SHAPE from SHADING

edited by

Berthold K. P. Horn

and

Michael J. Brooks

Contents

1

Introduction

The image of the world projected onto the retina is essentially two-dimensional. From this image we recover information about the shapes of objects in a three-dimensional world. How is this done? There are numerous cues available to help us recover the missing dimension. When we move around, for example, images of nearby objects are displaced more rapidly on the retina than are images of distant objects. This so-called *motion parallax* effect provides one important depth cue. Another is provided by *binocular disparity*, the difference between the relative positions of corresponding features resulting from the spatial separation of our two eyes. When we look at a moving picture projected on a screen, however, binocular stereo does not provide a useful cue, and when we look at a still photograph, even motion parallax can be of no help.

A passport photograph serves to identify an individual. Such a photograph, however, cannot simply be matched point for point with another image of the same individual, particularly if the second image is taken from a different viewpoint when different lighting conditions prevail. The two-dimensional distribution of brightness in the photograph itself does not provide the required information directly. Instead, we are apparently able to recover the three-dimensional shape, and the nature of the surface markings, from the two-dimensional image. It is this information that is used in matching the photograph with the person in front of us, not the brightness pattern itself.

Recognizing people is rather important, so we have no doubt developed specialized means for doing this; we can do it successfully even when there are many similarities between different faces. One piece of evidence that this is a highly specialized capability is that we perform rather poorly when we look at a person while standing on our head. More general purpose visual abilities tend not to be affected that much by the orientation of the image on the retina.

We do have the ability to recover the shapes of objects in general, whether we are familiar with them or not. Images produced by a scanning electron microscope (SEM), for example, are usually easy to interpret in terms of three-dimensional shape, even though the viewer may be unfamiliar with the objects portrayed. Similarly, images of hilly terrain produced by a synthetic aperture radar (SAR) system are immediately understood, even though SAR portrays the world at an unusual scale. Finally, the shape of the surface of a snow-covered glacier is apparent to anyone looking at it; yet there are few cues, and the shape may not be much like any shape one has ever seen before. Where then is the information that provides the hidden cue to shape?

1. Shading and the Recovery of Depth

The answer lies in the variation of brightness, or *shading*, often exhibited in a region of an image. In a photograph of a face, for example, there are variations in brightness, even though the reflecting properties of the skin presumably do not vary much from place to place. It may be concluded that shading effects arise primarily because some parts of a surface are oriented so as to reflect more of the incident light toward the viewer than are others.

Artists have used *chiaroscuro* for many centuries to convey the impression of three-dimensional shape. The map of Toscana that Leonardo da Vinci drew in 1502 or 1503, for example, contains oblique shaded views of relief forms illuminated from the left (see [Imhof 651]). Rembrandt van Rijn's concern with light and dark led, among other things, to more realistic use of shading in painting. Some super-realist art contains wonderful shading patterns, often almost as good as those in the photographs from which the art was copied! Art education includes exercises in rendering the shading on simple geometric shapes as well as more complex surfaces such as the folds in cloth draped over a body.

One might think at first that it would be natural to use shading in realistic depictions created while viewing the real world. But this would be far from correct: our visual system tries to interpret the brightness pattern on the retina as shading due to spatial fluctuations of surface orientation and spatial variations in the reflecting properties of the surface. So we "see" the surface as having some three-dimensional shape and some surface markings, not in terms of a two-dimensional pattern of light and dark. Attempting to depict what we see, we tend to pay attention mostly to the surface markings, leading to a "flat" depiction, devoid of shading. From time to time such a style has in fact found favor in the art world.

Shading is sometimes confused with shadowing. There are two kinds of shadows: self shadows (or attached shadows), and cast shadows. A portion of a surface is self-shaded when it is turned away from the source of light. A self-shadow edge provides localized cues to shape, since it is the locus of points where the rays graze the surface. A cast shadow on a surface results when another surface intercepts the light from the source. Cast shadows also provide localized cues to shape, although the shadow of a curved surface cast on another curved surface is very difficult to interpret. Shading, on the other hand, provides cues all over a surface, not just along special contours.

There are, however, some ways in which shading and shadowing are related. First of all, in the case of an extended light source (as opposed to a single point source), shading arises in part from the fact that the fraction of the light source that appears "below the horizon" at a particular point on the surface depends on surface orientation. That is, as far as a particular portion of the light source is concerned, some parts of the surface are self-shadowed. Smooth shading here arises in part from a superposition of these shadowing effects from all parts of the extended source. Secondly, we may be viewing an undulating surface from a distance so great that the shadows cast by the undulations are not resolved. The average brightness of a portion of the surface will depend in part on how much of it is shadowed. Thus shadowing of parts of the "microstructure" of the surface may contribute to the shading effects apparent on a macro scale.

It should be pointed out right away that the recovery of shape from shading is by no means trivial. We cannot simply associate a given image brightness with a particular surface orientation. The problem is that there are two degrees of freedom to surface orientation—it takes two numbers to specify the direction of a unit vector perpendicular to the surface. Since we have only one brightness measurement at each picture cell, we have one equation in two unknowns at every point in the image. Additional constraint must therefore be brought to bear. One way to provide the needed constraint is to assume that the surface is continuous and smooth, so that the surface orientations of neighboring surface patches are in fact not independent. Note that there is no magic at work here: we are not recovering a function of three variables given only a function of two variables. The distribution of some absorbing material in three-dimensional space cannot be recovered from a single two-dimensional projection. The techniques of *tomographic reconstruction* can be applied to that problem, but only if a large number of images taken from many different viewpoints are available. Why then are we able to learn so much about the three-dimensional world from merely two-dimensional images?

2. Our Special Visual World

Part of the answer is that we live in a very special visual world. In most cases we deal with opaque, cohesive bodies immersed in a transparent medium. Rays of light pass essentially unmolested through the medium and they do not penetrate the objects. This means that we can ignore the medium and that only the surfaces of the objects are of interest. Points on a surface can be specified by using two coordinates, so we have a mapping from a two-dimensional surface to a portion of a two-dimensional image. We are concerned here with inverting this mapping. Our difficulties would be much greater if we had to deal with partially translucent objects immersed in a partially absorbing medium. In this situation two-dimensional images would be very hard to interpret, as are microscopic images of biological specimens.

The mapping from surface orientation to image brightness is unique and can be determined for different surface materials and illuminating conditions. Thus the use of shading in computer graphics (which is very important for the realistic appearance of objects) is rather simple. Shaded overlays, which make topographic maps much easier to interpret, constitute a particular illustration of this. The problem of interest in this book may be considered an "inverse graphics" problem—one that is much harder than that of producing shading from shape.

Unless we see sharp discontinuities in brightness, or some other special cues, we assume that the surface of an object is homogeneous in its reflecting properties. The appearance of the surface will be altered if the surface properties do in fact vary from place to place. In this case a vision system that assumes uniform reflecting properties will recover a shape that is different from the actual one. This effect can be used to alter the apparent shape of a surface. For example, the counter shading found on many animals, where the underside is lighter than the side turned toward the light, may serve as a kind of camouflage. In this case, the variation in surface reflecting properties reduces or even cancels the shading one would normally expect to see, flattening the apparent shape and so reducing the ability of an observer to see the animal as a three-dimensional shape separate from the background.

Makeup exploits the same effect, usually in a flattering way. A cheek, for example, can be made to appear to recede more steeply than it actually does by applying a darker coloration to its side. A nose can be made to appear sharper by applying a thin line of light make-up along its ridge. Unless these alterations are done carefully, the illusion can disappear, as it often does when the viewpoint or lighting conditions are changed drastically.

Shading has sometimes been described as a "weak" cue, particularly when compared to motion parallax and binocular disparity. It is, however, an important cue to shape, especially when other cues are lacking. This is significant, for example, when we are viewing a smooth surface without any surface markings—if there were no shading we could only guess at the shape. Also, many other depth cues are absent when we look at a still photograph, as mentioned earlier, and in situations were we are too far away from objects for either motion or stereo to provide useful cues.

3. Basic Formulation of the Problem

Although it has been known for a long time that shading provides an important depth cue, only relatively recently has the shape-from-shading problem been properly formulated. No conjecture existed about whether there was enough information in an image to compute the shape and whether more than one surface could give rise to the same shading under given lighting conditions. Considerable further progress has been made, but not particularly rapidly. One reason for this is that the mathematical analysis of the general problem is highly complex.

The problem becomes tractable if a number of simplifying assumptions are made. The principal simplifications arise from the assumption that the viewer and the light sources are far enough away from the objects being viewed that the brightness of an oriented portion of the surface is independent of its spatial position. This means that brightness depends only on the orientation of the surface patch. It also means that we are dealing with *orthographic projection* instead of *perspective projection*, something that simplifies most vision problems (except for motion vision and photogrammetry, which depend on the effects of perspective to recover all of the components of an unknown displacement). The simplification engendered by these assumptions was noted, but not exploited, in the early work on shape from shading.

We clearly need a way to talk about the orientation of a surface patch. One way to do this, as already indicated, is to specify a unit vector, $\hat{\mathbf{n}}$, perpendicular to the local tangent plane. Another way is to specify the components p and q of the surface gradient. These are the partial derivatives of surface height z above some reference plane perpendicular to the optical axis, that is, $p = (\partial z/\partial x)$ and $q = (\partial z/\partial y)$. The two notations are connected by the equality

$$\hat{\mathbf{n}} = \frac{1}{\sqrt{1+p^2+q^2}}(-p, -q, 1)^T.$$

A third way of specifying surface orientation is to give the slope and the direction of steepest descent. The terms *slant* (σ) and *tilt* (τ) have been introduced for angles used in this fashion, which is unfortunate, since these terms are not mnemonic and are frequently confused (even the terms *dip* and *strike*, used in geology to specify the orientation of a sedimentary layer in the earth, would have been more suggestive). The relationship of slant and tilt to the unit normal is given by the equality

$$\hat{\mathbf{n}} = (\sin\sigma \cos\tau, \sin\sigma \sin\tau, \cos\sigma)^T.$$

Each notation for surface orientation has its own advantages, and the above three, as well as some others, are used in this book.

4. The Reflectance Map and Radiometry

If we do make the simplifying assumption that the viewer and the light sources are far from the object, we can introduce the *reflectance map*, a means of specifying the dependence of brightness on surface orientation. If we elect to use the unit surface normal $\hat{\mathbf{n}}$ as a way of specifying surface orientation, then we can write the brightness as a function of orientation in the form $R(\hat{\mathbf{n}})$. If we use p and q instead, we can use the form $R(p, q)$. A graphic representation of the reflectance map is possible if we plot contours of constant brightness in the *pq*-plane, also called *gradient space*. Sometimes the term reflectance map is reserved for this graphical representation.

Proper formulation of the shape-from-shading problem had to await a thorough understanding of the image formation process. Two issues are critical in regard to image formation: first, how the position of a point in the image is related to the position of the corresponding point in the environment; and second, what determines the brightness at a point in the image. The first issue is rather straightforward and well understood, involving as it does the well-known perspective projection. Much less attention has been paid to the second issue, which is of great importance in understanding vision, whether natural or artificial. The analysis of image formation was until quite recently hampered by a lack of clarity in certain aspects of radiometry.

It is generally thought that the definition of the unit of. brightness is the weakest part of the SI system of units (Systeme International d'Unites), because it involves the candela, the only unit still defined with reference to human sensations. The understanding of terms used for various measures of reflectance also was unsatisfactory until the U.S. National Bureau of Standards (NBS) introduced the *bidirectional* *reflectance distribution function* (BRDF) and defined a number of derived quantities (see [Nicodemus *et al.* 77]).

The microstructure of the material of a surface determines how much of the incoming light will be reemitted in various directions. Thus it is the microstructure that determines the BRDF. The reflectance map, needed for work on the shape-from-shading problem, can be determined if the BRDF of the surface material is known, along with the light source and viewer geometry. In practice the dependence of brightness on surface orientation is often determined experimentally using a calibration object of known shape, such as a sphere. Also, certain phenomenological models have found favor as approximations to real surface behavior. This includes the *ideal diffuser*, or Lambertian surface, which reflects all incident light and appears equally bright from all directions. The brightness of such a surface can be shown to be proportional to the cosine of the incident angle, the angle between the surface normal and the incident ray.

Note, by the way, that we eschew the term *image intensity*, since intensity is a term with a technical meaning in radiometry quite different from the intended one: the intensity of a point source of light is the power per unit solid angle radiated in a particular direction. The appropriate term for image brightness is *image irradiance*, the power falling on the image per unit area. The correct term for the brightness of a part of a surface in the scene being viewed is *scene radiance*, the power per unit solid angle per unit apparent area emitted from the surface. (The apparent area is the foreshortened area as seen by the viewer—it is the actual surface area times the cosine of the angle between the surface normal and the viewing direction.) We can use the common term brightness for both of these concepts without fear of confusion, since the two have been conveniently defined in such a way that they are intimately related: image irradiance is directly proportional to scene radiance in an optical imaging system.

5. History of the Problem

The earliest work on the quantitative use of shading information appears to have been in the mid-1960s on recovering the shape of parts of the lunar surface in preparation for the human exploration of the moon. This work used careful measurements of the reflective properties of the material in the maria of the moon, made mostly by Russian researchers. They discovered in the early part of this century that the brightness of this material was a function of the ratio of the cosine of the incident angle to the cosine of the emittance angle. (That is, in the case of the material in the maria of the moon, brightness is a function only of the *luminance longitude*, being independent of *luminance latitude*.) The reason the reflecting properties of the moon (and rocky planets) appear to us to be so unusual is that we view the surface from a great distance, and so what constitutes "microstructure" is very different from what it is when we look at smaller objects from nearby. Incidentally, anyone can tell that there is something odd about the reflecting properties of the lunar surface, since the full moon looks flat—more like a disk than a sphere. In fact, at full moon the lunar surface is more or less equally bright everywhere, if we ignore surface markings due to variations in surface *albedo*. An ordinary diffusely reflecting spherical surface would instead be bright in the middle and dark near the limb.

The unusual reflecting properties of the lunar material allow one to determine the slope of the surface, in the direction toward the subsolar point, from a local measurement of brightness. (The subsolar point is where the shadow of the camera is cast—that is, the point where a ray from the sun through the viewer intersects the surface.) This is because, in the case of the material in the maria of the moon, brightness happens to be a function of a linear combination of the components of the surface gradient, p and q. The slope at right angles to the direction toward the subsolar point is completely unconstrained. One can integrate the slope along a line through the subsolar point to generate a profile of the surface. Many such profiles, closely spaced, define the surface shape. This relatively simple method applies only in the case of the rather special reflectance properties of the material in the maria of the moon and does not generalize to other materials, such as Lambertian reflectors.

6. The Image Irradiance Equation

The general solution of the shape-from-shading problem revolves around the so-called *image irradiance equation* relating image irradiance to scene radiance:

$$E(x, y) = R(\hat{\mathbf{n}}(x, y)),$$

where $E(x, y)$ is the image irradiance at the point (x, y), while $R(\hat{\mathbf{n}}(x, y))$ is the radiance of a surface patch with unit normal $\hat{\mathbf{n}}(x, y)$. The unit surface normal at the point in the scene corresponding to the image point (x, y) is $\hat{\mathbf{n}}(x, y)$. Actually, image irradiance is not equal to scene radiance, only proportional to it, but the proportionality factor is usually ignored, because it is assumed that some calibration process normalizes one of these quantities so that it is commensurate with the other.

The image irradiance equation is a nonlinear first-order partial differential equation, as can be seen by noting that the normal can be expressed in terms of the two first-order partial derivatives p and q. As such, it can be solved using the method of characteristic strips, which reduces the partial differential equation to an equivalent set of five coupled ordinary differential equations (for x, y, and z, as well as p and q). A particular solution of these equations generates a so-called characteristic curve on the surface, along with surface orientation on that curve. The projection of such a curve into the image is called a *base characteristic*. The characteristic curve, along with the orientation, defines a *characteristic strip* on the surface. Many closely spaced strips define the shape of the surface.

Figure 1 shows the method of characteristic strip expansion at work. Figure 1(a) is a (rather coarsely quantized) gray-level image of a smooth surface illuminated by a light source near the viewing position. Even though the reader will most likely not have seen this before, the shape should be apparent from the shading in the image. Figure 1(b) shows the base characteristics computed by the numerical algorithm superimposed on the gray-level image. With the light source near the viewer, the reflectance map is rotationally symmetric and so the characteristics follow curves of steepest descent on the surface. Note how the base characteristics emerge from a singular point in the image (the brightest spot in this case) and how the algorithm used here sprouts new strips when existing ones stray too far apart from one another. The nested contours crossing the base characteristics connect points reached at the same stage in the expansion of the characteristics.

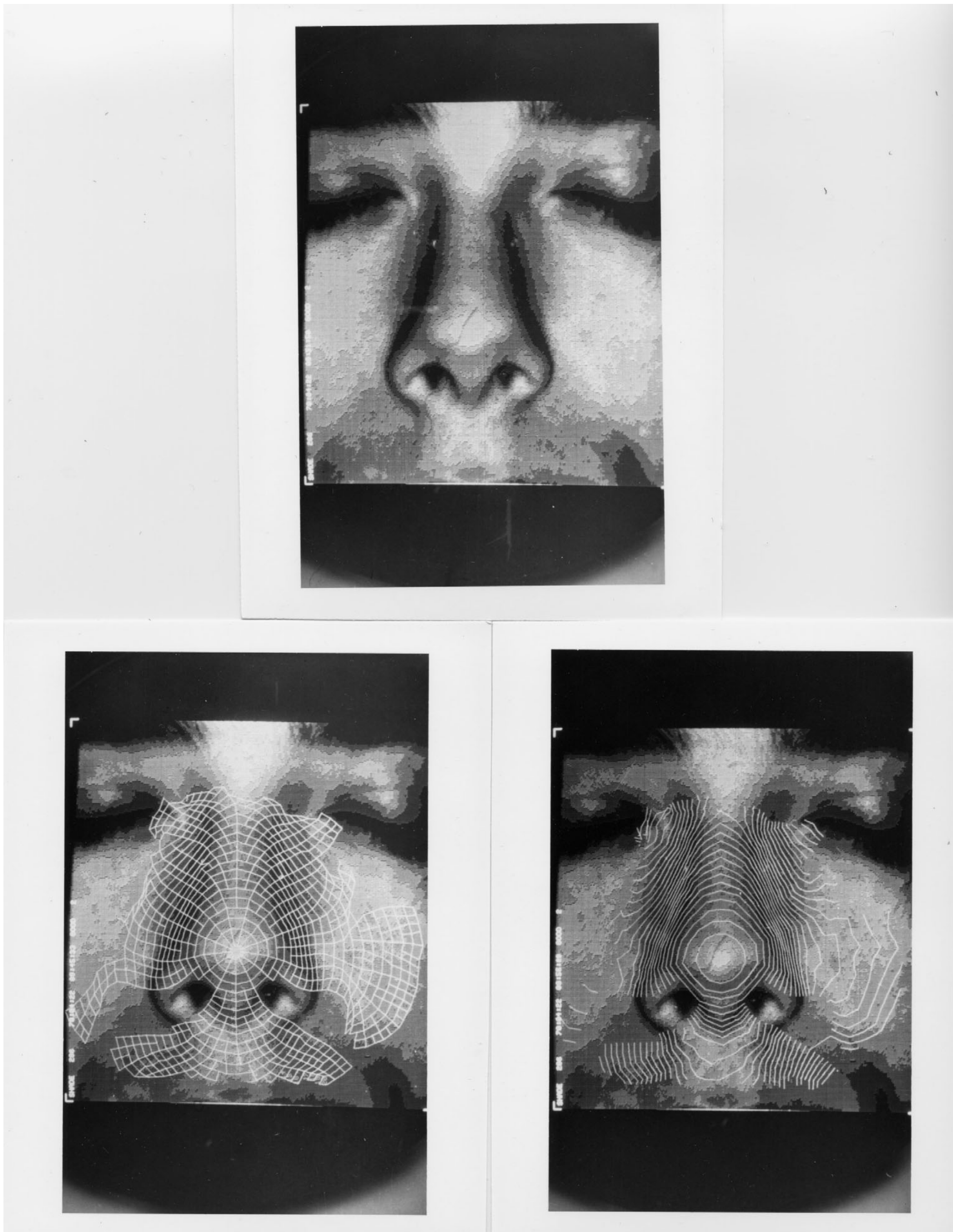

Figure 1. Recovery of shape from shading using the characteristic strip expansion method. A. Gray-level image. B. Base characteristics superimposed on the gray-level image. C. Contour map constructed from the three-dimensional characteristic strips.

Each contour is the locus of points at a fixed distance from the singular point. Finally, figure 1(c) shows contours of constant elevation above a reference plane obtained by interpolation from the three-dimensional characteristic strips. In this particular case, since the characteristics are lines of steepest descent, the contours are orthogonal to the base characteristics. The reader should verify that this contour map is in accord with the apparent shape of the object in the original image.

7. Iterative Solution on a Regular Grid of Points

Unfortunately, the method of characteristic strip expansion suggests neither reasonably likely biological schemes for solving the shape-from-shading problem, nor efficient and robust computational methods. This is why, from the very beginning, there has been a search for alternatives that are more like the methods used for solving linear second-order partial differential equations. These methods are iterative and can be implemented in parallel on a grid of locally interconnected cells.

Iterative approaches repeatedly make adjustments to surface orientation until the predicted shading, based on the estimated shape, matches that actually observed in the image. In most of these methods, the shape is specified not by height above a reference plane, but by surface orientation. Orientation estimates are stored for every point in a dense grid of points, usually one point for every picture cell. Now it is trivial to match the observed shading at a particular grid point by picking one of the infinite number of orientations that produces the observed image brightness there. The resulting field of surface orientations will most likely not correspond to a continuous surface, however.

It is possible to show that neighboring orientations cannot be chosen independently, since they have to correspond to some underlying surface. One cannot independently specify the two partial derivatives of a function of two variables. They must satisfy the condition that the two mixed derivatives of second order are the same. Thus we must have $z_{yx} = z_{xy}$, or equivalently, $p_y = q_x$. This condition is referred to as *integrability*, since one can recover the underlying surface $z(x, y)$ by evaluating line integrals of $(pdx + qdy)$ along arbitrary contours if the surface orientation information satisfies this constraint. Iterative methods make repeated adjustments to surface orientation to improve the match between brightness predicted from the estimated surface shape and brightness actually observed. Some methods attempt to do this in a way that maintains integrability of the estimated surface.

This turns out to be hard and so other methods instead ensure only that neighboring surface orientations remain similar. Such methods attempt to minimize the integral of a penalty term measuring "unsmoothness." With these simpler methods there is a problem when one wants to estimate the surface shape in terms of a depth map, since typically no shape corresponds exactly to the computed field of normals. Two approaches have been taken to solve this problem: one uses a projection onto the subspace of integrable solutions using Fourier Transform methods (see chapter 5 by Frankot and Chellappa), while the other involves a least-squares approach that leads to a linear second-order partial differential equation (see chapter 7 by Horn and Brooks).

8. Convergence of Iterative Methods

With few exceptions, there is so far only empirical evidence that the iterative schemes discussed above converge; it is not clear that they always do (but see chapter 12 by D. Lee). The solutions produced by these iterative schemes are also typically not quite accurate. There are two reasons for this: methods that enforce integrability can get stuck in local minima in their search for the global extremum, and methods that do not enforce integrability inherently trade off increased surface smoothness against departures from exact match of the shading information.

It has been claimed that the shape-from-shading problem is inherently *ill posed*. (An ill-posed problem is one that does not have any solution, does not have a unique solution, or has a solution that is very sensitive to the given data.) The shape-from-shading problem certainly is ill posed if one considers only an image patch that does not include a singular point or the projections of the limb of an object, since there are an infinite number of surfaces that yield the same shading pattern. However, the problem is not necessarily ill posed if singular points and information from the limbs of the objects are taken into account. (This is not unlike the case of a linear higher-order partial differential equation, which has a unique solution if appropriate boundary conditions are specified, but has an infinite number of solutions if they are not.)

The notion that the shape-from-shading problem is ill posed in general (even when singular points and limbs are included) probably has its origin in the use of an "unsmoothness" penalty term in methods that do not enforce integrability. The problem is in fact ill posed if the components of the surface gradient are treated as unrelated, since one can then choose orientations at each image point independently, and there are an infinite number of orientations that correspond to a particular observed brightness. To obtain a unique solution in this case, one has to apply a *regularization* method, one that selects a particular solution out of an infinite number of possible solutions. This is not necessary when one takes into account the fact that the surface orientation field is supposed to be integrable.

9. Existence and Uniqueness

Questions about the existence and uniqueness of the solution have, in fact, proved very difficult to answer. For example, suppose we are given the reflecting properties of the surface and the arrangement of the light sources. Then, is there always a surface shape that will generate, under these conditions, any given (arbitrary) image brightness pattern? The answer is not known. It may be that there are patterns that could not have been produced as the result of shading on any three-dimensional shape. If this is so, then the shape-from-shading problem has no solution in this case. People often appear to be able to tell that a particular pattern is not due to shading but to spatial variations in the reflecting properties of the surface. This may suggest an answer to the above question, or it may suggest something about our a priori assumptions about the world.

Typically there will be some unique surface orientation for which the brightness is a maximum (or a unique minimum, in some unusual cases). In the case of an ideal Lambertian surface illuminated by a single point source, for example, this occurs when the surface normal points directly toward the light source. A point in the image where this maximal brightness is observed is called a singular point. These image points have particular importance since the surface orientation at the corresponding point on the surface is immediately known (provided, of course, that the reflectance map is given).

Surface orientation can also be determined easily for points on the limb of an object. The *limb* is the locus of points on a smoothly curved surface where the rays to the viewer graze the surface. It separates those points that are visible to the viewer from those that are not.

In the case of a smoothly curved object, the silhouette is the projection of the limb. (Sharp edges of objects, where the surface normal does not vary continuously with position, may also project to form part of the silhouette.) The term *occluding boundary* is often used for the limb, since the surface on one side of the boundary occludes the background, which is visible in the image on the other side of the boundary. The surface normal at a point on the limb is clearly perpendicular to the viewing direction. Also, a plane can be constructed that is tangent to the object on the limb as well as the silhouette at the corresponding image point. This means that the surf ace normal on the limb is parallel to the normal to the silhouette at the corresponding point in the image.

In contrast to the special considerations that apply to points that are either singular or on the silhouette, the orientation at most points cannot be determined directly from image brightness. In fact, the brightness pattern in an arbitrary image region could in general arise from an infinite number of different surfaces. A very important issue, then, is how singular points and silhouettes constrain possible solutions. When is there only one shape that can give rise to the observed shading pattern? This is a mathematical problem that is just about impossible to solve unless details of the reflecting properties of the surface and the distribution of light sources are provided, and it is still very hard when this information is available. Only a small number of special cases have been successfully dealt with so far.

Now suppose that the position of the light source is also unknown. Can we determine both shape and light-source position from a shaded image? This problem is clearly less constrained than the basic shape-from-shading problem described above, and so we might expect that typically there may be several solutions.

10. Local Methods and Photometric Stereo

As mentioned above, shading information in an arbitrary image patch is, in general, infinitely ambiguous. If we make sufficiently strong assumptions about the surface, however, some useful information can be recovered from the first- and second-order variations of brightness within a patch. One assumption that leads to interesting results is that the surface is everywhere locally spherical. Another assumption of interest is that the surface is everywhere locally cylindrical. Such local methods have been explored only recently, but show promise.

A single measurement of image brightness provides only limited information about surface orientation. As explored above, one way of removing the local ambiguity is to take into account neighboring points. A quite different and simpler approach uses information from several registered images taken with different lighting. Two such images provide two constraints to recover the two unknown parameters of surface orientation at each point in the image. Since the corresponding image irradiance equations,

$$E_1(x, y) = R_1(\hat{\mathbf{n}}(x, y)) \quad \text{and} \quad E_2(x, y) = R_2(\hat{\mathbf{n}}(x, y)),$$

are typically nonlinear, however, several solutions may be found at a particular picture cell. Additional images can help remove this ambiguity and also allow one to recover further unknown parameters, such as the albedo of the surface.

How good are people at recovering shape from shading? One problem with a question like this is that it is hard to obtain quantitative information about the apparent shape seen. Various means have been devised for addressing this issue, including methods based on comparisons of estimated surface orientation with given reference orientations. A conclusion that may be drawn from the experiments performed so far is that people use shading information well, but they may not develop the kind of detailed quantitative representation of the shape used in machine vision systems. Since shading depends on surface orientation, and hence first partial derivatives of surface height, we can expect that lower spatial frequencies will have only a small effect on the image. Conversely, we should not be surprised if the largest errors in reconstruction of shape are in the lower-spatial-frequency components. Indeed it appears that people are good at recovering shape information corresponding to rapidly undulating surface features, while slow changes may be missed or misinterpreted.

11. The Papers Selected and the Bibliography

We have chosen not to place this collection of papers in chronological order, because there is no way to organize the book so that each part depends only on what comes earlier. Instead, we have simply arranged the chapters in alphabetical order of the authors' names. The subject of shape from shading is multifaceted as should be clear from the above, and we have attempted to include at least one contribution relevant to every major facet. The following is a brief introduction to each paper:

Chapter 2: Surface Descriptions from Stereo and Shading

by A. Blake, A. Zisserman, and G. Knowles

This is a paper with wide-ranging scope touching on many aspects of research on shape from shading. An important uniqueness result is presented. It is somewhat similar to the uniqueness theorem presented in chapter 4 by A. R. Bruss, except that the image boundary here is known to correspond to a particular curve in space, not necessarily a limb. This is important in integrating shading information with cues that generate shape information in this form, such as binocular stereo.

The paper also makes an interesting observation about the convergence properties of the approach presented in chapter 3 by Brooks and Horn. As the derivation of their method does not use a convex functional, extrema may exist that are not global minima. Additionally, an analysis is presented of the local approach described in chapter 15 by Pentland. The magnitude of the errors in surface slant and surface tilt are estimated.

The authors also consider the method described in chapter 10 by Koenderink and van Doorn for detecting parabolic lines from the directions of contours of constant brightness. The material is presented with novel simplicity, and an implementation is discussed. Finally, a battery of techniques is proposed that forms the basis for an assault on the shape-from-shading problem, with emphasis on reducing the amount of prerequisite information.

Chapter 3: Shape and Source from Shading
by M.J. Brooks and B.K.P. Horn

This paper deals with the situation in which a diffuse surface is illuminated by a light source in an unknown position. This is a special case of a situation in which the reflectance map is not known. Here it is assumed that its general form is given, but that some parameters are not specified. An iterative scheme is presented that alternately estimates the surface shape and the light source direction. The scheme is derived using the variational approach expounded in chapter 7, the companion paper by Hom and Brooks. The unit-normal vector notation is used for the first time in shape from shading schemes, and this leads to a particularly elegant formulation.

The method is extended to the case in which, in addition to the point light source, there is a distributed "sky" source. The iterative scheme itself is then generalized to arbitrary reflectance maps. All of the methods presented in this paper are simple to implement, but they do not enforce integrability.

Chapter 4: The Eikonal Equation: Some Results Applicable to Computer Vision
by A.R. Bruss

The author presents a uniqueness result for the situation in which the reflectance map is rotationally symmetric and the silhouette is a smooth, closed curve corresponding to the limb of the object. It is clearly important to know under what circumstances surface shape may be determined uniquely by image shading. This difficult problem finds its most serious treatment in this paper. (One of the few other attacks on this problem can be found in [Deift & Sylvester 81].)

Results in Bruss's paper are based on analysis of the Eikonal equation, studied in the context of wave propagation in optics:

$$F(x, y) = p^2 + q^2,$$

where, as mentioned before, p and q are the partial derivatives of height z with respect to x and y respectively. It can be shown that the analysis applies in general to the case of a rotationally symmetric reflectance map with monotonic dependence of brightness on slope. For example, a Lambertian surface illuminated by a point source at the viewer leads to an image irradiance equation of the form

$$E(x, y) = \frac{1}{\sqrt{1 + p^2 + q^2}}$$

which can be transformed into an Eikonal equation by letting

$$F(x, y) = \frac{1}{E(x, y)^2} - 1$$

The above problem is shown to have a unique solution (aside from a simple reversal of depth) if $F(x, y)$ equals zero at a single point in the domain, the height z vanishes to second order at that point, and the smooth closed silhouette corresponds to the limb of the object.

Chapter 5: A Method for Enforcing Integrability in Shape-from-Shading Algorithms
by R.T. Frankot and R. Chellappa

This paper addresses the issue of integrability of the surface normal field computed. Most iterative schemes that recover shape from shading do not attempt to enforce integrability, as discussed also in chapter 7 by Hom and Brooks. In Frankot and Chellappa's very original work, an orthogonal basis set of functions is used to describe surfaces. A standard iterative scheme is used to obtain successive estimates, at each image point, of the directional derivatives of surface height. These derivative estimates are then expressed in terms of a linear combination of orthogonal basis functions. In general, the estimated derivatives, and hence the associated expansion in terms of basis functions, will not satisfy the integrability condition. So, the critical step is to map these estimates to those corresponding to the "nearest" integrable field of surface orientations. This is done by finding the closest set of coefficients, in a least-squares sense, that also have the property of integrability.

A least-squares scheme for obtaining a depth map from inconsistent first partial derivatives is, by the way, mentioned in chapter 7 by Hom and Brooks. The depth map computed using their method could be numerically differentiated to obtain new, consistent estimates of the first partial derivatives. This approach provides an alternative to the projection scheme for assuring integrability described in this chapter by Frankot and Chellappa. (A companion paper explores the application of this technique to synthetic aperture radar imagery—see [Frankot & Chellappa 87].)

Since the method presented in this paper uses p and q, the first partial derivatives of height to represent surface orientation, it cannot deal with the limbs of objects. On the other hand, the method shows promise in dealing with situations in which boundary information is absent, while some low resolution height information is provided.

Chapter 6: Obtaining Shape from Shading Information
by B.K.P. Horn

This is the part of Horn's 1970 Ph.D. thesis that appeared in *The Psychology of Computer Vision* [Horn 75]. In it, the shape-from-shading problem is formulated and then solved using the characteristic strip approach. Methods for computing characteristic strips from noisy data are discussed and various special cases explored, including imaging in the scanning electron microscope and the recovery of lunar topography. While the importance of singular points is emphasized, little attention is paid to limbs of objects.

The mathematical notation in this paper is somewhat more complex than that in much of the later work, since the common assumptions of distant viewer and distant light sources are not made. (This belies the general impression that the shape-from-shading problem can only be dealt with when one assumes orthographic projection.) If one does assume a distant viewer and distant light sources, the problem becomes simpler, as pointed out in this paper. In fact, in this case, one can use 'the reflectance map, introduced later in a related paper [Horn 77]. The reflectance map is discussed in detail in chapter 8 by Horn and Sjoberg.

Chapter 7: The Variational Approach to Shape from Shading
by B.K.P. Horn and M.J. Brooks

In this paper the analysis of a number of existing and new iterative approaches to the shape-from-shading problem is unified using as a theme the minimization of some criterion function that measures the departure from exact match between observed and predicted image brightness, as well as some other properties of the solution, such as departure from smoothness or integrability.

Several iterative methods, such as the one developed earlier by Strat [Strat 79], are compared with that presented in chapter 9 by Ikeuchi and Horn. A least-squares method for recovering height from (possibly inconsistent) estimates of the partial derivatives is also described. The integrability condition is expressed in terms of the surface normal and its partial derivatives. It is found that strictly enforcing the integrability condition does not lead to convergent iterative schemes-in essence one is trying to simulate the solution of the characteristic strip equations on a grid. Methods involving a penalty term for departure from integrability, on the other hand, show definite promise. Using surface normals to describe surface orientation is shown to have a number of advantages, including the ability to deal better with the limbs of objects.

Chapter 8: Calculating the Reflectance Map
by B.K.P. Horn and R.W. Sjoberg

Here it is shown how the reflectance map may be computed if the bidirectional reflectance distribution function (BRDF) and the distribution of light sources is given. The reflectance map relates brightness to local surface orientation and is essential to the formulation of the image irradiance equation, which in tum is central to most approaches to the shape-from-shading problem.

In this paper, relevant parts of radiometry are reviewed and image formation analyzed carefully. Importantly, the paper adopts the new nomenclature of the National Bureau of Standards as set out in [Nicodemus *et al.* 77]. The computation of the reflectance map requires mathematical functions describing both lighting and surface reflection, and an appropriate integration over all incoming illumination directions. To illustrate the method, reflectance maps are generated for a variety of situations, encompassing collimated, uniform, and hemispherical sources striking both Lambertian and specularly reflecting surfaces. A large number of reflectance maps are also developed in a paper on hill shading, [Horn 81].

Chapter 9: Numerical Shape from Shading and Occluding Boundaries
by K. Ikeuchi and B.K.P. Hom

Here an iterative method for solution on a grid of points is developed for computing shape from shading. It is formulated in a way that makes it possible to incorporate information from the limbs of objects. The surface normal at a point on the limb of an object is parallel to the normal in the image plane to the corresponding point on the silhouette. This, along with the information from singular points, provides strong constraint on possible solutions. As shown by Bruss in chapter 4, there are situations in which the shape is uniquely determined by this information and the shading in the image. A difficulty arises, however, in using the boundary information in an iterative scheme where surface orientation is represented by the two partial derivatives of surface height, because the slope of the surf ace becomes infinite at the limb. In this paper the first iterative scheme able to use information from the silhouette is described. The method uses a novel parameterization for surface orientation, resulting from a *stereographic* projection (f and g) of the unit normal from the Gaussian sphere rather than the more common *gnomonic* projection (p and q).

This paper was the first to use the variational calculus in the derivation of iterative schemes for shape from shading. This approach was inspired by the use of the variational calculus in the estimation of optical flow, the first use of variational methods in machine vision [Horn & Schunck 81]. One disadvantage of the stereographic parameterization is that it is somewhat harder to express the condition of integrability than it is with the gnomonic parameterization. In part because of this, the authors here elected to ignore integrability, instead minimizing the integral of an "unsmoothness" penalty term. This leads to a simple iterative scheme, but also unfortunately suggested to some that the shape-from-shading problem is inherently ill posed and that it needs to be regularized. The issue of integrability is addressed in chapter 7 by Horn and Brooks and in chapter 5 by Frankot and Chellappa.

Chapter 10: Photometric Invariants Related to Solid Shape
by J.J. Koenderink and A.J. van Doorn

The authors deal with the relationship between patterns of the contours of constant brightness in the image and the differential geometry of the surface being imaged. They consider Lambertian surfaces and show that the parabolic curves on the surface are of particular importance. (The Gaussian curvature is zero at each point on a parabolic curve, so that these curves separate elliptic regions from hyperbolic regions.) It turns out, for example, that the contours of constant brightness cut the parabolic lines at a fixed angle, independently of the light source position. The parabolic curves are shown to be important in another respect: certain singularities travel along the parabolic curves as the light source is moved.

Patterns of the constant brightness contours are derived for certain canonical surface undulations, such as furrows, dimples, and a shape like a hat with the brim turned down. A particularly exciting aspect of this work is the way it relates classes of surfaces to types of brightness patterns. Most other work on shape from shading deals with methods for recovering specific solutions to particular brightness patterns. The discussion of surface undulations in this paper is aided by appealing to a mapping onto the Gaussian sphere. The authors conjecture that human observers can interpret shading information because of certain invariants of the pattern of constant brightness contours.

Chapter 11: Improved Methods of Estimating Shape from Shading using the Light Source Coordinate System
by C.-H. Lee and A. Rosenfeld

Along with Pentland (see chapter 15), the authors assume that a portion of the surface is spherical.

If there is a single light source at the viewer, then the reflectance map is rotationally symmetric, and slope can be determined directly from the brightness, although the direction of steepest ascent cannot be recovered. The reflectance map of a surface illuminated by a point light source in an arbitrary position can be transformed to a rotationally symmetric form in a light source coordinate system. The authors find a way of applying this basic idea and show that two points on the surface have the same tilt in a coordinate system aligned with the direction of the incoming light if and only if the direction of the brightness gradient is the same at corresponding image points. This relationship can be used to recover the surface shape if the direction of the illumination is known. A coordinate transformation can be used to translate the result into a coordinate system aligned with the viewing direction. This leads to estimates for slant and tilt that are based on first-order derivatives of image brightness—in contrast to the earlier method of Pentland, which uses second-order derivatives. Lee and Rosenfeld also present a method for estimating the direction of the light source that may be compared with the method given in chapter 15 by Pentland.

Chapter 12: A Provably Convergent Algorithm for Shape from Shading
by D. Lee

This paper deals with the problem of convergence of iterative schemes for recovering shape from shading, such as the one discussed in chapter 9 by Ikeuchi and Horn. These algorithms are based on a variational approach with a penalty term for departure from smoothness. Such methods involve a parameter that determines the tradeoff between errors in matching brightness and departure from smoothness. The author shows that a particular iterative scheme he developed converges to a unique solution for certain ranges of this parameter. This is an important result since iterative schemes are the most commonly used, and there was, before this paper appeared, only empirical evidence that some of them might converge.

The author also notes that, at least for some values of the parameter, the estimated surface is too smooth and departs noticeably from the true surface. The algorithm considered here does not enforce integrability.

Chapter 13: Recovering Three-Dimensional Shape from a Single Image of Curved Objects
by J. Malik and D. Maydan

The authors of this paper address the problem of recovering the shapes of surfaces that are only piecewise smooth. Almost all of the work on shape from shading has focused on smoothly curved objects, where surface normals vary continuously with position on the surface. There has also been some work on recovering the shapes of polyhedral objects—that is, objects bounded by planar faces. With few exceptions, however, the brightness of the faces of the polyhedra has been ignored. It turns out that it is impossible to completely recover the shape of a polyhedron from a single line-drawing of the object—there are always at least three degrees of freedom unspecified. Only when the brightness of the surfaces is also used is a unique solution possible (see, for example, [Sugihara 86]). The discussion above applies only to objects with planar faces.

Malik and Maydan approach the problem of recovering the shape of an object that is composed of smoothly curved surfaces that intersect along a number of sharp edges. They cleverly combine the methods of shape from shading for smooth surfaces and methods used in labeling line-drawings. The task is made more challenging by the fact that the label of a line may change along its length, something that does not happen when one considers line drawings of polyhedra.

Chapter 14: Perception of Solid Shape from Shading
by E. Mingolla and J.T. Todd

The authors explore human capabilities for estimating surface orientation in shaded images. Their experiments involved synthetic images of ellipsoids with varying axes and varying surface reflectance properties. The perception of surface shape was monitored by having the viewer compare estimated surface orientations with the orientations of known planar patches. The authors conclude that specular highlights and cast shadows have little influence on performance and that the observer need not know where the light source is. Apparently, perception of shape is distorted by a tendency to see the ellipsoids with axes aligned with the display surface. The authors provide some evidence that shape may not be recovered by the sort of local method favored by Pentland in chapter 15. They also suggest that the approach of Koenderink and van Doorn in chapter 10 may be helpful in understanding human performance.

This is an important paper, since there has been little in the way of quantitative psychophysical experimentation exploring human capabilities in this domain. There are only a few other pieces of work on this subject (see [Bilthoff & Mallot 87] and [Ramachandran 88b]).

Chapter 15: Local Shading Analysis
by A.P. Pentland

The author makes a case for the recovery of shape information from brightness patterns in small patches, and so argues against the reliance on information from singular points and the image projections of limbs of objects. To overcome the ambiguity inherent in shading, the assumption is made that the surface is locally (at least approximately) spherical. This single additional constraint is powerful enough to allow a solution based on first and second partial derivatives of brightness. In essence, the surface orientation is recovered by matching these derivatives with the brightness derivatives in an image of a sphere. (In some instances, a combination of brightness derivatives may be found that does not correspond to that on any part of the sphere—in these cases it is clear that the assumption made about the surface is not reasonable.)

Most shape-from-shading methods require knowledge of the reflectance map, which indirectly implies that one knows where the light sources are. In practice this information may not be explicitly provided. In this paper a method is proposed for recovering the direction of a single light source based on the statistics of the distribution of brightness derivative patterns in the image.

Chapter 16: Radarclinometry for the Venus Radar Mapper
by R.L. Wildey

The author uses a different strong constraint to allow the recovery of surface shape from local information: the surface is assumed to be locally cylindrical. The direction of the axis of the cylinder is determined by analysis of the derivatives of brightness in a small patch. It is illuminating to compare this approach with Pentland's presented in chapter 15.

What makes this paper particularly interesting is that it illustrates a parallel evolution of "photoclinometric" methods in the astrogeological community and "shape-from-shading" methods in the field of machine vision. The two groups apparently were unaware of each other's efforts until recently, and developed somewhat different terminology. This paper explores a particularly important application of such methods: the recovery of the shapes of the surfaces of other planets. It also exploits the similarity between shading in ordinary optical images and in images obtained using synthetic aperture radar, a topic mentioned also in chapter 5 by Frankot and Chellappa. The term *radarclinometry* is used to describe the recovery of surface slope information from synthetic aperture radar images.

Chapter 17: Photometric Method for Determining Surface Orientation from Multiple Images
by R.J. Woodham

Woodham developed the photometric stereo method described here while working on his Ph.D. thesis concerned with automated inspection. The basic idea is to get around the ambiguity inherent in a single measurement of image brightness by taking more measurements under different lighting conditions, rather than exploiting surface continuity or smoothness. Although this multiple exposure approach does not help explain how people use shading, since we do not usually have the opportunity to change the lighting at will, it does lead to a method of great practical importance. The reason is that the recovery of surface orientation is completely local and very simple, involving little more than table lookup. Calibration for different surface materials and different lighting conditions is also straightforward, requiring only observation of an object of known shape, such as a sphere, and construction of the lookup table.

This method may well spark the next revolution in the application of machine vision techniques to industrial problems, now that the limitations of binary image processing methods and edge-based processing are becoming apparent to most users. Photometric stereo was used, for example, in the system described in [Horn & Ikeuchi 84] for picking parts out of a pile of parts. Woodham, by the way, has written several papers on shape from shading and related methods, some of which include excellent surveys [Woodham 79a, 79b, 81, 84, 87].

Bibliography

Following the last chapter of the book are a comprehensive bibliography and a discussion of how the various references relate to different aspects of the shape-from-shading problem. The bibliography is surprisingly short. It seems that the problem of shape from shading is substantially more difficult than many others in the vision area, and so many researchers have shied away from it. It is encouraging, however, to see a renewed interest in the subject, as indicated by the large number of entries with recent dates.

The field of shape from shading appears to be maturing, judging by a number of significant recent pieces of work. People working on different aspects of the problem, even in quite different disciplines, are beginning to become more aware of each other's work. We hope that this collection of important papers will help speed up this process, and inspire newcomers to enter the field.

12. Acknowledgements

We thank Robert J. Woodham, Katsushi lkeuchi and Azriel Rosenfeld for suggestions regarding entries in the bibliography, Jean-Pierre Schott for converting several of the chapters into TEX format, Richard Kozera for help with proofreading, and Teresa A. Ehling for putting this book on an accelerated production schedule.