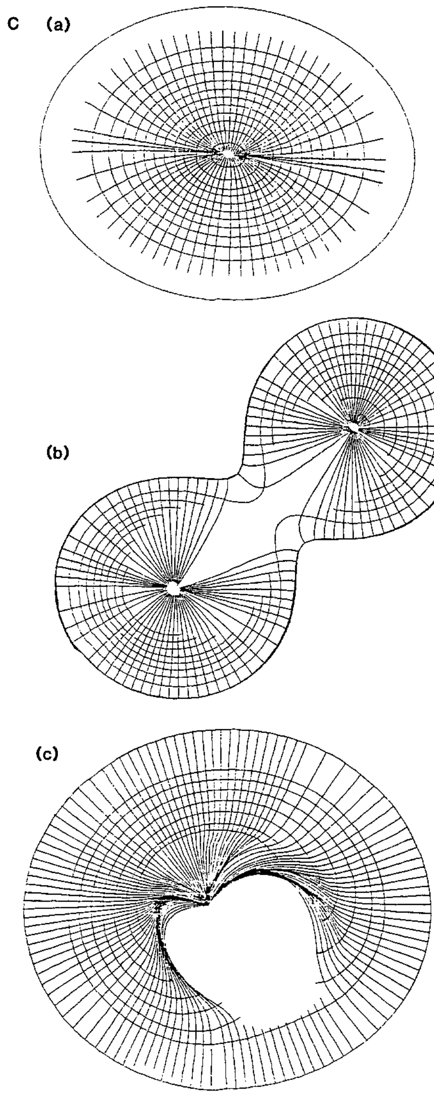

FIG. 3.—*Continued.*

topological investigation of surfaces providing the relationships that exist between local maxima, minima, saddle points, and the regions surrounding them is most helpful. The results of Maxwell and Cayley on passages between "hills and dales," [3, 12] could indeed provide the basis for an intelligent algorithm incorporating topological reasoning with partial differential equations based shape-from-shading processes as discussed in this paper.

ACKNOWLEDGMENTS

I wish to thank Professor Ronald Bracewell of Stanford University for a discussion in which he posed the question whether a surface may be recovered from the cosines of the angles between the normals and the vertical direction. Figure 3 is taken from the M.Sc. project of Mr. Marc Smadja, at the Technion. I thank him for implementing and testing the algorithms discussed in this paper. Also, I wish to thank the anonymous referees for their very useful and constructive comments.

REFERENCES

1. M. J. Brooks, Surface normals from closed path, in *Proceedings 6th IJCAI, Tokyo, Japan*, 1979, pp. 98–101.
2. A. R. Bruss, The eikonal equation: Some results applicable to computer vision, *J. Math. Phys.* **23**, No. 5, 1982, 890–896.
3. A. Cayley, On contour and slope lines, *Philos. Mag.* October 1859, 264–268.
4. H. B. Griffith, *Surfaces*, 2nd ed., Cambridge Univ. Press, Cambridge, 1981.
5. B. K. P. Horn, Obtaining shape from shading information, in *The Psychology of Computer Vision*, (P. H. Winston, Ed.), pp. 115–155, McGraw–Hill, New York, 1975.
6. B. K. P. Horn, Understanding image intensities, *Artif. Intell.* **8**, 1977, 201–231.
7. B. K. P. Horn and M. J. Brooks. The variational approach to shape from shading, *Comput. Vision Graphics Image Process.* **33**, 1986, 174–208.
8. K. Ikeuchi and B. K. P. Horn, Numerical shape from shading and occluding boundaries, *Artif. Intell.* **17**, 1981, 141–184.
9. F. John, Partial differential equations, in *Applied Mathematical Sciences*, 4th ed., Springer Verlag, New York, 1982.
10. H. C. Longuet-Higgins, A computer algorithm for reconstructing a scene from two projections, *Nature* **293**, 1981, 133–135.
11. D. Marr, *Vision*, Freeman, San Francisco, 1982.
12. J. C. Maxwell, On hills and dales, *Philos. Mag.* (Maxwell's Collected Works), 1870, 233–240.
13. K. Prazdny, On the information in optical flows, *Comput. Vision Graphics Image Process.* **22**, No. 9, 1983, 239–259.
14. V. I. Smirnov, *A Course in Higher Mathematics*, Vol IV, Chap. III, Pergamon, Oxford, 1964.
15. G. Smith, The recovery of surface orientation from image irradiance, in *Proceedings DARPA Image Understanding Workshop, Palo Alto, September*, 1982.
16. D. Terzopoulos Efficient multiresolution algorithms for computing lightness, shape-from-shading, and optical flow, in *Proceedings 4th National Conference of Artificial Intelligence, Austin, Texas*, 1984, pp. 314–317.
17. P. H. Winston (Ed.), *The Psychology of Computer Vision*, McGraw–Hill, New York 1975.
18. R. J. Woodham, Photometric method for determining surface orientation from multiple images, *Opt. Eng.* **19**, 1980, 139–144.
19. R. J. Woodham, Analyzing images of curved surfaces, *Artif. Intell.* **17**, 1981, 117–140.
20. Special issue, *Artif. Intell.* **17**, 1981.