10-18 We might want to take more than two brightness measurements in order to improve accuracy when using the photometric stereo method. Imagine that n light sources are used in turn to obtain n images. Suppose that the surface under consideration is Lambertian and that the direction to the i^{th} source is given by the unit vector $\hat{\mathbf{s}}_i$. This time we assume that the surface can have an albedo ρ different from one. At each point in the image, we wish to find the unit surface normal $\hat{\mathbf{n}}$ that minimizes

$$\sum_{i=1}^{n} (\rho\,\hat{\mathbf{n}} \cdot \hat{\mathbf{s}}_i - E_i)^2 ,$$

where E_i is the i^{th} measurement of brightness at that point.

(a) Show that the vector that minimizes the error sum is

$$\rho\,\hat{\mathbf{n}} = \left[\sum_{i=1}^{n} \hat{\mathbf{s}}_i \hat{\mathbf{s}}_i^T\right]^{-1} \sum_{i=1}^{n} E_i \hat{\mathbf{s}}_i ,$$

where $\mathbf{a}\mathbf{b}^T$ is the *dyadic product* of the vectors $\mathbf{a}$ and $\mathbf{b}$,

$$\begin{pmatrix} a_x b_x & a_x b_y & a_x b_z \\ a_y b_x & a_y b_y & a_y b_z \\ a_z b_x & a_z b_y & a_z b_z \end{pmatrix} ,$$

and $[\,]^{-1}$ indicates the inverse of the matrix.

(b) What is the smallest number n of measurements needed to guarantee that the indicated matrix inverse exists? Warning: This part of the problem is nontrivial.

(c) How does all this change if we assume that the albedo is one? Hint: You may need to introduce a Lagrange multiplier to enforce this constraint, as explained in the appendix.

11

Reflectance Map: Shape from Shading

In the previous chapter we introduced the reflectance map and the image irradiance equation, and we used them to recover surface orientation from registered images taken under different lighting conditions. In this chapter we concentrate on the recovery of surface shape from a single image. This is a more difficult problem that will require the development of more advanced tools. We first examine the case of a linear reflectance map. It turns out that, under point-source illumination, the reflectance maps of the surface material in the maria of the moon and on rocky planets such as Mercury are functions of linear combinations of the components of the gradient. Next, we consider the shape-from-shading problem when the reflectance map is rotationally symmetric. This applies, for example, to images taken with the scanning electron microscope. We then solve the general case.

The image irradiance equation can be viewed as a nonlinear first-order partial differential equation. The traditional methods for solving such equations depend on growing characteristic strips. This is a sequential process. We are more interested in methods that ultimately lead to parallel algorithms. Consequently, we formulate a minimization problem that leads to a relaxation algorithm on a grid. We choose to minimize the integral of the difference between the observed brightness and that predicted for the estimated shape.

It is, of course, very important to know whether a solution to these problems exists and whether there is more than one solution. Unfortunately, these existence and uniqueness questions are difficult to decide without detailed assumptions about the reflectance map. We briefly explore what is known in this regard and then finish the chapter by showing how the ideas developed here can be applied to improve the results obtained by means of the photometric stereo method discussed in the previous chapter.

11.1 Recovering Shape from Shading

How can we recover the shape of a surface from a single image? Different parts of the surface are oriented differently and thus will appear with different brightnesses. We can take advantage of this spatial variation of brightness, referred to as *shading*, in estimating the orientation of surface patches. Measurement of brightness at a single point in the image, however, only provides one constraint, while surface orientation has two degrees of freedom. Without additional information, we cannot recover the orientation of a surface patch from the image irradiance equation

$$E(x, y) = R(p, q).$$

We have already discussed one method for introducing another constraint: the use of additional images taken under different lighting conditions.

11.1.1 Growing a Solution

But what if we have only one image? People can estimate the shapes of facial features using a single picture reproduced in a magazine. This suggests that there is enough information or that we implicitly introduce additional assumptions. Many surfaces are smooth, lacking discontinuities in depth. Also, there are often no discontinuities in the partial derivatives. An even wider class of objects have piecewise-smooth surfaces, with departures from smoothness concentrated along edges.

The assumption of smoothness provides a strong constraint. Neighboring patches of the surface cannot assume arbitrary orientations. They have to fit together to make a continuous, smooth surface. Thus a global method exploiting a smoothness constraint can be envisioned.

11.1.2 Linear Reflectance Maps

To begin with, we consider some special cases. Suppose that

$$R(p, q) = f(ap + bq),$$

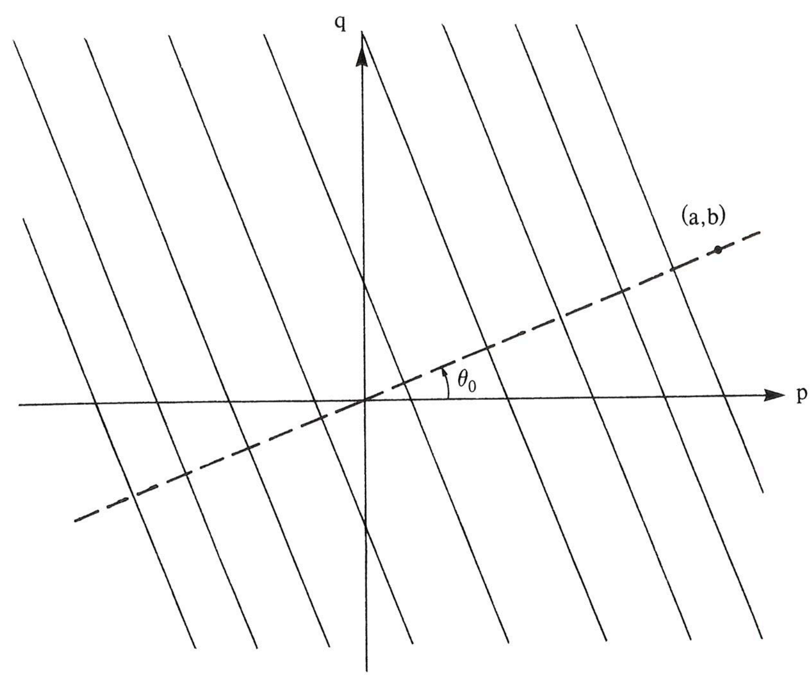

Figure 11-1. A reflectance map that is a function of a linear combination of the components of the gradient is particularly simple. The contours of constant brightness are parallel straight lines in gradient space.

where a and b are constants (figure 11-1).

Here f is a strictly monotonic function that has an inverse, f^{-1} (figure 11-2). From the image irradiance equation we then have

$$ap + bq = f^{-1}\big(E(x, y)\big).$$

We cannot determine the gradient (p, q) at a particular image point from a measurement of image brightness alone, but we do have one equation that constrains its possible values.

The slope of the surface, in a direction that makes an angle θ with the x-axis, is

$$m(\theta) = p\cos\theta + q\sin\theta.$$

This is the directional derivative. Now choose a particular direction θ_0 (figure 11-1), where $\tan\theta_0 = b/a$, that is,

$$\cos\theta_0 = a/\sqrt{a^2 + b^2} \qquad \text{and} \qquad \sin\theta_0 = b/\sqrt{a^2 + b^2}.$$

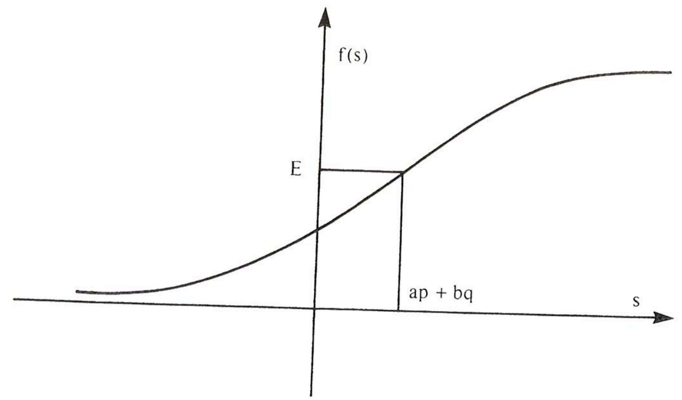

Figure 11-2. If the function f is continuous and monotonic, an inverse can be found and $s = ap+bq$ can be recovered from the brightness measurement $E(x,y)$.

The slope in this direction is

$$m(\theta_0) = \frac{ap+bq}{\sqrt{a^2+b^2}} = \frac{1}{\sqrt{a^2+b^2}} f^{-1}\big(E(x,y)\big).$$

Thus we can determine the slope in a particular direction. Note that we know nothing about the slope in the direction at right angles to this, however.

Starting at a particular image point we can take a small step of length $\delta\xi$, producing a change in z of $\delta z = m\,\delta\xi$. Thus

$$\frac{dz}{d\xi} = \frac{1}{\sqrt{a^2+b^2}} f^{-1}\big(E(x,y)\big),$$

where

$$x(\xi) = x_0 + \xi\cos\theta \qquad \text{and} \qquad y(\xi) = y_0 + \xi\sin\theta.$$

Suppose that we start the solution at the point $(x_0, y_0, z_0)^T$ on the surface. Integrating the differential equation for z derived above, we obtain

$$z(\xi) = z_0 + \frac{1}{\sqrt{a^2+b^2}} \int_0^{\xi} f^{-1}\big(E(x,y)\big)\, d\xi,$$

where x and y in the integrand are the linear functions of ξ given above. In this fashion we obtain a profile of the surface along a line in the special direction defined above (one of the straight lines in figure 11-3). The profile is called a *characteristic curve.* In practice, of course, the integrand will not be given as a formula, so that numerical integration is called for.

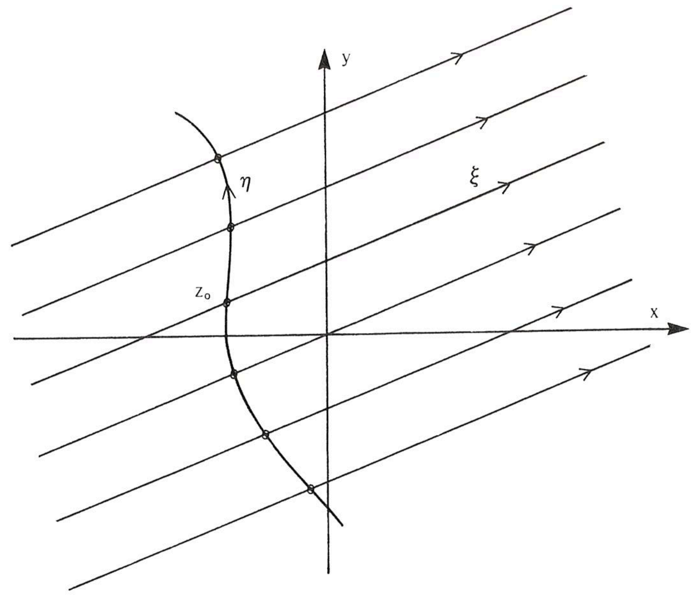

Figure 11-3. The base characteristics are parallel straight lines when the reflectance map is a function of a linear combination of the components of the gradient. The surface can be recovered by integration along these lines, provided the height $z_0(\eta)$ along some initial curve is given.

We cannot determine the absolute distance to the surface—the constant of integration—since the absolute distance does not influence the shading, only variations in depth do. If we require information about absolute distance, we shall need to know the value z_0 at one point. The shape can be recovered without this additional information, however.

Now suppose that we are given initial information not just at a point, but as a profile $z(\eta)$ along some curve that is nowhere parallel to the special direction (a,b) (figure 11-3). Then we can integrate along lines starting at points of this initial curve. The whole surface can be explored in this way if the initial curve extends far enough. The general case, to be explored later, is similar in that the surface is determined by integration along special curves in the image. The general case differs, however, in that these curves are not predetermined straight lines.

The special case discussed here is of practical importance because the material in the maria of the moon has reflectance properties that can be closely approximated by some function of $\cos\theta_i/\cos\theta_e$, as already mentioned. The reflectance map, in this case, is a function of a linear combination of p and q. This was the version of the shape-from-shading problem that first received attention. We use orthographic projection for simplicity here, but the method can be extended to the case of perspective projection.

11.1.3 Rotationally Symmetric Reflectance Maps

If the light source is distributed in a rotationally symmetric fashion about the viewer, then the reflectance map is rotationally symmetric, too. That is, we can write

$$R(p,q) = f(p^2+q^2)$$

for some f. One situation leading to a rotationally symmetric reflectance map is provided by a hemispherical sky, if we assume that the viewer is looking straight down from above. Another example is that of a point source at essentially the same place as the viewer.

Now suppose that the function f is strictly monotonic and differentiable, with inverse f^{-1}. From the image irradiance equation we obtain

$$p^2+q^2 = f^{-1}\big(E(x,y)\big).$$

The direction of steepest ascent makes an angle θ_s with the x-axis, where $\tan\theta_s = q/p$, so that

$$\cos\theta_s = p/\sqrt{p^2+q^2} \quad \text{and} \quad \sin\theta_s = q/\sqrt{p^2+q^2}.$$

The slope in the direction of steepest ascent is

$$m(\theta_s) = \sqrt{p^2+q^2} = \sqrt{f^{-1}\big(E(x,y)\big)}.$$

Thus in this case we can find the slope of the surface, given its brightness, but we cannot find the direction of steepest ascent.

Suppose we did know the direction of steepest ascent, given by (p,q). Then we could take a small step of length $\delta\xi$ in the direction of steepest ascent. The changes in x and y would be given by

$$\delta x = \frac{p}{\sqrt{p^2+q^2}}\delta\xi \quad \text{and} \quad \delta y = \frac{q}{\sqrt{p^2+q^2}}\delta\xi.$$

The change in z would be

$$\delta z = m\,\delta\xi = \sqrt{p^2+q^2}\,\delta\xi = \sqrt{f^{-1}\big(E(x,y)\big)}\,\delta\xi.$$

To simplify these equations, we could take a step of length $\sqrt{p^2+q^2}\,\delta\xi$ rather than $\delta\xi$. Then

$$\delta x = p\,\delta\xi, \quad \delta y = q\,\delta\xi, \quad \delta z = (p^2+q^2)\,\delta\xi = f^{-1}\big(E(x,y)\big)\,\delta\xi.$$

The problem with this approach is that we need to determine the values of p and q at the new point in order to continue the solution. We need to develop equations for the changes δp and δq in p and q, respectively.

Before we address this issue, let us look at the image brightness gradient $(E_x, E_y)^T$. We know that a planar surface patch gives rise to a region of uniform brightness in the image. Thus a nonzero brightness gradient can occur only where the surface is curved. To find the brightness gradient, we differentiate the image irradiance equation

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

with respect to x and y. Let r, s, and t be the second partial derivatives of z with respect to x and y as defined by

$$r = \frac{\partial^2 z}{\partial x^2}, \quad \frac{\partial^2 z}{\partial x \partial y} = s = \frac{\partial^2 z}{\partial y \partial x}, \quad t = \frac{\partial^2 z}{\partial y^2}.$$

Then, using the chain rule for differentiation, we obtain

$$E_x = 2(p\,r + q\,s)f' \quad \text{and} \quad E_y = 2(p\,s + q\,t)f',$$

where $f'(s)$ is the derivative of $f(s)$ with respect to its single argument s.

Now we return to the problem of determining the changes δp and δq occasioned by the step $(\delta x, \delta y)$ in the image plane. We find

$$\delta p = r\,\delta x + s\,\delta y \quad \text{and} \quad \delta q = s\,\delta x + t\,\delta y$$

by simple differentiation. In our case $\delta x = p\,\delta\xi$ and $\delta y = q\,\delta\xi$, so that

$$\delta p = (p\,r + q\,s)\,\delta\xi \quad \text{and} \quad \delta q = (p\,s + q\,t)\,\delta\xi,$$

or

$$\delta p = \frac{E_x}{2f'}\delta\xi \quad \text{and} \quad \delta q = \frac{E_y}{2f'}\delta\xi.$$

In the limit as $\delta\xi \to 0$, we obtain the differential equations

$$\dot{x} = p, \quad \dot{y} = q, \quad \dot{z} = p^2+q^2,$$

$$\dot{p} = \frac{E_x}{2f'}, \quad \dot{q} = \frac{E_y}{2f'},$$

where the dots denote differentiation with respect to ξ. Given starting values, this set of five ordinary differential equations can be solved numerically to produce a curve on the surface of the object. Curves generated in this fashion are called *characteristic curves*, and in this particular case they happen to be the curves of steepest ascent. These curves are everywhere perpendicular to the contours of constant height. In the case treated previously, in which the reflectance map was a linear function of p and q, the characteristic curves were parallel planar sections of the surface.

By differentiating $\dot{x} = p$ and $\dot{y} = q$ one more time with respect to ξ, we obtain the alternate formulation

$$\ddot{x} = \frac{E_x}{2f'}, \quad \ddot{y} = \frac{E_y}{2f'}, \quad \dot{z} = f^{-1}\big(E(x,y)\big).$$

Naturally, these equations can only be solved numerically, since E_x and E_y are image brightness measurements, not functions of x and y given in closed form.

The special case discussed above is of practical importance, since scanning electron microscopes produce images analogous to those produced in an optical system with a light source disposed around the viewer in a rotationally symmetric fashion. In such a device a focused beam of electrons strikes a surface in a position determined by two orthogonal deflection coils. Secondary electrons are generated as a result of collisions between the incident primary electrons and the atoms in the material. Some of these escape and are collected by an electrode. Secondary electrons generated deep inside the material have less of a chance to escape than those generated near the surface. The secondary electron flux is thus lowest when the beam strikes the surface at right angles and is highest at grazing incidence. The probing beam scans out a raster, while the brightness of a cathode ray tube scanned in the same fashion is modulated in proportion to the secondary electron current. The result is a (highly magnified) picture of the surface. People find such pictures easy to interpret, because they exhibit shading due to the dependence of brightness on surface orientation. The only strange thing about these images is that surface patches perpendicular to the viewer appear darkest, not brightest, in a scanning electron microscope picture.

11.1.4 The General Case

Suppose that we have the coordinates of a particular point on the surface and that we wish to extend the solution from this point. Taking a small step $(\delta x, \delta y)$, we note once more that the change in depth is given by

$$\delta z = p\,\delta x + q\,\delta y,$$

where p and q are the first partial derivatives of z with respect to x and y (figure 11-4). We cannot proceed unless p and q are also known. Unfortunately, the image irradiance equation provides only one constraint; this is not enough information to allow a solution for both p and q.

Suppose for the moment that we did know p and q at the given point. Then we could extend the solution from (x, y) to $(x + \delta x, y + \delta y)$. But to continue from there we would need the new values of p and q at that point (figure 11-5). Now the changes in p and q can be computed using

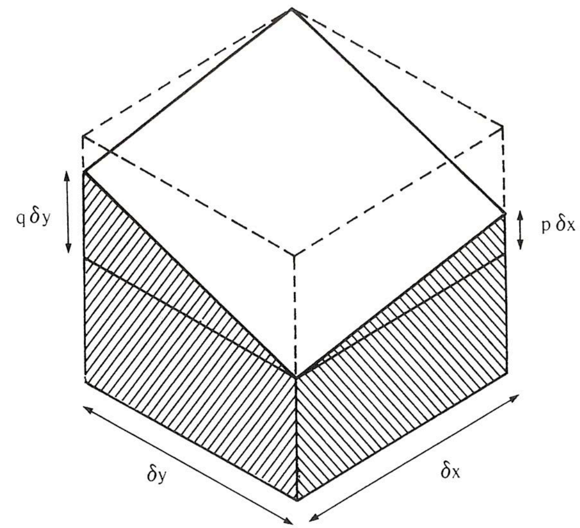

Figure 11-4. The change in height, δz, is the sum of $p\,\delta x$, the change in height due to a small step in the x-direction, and $q\,\delta y$, the change in height due to a small step in the y-direction.

$$\delta p = r\,\delta x + s\,\delta y \quad \text{and} \quad \delta q = s\,\delta x + t\,\delta y,$$

where r, s, and t are the second partial derivatives of z with respect to x and y. This can be written in a more compact form as

$$\begin{pmatrix} \delta p \\ \delta q \end{pmatrix} = \mathbf{H} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix},$$

where $\mathbf{H}$ is the *Hessian matrix* of second partial derivatives:

$$\mathbf{H} = \begin{pmatrix} r & s \\ s & t \end{pmatrix}.$$

The Hessian provides information on the curvature of the surface. For small surface inclinations, its determinant is the Gaussian curvature, to be introduced later. Also, the *trace* of the Hessian (the sum of its diagonal elements) is the Laplacian of depth, which for small surface inclinations is twice the so-called *mean curvature*. We shall explore surface curvature in chapter 16, where we discuss extended Gaussian images.

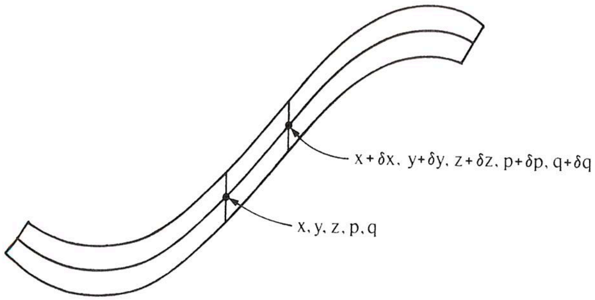

Figure 11-5. The solution of the shape-from-shading problem is determined by solving five differential equations for x, y, z, p, and q. The result is a characteristic strip, a curve in space along which surface orientation is known.

To use the Hessian matrix for computing the changes in p and q, we need to know its components, the second partial derivatives of z. To keep track of them we would need still higher derivatives. We could go on differentiating ad infinitum. Note, however, that we have not yet used the image irradiance equation! Differentiating it with respect to x and y, and using the chain rule, we obtain

$$E_x = r\,R_p + s\,R_q \qquad \text{and} \qquad E_y = s\,R_p + t\,R_q,$$

or

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \mathbf{H} \begin{pmatrix} R_p \\ R_q \end{pmatrix},$$

where the Hessian $\mathbf{H}$ once again makes an appearance. This is a relationship between the gradient $(E_x, E_y)^T$ in the image and the gradient $(R_p, R_q)^T$ in the reflectance map. We cannot solve for $\mathbf{H}$, since we have only two equations and three unknowns r, s, and t, but fortunately we do not need the individual elements of $\mathbf{H}$. While we cannot continue the solution in an arbitrary direction, we can do so in a specially chosen direction. This is the key idea. Let

$$\begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = \begin{pmatrix} R_p \\ R_q \end{pmatrix} \delta\xi,$$

where $\delta\xi$ is a small quantity. Then

$$\begin{pmatrix} \delta p \\ \delta q \end{pmatrix} = \mathbf{H} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = \mathbf{H} \begin{pmatrix} R_p \\ R_q \end{pmatrix} \delta\xi,$$

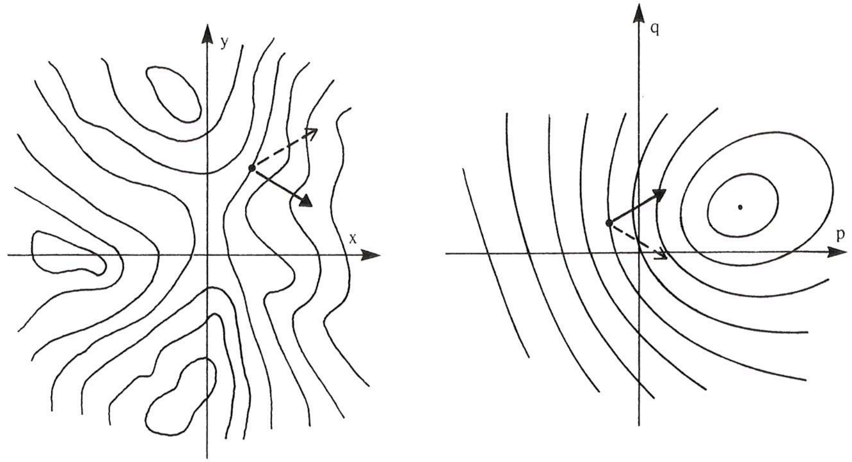

Figure 11-6. Curiously, the step taken in pq-space is parallel to the gradient of $E(x, y)$, while the step taken in xy-space is parallel to the gradient of $R(p, q)$.

or

$$\begin{pmatrix} \delta p \\ \delta q \end{pmatrix} = \begin{pmatrix} E_x \\ E_y \end{pmatrix} \delta\xi.$$

Thus, if the direction of the change in the image plane is parallel to the gradient of the reflectance map, then the change in (p, q) can be computed. The direction of the change in gradient space is parallel, in turn, to the gradient in the image (figure 11-6). We can summarize all this in five ordinary differential equations:

$$\dot{x} = R_p, \qquad \dot{y} = R_q, \qquad \dot{z} = p\,R_p + q\,R_q,$$

$$\dot{p} = E_x, \qquad \dot{q} = E_y,$$

where the dots denote differentiation with respect to ξ. A solution of these differential equations is a curve on the surface. The parameter ξ will vary along this curve. By rescaling the equations, we can easily arrange for ξ to be any function of length along the curve.

11.2 Characteristic Curves and Initial Curves

The curves traced out by the solutions of the five ordinary differential equations are called *characteristic curves*, and their projections in the image are called *base characteristics*. The solutions for x, y, z, p, and q actually form a *characteristic strip*, since they define not only a curve in space but surface orientation along this curve as well (figure 11-5).

To obtain the whole surface we must patch together characteristic strips. Each requires a point where initial values are given in order to start the solution. If we are given an initial curve on the surface, a solution for the surface can be obtained as long as this curve is nowhere parallel to any of the characteristics. On this curve, starting values of p and q can be obtained using the image irradiance equation,

$$E(x, y) = R(p, q),$$

and the known derivatives of z along the curve. Suppose, for example, that the initial curve is given in terms of a parameter η, as $x(\eta)$, $y(\eta)$, and $z(\eta)$. Then, along this curve,

$$\frac{\partial z}{\partial \eta} = p\frac{\partial x}{\partial \eta} + q\frac{\partial y}{\partial \eta}.$$

We have just derived the method of characteristic strip expansion for solving first-order partial differential equations. In our case the relevant equation is the image irradiance equation, a (possibly very nonlinear) first-order partial differential equation.

Figure 11-7 shows a digitized picture of a face, the face with base characteristics superimposed, and the face with a contour map of the recovered shape.

11.3 Singular Points

We are normally not given an initial curve along with the image of an object. How much can we tell about shape in the absence of such auxiliary information? Are there any points where surface orientation can be determined directly? Suppose that $R(p, q)$ has a unique isolated maximum at (p_0, q_0); that is,

$$R(p, q) < R(p_0, q_0) \qquad \text{for all } (p, q) \neq (p_0, q_0).$$

Also assume that at some point (x_0, y_0) in the image, $E(x_0, y_0) = R(p_0, q_0)$. Then it is clear that at this point the gradient (p, q) is uniquely determined to be (p_0, q_0). It would seem, then, that we could start the solution at such a singular point. Unfortunately, at a maximum of $R(p, q)$ the partial derivatives R_p and R_q are zero. Thus the solution will not move from such a point because $\dot{x}$ and $\dot{y}$ are zero. One way to bypass this apparent impasse is to construct a small "cap" at this point and start the solution at the edge of this cap, as we shall show in the next section.

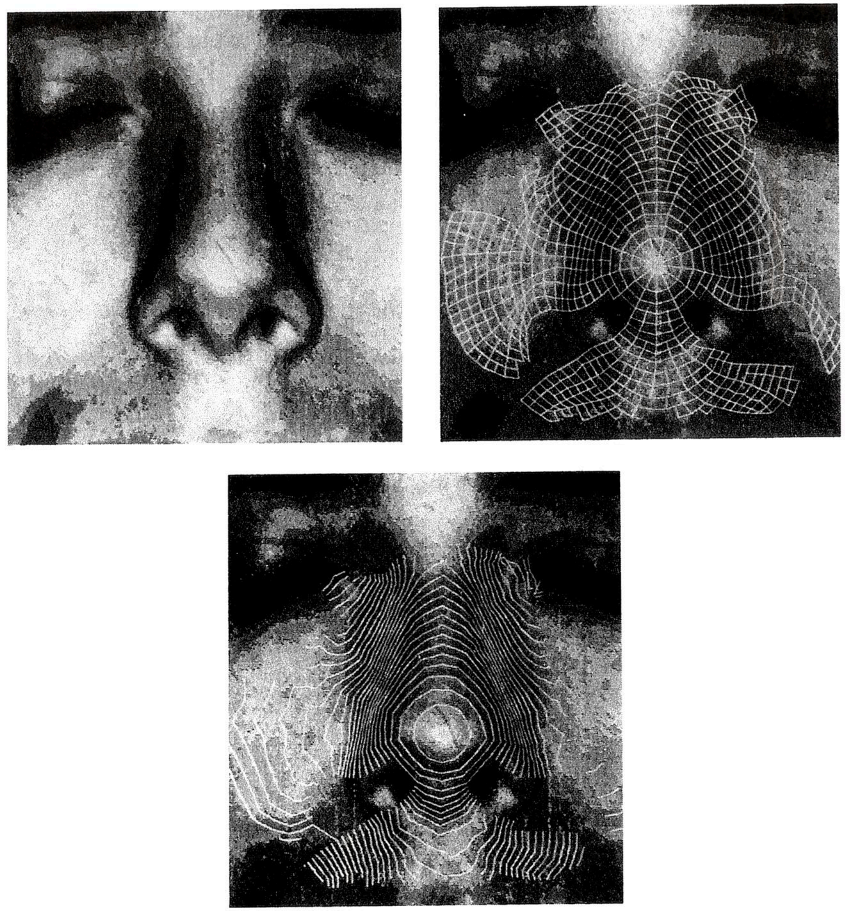

Figure 11-7. The shape-from-shading method is applied here to the recovery of the shape of a nose. The first picture shows the (crudely quantized) gray-level image available to the program. The second picture shows the base characteristics superimposed, while the third shows a contour map computed from the elevations found along the characteristic curves.

11.4 Power Series near a Singular Point

To observe what happens near a singular point, consider the reflectance map

$$R(p, q) = \frac{1}{2}(p^2 + q^2).$$