CHAPTER 2

Light and Shading

The brightness of a pixel in the image is a function of the brightness of the surface patch in the scene that projects to the pixel. In turn, the brightness of the patch depends on how much incident light arrives at the patch and on the fraction of the incident light that gets reflected (Models in Section 2.1).

This means that the brightness of a pixel is profoundly ambiguous. Surprisingly, people can disentangle these effects quite accurately. Often, but not always, people can tell whether objects are in bright light or in shadow, and do not perceive objects in shadow as having dark surfaces. People can usually tell whether changes of brightness are caused by changes in reflection or by shading (cinemas wouldn't work if we got it right all the time, however). Typically, people can tell that shading comes from the geometry of the object, but sometimes get shading and markings mixed up. For example, a streak of dark makeup under a cheekbone will often look like a shading effect, making the face look thinner. Quite simple models of shading (Section 2.1) support a range of inference procedures (Section 2.2). More complex models are needed to explain some important effects (Section 2.1.4), but make inference very difficult indeed (Section 2.4).

2.1 MODELLING PIXEL BRIGHTNESS

Three major phenomena determine the brightness of a pixel: the response of the camera to light, the fraction of light reflected from the surface to the camera, and the amount of light falling on the surface. Each can be dealt with quite straightforwardly.

Camera response: Modern cameras respond linearly to middling intensities of light, but have pronounced nonlinearities for darker and brighter illumination. This allows the camera to reproduce the very wide dynamic range of natural light without saturating. For most purposes, it is enough to assume that the camera response is linearly related to the intensity of the surface patch. Write $\boldsymbol{X}$ for a point in space that projects to $\boldsymbol{x}$ in the image, $I_{patch}(\boldsymbol{X})$ for the intensity of the surface patch at $\boldsymbol{X}$, and $I_{camera}(\boldsymbol{x})$ for the camera response at $\boldsymbol{x}$. Then our model is:

$$I_{camera}(\boldsymbol{x}) = kI_{patch}(\boldsymbol{x}),$$

where k is some constant to be determined by calibration. Generally, we assume that this model applies and that k is known if needed. Under some circumstances, a more complex model is appropriate; we discuss how to recover such models in Section 2.2.1.

Surface reflection: Different points on a surface may reflect more or less of the light that is arriving. Darker surfaces reflect less light, and lighter surfaces reflect more. There is a rich set of possible physical effects, but most can be ignored. Section 2.1.1 describes the relatively simple model that is sufficient for almost all

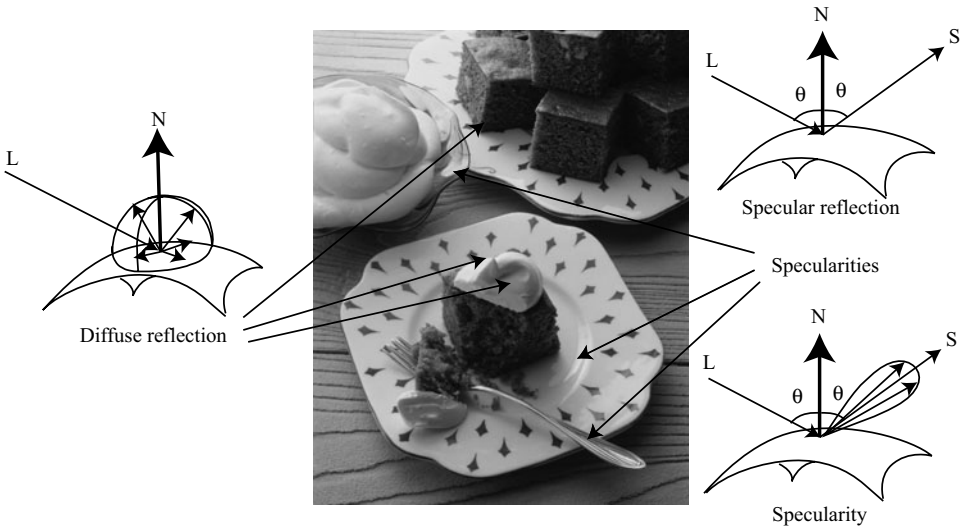

FIGURE 2.1: The two most important reflection modes for computer vision are diffuse reflection (**left**), where incident light is spread evenly over the whole hemisphere of outgoing directions, and specular reflection (**right**), where reflected light is concentrated in a single direction. The specular direction $\boldsymbol{S}$ is coplanar with the normal and the source direction ($\boldsymbol{L}$), and has the same angle to the normal that the source direction does. Most surfaces display both diffuse and specular reflection components. In most cases, the specular component is not precisely mirror like, but is concentrated around a range of directions close to the specular direction (**lower right**). This causes specularities, where one sees a mirror like reflection of the light source. Specularities, when they occur, tend to be small and bright. In the photograph, they appear on the metal spoon and on the plate. Large specularities can appear on flat metal surfaces (arrows). Most curved surfaces (such as the plate) show smaller specularities. Most of the reflection here is diffuse; some cases are indicated by arrows. *Martin Brigdale © Dorling Kindersley, used with permission.*

purposes in computer vision.

Illumination: The amount of light a patch receives depends on the overall intensity of the light, and on the geometry. The overall intensity could change because some *luminaires* (the formal term for light sources) might be shadowed, or might have strong directional components. Geometry affects the amount of light arriving at a patch because surface patches facing the light collect more radiation and so are brighter than surface patches tilted away from the light, an effect known as *shading*. Section 2.1.2 describes the most important model used in computer vision; Section 2.3 describes a much more complex model that is necessary to explain some important practical difficulties in shading inference.

2.1.1 Reflection at Surfaces

Most surfaces reflect light by a process of *diffuse reflection*. Diffuse reflection scatters light evenly across the directions leaving a surface, so the brightness of a diffuse surface doesn't depend on the viewing direction. Examples are easy to identify with

this test: most cloth has this property, as do most paints, rough wooden surfaces, most vegetation, and rough stone or concrete. The only parameter required to describe a surface of this type is its *albedo*, the fraction of the light arriving at the surface that is reflected. This does not depend on the direction in which the light arrives or the direction in which the light leaves. Surfaces with very high or very low albedo are difficult to make. For practical surfaces, albedo lies in the range 0.05 – 0.90 (see Brelstaff and Blake (1988*b*), who argue the dynamic range is closer to 10 than the 18 implied by these numbers). Mirrors are not diffuse, because what you see depends on the direction in which you look at the mirror. The behavior of a perfect mirror is known as *specular reflection*. For an ideal mirror, light arriving along a particular direction can leave only along the *specular direction*, obtained by reflecting the direction of incoming radiation about the surface normal (Figure 2.1). Usually some fraction of incoming radiation is absorbed; on an ideal specular surface, this fraction does not depend on the incident direction.

If a surface behaves like an ideal specular reflector, you could use it as a mirror, and based on this test, relatively few surfaces actually behave like ideal specular reflectors. Imagine a near perfect mirror made of polished metal; if this surface suffers slight damage at a small scale, then around each point there will be a set of small facets, pointing in a range of directions. In turn, this means that light arriving in one direction will leave in several different directions because it strikes several facets, and so the specular reflections will be blurred. As the surface becomes less flat, these distortions will become more pronounced; eventually, the only specular reflection that is bright enough to see will come from the light source. This mechanism means that, in most shiny paint, plastic, wet, or brushed metal surfaces, one sees a bright blob—often called a *specularity*—along the specular direction from light sources, but few other specular effects. Specularities are easy to identify, because they are small and very bright (Figure 2.1; Brelstaff and Blake (1988*b*)). Most surfaces reflect only some of the incoming light in a specular component, and we can represent the percentage of light that is specularly reflected with a *specular albedo*. Although the diffuse albedo is an important material property that we will try to estimate from images, the specular albedo is largely seen as a nuisance and usually is not estimated.

2.1.2 Sources and Their Effects

The main source of illumination outdoors is the sun, whose rays all travel parallel to one another in a known direction because it is so far away. We model this behavior with a *distant point light source*. This is the most important model of lighting (because it is like the sun and because it is easy to use), and can be quite effective for indoor scenes as well as outdoor scenes. Because the rays are parallel to one another, a surface that faces the source cuts more rays (and so collects more light) than one oriented along the direction in which the rays travel. The amount of light collected by a surface patch in this model is proportional to the cosine of the angle θ between the illumination direction and the normal (Figure 2.2). The figure yields *Lambert's cosine law*, which states the brightness of a diffuse patch illuminated by a distant point light source is given by

$$I = \rho I_0 \cos\theta,$$

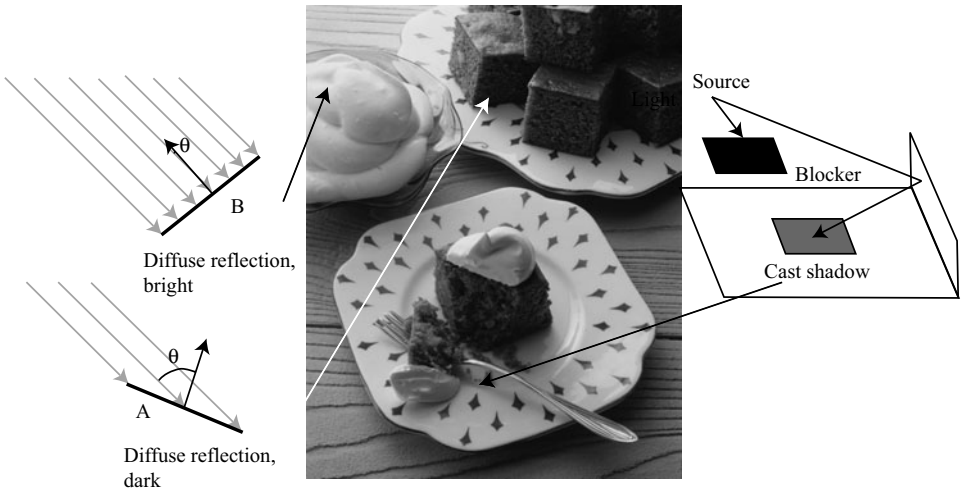

FIGURE 2.2: The orientation of a surface patch with respect to the light affects how much light the patch gathers. We model surface patches as illuminated by a distant point source, whose rays are shown as light arrowheads. Patch A is tilted away from the source (θ is close to 90^0) and collects less energy, because it cuts fewer light rays per unit surface area. Patch B, facing the source (θ is close to 0^0), collects more energy, and so is brighter. Shadows occur when a patch cannot see a source. The shadows are not dead black, because the surface can see interreflected light from other surfaces. These effects are shown in the photograph. The darker surfaces are turned away from the illumination direction. *Martin Brigdale © Dorling Kindersley, used with permission.*

where I_0 is the intensity of the light source, θ is the angle between the light source direction and the surface normal, and ρ is the diffuse albedo. This law predicts that bright image pixels come from surface patches that face the light directly and dark pixels come from patches that see the light only tangentially, so that the shading on a surface provides some shape information. We explore this cue in Section 2.4.

If the surface cannot see the source, then it is in *shadow*. Since we assume that light arrives at our patch only from the distant point light source, our model suggests that shadows are deep black; in practice, they very seldom are, because the shadowed surface usually receives light from other sources. Outdoors, the most important such source is the sky, which is quite bright. Indoors, light reflected from other surfaces illuminates shadowed patches. This means that, for example, we tend to see few shadows in rooms with white walls, because any shadowed patch receives a lot of light from the walls. These *interreflections* also can have a significant effect on the brightness surfaces that are not in shadow. Interreflection effects are sometimes modelled by adding a constant *ambient illumination* term to the predicted intensity. The ambient term ensures that shadows are not too dark, but this is not a particularly successful model of the spatial properties of interreflections. More detailed models require some familiarity with radiometric terminology, but they are important in some applications; we have confined this topic to Section 2.3.

2.1.3 The Lambertian+Specular Model

For almost all purposes, it is enough to model all surfaces as being diffuse with specularities. This is the *lambertian+specular model*. Specularities are relatively seldom used in inference (Section 2.2.2 sketches two methods), and so there is no need for a formal model of their structure. Because specularities are small and bright, they are relatively easy to identify and remove with straightforward methods (find small bright spots, and replace them by smoothing the local pixel values). More sophisticated specularity finders use color information (Section 3.5.1). Thus, to apply the lambertian+specular model, we find and remove specularities, and then use Lambert's law (Section 2.1.2) to model image intensity.

We must choose which source effects to model. In the simplest case, a *local shading model*, we assume that shading is caused only by light that comes from the luminaire (i.e., that there are no interreflections).

Now assume that the luminaire is an infinitely distant source. For this case, write $\boldsymbol{N}(x)$ for the unit surface normal at $\boldsymbol{x}$, $\boldsymbol{S}$ for a vector pointing from $\boldsymbol{x}$ toward the source with length I_o (the source intensity), $\rho(\boldsymbol{x})$ for the albedo at $\boldsymbol{x}$, and $Vis(\boldsymbol{S}, \boldsymbol{x})$ for a function that is 1 when $\boldsymbol{x}$ can see the source and zero otherwise. Then, the intensity at $\boldsymbol{x}$ is

$$\begin{array}{ccccccc} I(\boldsymbol{x}) & = & \rho(\boldsymbol{x})\,(\boldsymbol{N}\cdot\boldsymbol{S})\,\mathrm{Vis}(\boldsymbol{S},\,\boldsymbol{x}) & + & \rho(\boldsymbol{x})A & + & M \\ \\ \begin{array}{c}\text{Image}\\ \text{intensity}\end{array} & = & \begin{array}{c}\text{Diffuse}\\ \text{term}\end{array} & + & \begin{array}{c}\text{Ambient}\\ \text{term}\end{array} & + & \begin{array}{c}\text{Specular (mirror-like)}\\ \text{term}\end{array} \end{array}$$

This model can still be used for a more complex source (for example, an area source), but in that case it is more difficult to determine an appropriate $\boldsymbol{S}(\boldsymbol{x})$.

2.1.4 Area Sources

An *area source* is an area that radiates light. Area sources occur quite commonly in natural scenes—an overcast sky is a good example—and in synthetic environments—for example, the fluorescent light boxes found in many industrial ceilings. Area sources are common in illumination engineering, because they tend not to cast strong shadows and because the illumination due to the source does not fall off significantly as a function of the distance to the source. Detailed models of area sources are complex (Section 2.3), but a simple model is useful to understand shadows. Shadows from area sources are very different from shadows cast by point sources. One seldom sees dark shadows with crisp boundaries indoors. Instead, one could see no visible shadows, or shadows that are rather fuzzy diffuse blobs, or sometimes fuzzy blobs with a dark core (Figure 2.3). These effects occur because rooms tend to have light walls and diffuse ceiling fixtures, which act as area sources. As a result, the shadows one sees are area source shadows.

To compute the intensity at a surface patch illuminated by an area source, we can break the source up into infinitesimal source elements, then sum effects from each element. If there is an occluder, then some surface patches may see none of the source elements. Such patches will be dark, and lie in the *umbra* (a Latin word meaning "shadow"). Other surface patches may see some, but not all, of the source elements. Such patches may be quite bright (if they see most of the elements), or

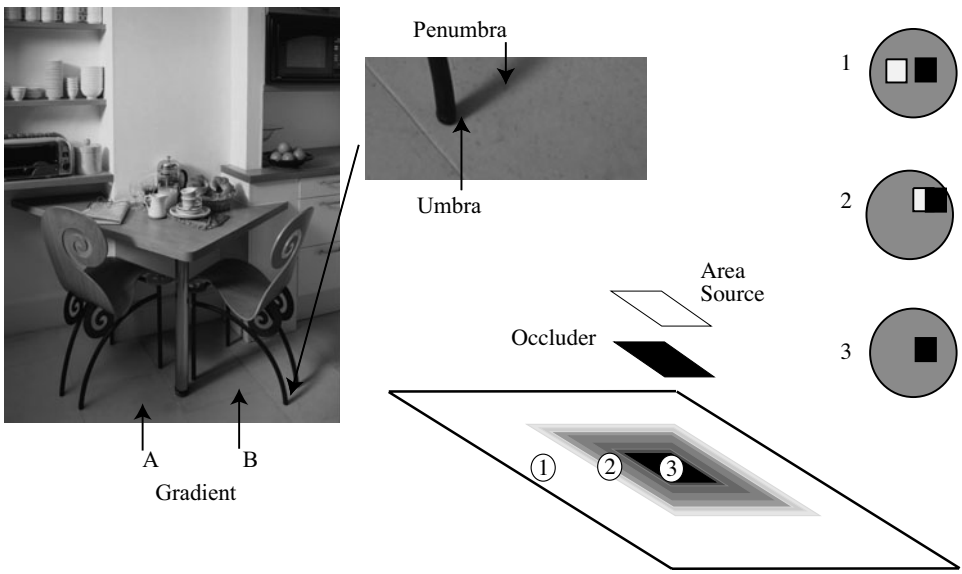

FIGURE 2.3: Area sources generate complex shadows with smooth boundaries, because from the point of view of a surface patch, the source disappears slowly behind the occluder. **Left:** a photograph, showing characteristic area source shadow effects. Notice that A is much darker than B; there must be some shadowing effect here, but there is no clear shadow boundary. Instead, there is a fairly smooth gradient. The chair leg casts a complex shadow, with two distinct regions. There is a core of darkness (the *umbra*—where the source cannot be seen at all) surrounded by a partial shadow (*penumbra*—where the source can be seen partially). A good model of the geometry, illustrated **right**, is to imagine lying with your back to the surface looking at the world above. At point 1, you can see all of the source; at point 2, you can see some of it; and at point 3, you can see none of it. *Peter Anderson © Dorling Kindersley, used with permission.*

relatively dark (if they see few elements), and lie in the *penumbra* (a compound of Latin words meaning "almost shadow"). One way to build intuition is to think of a tiny observer looking up from the surface patch. At umbral points, this observer will not see the area source at all whereas at penumbral points, the observer will see some, but not all, of the area source. An observer moving from outside the shadow, through the penumbra and into the umbra will see something that looks like an eclipse of the moon (Figure 2.3). The penumbra can be large, and can change quite slowly from light to dark. There might even be no umbral points at all, and, if the occluder is sufficiently far away from the surface, the penumbra could be very large and almost indistinguishable in brightness from the unshadowed patches. This is why many objects in rooms appear to cast no shadow at all (Figure 2.4).

2.2 INFERENCE FROM SHADING

Shading can be used to infer a variety of properties of the visual world. Successful inference often requires that we calibrated the camera radiometrically, so that we know how pixel values map to radiometric values (Section 2.2.1). As Figure 2.1

FIGURE 2.4: The photograph on the **left** shows a room interior. Notice the lighting has some directional component (the vertical face indicated by the arrow is dark, because it does not face the main direction of lighting), but there are few visible shadows (for example, the chairs do not cast a shadow on the floor). On the **right**, a drawing to show why; here there is a small occluder and a large area source. The occluder is some way away from the shaded surface. Generally, at points on the shaded surface the incoming hemisphere looks like that at point 1. The occluder blocks out some small percentage of the area source, but the amount of light lost is too small to notice (compare figure 2.3). *Jake Fitzjones © Dorling Kindersley, used with permission.*

suggests, specularities are a source of information about the shape of a surface, and Section 2.2.2 shows how this information can be interpreted. Section 2.2.3 shows how to recover the albedoes of surfaces from images. Finally, Section 2.2.4 shows how multiple shaded images can be used to recover surface shape.

2.2.1 Radiometric Calibration and High Dynamic Range Images

Real scenes often display a much larger range of intensities than cameras can cope with. Film and charge-coupled devices respond to energy. A property called *reciprocity* means that, if a scene patch casts intensity E onto the film, and if the shutter is open for time Δt, the response is a function of $E\Delta t$ alone. In particular, we will get the same outcome if we image one patch of intensity E for time Δt and another patch of intensity E/k for time $k\Delta t$. The actual response that the film produces is a function of $E\Delta t$; this function might depend on the imaging system, but is typically somewhat linear over some range, and sharply non-linear near the top and bottom of this range, so that the image can capture very dark and very light patches without saturation. It is usually monotonically increasing.

There are a variety of applications where it would be useful to know the actual radiance (equivalently, the intensity) arriving at the imaging device. For example, we might want to compare renderings of a scene with pictures of the scene, and to do that we need to work in real radiometric units. We might want to use pictures of a scene to estimate the lighting in that scene so we can postrender new objects into the scene, which would need to be lit correctly. To infer radiance, we must determine the film response, a procedure known as *radiometric calibration*. As we

shall see, doing this will require more than one image of a scene, each obtained at different exposure settings. Imagine we are looking at a scene of a stained glass window lit from behind in a church. At one exposure setting, we would be able to resolve detail in the dark corners, but not on the stained glass, which would be saturated. At another setting, we would be able to resolve detail on the glass, but the interior would be too dark. If we have both settings, we may as well try to recover radiance with a very large dynamic range—producing a *high dynamic range image*.

Now assume we have multiple registered images, each obtained using a different exposure time. At the i, j'th pixel, we know the image intensity value $I_{ij}^{(k)}$ for the k'th exposure time, we know the value of the k'th exposure time Δt_k, and we know that the intensity of the corresponding surface patch E_{ij} is the same for each exposure, but we do not know the value of E_{ij}. Write the camera response function f, so that

$$I_{ij}^{(k)} = f(E_{ij}\Delta t_k).$$

There are now several possible approaches to solve for f. We could assume a parametric form—say, polynomial—then solve using least squares. Notice that we must solve not only for the parameters of f, but also for E_{ij}. For a color camera, we solve for calibration of each channel separately. Mitsunaga and Nayar (1999) have studied the polynomial case in detail. Though the solution is not unique, ambiguous solutions are strongly different from one another, and most cases are easily ruled out. Furthermore, one does not need to know exposure times with exact accuracy to estimate a solution, as long as there are sufficient pixel values; instead, one estimates f from a fixed set of exposure times, then estimates the exposure times from f, and then re-estimates. This procedure is stable.

Alternatively, because the camera response is monotonic, we can work with its inverse $g = f^{-1}$, take logs, and write

$$\log g(I_{ij}^{(k)}) = \log E_{ij} + \log \Delta t_k.$$

We can now estimate the values that g takes at each point and the E_{ij} by placing a smoothness penalty on g. In particular, we minimize

$$\sum_{i,j,k} (\log g(I_{ij}^{(k)}) - (\log E_{ij} + \log \Delta t_k))^2 + \text{smoothness penalty on } g$$

by choice of g. Debevec and Malik (1997) penalize the second derivative of g. Once we have a radiometrically calibrated camera, estimating a high dynamic range image is relatively straightforward. We have a set of registered images, and at each pixel location, we seek the estimate of radiance that predicts the registered image values best. In particular, we assume we know f. We seek an E_{ij} such that

$$\sum_k w(I_{ij})(I_{ij}^{(k)} - f(E_{ij}\Delta t_k))^2$$

is minimized. Notice the weights because our estimate of f is more reliable when I_{ij} is in the middle of the available range of values than when it is at larger or smaller values.

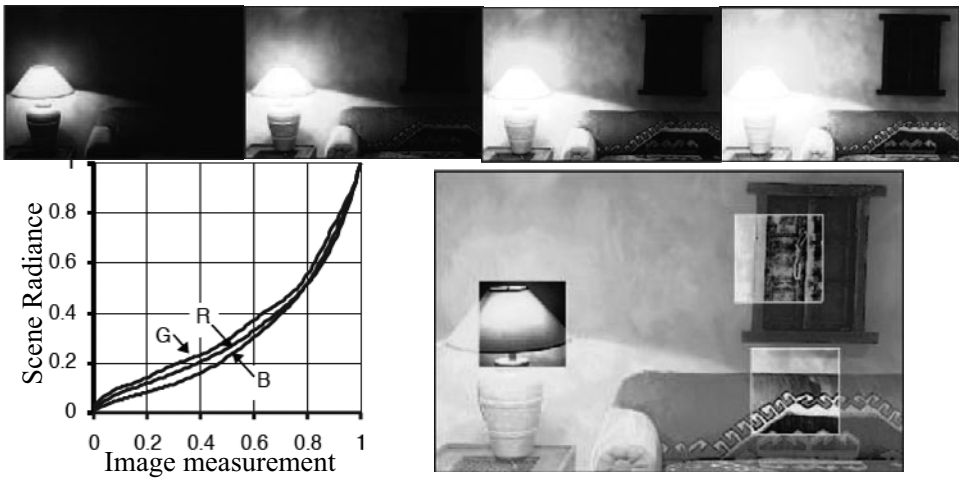

FIGURE 2.5: It is possible to calibrate the radiometric response of a camera from multiple images obtained at different exposures. The **top** row shows four different exposures of the same scene, ranging from darker (shorter shutter time) to lighter (longer shutter time). Note how, in the dark frames, the lighter part of the image shows detail, and in the light frames, the darker part of the image shows detail; this is the result of non-linearities in the camera response. On the **bottom left**, we show the inferred calibration curves for each of the R, G, and B camera channels. On the **bottom right**, a composite image illustrates the results. The dynamic range of this image is far too large to print; instead, the main image is normalized to the print range. Overlaid on this image are boxes where the radiances in the box have also been normalized to the print range; these show how much information is packed into the high dynamic range image. *This figure was originally published as Figure 7 of "Radiometric Self Calibration," by T. Mitsunaga and S. Nayar, Proc. IEEE CVPR 1999, © IEEE, 1999.*

2.2.2 The Shape of Specularities

Specularities are informative. They offer hints about the color of illumination (see Chapter 3) and offer cues to the local geometry of a surface. Understanding these cues is a useful exercise in differential geometry. We consider a smooth specular surface and assume that the brightness reflected in the direction $\boldsymbol{V}$ is a function of $\boldsymbol{V} \cdot \boldsymbol{P}$, where $\boldsymbol{P}$ is the specular direction. We expect the specularity to be small and isolated, so we can assume that the source direction $\boldsymbol{S}$ and the viewing direction $\boldsymbol{V}$ are constant over its extent. Let us further assume that the specularity can be defined by a threshold on the specular energy, i.e., $\boldsymbol{V} \cdot \boldsymbol{P} \geq 1 - \varepsilon$ for some constant ε, denote by $\boldsymbol{N}$ the unit surface normal, and define the *half-angle direction* as $\boldsymbol{H} = (\boldsymbol{S} + \boldsymbol{V})/2$ (Figure 2.6(left)). Using the fact that the vectors $\boldsymbol{S}$, $\boldsymbol{V}$ and $\boldsymbol{P}$ have unit length and a whit of plane geometry, it can easily be shown that the boundary of the specularity is defined by (see exercises)

$$1 - \varepsilon = \boldsymbol{V} \cdot \boldsymbol{P} = 2\frac{(\boldsymbol{H} \cdot \boldsymbol{N})^2}{(\boldsymbol{H} \cdot \boldsymbol{H})} - 1. \tag{2.1}$$

Because the specularity is small, the second-order structure of the surface

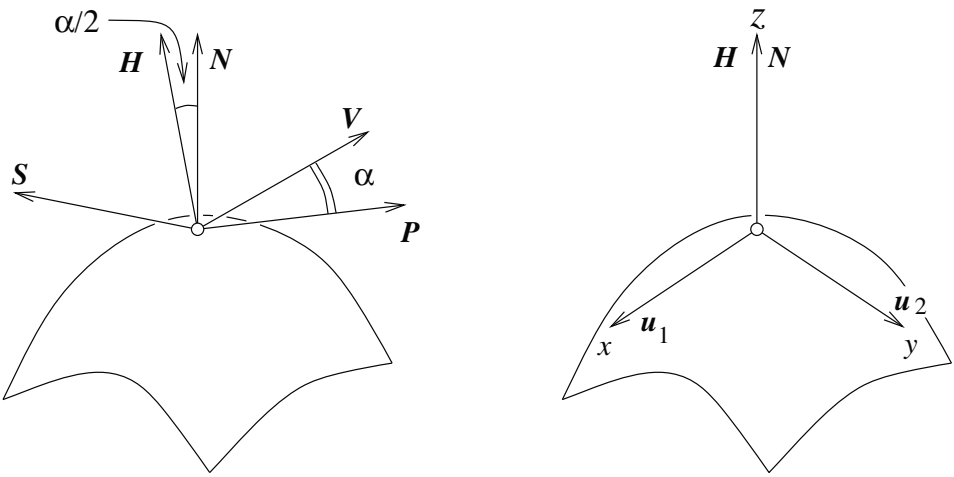

FIGURE 2.6: A specular surface viewed by a distant observer. We establish a coordinate system at the brightest point of the specularity (where the half-angle direction is equal to the normal) and orient the system using the normal and principal directions.

will allow us to characterize the shape of its boundary as follows: there is some point on the surface inside the specularity (in fact, the brightest point) where $\boldsymbol{H}$ is parallel to $\boldsymbol{N}$. We set up a coordinate system at this point, oriented so that the z-axis lies along $\boldsymbol{N}$ and the x- and y-axes lie along the principal directions $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ (Figure 2.6(right)). As noted earlier, the surface can be represented up to second order as $z = -1/2(\kappa_1 x^2 + \kappa_2 y^2)$ in this frame, where κ_1 and κ_2 are the principal curvatures. Now, let us define a *parametric surface* as a differentiable mapping $\boldsymbol{x} : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ associating with any couple $(u, v) \in U$ the coordinate vector $(x, y, z)^T$ of a point in some fixed coordinate system. It is easily shown (see exercises) that the normal to a parametric surface is along the vector $\frac{\partial}{\partial u}\boldsymbol{x} \times \frac{\partial}{\partial v}\boldsymbol{x}$. Our second-order surface model is a parametric surface parameterized by x and y, thus its unit surface normal is defined in the corresponding frame by

$$\boldsymbol{N}(x, y) = \frac{1}{\sqrt{1 + \kappa_1^2 x^2 + \kappa_2^2 y^2}} \begin{pmatrix} \kappa_1 x \\ \kappa_2 y \\ 1 \end{pmatrix},$$

and $\boldsymbol{H} = (0, 0, 1)^T$. Because $\boldsymbol{H}$ is a constant, we can rewrite Eq. (2.1) as $\kappa_1^2 x^2 + \kappa_2^2 y^2 = \zeta$, where ζ is a constant depending on ε. In particular, the shape of the specularity on the surface contains information about the second fundamental form. The specularity will be an ellipse, with major and minor axes oriented along the principal directions, and an eccentricity equal to the ratio of the principal curvatures. Unfortunately, the shape of the specularity on the surface is not, in general, directly observable, so this property can be exploited only when a fair amount about the viewing and illumination setup is known (Healey and Binford 1986).

Although we cannot get much out of the shape of the specularity in the image, it is possible to tell a convex surface from a concave one by watching how a

specularity moves as the view changes (you can convince yourself of this with the aid of a spoon).[1] Let us consider a point source at infinity and assume that the specular lobe is very narrow, so the viewing direction and the specular direction coincide. Initially, the specular direction is $\boldsymbol{V}$, and the specularity is at the surface point P; after a small eye motion, $\boldsymbol{V}$ changes to $\boldsymbol{V}'$, while the specularity moves to the close-by point P' (Figure 2.7).

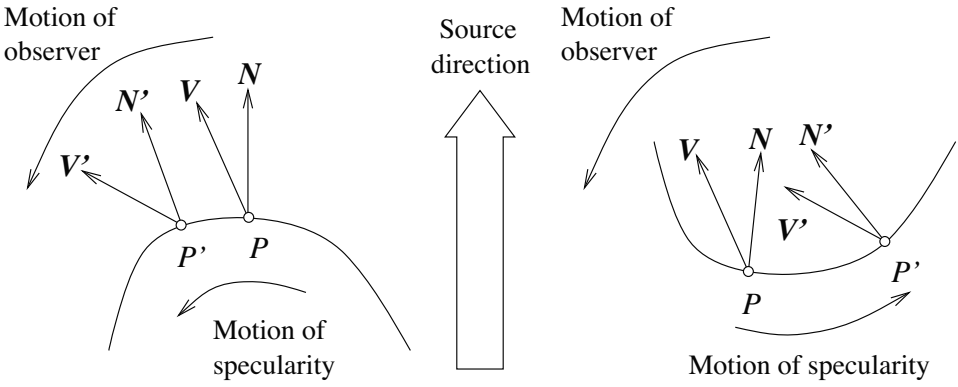

FIGURE 2.7: Specularities on convex and concave surfaces behave differently when the view changes. With an appropriate choice of source direction and motion, this could be used to obtain the signs of the principal curvatures.

The quantity of interest is $\delta a = (\boldsymbol{V}' - \boldsymbol{V}) \cdot \boldsymbol{t}$, where $\boldsymbol{t} = \frac{1}{\delta s} \overrightarrow{PP'}$ is tangent to the surface, and δs is the (small) distance between P and P': if δa is positive, then the specularity moves in the direction of the view (back of the spoon), and if it is negative, the specularity moves against the direction of the view (bowl of the spoon). By construction, we have $\boldsymbol{V} = 2(\boldsymbol{S} \cdot \boldsymbol{N})\boldsymbol{N} - \boldsymbol{S}$, and

$$\begin{aligned} \boldsymbol{V}' &= 2(\boldsymbol{S} \cdot \boldsymbol{N}')\boldsymbol{N}' - \boldsymbol{S} = 2(\boldsymbol{S} \cdot (\boldsymbol{N} + \delta\boldsymbol{N}))(\boldsymbol{N} + \delta\boldsymbol{N}) - \boldsymbol{S} \\ &= \boldsymbol{V} + 2(\boldsymbol{S} \cdot \delta\boldsymbol{N})\boldsymbol{N} + 2(\boldsymbol{S} \cdot \boldsymbol{N})\delta\boldsymbol{N} + 2(\boldsymbol{S} \cdot \delta\boldsymbol{N})\delta\boldsymbol{N}, \end{aligned}$$

where $\delta\boldsymbol{N} \stackrel{\text{def}}{=} \boldsymbol{N}' - \boldsymbol{N} = \delta s \, d\boldsymbol{N}(\boldsymbol{t})$. Because $\boldsymbol{t}$ is tangent to the surface in P, ignoring second-order terms yields

$$\delta a = (\boldsymbol{V} - \boldsymbol{V}') \cdot \boldsymbol{t} = 2(\boldsymbol{S} \cdot \boldsymbol{N})(\delta\boldsymbol{N} \cdot \boldsymbol{t}) = 2(\boldsymbol{S} \cdot \boldsymbol{N})(\delta s)(\mathrm{II}(\boldsymbol{t}, \boldsymbol{t})).$$

Thus, for a concave surface, the specularity always moves against the view, and for a convex surface, it always moves with the view. Things are more complex with hyperbolic surfaces; the specularity may move with the view, against the view, or perpendicular to the view (when $\boldsymbol{t}$ is an asymptotic direction).

[1] Of course, there is a simpler way to distinguish (by sight) the concave bowl of a spoon from its convex back: it is just the side where your reflection is upside down. This property of concave mirrors was demonstrated to one of the authors by one of his friends at a dinner party, causing much consternation since the author in question was at the time bragging about differential geometry, and his friend, who does not pretend to know anything about mathematics, just looked at herself in the spoon.

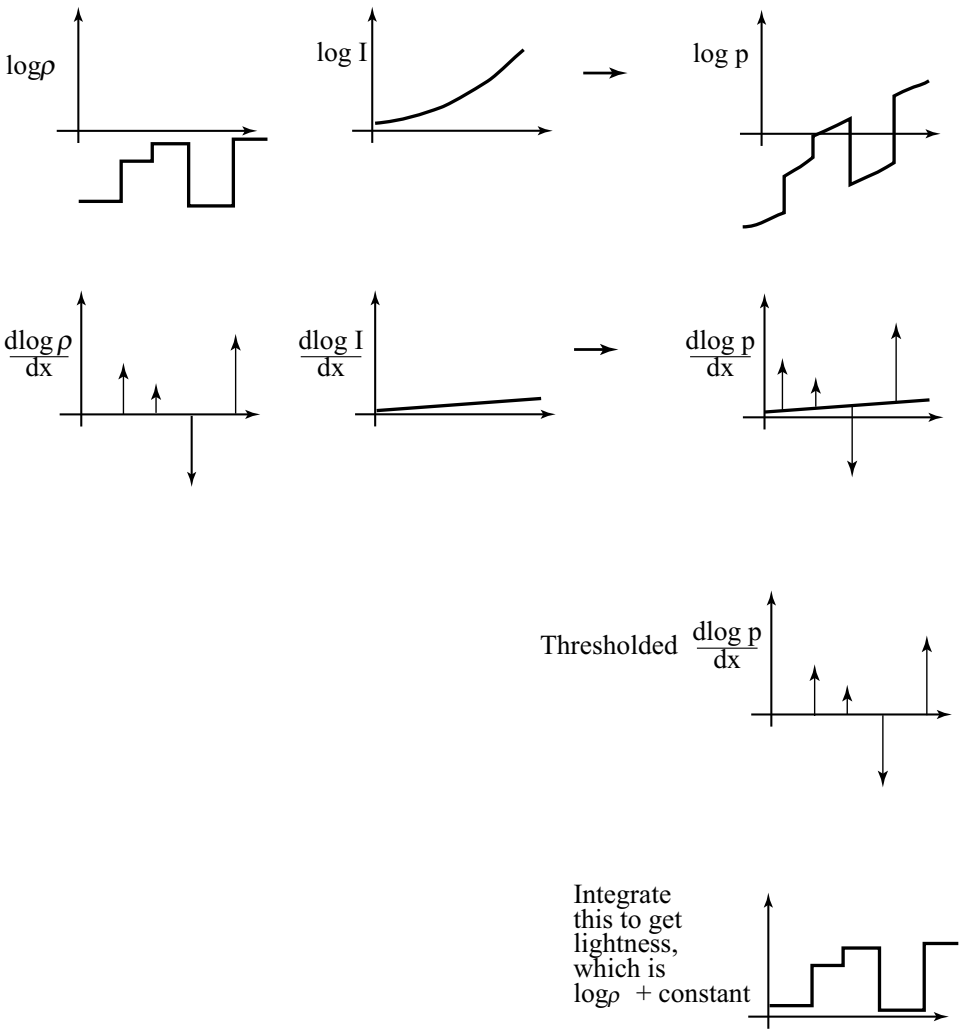

FIGURE 2.8: The lightness algorithm is easiest to illustrate for a 1D image. In the **top row**, the graph on the **left** shows $\log \rho(x)$, that in the **center** $\log I(x)$, and that on the **right** their sum, which is $\log C$. The log of image intensity has large derivatives at changes in surface reflectance and small derivatives when the only change is due to illumination gradients. Lightness is recovered by differentiating the log intensity, thresholding to dispose of small derivatives, and integrating at the cost of a missing constant of integration.

2.2.3 Inferring Lightness and Illumination

If we could estimate the albedo of a surface from an image, then we would know a property of the surface itself, rather than a property of a picture of the surface. Such properties are often called *intrinsic representations*. They are worth estimating, because they do not change when the imaging circumstances change. It might seem

that albedo is difficult to estimate, because there is an ambiguity linking albedo and illumination; for example, a high albedo viewed under middling illumination will give the same brightness as a low albedo viewed under bright light. However, humans can report whether a surface is white, gray, or black (the *lightness* of the surface), despite changes in the intensity of illumination (the *brightness*). This skill is known as *lightness constancy*. There is a lot of evidence that human lightness constancy involves two processes: one process compares the brightness of various image patches and uses this comparison to determine which patches are lighter and which darker; the second establishes some form of absolute standard to which these comparisons can be referred (e.g. Gilchrist *et al.* (1999)).

Current lightness algorithms were developed in the context of simple scenes. In particular, we assume that the scene is flat and frontal; that surfaces are diffuse, or that specularities have been removed; and that the camera responds linearly. In this case, the camera response C at a point $\boldsymbol{x}$ is the product of an illumination term, an albedo term, and a constant that comes from the camera gain:

$$C(\boldsymbol{x}) = k_c I(\boldsymbol{x})\rho(\boldsymbol{x}).$$

If we take logarithms, we get

$$\log C(\boldsymbol{x}) = \log k_c + \log I(\boldsymbol{x}) + \log \rho(\boldsymbol{x}).$$

We now make a second set of assumptions:

- First, we assume that albedoes change only quickly over space. This means that a typical set of albedoes will look like a collage of papers of different grays. This assumption is quite easily justified: There are relatively few continuous changes of albedo in the world (the best example occurs in ripening fruit), and changes of albedo often occur when one object occludes another (so we would expect the change to be fast). This means that spatial derivatives of the term $\log \rho(\boldsymbol{x})$ are either zero (where the albedo is constant) or large (at a change of albedo).

- Second, illumination changes only slowly over space. This assumption is somewhat realistic. For example, the illumination due to a point source will change relatively slowly unless the source is very close, so the sun is a particularly good source for this method, as long as there are no shadows. As another example, illumination inside rooms tends to change very slowly because the white walls of the room act as area sources. This assumption fails dramatically at shadow boundaries, however. We have to see these as a special case and assume that either there are no shadow boundaries or that we know where they are.

We can now build algorithms that use our model. The earliest algorithm is the Retinex algorithm of Land and McCann (1971); this took several forms, most of which have fallen into disuse. The key insight of Retinex is that small gradients are changes in illumination, and large gradients are changes in lightness. We can use this by differentiating the log transform, throwing away small gradients, and integrating the results (Horn 1974); these days, this procedure is widely known as

Retinex. There is a constant of integration missing, so lightness ratios are available, but absolute lightness measurements are not. Figure 2.8 illustrates the process for a one-dimensional example, where differentiation and integration are easy.

This approach can be extended to two dimensions as well. Differentiating and thresholding is easy: at each point, we estimate the magnitude of the gradient; if the magnitude is less than some threshold, we set the gradient vector to zero; otherwise, we leave it alone. The difficulty is in integrating these gradients to get the log albedo map. The thresholded gradients may not be the gradients of an image because the mixed second partials may not be equal (integrability again; compare with Section 2.2.4).

Form the gradient of the log of the image
At each pixel, if the gradient magnitude is below
  a threshold, replace that gradient with zero
Reconstruct the log-albedo by solving the minimization
  problem described in the text
Obtain a constant of integration
Add the constant to the log-albedo, and exponentiate

Algorithm 2.1: Determining the Lightness of Image Patches.

The problem can be rephrased as a minimization problem: choose the log albedo map whose gradient is most like the thresholded gradient. This is a relatively simple problem because computing the gradient of an image is a linear operation. The x-component of the thresholded gradient is scanned into a vector $\boldsymbol{p}$, and the y-component is scanned into a vector $\boldsymbol{q}$. We write the vector representing log-albedo as $\boldsymbol{l}$. Now the process of forming the x derivative is linear, and so there is some matrix $\mathcal{M}_x$, such that $\mathcal{M}_x\boldsymbol{l}$ is the x derivative; for the y derivative, we write the corresponding matrix $\mathcal{M}_y$.

The problem becomes finding the vector $\boldsymbol{l}$ that minimizes

$$| \mathcal{M}_x\boldsymbol{l} - \boldsymbol{p} |^2 + | \mathcal{M}_y\boldsymbol{l} - \boldsymbol{q} |^2 .$$

This is a quadratic minimization problem, and the answer can be found by a linear process. Some special tricks are required because adding a constant vector to $\boldsymbol{l}$ cannot change the derivatives, so the problem does not have a unique solution. We explore the minimization problem in the exercises.

The constant of integration needs to be obtained from some other assumption. There are two obvious possibilities:

- we can assume that the *brightest patch is white*;
- we can assume that the *average lightness is constant*.

We explore the consequences of these models in the exercises.

More sophisticated algorithms are now available, but there were no quantitative studies of performance until recently. Grosse *et al.* built a dataset for

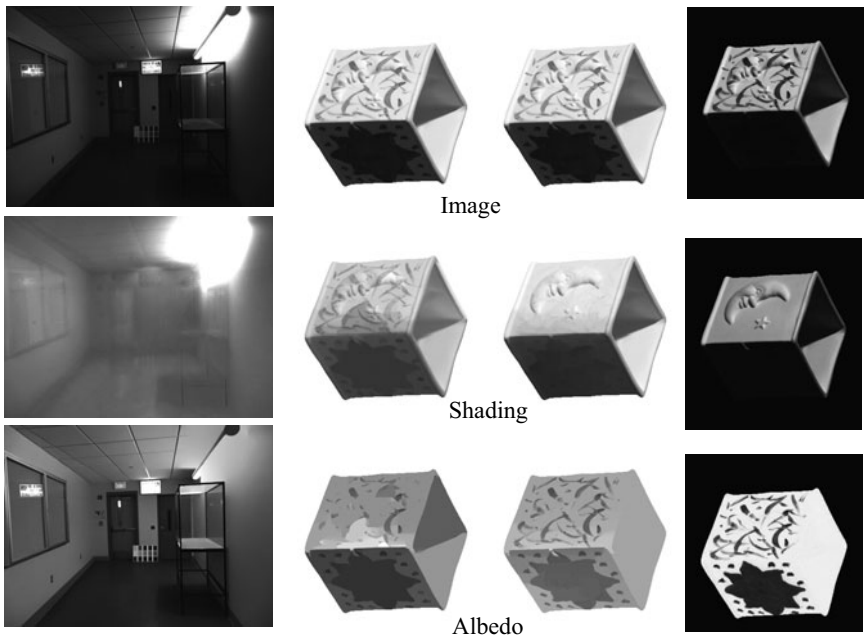

FIGURE 2.9: Retinex remains a strong algorithm for recovering albedo from images. Here we show results from the version of Retinex described in the text applied to an image of a room (**left**) and an image from a collection of test images due to Grosse *et al.* (2009). The **center-left** column shows results from Retinex for this image, and the **center-right** column shows results from a variant of the algorithm that uses color reasoning to improve the classification of edges into albedo versus shading. Finally, the **right** column shows the correct answer, known by clever experimental methods used when taking the pictures. This problem is very hard; you can see that the albedo images still contain some illumination signal. *Part of this figure courtesy Kevin Karsch, U. Illinois. Part of this figure was originally published as Figure 3 of "Ground truth dataset and baseline evaluations for intrinsic image algorithms," by R. Grosse, M. Johnson, E. Adelson, and W. Freeman, Proc. IEEE ICCV 2009, © IEEE, 2009.*

evaluating lightness algorithms, and show that a version of the procedure we describe performs extremely well compared to more sophisticated algorithms (2009). The major difficulty with all these approaches is caused by shadow boundaries, which we discuss in Section 3.5.2.

2.2.4 Photometric Stereo: Shape from Multiple Shaded Images

It is possible to reconstruct a patch of surface from a series of pictures of that surface taken under different illuminants. First, we need a camera model. For simplicity, we choose a camera situated so that the point (x, y, z) in space is imaged to the point (x, y) in the camera (the method we describe works for the other camera models described in Chapter 1).

In this case, to measure the shape of the surface, we need to obtain the depth to the surface. This suggests representing the surface as $(x, y, f(x, y))$—a

representation known as a *Monge patch* after the French military engineer who first used it (Figure 2.10). This representation is attractive because we can determine a unique point on the surface by giving the image coordinates. Notice that to obtain a measurement of a solid object, we would need to reconstruct more than one patch because we need to observe the back of the object.

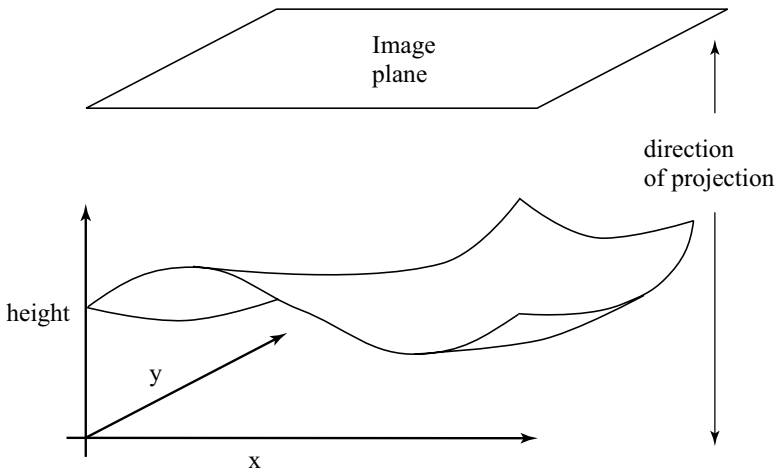

FIGURE 2.10: A Monge patch is a representation of a piece of surface as a height function. For the photometric stereo example, we assume that an orthographic camera—one that maps (x, y, z) in space to (x, y) in the camera—is viewing a Monge patch. This means that the shape of the surface can be represented as a function of position in the image.

Photometric stereo is a method for recovering a representation of the Monge patch from image data. The method involves reasoning about the image intensity values for several different images of a surface in a fixed view illuminated by different sources. This method recovers the height of the surface at points corresponding to each pixel; in computer vision circles, the resulting representation is often known as a *height map*, *depth map*, or *dense depth map*.

Fix the camera and the surface in position, and illuminate the surface using a point source that is far away compared with the size of the surface. We adopt a local shading model and assume that there is no ambient illumination (more about this later) so that the brightness at a point $\boldsymbol{x}$ on the surface is

$$B(\boldsymbol{x}) = \rho(\boldsymbol{x})\boldsymbol{N}(\boldsymbol{x}) \cdot \boldsymbol{S}_1,$$

where $\boldsymbol{N}$ is the unit surface normal and $\boldsymbol{S}_1$ is the source vector. We can write $B(x, y)$ for the radiosity of a point on the surface because there is only one point on the surface corresponding to the point (x, y) in the camera. Now we assume that the response of the camera is linear in the surface radiosity, and so have that the

value of a pixel at (x, y) is

$$\begin{aligned} I(x,y) &= kB(\boldsymbol{x}) \\ &= kB(x,y) \\ &= k\rho(x,y)\boldsymbol{N}(x,y) \cdot \boldsymbol{S}_1 \\ &= \boldsymbol{g}(x,y) \cdot \boldsymbol{V}_1, \end{aligned}$$

where $\boldsymbol{g}(x,y) = \rho(x,y)\boldsymbol{N}(x,y)$ and $\boldsymbol{V}_1 = k\boldsymbol{S}_1$, where k is the constant connecting the camera response to the input radiance.

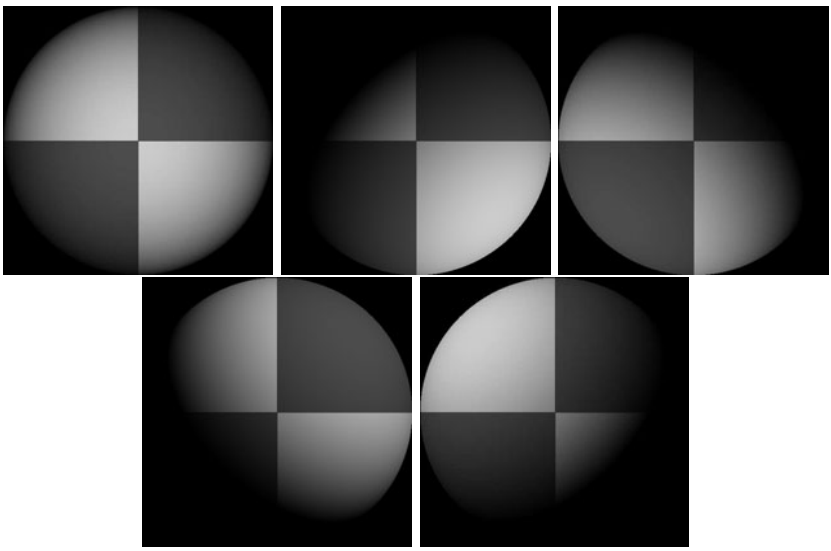

FIGURE 2.11: Five synthetic images of a sphere, all obtained in an orthographic view from the same viewing position. These images are shaded using a local shading model and a distant point source. This is a convex object, so the only view where there is no visible shadow occurs when the source direction is parallel to the viewing direction. The variations in brightness occuring under different sources code the shape of the surface.

In these equations, $\boldsymbol{g}(x,y)$ describes the surface, and $\boldsymbol{V}_1$ is a property of the illumination and of the camera. We have a dot product between a vector field $\boldsymbol{g}(x,y)$ and a vector $\boldsymbol{V}_1$, which could be measured; with enough of these dot products, we could reconstruct $\boldsymbol{g}$ and so the surface.

Now if we have n sources, for each of which $\boldsymbol{V}_i$ is known, we stack each of these $\boldsymbol{V}_i$ into a known matrix $\mathcal{V}$, where

$$\mathcal{V} = \begin{pmatrix} \boldsymbol{V}_1^T \\ \boldsymbol{V}_2^T \\ \dots \\ \boldsymbol{V}_n^T \end{pmatrix}.$$

For each image point, we stack the measurements into a vector

$$\boldsymbol{i}(x,y) = \{I_1(x,y), I_2(x,y), \dots, I_n(x,y)\}^T.$$

Notice that we have one vector per image point; each vector contains all the image brightnesses observed at that point for different sources. Now we have

$$\boldsymbol{i}(x, y) = \mathcal{V}\boldsymbol{g}(x, y),$$

and $\boldsymbol{g}$ is obtained by solving this linear system—or rather, one linear system per point in the image. Typically, $n > 3$, so that a least-squares solution is appropriate. This has the advantage that the residual error in the solution provides a check on our measurements.

Substantial regions of the surface might be in shadow for one or the other light (see Figure 2.11). We assume that all shadowed regions are known, and deal only with points that are not in shadow for any illuminant. More sophisticated strategies can infer shadowing because shadowed points are darker than the local geometry predicts.

We can extract the albedo from a measurement of $\boldsymbol{g}$ because $\boldsymbol{N}$ is the unit normal. This means that $|\boldsymbol{g}(x, y)| = \rho(x, y)$. This provides a check on our measurements as well. Because the albedo is in the range zero to one, any pixels where $|\boldsymbol{g}|$ is greater than one are suspect—either the pixel is not working or $\mathcal{V}$ is incorrect. Figure 2.12 shows albedo recovered using this method for the images shown in Figure 2.11.

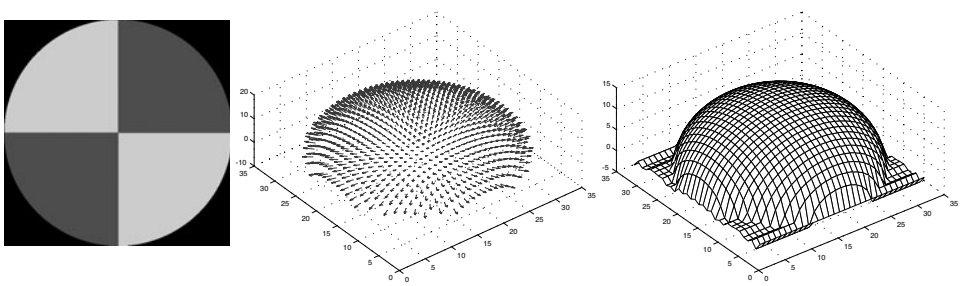

FIGURE 2.12: The image on the **left** shows the magnitude of the vector field $\boldsymbol{g}(x, y)$ recovered from the input data of Figure 2.11 represented as an image—this is the reflectance of the surface. The **center** figure shows the normal field, and the **right** figure shows the height field.

We can extract the surface normal from $\boldsymbol{g}$ because the normal is a unit vector

$$\boldsymbol{N}(x, y) = \frac{\boldsymbol{g}(x, y)}{|\boldsymbol{g}(x, y)|}.$$

Figure 2.12 shows normal values recovered for the images of Figure 2.11.

The surface is $(x, y, f(x, y))$, so the normal as a function of (x, y) is

$$\boldsymbol{N}(x, y) = \frac{1}{\sqrt{1 + \frac{\partial f}{\partial x}^2 + \frac{\partial f}{\partial y}^2}} \left\{ \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, 1 \right\}^T.$$

To recover the depth map, we need to determine $f(x, y)$ from measured values of the unit normal.

Obtain many images in a fixed view under different illuminants
Determine the matrix $\mathcal{V}$ from source and camera information

Inferring albedo and normal:
For each point in the image array that is not shadowed
  Stack image values into a vector $\boldsymbol{i}$
  Solve $\mathcal{V}\boldsymbol{g} = \boldsymbol{i}$ to obtain $\boldsymbol{g}$ for this point
  Albedo at this point is $|\boldsymbol{g}|$
  Normal at this point is $\frac{\boldsymbol{g}}{|\boldsymbol{g}|}$
  p at this point is $\frac{N_1}{N_3}$
  q at this point is $\frac{N_2}{N_3}$
end
Check: is $(\frac{\partial p}{\partial y} - \frac{\partial q}{\partial x})^2$ small everywhere?

Integration:
Top left corner of height map is zero
For each pixel in the left column of height map
  height value = previous height value + corresponding q value
end
For each row
  For each element of the row except for leftmost
    height value = previous height value + corresponding p value
  end
end

Algorithm 2.2: Photometric Stereo.

Assume that the measured value of the unit normal at some point (x, y) is $(a(x, y), b(x, y), c(x, y))$. Then

$$\frac{\partial f}{\partial x} = \frac{a(x,y)}{c(x,y)} \text{ and } \frac{\partial f}{\partial y} = \frac{b(x,y)}{c(x,y)}.$$

We have another check on our data set, because

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x},$$

so we expect that

$$\frac{\partial \left(\frac{a(x,y)}{c(x,y)}\right)}{\partial y} - \frac{\partial \left(\frac{b(x,y)}{c(x,y)}\right)}{\partial x}$$

should be small at each point. In principle it should be zero, but we would have to estimate these partial derivatives numerically and so should be willing to accept small values. This test is known as a test of *integrability*, which in vision applications always boils down to checking that mixed second partials are equal.

FIGURE 2.13: Photometric stereo could become the method of choice to capture complex deformable surfaces. On the **top**, three images of a garment, lit from different directions, which produce the reconstruction shown on the **top right**. A natural way to obtain three different images at the same time is to use a color camera; if one has a red light, a green light, and a blue light, then a single color image frame can be treated as three images under three separate lights. On the **bottom**, an image of the garment captured in this way, which results in the photometric stereo reconstruction on the **bottom right**. *This figure was originally published as Figure 6 of "Video Normals from Colored Lights," G. J. Brostow, C. Hernández, G. Vogiatzis, B. Stenger, and R. Cipolla, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011* © *IEEE, 2011.*

Assuming that the partial derivatives pass this sanity test, we can reconstruct the surface up to some constant depth error. The partial derivative gives the change in surface height with a small step in either the x or the y direction. This means we can get the surface by summing these changes in height along some path. In particular, we have

$$f(x,y) = \oint_C \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) \cdot d\boldsymbol{l} + c,$$

where C is a curve starting at some fixed point and ending at (x, y), and c is a constant of integration, which represents the (unknown) height of the surface at the start point. The recovered surface does not depend on the choice of curve (exercises). Another approach to recovering shape is to choose the function $f(x, y)$ whose partial derivatives most look like the measured partial derivatives. Figure 2.12 shows the reconstruction obtained for the data shown in Figure 2.11.

Current reconstruction work tends to emphasize geometric methods that reconstruct from multiple views. These methods are very important, but often require feature matching, as we shall see in Chapters 7 and 8. This tends to mean that it is hard to get very high spatial resolution, because some pixels are consumed in resolving features. Recall that resolution (which corresponds roughly to the spatial frequencies that can be reconstructed accurately) is not the same as ac-

curacy (which involves a method providing the right answers for the properties it estimates). Feature-based methods are capable of spectacularly accurate reconstructions. Because photometric cues have such spatial high resolution, they are a topic of considerable current interest. One way to use photometric cues is to try and match pixels with the same brightness across different cameras; this is difficult, but produces impressive reconstructions. Another is to use photometric stereo ideas. For some applications, photometric stereo is particularly atractive because one can get reconstructions from a single view direction—this is important, because we cannot always set up multiple cameras. In fact, with a trick, it is possible to get reconstructions from a single frame. A natural way to obtain three different images at the same time is to use a color camera; if one has a red light, a green light and a blue light, then a single color image frame can be treated as three images under three separate lights, and photometric stereo methods apply. In turn, this means that photometric stereo methods could be used to recover high-resolution reconstructions of deforming surfaces in a relatively straightforward way. This is particularly useful when it is difficult to get many cameras to view the object. Figure 2.13 shows one application to reconstructing cloth in video (from Brostow *et al.* (2011)), where multiple view reconstruction is complicated by the need to synchronize frames (alternatives are explored in, for example, White *et al.* (2007) or Bradley *et al.* (2008*b*)).

2.3 MODELLING INTERREFLECTION

The difficulty with a local shading model is that it doesn't account for all light. The alternative is a *global shading model*, where we account for light arriving from other surfaces as well as from the luminaire. As we shall see, such models are tricky to work with. In such models, each surface patch receives power from all the radiating surfaces it can see. These surfaces might radiate power that they generate internally because they are luminaires, or they might simply reflect power. The general form of the model will be:

$$\begin{pmatrix} \text{Power leaving} \\ \text{a patch} \end{pmatrix} = \begin{pmatrix} \text{Power generated} \\ \text{by that patch} \end{pmatrix} + \begin{pmatrix} \text{Power received from} \\ \text{other patches and reflected} \end{pmatrix}$$

This means we need to be able to model the power received from other patches and reflected. We will develop a model assuming that all surfaces are diffuse. This leads to a somewhat simpler model, and describes all effects that are currently of interest to vision (it is complicated, but not difficult, to build more elaborate models). We will also need some radiometric terminology.

2.3.1 The Illumination at a Patch Due to an Area Source

The appropriate unit for illumination is **radiance**, defined as

> the power (amount of energy per unit time) traveling at some point in a specified direction, per unit area *perpendicular to the direction of travel*, per unit solid angle.

The units of radiance are watts per square meter per steradian ($Wm^{-2}sr^{-1}$). The definition of radiance might look strange, but it is consistent with the most basic

phenomenon in radiometry: the amount of energy a patch collects from a source depends both on how large the source looks from the patch *and* on how large the patch looks from the source.

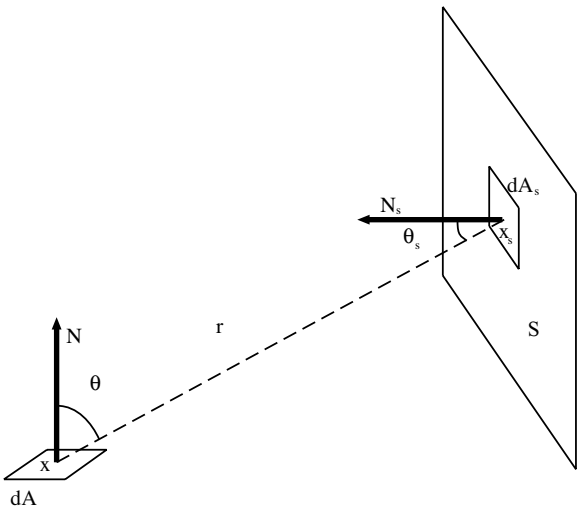

FIGURE 2.14: A patch with area dA views an area source S. We compute the power received by the patch by summing the contributions of each element on S, using the notation indicated in this figure.

It is important to remember that the square meters in the units for radiance are *foreshortened* (i.e., perpendicular to the direction of travel) to account for this phenomenon. Assume we have two elements, one at $\boldsymbol{x}$ with area dA and the other at $\boldsymbol{x}_s$ with area dA_s. Write the angular direction from $\boldsymbol{x}$ to $\boldsymbol{x}_s$ as $\boldsymbol{x} \to \boldsymbol{x}_s$, and define the angles θ and θ_s as in Figure 2.14. Then the solid angle subtended by element 2 at element 1 is

$$d\omega_{2(1)} = \frac{\cos\theta_s dA_s}{r^2},$$

so the power leaving $\boldsymbol{x}$ toward $\boldsymbol{x}_s$ is

$$\begin{aligned} d^2P_{1\to 2} &= (\text{radiance})(\text{foreshortened area})(\text{solid angle}) \\ &= L(\boldsymbol{x}, \boldsymbol{x} \to \boldsymbol{x}_s)(\cos\theta dA)(d\omega_{2(1)}) \\ &= L(\boldsymbol{x}, \boldsymbol{x} \to \boldsymbol{x}_s)\left(\frac{\cos\theta\cos\theta_s}{r^2}\right) dA_s dA. \end{aligned}$$

By a similar argument, the same expression yields the power arriving at $\boldsymbol{x}$ from $\boldsymbol{x}_2$; this means that, in a vacuum, *radiance is constant along (unoccluded) straight lines.*

We can now compute the power that an element dA collects from an area source, by summing the contributions of elements over that source. Using the notation of Figure 2.14, we get

$$dP_{S\to dA} = \left(\int_S L(\boldsymbol{x}_s, \boldsymbol{x}_s \to \boldsymbol{x})\left(\frac{\cos\theta_s\cos\theta}{r^2}\right) dA_s\right) dA.$$

To get a more useful area source model, we need further units.

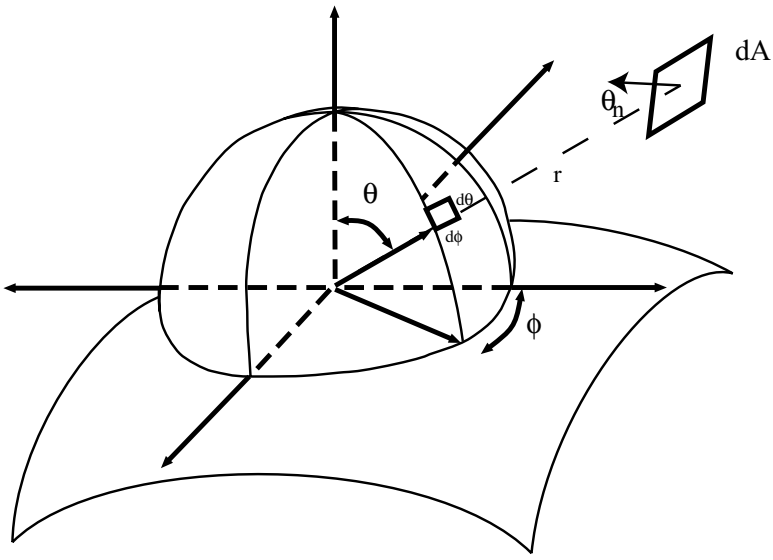

FIGURE 2.15: A hemisphere on a patch of surface, to show our angular coordinates for computing radiometric quantities. The coordinate axes are there to help you see the drawing as a 3D surface. An infinitesimal patch of surface with area dA which is distance r away is projected onto the unit hemisphere centered at the relevant point; the resulting area is the solid angle of the patch, marked as $d\theta d\phi$. In this case, the patch is small so that the area and hence the solid angle is $(1/r^2)dA\cos\theta_n$, where θ_n is the angle of inclination of the patch.

2.3.2 Radiosity and Exitance

We are dealing with diffuse surfaces, and our definition of a diffuse surface is that the intensity (formally, the radiance) leaving the surface is independent of the direction in which it leaves. There is no point in describing the intensity of such a surface with radiance (which explicitly depends on direction). The appropriate unit is *radiosity*, defined as

> the total power leaving a point on a surface per unit area on the surface.

Radiosity, which is usually written as $B(\boldsymbol{x})$, has units watts per square meter (Wm^{-2}). To obtain the radiosity of a surface at a point, we can sum the radiance leaving the surface at that point over the whole exit hemisphere. Thus, if $\boldsymbol{x}$ is a point on a surface emitting radiance $L(\boldsymbol{x}, \theta, \phi)$, the radiosity at that point is

$$B(\boldsymbol{x}) = \int_{\Omega} L(\boldsymbol{x}, \theta, \phi) \cos\theta d\omega,$$

where Ω is the exit hemisphere, $d\omega$ is unit solid angle, and the term $\cos\theta$ turns foreshortened area into area (look at the definitions of radiance and of radiosity again). We could substitute $d\omega = \sin\theta d\theta d\phi$, using the units of Figure 2.15.

Consider a surface element as in Figure 2.14. We have computed how much power it receives from the source as a function of the source's radiance. The surface element is diffuse, and its albedo is $\rho(\boldsymbol{x})$. The albedo is the fraction of incoming power that the surface radiates, so the radiosity due to power received from the area source is

$$B(\boldsymbol{x}) = \frac{dP_{S \to dA}}{dA} = \rho(\boldsymbol{x}) \left(\int_S L(\boldsymbol{x}_s, \boldsymbol{x}_s \to \boldsymbol{x}) \left(\frac{\cos\theta_s \cos\theta}{r^2} \right) dA_s \right).$$

Now if a point $\boldsymbol{u}$ on a surface has radiosity $B(\boldsymbol{u})$, what is the radiance leaving the surface in some direction? We write L for the radiance, which is independent of angle, and we must have

$$B(\boldsymbol{u}) = \int_\Omega L(\boldsymbol{x}, \theta, \phi) \cos\theta d\omega = L(\boldsymbol{u}) \int_\Omega \cos\theta d\omega = L(\boldsymbol{u})\pi.$$

This means that if the area source has radiosity $B(\boldsymbol{x}_s)$, then the radiosity at the element *due to the power received from the area source* is

$$\begin{aligned} B(\boldsymbol{x}) &= \rho \left(\int_S L(\boldsymbol{x}_s, \boldsymbol{x}_s \to \boldsymbol{x}) \left(\frac{\cos\theta_s \cos\theta}{r^2} \right) dA_s \right) \\ &= \rho \left(\int_S \frac{B(\boldsymbol{x})}{\pi} \left(\frac{\cos\theta_s \cos\theta}{r^2} \right) dA_s \right) \\ &= \frac{\rho}{\pi} \left(\int_S B(\boldsymbol{x}) \left(\frac{\cos\theta_s \cos\theta}{r^2} \right) dA_s \right). \end{aligned}$$

Our final step is to model illumination generated internally in a surface—light generated by a luminaire, rather than reflected from a surface. We assume there are no directional effects in the luminaire and that power is uniformly distributed across outgoing directions (this is the least plausible component of the model, but is usually tolerable). We use the unit *exitance*, which is defined as

> the total power internally generated power leaving a point on a surface per unit area on the surface.

2.3.3 An Interreflection Model

We can now write a formal model of interreflections for diffuse surfaces by substituting terms into the original expression. Recall that radiosity is power per unit area, write $E(\boldsymbol{x})$ for exitance at the point $\boldsymbol{x}$, write $\boldsymbol{x}_s$ for a coordinate that runs over all surface patches, $\mathcal{S}$ for the set of all surfaces, dA for the element of area at $\boldsymbol{x}$, $\mathrm{V}(\boldsymbol{x}, \boldsymbol{x}_s)$ for a function that is one if the two points can see each other and zero otherwise, and $\cos\theta$, $\cos\theta_s$, r, as in Figure 2.14. We obtain

$$\begin{array}{ccccc} \text{Power leaving a patch} & = & \text{Power generated by that patch} & + & \text{Power received from other patches and reflected} \\ B(\boldsymbol{x})dA & = & E(\boldsymbol{x})dA & + & \frac{\rho(\boldsymbol{x})}{\pi} \int_{\mathcal{S}} \left[\begin{array}{c} \frac{\cos\theta \cos\theta_s}{r^2} \\ \times \\ \mathrm{V}(\boldsymbol{x}, \boldsymbol{x}_s) \end{array} \right] B(\boldsymbol{x_s}) dA_s dA \end{array}$$

and so, dividing by area, we have

$$B(\boldsymbol{x}) \quad = \quad E(\boldsymbol{x}) \quad + \quad \frac{\rho(\boldsymbol{x})}{\pi} \int_{\mathcal{S}} \left[\frac{\cos\theta \cos\theta_s}{r^2} \text{Vis}(\boldsymbol{x}, \boldsymbol{x}_s) \right] B(\boldsymbol{x}_s) dA_s.$$

It is usual to write

$$K(\boldsymbol{x}, \boldsymbol{x}_s) = \frac{\cos\theta \cos\theta_s}{\pi r^2}$$

and refer to K as the interreflection kernel. Substituting gives

$$B(\boldsymbol{x}) = E(\boldsymbol{x}) + \rho(\boldsymbol{x}) \int_{\mathcal{S}} K(\boldsymbol{x}, \boldsymbol{x}_s) \text{Vis}(\boldsymbol{x}, \boldsymbol{x}_s) B(\boldsymbol{x}_s) dA_{\boldsymbol{x}_s}$$

an equation where the solution appears inside the integral. Equations of this form are known as Fredholm integral equations of the second kind. This particular equation is a fairly nasty sample of the type because the interreflection kernel generally is not continuous and may have singularities. Solutions of this equation can yield quite good models of the appearance of diffuse surfaces, and the topic supports a substantial industry in the computer graphics community (good places to start for this topic are Cohen and Wallace (1993) or Sillion (1994)). The model produces good predictions of observed effects (Figure 2.16).

2.3.4 Qualitative Properties of Interreflections

Interreflections are a problem, because they are difficult to account for in our illumination model. For example, photometric stereo as we described it uses the model that light at a surface patch comes only from a distant light source. One could refine the method to take into account nearby light sources, but it is much more difficult to deal with interreflections. Once one accounts for interreflections, the brightness of each surface patch could be affected by the configuration of every other surface patch, making a very nasty global inference problem. While there have been attempts to build methods that can infer shape in the presence of interreflections (Nayar *et al.* 1991*a*), the problem is extremely difficult. One source of difficulties is that one may need to account for every radiating surface in the solution, even distant surfaces one cannot see.

An alternative strategy to straightforward physical inference is to understand the qualitative properties of interreflected shading. By doing so, we may be able to identify cases that are easy to handle, the main types of effect, and so on. The effects can be quite large. For example, Figure 2.17 shows views of the interior of two rooms. One room has black walls and contains black objects. The other has white walls and contains white objects. Each is illuminated (approximately!) by a distant point source. Given that the intensity of the source is adjusted appropriately, the local shading model predicts that these pictures would be indistinguishable. In fact, the black room has much darker shadows and crisper boundaries at the creases of the polyhedra than the white room. This is because surfaces in the black room reflect less light onto other surfaces (they are darker), whereas in the white room other surfaces are significant sources of radiation. The sections of the camera response to the radiosity (these are proportional to radiosity for diffuse surfaces) shown in the figure are hugely different qualitatively. In the black room, the radiosity is constant in patches, as a local shading model would predict, whereas in the white room slow

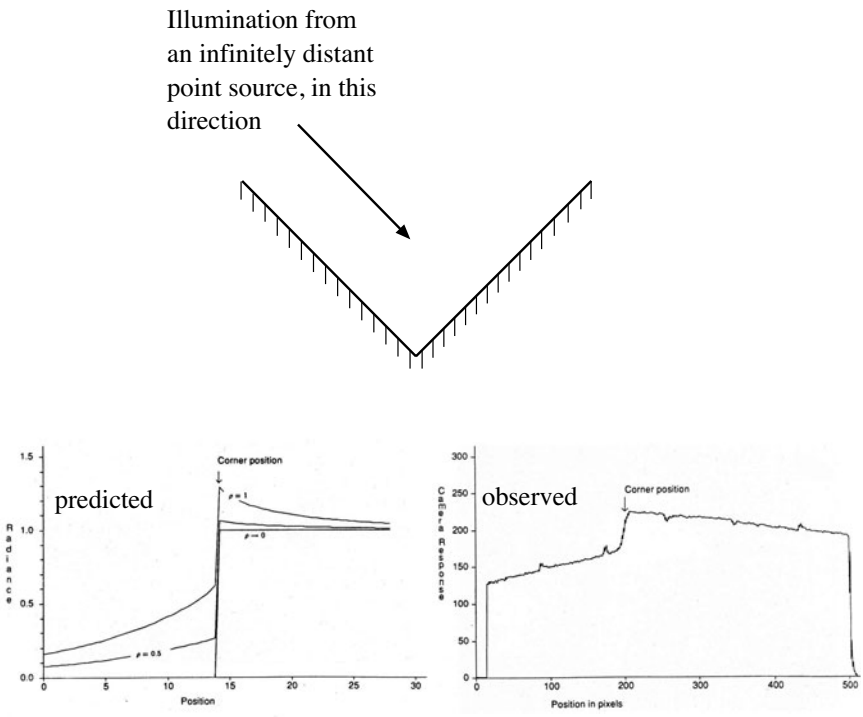

FIGURE 2.16: The model described in the text produces quite accurate qualitative predictions for interreflections. The **top** figure shows a concave right-angled groove illuminated by a point source at infinity where the source direction is parallel to the one face. On the **left** of the bottom row is a series of predictions of the radiosity for this configuration. These predictions have been scaled to lie on top of one another; the case $\rho \to 0$ corresponds to the local shading model. On the **right**, an observed image intensity for an image of this form for a corner made of white paper, showing the roof-like gradient in radiosity associated with the edge. A local shading model predicts a step. *This figure was originally published as Figures 5 and 7 of "Mutual Illumination," by D.A. Forsyth and A.P. Zisserman, Proc. IEEE CVPR, 1989,* © *IEEE, 1989.*

image gradients are quite common; these occur in concave corners, where object faces reflect light onto one another.

First, interreflections have a characteristic smoothing effect. This is most obviously seen when one tries to interpret a stained glass window by looking at the pattern it casts on the floor; this pattern is almost always a set of indistinct colored blobs. The effect is seen most easily with the crude model illustrated in Figure 2.18. The geometry consists of a patch with a frontal view of an infinite plane, which is a unit distance away and carries a radiosity $\sin \omega x$. There is no reason to vary the distance of the patch from the plane, because interreflection problems have scale invariant solutions, which means that the solution for a patch two units away can be obtained by reading our graph at 2ω. The patch is small enough that its contribution to the plane's radiosity can be ignored. If the patch is

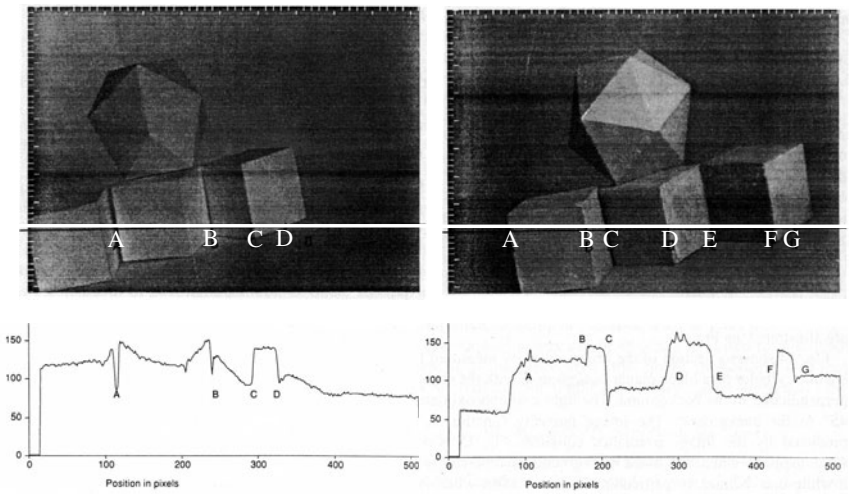

FIGURE 2.17: The column on the **left** shows data from a room with matte black walls and containing a collection of matte black polyhedral objects; that on the **right** shows data from a white room containing white objects. The images are qualitatively different, with darker shadows and crisper boundaries in the black room and bright reflexes in the concave corners in the white room. The graphs show sections of the image intensity along the corresponding lines in the images. *This figure was originally published as Figures 17, 18, 19, and 20 of "Mutual Illumination," by D.A. Forsyth and A.P. Zisserman, Proc. IEEE CVPR, 1989,* © *IEEE, 1989.*

slanted by σ with respect to the plane, it carries radiosity that is nearly periodic, with spatial frequency $\omega \cos \sigma$. We refer to the amplitude of the component at this frequency as the gain of the patch and plot the gain in Figure 2.18. The important property of this graph is that high spatial frequencies have a difficult time jumping the gap from the plane to the patch. This means that shading effects with high spatial frequency and high amplitude generally cannot come from distant surfaces (unless they are abnormally bright).

The extremely fast fall-off in amplitude with spatial frequency of terms due to distant surfaces means that, if one observes a high-amplitude term at a high spatial frequency, *it is very unlikely to have resulted from the effects of distant, passive radiators* (because these effects die away quickly). There is a convention, which we see in Section 2.2.3, that classifies effects in shading as due to reflectance if they are fast ("edges") and the dynamic range is relatively low and due to illumination otherwise. We can expand this convention. There is a mid range of spatial frequencies that are largely unaffected by mutual illumination from distant surfaces because the gain is small. Spatial frequencies in this range cannot be transmitted by distant passive radiators unless these radiators have improbably high radiosity. As a result, spatial frequencies in this range can be thought of as *regional properties*, which can result only from interreflection effects within a region.

The most notable regional properties are probably *reflexes*— small bright patches that appear mainly in concave regions (illustrated in Figure 2.19). A second

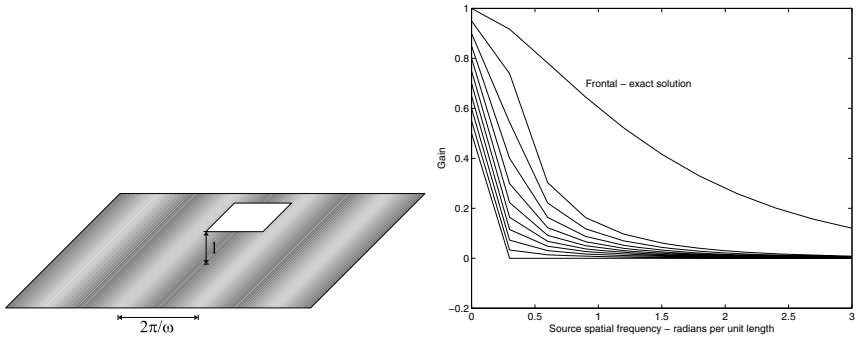

FIGURE 2.18: A small patch views a plane with sinusoidal radiosity of unit amplitude. This patch has a (roughly) sinusoidal radiosity due to the effects of the plane. We refer to the amplitude of this component as the *gain of the patch*. The graph shows numerical estimates of the gain for patches at 10 equal steps in slant angle, from 0 to $\pi/2$, as a function of spatial frequency *on the plane*. The gain falls extremely fast, meaning that large terms at high spatial frequencies must be regional effects, rather than the result of distant radiators. This is why it is hard to determine the pattern in a stained glass window by looking at the floor at the foot of the window. *This figure was originally published as Figures 1 and 2 from "Shading Primitives: Finding Folds and Shallow Grooves," J. Haddon and D.A. Forsyth, Proc. IEEE ICCV, 1998* © *IEEE, 1998.*

important effect is *color bleeding*, where a colored surface reflects light onto another colored surface. This is a common effect that people tend not to notice unless they are consciously looking for it. It is quite often reproduced by painters.

2.4 SHAPE FROM ONE SHADED IMAGE

There is good evidence that people get some perception of shape from the shading pattern in a single image, though the details are uncertain and quite complicated (see the notes for a brief summary). You can see this evidence in practice: whenever you display a reconstruction of a surface obtained from images, it is a good idea to shade that reconstruction using image pixels, because it always looks more accurate. In fact, quite bad reconstructions can be made to look good with this method. White and Forsyth (2006) use this trick to replace surface albedos in movies; for example, they can change the pattern on a plastic bag in a movie. Their method builds and tracks very coarse geometric reconstructions, uses a form of regression to recover the original shading pattern of the object, and then shades the coarse geometric reconstruction using the original shading pattern (Figure 2.20). In this figure, the pictures look plausible, not because the reconstruction is good (it isn't), but because the shading pattern masks the errors in geometric reconstruction.

The cue to shape must come from the fact that a surface patch that faces the light source is brighter than one that faces away from the source. But going from this observation to a working algorithm remains an open question. The key seems to be an appropriate use of the *image irradiance equation*. Assume we have a surface in the form $(x, y, f(x, y))$ viewed orthographically along the z-axis. Assume that the surface is diffuse, and its albedo is uniform and known. Assume also that

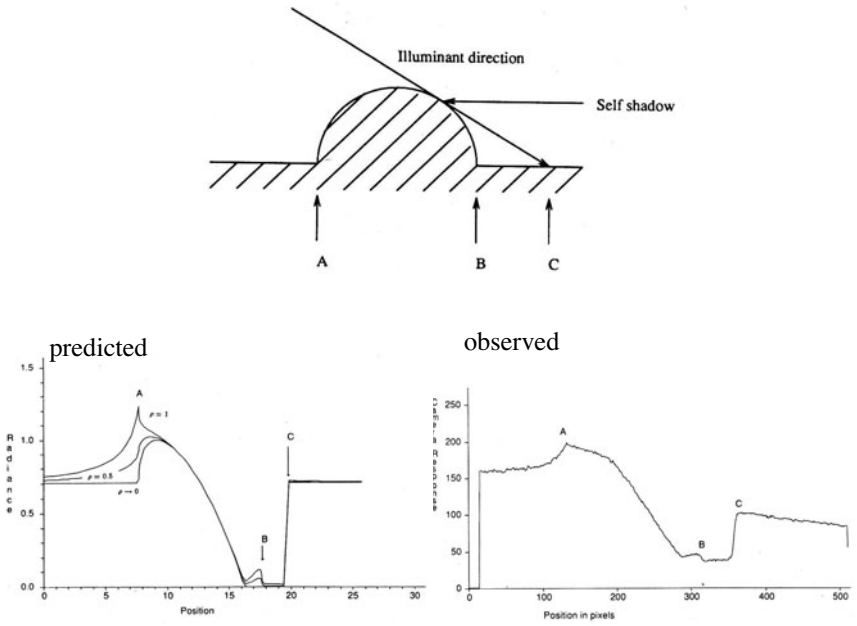

FIGURE 2.19: Reflexes occur quite widely; they are usually caused by a favorable view of a large reflecting surface. In the geometry shown on the **top**, the shadowed region of the cylindrical bump sees the plane background at a fairly favorable angle: if the background is large enough, nearly half the hemisphere of the patch at the base of the bump is a view of the plane. This means there will be a reflex with a large value attached to the edge of the bump and inside the cast shadow region (which a local model predicts as black). There is another reflex on the other side, too, as the series of solutions (again normalized for easy comparison) on the **left** show. On the **right**, an observation of this effect in a real scene. *This figure was originally published as Figures 24 and 26 of "Mutual Illumination," by D.A. Forsyth and A.P. Zisserman, Proc. IEEE CVPR, 1989,* © *IEEE, 1989.*

the model of Section 2.1.3 applies, so that the shading at a point with normal $\boldsymbol{N}$ is given by some function $R(\boldsymbol{N})$ (the function of our model is $R(\boldsymbol{N}) = \boldsymbol{N} \cdot \boldsymbol{S}$, but others could be used). Now the normal of our surface is a function of the two first partial derivatives

$$p = \frac{\partial f}{\partial x}, \; q = \frac{\partial f}{\partial y}$$

so we can write $R(p, q)$. Assume that the camera is radiometrically calibrated, so we can proceed from image values to intensity values. Write the intensity at x, y as $I(x, y)$. Then we have

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

This is a first order partial differential equation, because p and q are partial derivatives of f. In principle, we could set up some boundary conditions and solve this equation. Doing so reliably and accurately for general images remains outside our competence, 40 years after the problem was originally posed by Horn (1970*a*).

FIGURE 2.20: On the **left**, an original frame from a movie sequence of a deforming plastic bag. On the **right**, two frames where the original texture has been replaced by another. The method used is a form of regression; its crucial property is that it has a very weak geometric model, but is capable of preserving the original shading field of the image. If you look closely at the *albedo* (i.e., the black pattern) of the bag, you may notice that it is inconsistent with the wrinkles on the bag, but because the shading has been preserved, the figures look quite good. This is indirect evidence that shading is a valuable cue to humans. Little is known about how this cue is to be exploited, however. *This figure was originally published as Figure 10 of "Retexturing single views using texture and shading," by R. White and D.A. Forsyth, Proc. European Conference on Computer Vision. Springer Lecture Notes in Computer Science, Volume 3954, 2006* © *Springer 2006.*

There are a variety of difficulties here. The physical model is a poor model of what actually happens at surfaces because any particular patch is illuminated by other surface patches, as well as by the source. We expect to see a rich variety of geometric constraints on the surface we reconstruct, and it is quite difficult to formulate shape from shading in a way that accomodates these constraints and still has a solution. Shading is a worthwhile cue to exploit, because we can observe shading at extremely high spatial resolutions, but this means we must work with very high dimensional models to reconstruct. Some schemes for shading reconstruction can be unstable, but there appears to be no theory to guide us to stable schemes. We very seldom actually see isolated surfaces of known albedo, and there are no methods that are competent to infer both shading and albedo, though there is some reason to hope that such methods can be built. We have no theory that is capable of predicting the errors in shading-based reconstructions from first principles. All this makes shape inference from shading in a single image one of the most frustrating open questions in computer vision.

2.5 NOTES

Horn started the systematic study of shading in computer vision, with important papers on recovering shape from a local shading model using a point source (in (Horn 1970*b*), (Horn 1975)), with a more recent account in Horn (1990).

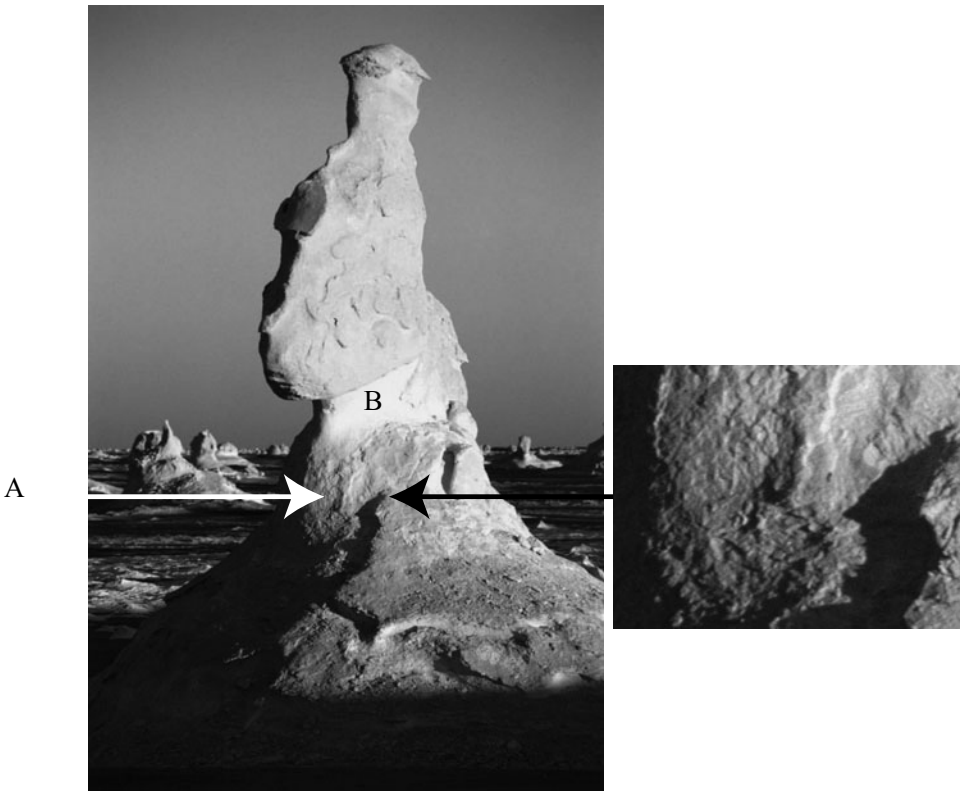

FIGURE 2.21: This picture shows two important mechanisms by which it might be possible to infer surface shape from single images. First, patches that face away from the light (like A, on the **left**) are darker than those that face the light (B). Second, shadows pick out relief—for example, small dents in a surface (more easily seen in the detail patch on the **right**), have a bright face facing the light and a dark face which is in shadow. *Peter Wilson © Dorling Kindersley, used with permission.*

Models of Shading

The first edition of this book contained more formal radiometry, which was widely disliked (and for good reason; making the subject exciting is beyond our skills). We've cut this down, and tried to avoid using the ideas, but point those who really want to know more toward that earlier edition. We strongly recommend François Sillion's excellent book (Sillion 1994) for its clear account of radiometric calculations. There are a variety of more detailed publications for reference (Nayar *et al.* 1991*c*). Our discussion of reflection is thoroughly superficial. The specular plus diffuse model appears to be originally due to Cook, Torrance, and Sparrow (Torrance and Sparrow 1967, Cook and Torrance 1987). A variety of modifications of this model appear in computer vision and computer graphics. Reflection models can be derived by combining a statistical description of surface roughness with electromagnetic considerations (e.g., Beckmann and Spizzichino (1987)) or by

adopting scattering models (as in the work of Torrance and Sparrow (1967) and of Cook and Torrance (1987)).

It is commonly believed that rough surfaces are Lambertian. This belief has a substantial component of wishful thinking because rough surfaces often have local shadowing effects that make the radiance reflected quite strongly dependent on the illumination angle. For example, a stucco wall illuminated at a near grazing angle shows a clear pattern of light and dark regions where facets of the surface face toward the light or are shadowed. If the same wall is illuminated along the normal, this pattern largely disappears. Similar effects at a finer scale are averaged to endow rough surfaces with measurable departures from a Lambertian model (for details, see Koenderink *et al.* (1999), Nayar and Oren (1993), (1995), Oren and Nayar (1995), and Wolff *et al.* (1998)).

Inference from Shading

Registered images are not essential for radiometric calibration. For example, it is sufficient to have two images where we believe the histogram of E_{ij} values is the same (Grossberg and Nayar 2002). This occurs, for example, when the images are of the same scene, but are not precisely registered. Patterns of intensity around edges also can reveal calibration (Lin *et al.* 2004).

There has not been much recent study of lightness constancy algorithms. The basic idea is due to Land and McCann (1971). Their work was formalized for the computer vision community by Horn (1974). A variation on Horn's algorithm was constructed by Blake (1985). This is the lightness algorithm we describe. It appeared originally in a slightly different form, where it was called the *Retinex* algorithm (Land and McCann 1971). Retinex was originally intended as a color constancy algorithm. It is surprisingly difficult to analyze (Brainard and Wandell 1986).

Retinex estimates the log-illumination term by subtracting the log-albedo from the log-intensity. This has the disadvantage that we do not impose any structural constraints on illumination. This point has largely been ignored, because the main focus has been on albedo estimates. However, albedo estimates are likely to be improved by balancing violations of albedo constraints with those of illumination constraints.

Lightness techniques are not as widely used as they should be, particularly given that there is some evidence that they produce useful information on real images (Brelstaff and Blake 1987). Classifying illumination versus albedo simply by looking at the magnitude of the gradient is crude, and ignores important cues. Sharp shading changes occur at shadow boundaries or normal discontinuities, but using chromaticity (Funt *et al.* 1992) or multiple images under different lighting conditions (Weiss 2001) yields improved estimates. One can learn to distinguish illumination from albedo (Freeman *et al.* 2000). Discriminative methods to classify edges into albedo or shading help (Tappen *et al.* 2006*b*) and chromaticity cues can contribute (Farenzena and Fusiello 2007). Shading and albedo are sometimes known as *intrinsic images*. Tappen *et al.* (2006*a*) regress local intrinsic image patches against the image, exploiting the constraint that patches join up. When more than one image is available, recent methods can recover quite complex surface

properties (Romeiro *et al.* 2008). When geometry is available, Yu *et al.* (1999) showed significant improvements in lightness recovery are possible.

In its original form, photometric stereo is due to Woodham. There are a number of variants of this useful idea (Horn *et al.* (1978), Woodham (1979), (1980), (1989), (1994), Woodham *et al.* (1991)). Current methods for photometric stereo require at least two unshadowed views; see Hernandez *et al.* (2008) which describes methods to cope in this case. There are a variety of variations on photometric stereo. Color photometric stereo seems to date to Petrov (1987), with a variant in Petrov (1991).

Photometric stereo depends only on adopting a local shading model. This model need not be a Lambertian surface illuminated by a distant point source. If the brightness of the surface is a known function of the surface normal satisfying a small number of constraints, photometric stereo is still possible. This is because the intensity of a pixel in a single view determines the normal up to a one-parameter family. This means that two views determine the normal. The simplest example of this case occurs for a surface of known albedo illuminated by a distant point source.

In fact, if the radiosity of the surface is a k-parameter function of the surface normal, photometric stereo is still possible. The intensity of the pixel in a single view determines the normal up to a $k+1$ parameter family, and $k+1$ views give the normal. For this approach to work, the brightness needs to be given by a function for which our arithmetic works (e.g., if the brightness of the surface is a constant function of the surface normal, it isn't possible to infer any constraint on the normal from the brightness). One can then recover shape and reflectance maps simultaneously (Garcia-Bermejo *et al.* (1996); Mukawa (1990); Nayar *et al.* (1990); Tagare and de Figueiredo (1992); (1993)).

A converse to photometric stereo might be as follows: Assume we have a diffuse sphere, immersed in an environment where illumination depends only on direction. What can we determine about the illumination field from the surface brightness? The answer is very little, because diffuse surfaces engage in a form of averaging that heavily smoothes the illumination field (Ramamoorthi and Hanrahan 2001). This is valuable because it suggests that complex representations of the directional properties illumination aren't required in a diffuse world. For example, this result allowed Jacobs (1981) to produce a form of photometric stereo that requires no illuminant information, using sufficient images.

Interreflections

The effects of global shading are often ignored in the shading literature, which causes a reflex response of hostility in one of the authors. The reason to ignore interreflections is that they are extremely hard to analyze, particularly from the perspective of inferring object properties given the output of a global shading model. If interreflection effects do not change the output of a method much, then it is probably all right to ignore them. Unfortunately, this line of reasoning is seldom pursued because it is quite difficult to show that a method is stable under interreflections. The discussion of spatial frequency issues follows Haddon and Forsyth (1998*a*), after an idea of Koenderink and van Doorn (1983). Apart from this, there is not much knowledge about the overall properties of interreflected shading, which is an

important gap in our knowledge. An alternative strategy is to iteratively reestimate shape using a rendering model (Nayar *et al.* 1991*b*).

Horn is also the first author to indicate the significance of global shading effects (Horn 1977). Koenderink and van Doorn (1983) noted that the radiosity under a global model is obtained by taking the radiosity under a local model, and applying a linear operator. One then studies that operator; in some cases, its eigenfunctions (often called *geometrical modes*) are informative. Forsyth and Zisserman (1989, 1990, 1991) then demonstrated a variety of the qualitative effects due to interreflections.

Shape from One Shaded Image

Shape from shading is an important puzzle. Comprehensive surveys include (Horn and Brooks (1989); Zhang *et al.* (1999); Durou *et al.* (2008*b*)). In practice, despite the ongoing demand for high-resolution shape reconstructions, shape-from-shading has been a disappointment. This may be because, as currently formulated, it solves a problem that doesn't commonly occur. Image irradiance equation methods are formulated to produce reconstructions when there is very little geometric data, but it is much more common to want to improve the resolution of a method that already produces quite rich geometric data.

Methods are either too fragile or the reconstructions too poor for the method to be useful. Some of this may be due to the effects of interreflections. Another source of difficulty could be the compromises that need to be made to obtain a solution in the presence of existence difficulties. Most reconstructions shown in the literature are poor. In a comparative review, Zhang *et al.* (2002) summarize: "All the SFS algorithms produce generally poor results when given synthetic data . . . Results are even worse on real images, and . . . [r]esults on synthetic data are not generally predictive of results on real data." More recently, Tankus *et al.* (2005) showed good looking reconstructions of various body structures from endoscopic images, but cannot compare with veridical information. Prados and Faugeras show a good-looking reconstruction of a face, but cannot compare with veridical information (Prados and Faugeras (2005*a*); (2005*b*)). Durou *et al.* (2008*a*), in a recent comparative review, show some fair reconstructions on both synthetic and real data. However, on quite simple shapes methods still produce reconstructions with profound eccentricities.

These problems have driven a search for methods that do not require a reconstruction. Some local features of a shading field—*shading primitives*—are revealing because some geometric structures generate about the same shading pattern whatever the illumination. For example, a pit in a surface will always be dark; grooves and folds tend to appear as a thin, light band next to a thin, dark band; and the shading on a cylinder is usually either a dark band next to a light band, or a light band with dark band on either side. This idea originates with Koenderink and Doorn (1983), and is expounded in (Haddon and Forsyth 1997, Haddon and Forsyth 1998*b*, Han and Zhu 2005, Han and Zhu 2007). On a larger spatial scale, the family of shading patterns that can be produced by a particular object—the *illumination cone*—is smaller than one might expect (Basri and Jacobs 2003, Belhumeur and Kriegman 1998, Georghiades *et al.* 2001), allowing illumination invariant de-

tection by detecting elements of such cones or by matching with an image distance that discounts changes in illumination (Chen *et al.* 2000, Jacobs *et al.* 1998).

PROBLEMS

2.1. We see a diffuse sphere centered at the origin, with radius one and albedo ρ, in an orthographic camera, looking down the z-axis. This sphere is illuminated by a distant point light source whose source direction is $(0, 0, 1)$. There is no other illumination. Show that the shading field in the camera is

$$\rho\sqrt{1 - x^2 - y^2}$$

2.2. What shapes can the shadow of a sphere take if it is cast on a plane and the source is a point source?

2.3. We have a square area source and a square occluder, both parallel to a plane. The source is the same size as the occluder, and they are vertically above one another with their centers aligned.
(a) What is the shape of the umbra?
(b) What is the shape of the outside boundary of the penumbra?

2.4. We have a square area source and a square occluder, both parallel to a plane. The edge length of the source is now twice that of the occluder, and they are vertically above one another with their centers aligned.
(a) What is the shape of the umbra?
(b) What is the shape of the outside boundary of the penumbra?

2.5. We have a square area source and a square occluder, both parallel to a plane. The edge length of the source is now half that of the occluder, and they are vertically above one another with their centers aligned.
(a) What is the shape of the umbra?
(b) What is the shape of the outside boundary of the penumbra?

2.6. A small sphere casts a shadow on a larger sphere. Describe the possible shadow boundaries that occur.

2.7. Explain why it is difficult to use shadow boundaries to infer shape, particularly if the shadow is cast onto a curved surface.

2.8. As in Figure 2.18, a small patch views an infinite plane at unit distance. The patch is sufficiently small that it reflects a trivial quantity of light onto the plane. The plane has radiosity $B(x, y) = 1 + \sin ax$. The patch and the plane are parallel to one another. We move the patch around parallel to the plane, and consider its radiosity at various points.
(a) Show that if one translates the patch, its radiosity varies periodically with its position in x.
(b) Fix the patch's center at $(0, 0)$; determine a *closed form* expression for the radiosity of the patch at this point as a function of a. You'll need a table of integrals for this (if you don't, you're entitled to feel very pleased with yourself).

2.9. If one looks across a large bay in the daytime, it is often hard to distinguish the mountains on the opposite side; near sunset, they are clearly visible. This phenomenon has to do with scattering of light by air—a large volume of air is actually a source. Explain what is happening. We have modeled air as a vacuum and asserted that no energy is lost along a straight line in a vacuum. Use your explanation to give an estimate of the kind of scales over which that model is acceptable.