Shading Ambiguity: Reflectance and Illumination

Michael D'Zmura

An object's shading is determined by its shape and position, by its surface reflectance properties, and by the pattern of illumination. Many studies of shading focus on the recovery of one of these factors by introducing constraints on the other two; one might, for instance, constrain reflectances to be Lambertian and lights to be point sources in a shape-from-shading scheme (Horn, 1975, 1986; Pentland, 1982, 1984, 1988). Yet such constraints are not met in the daily operation of the human visual system. For example, while a point source might suffice as a first-order model of illumination for outdoor viewing on sunny days, it is inadequate to describe spatially extended patterns of illumination under forest canopies, on overcast, wintry days, or within almost any indoor scene. Likewise, we encounter surfaces with many varieties of gloss in daily viewing, so that it is unreasonable to suppose that surfaces are Lambertian or that they have some fixed degree of specularity.

It could be the case that each pattern of shading comes from one specific combination of shape, reflectance, and lighting. In this event, we would find little fault with a visual system that places great stock in quantitative estimates of shape, surface gloss, and lighting from shading information. Yet patterns of shading are not uniquely determined by physical factors, rather infinitely many combinations of shape, reflectance, and lighting give rise to the same shading. This allows photographic reproduction: The reflectance of a print or the light reaching a screen from a slide are modulated to mimic the shading of objects viewed from the position of a camera.

The present concern is with a more constrained situation, namely with the shading of objects of uniform reflectance under more natural lighting conditions. To see intuitively that shading can be ambiguous under these circumstances, consider the shading of a shiny yellow sphere. When lit by a distant point source, the sphere provides a shading pattern with a spatially concentrated highlight. We can blur the highlight by diffusing the

point source. We can also blur the highlight by reducing the gloss of the surface or by reducing the surface curvature in the area of the highlight. The pattern of shading associated with the blurred highlight is physically ambiguous: It can be caused by diffusing the point source, decreasing the degree of surface specularity, or reducing the surface curvature, or some combination of the three.

I present in this chapter a study of the trade-off between the spatial properties of a surface's reflectance and the spatial pattern of its illumination. Two intuitions about highlights suggest that this form of shading ambiguity is most easily characterized in the Fourier domain. The first is that a shiny surface transfers high-frequency components in a pattern of incident illumination better than a dull surface. For instance, details in an incident light pattern can be seen in the light reflected by a mirror but not in the light reflected by a matte surface. The second intuition is that surfaces act as lowpass filters of patterns of incident illumination (Cabral, Max & Springmeyer, 1987): With the exception of mirrors, surfaces blur point sources of incident illumination, and there is no "ringing" in the resulting highlights.

Reflectance functions are formally similar to "transfer" functions that one finds in much of the vision literature; they differ in that their operation on an incident light pattern to produce a reflected light pattern is described using spherical coordinates rather than a time coordinate or Cartesian coordinates on the plane. Fourier analysis on the sphere involves the spherical harmonics; these functions play a role similar to that of sinusoids in frequency-domain analysis on the line or on the plane. By expressing reflectance functions and light patterns in terms of the spherical harmonics, we can describe reflection in a way that captures the frequency-domain intuitions. The analysis reveals the structures of different kinds of reflectances, shows how to handle gracefully the reflection of spatially extended light sources, and leads readily to several results on shading ambiguity. These include solutions to probbems of (1) determining the set of illumination patterns that, for a given surface, give rise to identical reflected light patterns; (2) characterizing illumination patterns in terms of their ability to reveal the spatial properties of reflectance functions; (3) finding pairs of reflectances and illuminants that give rise to identical reflected light patterns, and (4) choosing the lighting of an arbitrary collection of surfaces, seen from a given viewpoint, so that the resulting visual image matches some given image as well as possible.

After briefly reviewing the space-domain properties of lights and surface reflectances, I introduce the basic tools in linear systems analysis on the sphere, namely rotations and the spherical harmonics, in a way that I hope is useful to the reader. I then examine the reflection of light from surfaces in the Fourier domain and point out results on shading ambiguity that follow from this particular representation of the problem.

Reflectance and Illumination

Space Domain

The spatial properties of a surface's reflectance are represented by a bidirectional reflectance function (Horn & Sjoberg, 1979; Nicodemus et al., 1977). This function states what fraction of light at some wavelength, incident on a differential area from some given direction relative to the surface normal, is reflected towards some particular exitant direction. The reflectance depends not only on surface roughness (Bennett & Mattsson, 1989), which is affected by processes such as polishing, but also on surface material so that, for instance, the reflectances of dielectrics such as plastic have a form distinct from those for metals. Reflectance models have been introduced for use in computer graphics that appear to capture well the spatial properties of many materials' reflectance functions (Blinn, 1977/1988; Cook & Torrance, 1982/1987; Phong, 1975). The following brief review of reflectance in the space domain identifies those properties of reflectances that make them amenable to analysis in the frequency domain.

Spherical Coordinate System

To quantify reflectance, let us erect a spherical coordinate system about the normal to some differential area of the surface. As shown in figure 14.1a, we identify the directions of incident and exitant lights with points on the hemisphere whose equator matches the plane tangent to the surface and whose north pole matches the direction of the normal. A direction is specified by its elevation θ, $0 \leqslant \theta \leqslant \pi$, and its azimuth ϕ, $0 \leqslant \phi < 2\pi$. Such directions stand in one-to-one correspondence with unit vectors defined in a Cartesian coordinate system in which the

z-axis is aligned with the normal and the x- and y-axes lie in the plane tangent to the surface; a direction specified by a unit vector with Cartesian coordinates $(u_x\, u_y\, u_z)$ has spherical coordinates $(\theta\, \phi) = (\arccos(u_z)\ \arctan(u_y/u_x))$.

We form a flat representation of the surface of a sphere familiar as the Mercator projection by directly mapping the parameters θ and ϕ onto the plane. As shown in figure 14.1b, the north pole or surface normal direction ($\theta = 0$) is represented along the top of the map; the equator or surface horizon ($\theta = \pi/2$, variable ϕ) is represented by the line across the middle, while the underside of the opaque surface is represented by all directions that lie below the surface horizon ($\pi/2 < \theta \leqslant \pi$).

Incident Lights: Intensity and Irradiance

Taking incident lights to be unpolarized and incoherent, we can describe a pattern of incident light intensity using a real-valued function $I(\theta, \phi)$ of the spherical coordinates set up about the surface normal (the dependence of the intensity on wavelength is suppressed). A bidirectional reflectance function operates on incident irradiance, a function of the incident intensity, to produce the reflected light (Horn & Sjoberg, 1979; Nicodemus et al., 1977). To calculate an irradiance from an intensity, note that a light source of fixed intensity provides energy to the differential area of surface in a way that falls off as $\cos(\theta)$, as shown in figure 14.2. For an opaque surface, furthermore, incident intensities at directions below the surface horizon must be set to zero, so that incident irradiance $I'(\theta, \phi)$ is given in terms of incident intensity as follows:

$$I'(\theta, \phi) = \max(\cos(\theta), 0) I(\theta, \phi). \qquad (1)$$

Reflectance Properties

A reflectance operates on an incident pattern of irradiance $I'(\theta, \phi)$ to produce an exitant pattern of radiance, namely a real-valued function $E(\theta, \phi)$ that describes the amount of light returned towards a viewer at all possible directions.

The results of reflecting an incident light with three types of reflectance are shown in figure 14.3 using the Mercator projection. The top left panel shows the pattern of incident irradiance: a blurred point source that is rotationally symmetric about the direction $(\theta\, \phi) = (\pi/4\ \pi/4)$. The bottom left panel shows the radiance exiting from a perfect mirror that simply rotates the incident irradiance about the normal by π radians: $E(\theta, \phi) = I'(\theta, \phi + \pi)$. The perfect mirror is evidently an "all pass" filter, because all details of the incident irradiance are represented faithfully in the exitant light. The middle panel shows the light reflected by a perfectly matte surface, which suppresses all spatial detail in the incident irradiance: $E(\theta, \phi) = c$, a constant. The right panel, finally, shows the radiance pattern provided by a "mixed" reflectance with both specular and diffuse properties.

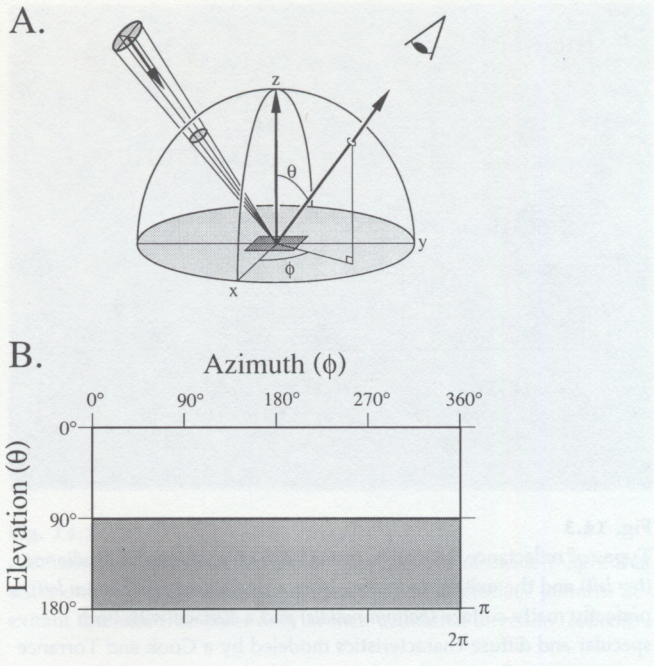

Fig. 14.1
(A) Spherical coordinate system for describing lights and reflectances. (B) Mercator projection for depicting functions of direction about the origin.

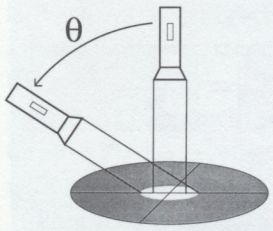

Fig. 14.2
Cosine dependence of irradiance on intensity. The light energy reaching an area from a source of fixed intensity (shown as a flashlight) falls off as the cosine of the angle θ between the surface normal and the source; intensities behind an opaque surface provide zero irradiance.

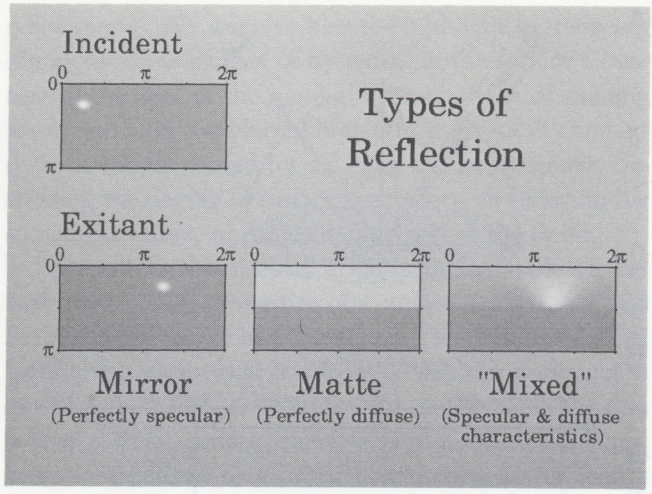

Fig. 14.3
Types of reflectance. Mercator projections of a pattern of irradiance (*top left*) and the exitant radiances from a perfect mirror (*bottom left*), a perfectly matte surface (*bottom middle*) and a surface with both specular and diffuse characteristics modeled by a Cook and Torrance (1982/1987) reflectance function.

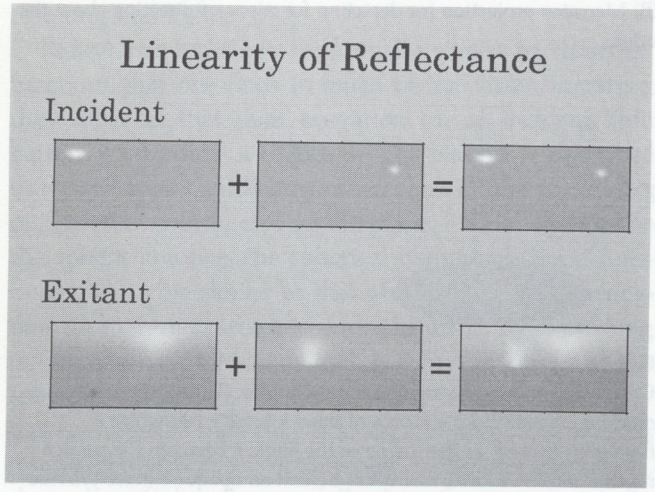

Fig. 14.4
Linearity of reflectance. On the left are shown in Mercator projection an incident pattern of irradiance and the resulting exitant radiance found with a Cook and Torrance (1982/1987) reflectance. In the middle are shown another pattern of irradiance and the resulting exitant radiance from the same surface. The radiance exiting the surface when lit by the sum of the two incident irradiances (*bottom right*) can be found either by adding the two component exitant radiances (*bottom left* and *middle*) or by reflecting directly the sum of the incident irradiances (*top right*).

Linearity

All real surface reflectances are like the "mixed" reflectance of figure 14.3 in that they lie between the extreme mirror and matte cases. We can more easily understand how they filter incident light patterns by relying on their linearity. Illustrated in figure 14.4, this property has the formal statement:

$$E_1 = R(I'_1) \text{ implies } sE_1 = R(sI'_1), \text{ and} \tag{2a}$$

$$E_1 = R(I'_1) \text{ and } E_2 = R(I'_2) \text{ imply that } E_1 + E_2 = R(I'_1 + I'_2), \tag{2b}$$

for all incident irradiances I', exitant radiances E, scalars s, and reflectance functions R. As a consequence of linearity, we can express the operator R as a function of the coordinates $(\theta_i \phi_i)$ and $(\theta_e \phi_e)$ of the incident and exitant directions, respectively:

$$E(\theta_e, \phi_e) = R(\theta_e, \phi_e, \theta_i, \phi_i)'(\theta_i, \phi_i). \tag{3}$$

Symmetry

A second property of surface reflectances is that they satisfy the Helmholtz reciprocity law, which relies on the reversibility of light paths to impose a spatial symmetry on incident and exitant directions. This symmetry is illustrated in figure 14.5; formally, we have for any pair of incident $(\theta_i \phi_i)$ and exitant $(\theta_e \phi_e)$ directions

$$R(\theta_i, \phi_i, \theta_e, \phi_e) = R(\theta_e, \phi_e, \theta_i, \phi_i). \tag{4}$$

That reflectances are symmetric may seem an innocuous property at the moment, yet real symmetric linear operators have real eigenvalues, and this helps considerably in the analysis below.

Shift variation

A third property of many reflectances is that the pattern of reflected light depends on the position of the incident light. The most common manifestation of this shift variation is shown in figure 14.6; if we take an incident light and rotate it in a way that alters its elevation, then the reflected light will change not only in its positioning but also in its pattern. However, many such reflectances are *azimuthally* isotropic (figure 14.7): If the shift (rotation) changes the azimuth but not the elevation of the incident light source, then the exitant light changes its position in azimuth but not in its pattern.

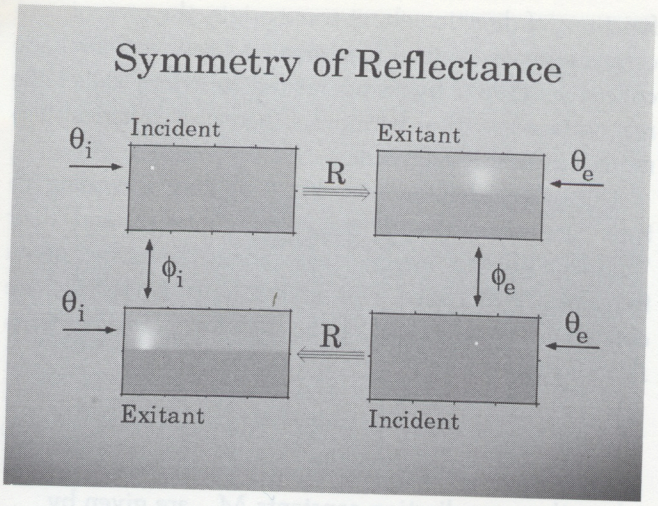

Fig. 14.5
Symmetry of reflectance. At top left is shown a point source irradiance in direction $(\theta_i \phi_i)$ which, when reflected, produces the exitant radiance at top right. At bottom right is shown a point source irradiance in direction $(\theta_e \phi_e)$ that produces an exitant radiance at bottom left. The number $R(\theta_e, \phi_e, \theta_i, \phi_i)$ that describes how much light exits in direction $(\theta_e \phi_e)$ when the surface is lit by the point source at $(\theta_i \phi_i)$ (*top panels*) is identical to the number $R(\theta_i, \phi_i, \theta_e, \phi_e)$ that describes how much light exits in direction $(\theta_i \phi_i)$ when the surface is lit by the point source at $(\theta_e \phi_e)$ (*bottom panels*).

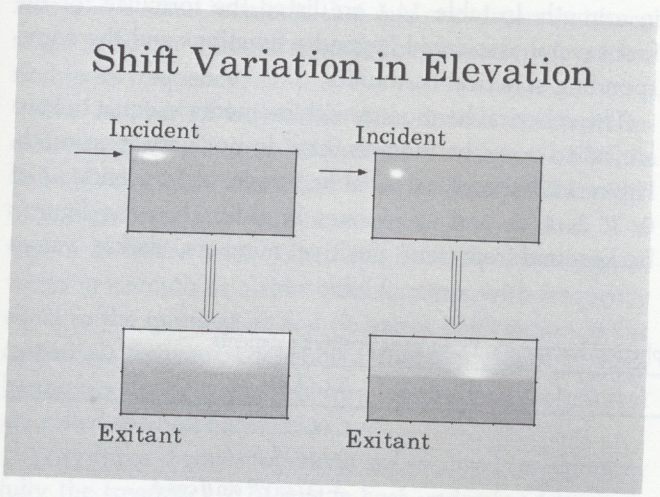

Fig. 14.6
Shift variation in elevation. Two patterns of incident light (*top*) that differ only in their elevation may, when reflected, produce patterns of exitant radiance (*bottom*) that differ in both position *and* pattern.

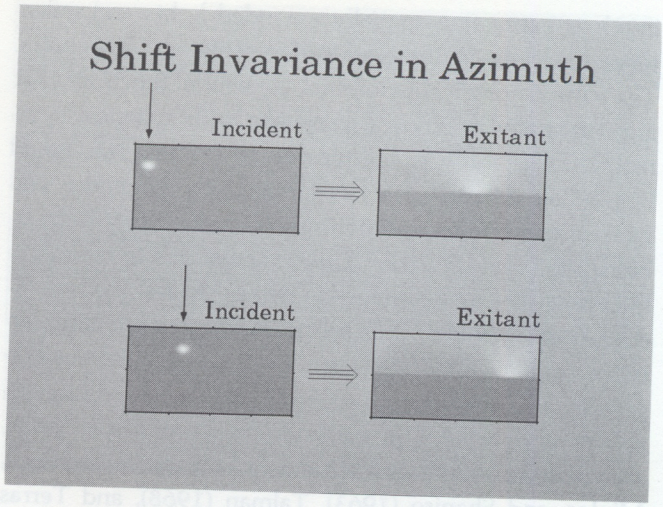

Fig. 14.7
Azimuthally isotropic reflectances. Two patterns of incident light that differ only in their azimuthal position (*top*) often produce patterns of exitant radiance that differ only in their azimuthal direction (*bottom*).

The reflectance model used in computer graphics to mimic reflectances of this sort is Cook and Torrance's (1982/1987). This model's shift variation in elevation, based on the physical model of Torrance and Sparrow (1967), (1) allows the reflectance to provide peak specularity towards a direction with an elevation (measured away from the surface normal) that is greater than that of the perfect specular direction (Cook & Torrance, 1982/1987), and (2) allows the surface to possess "sheen" (Hunter, 1975). These effects are most noticeable for lights with near-grazing angles of incidence.

Reflectances modelled by Cook and Torrance fall into a class midway between the simplest type of reflectance, a fully shift-invariant system, and the most general sort of reflectance, which is isotropic in neither elevation nor azimuth. The simplest type of reflectance often appears in computer graphics under the guise of the Phong model (Blinn, 1977/1988; Phong, 1975). The most general reflectances are "directional" (Hunter, 1975) and are typical of scored surfaces; these have not yet found widespread use in graphics. This three-way classification of reflectances according to shift variation finds a ready expression in the Fourier domain.

Fourier Domain

Because reflectance functions are linear, it is natural to examine them in the Fourier domain which, on the sphere,

involves expressing spatially extended lights and reflectance functions in terms of surface spherical harmonics. The surface spherical harmonics are special functions that arise naturally in physical problems with spherical symmetry. Their study in the nineteenth century was based largely on detailed analysis initiated by Legendre and Laplace in work on gravitation; Hobson (1965), MacRobert (1967), and Walker (1988) follow this approach in their texts. It was recognized in the first half of the twentieth century that almost all special functions arise naturally in the study of continuous groups of transformations (Lie groups). The link between spherical harmonics and rotation groups is developed in many books; several such texts, accessible to nonmathematicians, are those of Gel'Fand, Minlos, and Shapiro (1963), Talman (1968), and Terras (1985). This modern approach stresses matrix representations of operators; Hoffman and Kunze (1961) provide one of many good introductions to linear algebra. More closely related to the present topic is work in computer graphics by Cabral and colleagues (1987), who transformed lights and reflectances into the Fourier domain to derive reflectance functions from surface bump maps. Press and coworkers (1988) and Swarztrauber (1979) present relevant numerical methods.

Spherical Harmonics

The surface spherical harmonics play a role similar to that played by sinusoids in Fourier series expansions of periodic functions. In particular, the surface spherical harmonics provide a complete orthonormal basis with which to express a function defined on the surface of the sphere. While surface spherical harmonics are complex-valued functions of direction about some origin, the present concern is with real-valued functions, so I follow Cabral and colleagues (1987) by using related functions called the real surface spherical harmonics (hereafter spherical harmonics or simply harmonics).

The spherical harmonics $Y_{\ell m}(\theta, \phi)$ vary according to their "frequency" ℓ, a nonnegative integer, and their "moment" m, an integer that ranges for some given frequency ℓ between $-\ell$ and ℓ. The harmonics are defined in terms of the associated Legendre functions $P_{\ell m}(x)$ as follows:

$$Y_{\ell m}(\theta, \phi) = \begin{cases} M_{\ell m} P_{\ell m}(\cos\theta) \cos m\phi, & \text{for } m \geqslant 0 \\ M_{\ell |m|} P_{\ell |m|}(\cos\theta) \sin |m| \phi, & \text{for } m < 0, \end{cases} \tag{5a}$$

where the normalization constants $M_{\ell m}$ are given by

$$M_{\ell m} = \begin{cases} \sqrt{\dfrac{(2\ell + 1)}{2\pi} \dfrac{(\ell - |m|)!}{(\ell + |m|)!}}, & \text{for } m \neq 0 \\ \sqrt{\dfrac{(2\ell + 1)}{4\pi}}, & \text{for } m = 0. \end{cases} \tag{5b}$$

The spherical harmonics are separable functions of elevation and azimuth: an associated Legendre function that takes $\cos\theta$ as its argument describes variation in elevation, while a sinusoid of frequency m describes variation in azimuth. In table 14.1 are listed the formulae for the first several associated Legendre functions and the corresponding spherical harmonics.

The spherical harmonics with moments m equal to zero are called zonal harmonics; they do not vary in azimuth. Figure 14.8 shows six zonal harmonics at frequencies ℓ of 0, 1, 2, 4, 8, and 16. Values brighter than the figure's background represent positive numbers, darker values

Table 14.1
Formulae for the associated Legendre functions $P_{\ell m}(x)$ and the real surface spherical harmonics $Y_{\ell m}(\theta, \phi)$ at positive moments

Frequency, moment	$P_{\ell m}(x)$	$Y_{\ell m}(\theta, \phi)$
$\ell = 0, m = 0$(z)*	1	$(1/4\pi)^{1/2}$
$\ell = 1, m = 0$(z)	x	$(3/4\pi)^{1/2} \cos\theta$
$\ell = 1, m = 1$(s)	$-(1 - x^2)^{1/2}$	$-(3/4\pi)^{1/2} \sin\theta \cos\phi$
$\ell = 2, m = 0$(z)	$1/2(3x^2 - 1)$	$(5/16\pi)^{1/2}(3\cos^2\theta - 1)$
$\ell = 2, m = 1$(t)	$-3(1 - x^2)^{1/2}x$	$-(15/4\pi)^{1/2} \cos\theta \sin\theta \cos\phi$
$\ell = 2, m = 2$(s)	$3(1 - x^2)$	$(15/16\pi)^{1/2} \sin^2\theta \cos 2\phi$
$\ell = 3, m = 0$(z)	$1/2(5x^3 - 3x)$	$(7/16\pi)^{1/2}(5\cos^3\theta - 3\cos\theta)$
$\ell = 3, m = 1$(t)	$-(3/2)(1 - x^2)^{1/2}(5x^2 - 1)$	$-(21/32\pi)^{1/2}(5\cos^2\theta - 1)\sin\theta\cos\phi$
$\ell = 3, m = 2$(t)	$15(1 - x^2)x$	$(105/16\pi)^{1/2} \sin^2\theta \cos\theta \cos 2\phi$
$\ell = 3, m = 3$(s)	$-15(1 - x^2)^{3/2}$	$-(35/32\pi)^{1/2} \sin^3\theta \cos 3\phi$

*The spherical harmonics $Y_{\ell m}(\theta, \phi)$ are zonal (z), sectoral (s) or tesseral (t).

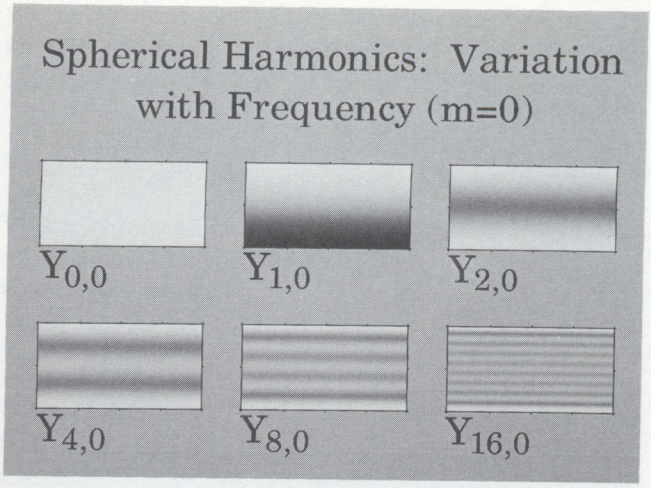

Fig. 14.8
Zonal harmonics at frequencies 0 (*top left*), 1 (*top middle*), 2 (*top right*), 4 (*bottom left*), 8 (*bottom middle*), and 16 (*bottom right*). Points brighter than the background indicate directions where the harmonic takes a positive value; points darker than the background indicate directions where a function has a negative value.

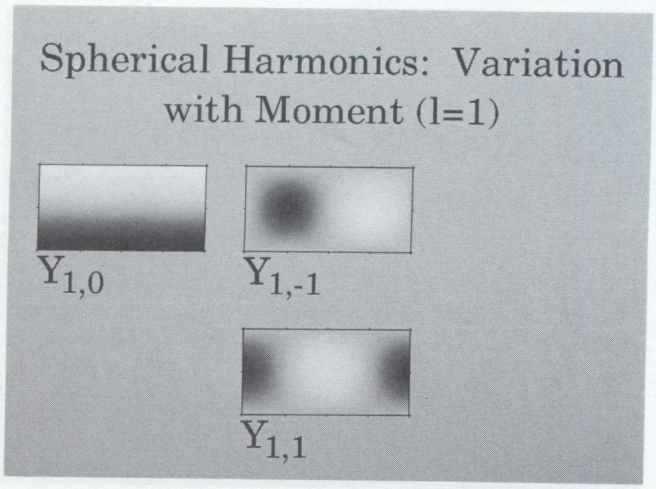

Fig. 14.9
The three harmonics $Y_{1,0}$ (*top left*), $Y_{1,-1}$ (*top middle*), and $Y_{1,1}$ (*bottom middle*).

negative. The figure shows that the harmonic with frequency zero is simply a constant function of direction and that the number of cycles across the surface of the sphere increases with frequency. More precisely, the number of zeroes in a zonal harmonic that one encounters in traversing a great circle from the north to the south pole is identical to its frequency.

The number of harmonics with distinct moments increases with increasing frequency: $-\ell \leqslant m \leqslant \ell$. While there is only one harmonic at $\ell = 0$, there are three distinct harmonics at $\ell = 1$ at moments -1, 0, and 1, as shown in figure 14.9. When $|m| = \ell$, the harmonic (*a*) varies in azimuth as a sinusoidal function with frequency equal to the moment m, and (*b*) varies in elevation as the (spherical) frequency-dependent function $(\sin\theta)^{\ell}$. These harmonics divide the sphere into longitudinal sectors and are called sectoral harmonics.

Why three distinct functions are required to represent fully the frequency ℓ of one is best understood by considering what is meant by frequency. Frequency labels a subspace invariant under a group of transformations. On the line, translating a sinusoid alters its phase but not its frequency, so that frequency is preserved under actions by the group of translations. A sine and a cosine at a particular frequency provide a basis for the invariant subspace. On the sphere, frequency labels a subspace invariant under transformations by the group of rotations. The single, constant function $Y_{0,0}(\theta, \phi)$ occupies the subspace at frequency zero; an arbitrary rotation of this function about the origin returns the same function, so that the subspace at frequency zero is invariant under rotations. That this is the case for the subspace spanned by the three harmonics at a frequency of one (see figure 14.9) is seen intuitively by noting that these harmonics are "dipoles" oriented along the x-axis ($Y_{1,1}$), the y-axis ($Y_{1,-1}$), and the z-axis ($Y_{1,0}$). It turns out that such a dipole, oriented along an *arbitrary* axis, is a linear combination of the three basic dipoles. Rotating a dipole given by some particular linear combination of the three basis functions produces another dipole, given again by some combination of these basis functions, so that the subspace spanned by the three functions is invariant under rotation.

That frequency is preserved under rotation is perhaps more difficult to intuit (but nevertheless true) for the five harmonics that represent the frequency ℓ of two, shown in figure 14.10. This set includes two tesseral harmonics ($Y_{2,-1}$ and $Y_{2,1}$); these are neither zonal ($m = 0$) nor sectoral ($|m| = \ell$) and, again, vary separably in both elevation and azimuth.

Spherical Harmonic Transform

The (real surface) spherical harmonics form a complete orthonormal set of real-valued functions on the surface of the sphere, so that any incident or exitant light may be

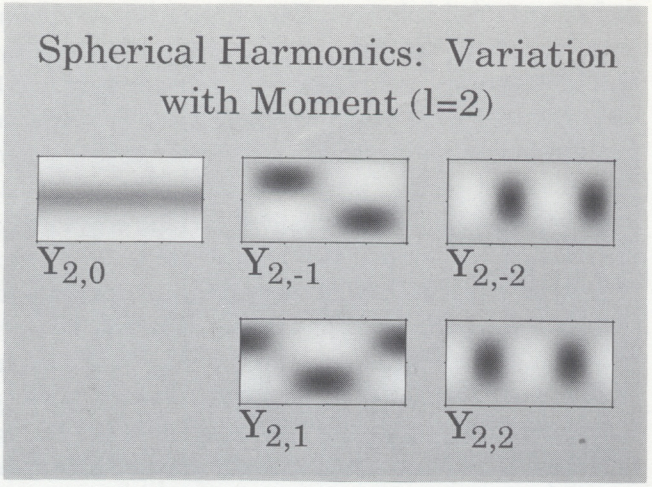

Fig. 14.10
The five harmonics $Y_{2,0}$ (*left*), $Y_{2,-1}$ and $Y_{2,1}$ (*middle top* and *bottom*, respectively), and $Y_{2,-2}$ and $Y_{2,2}$ (*right top* and *bottom*, respectively).

expressed in the Fourier domain as an appropriate linear combination of harmonics. In the case of an incident irradiance we have the series expansion

$$I'(\theta_i, \phi_i) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a'_{\ell m} Y_{\ell m}(\theta_i, \phi_i), \tag{6a}$$

where

$$a'_{\ell m} = \int_0^{2\pi} \int_0^{\pi} I'(\theta_i, \phi_i) Y_{\ell m}(\theta_i, \phi_i) \sin\theta_i \, d\theta_i \, d\phi_i. \tag{6b}$$

In words, one determines the coefficients $a'_{\ell m}$ of the transform of the incident irradiance by integrating the product of the function I' and each individual spherical harmonic $Y_{\ell m}$ over the surface of the sphere.[1] In numerical applications, this spherical harmonic transform is carried out for frequencies ℓ up to some finite cutoff frequency L_{max}. Swarztrauber (1979) discusses methods and sampling grids with which one can more rapidly transform functions than suggested by equation 6; in particular, one can rely on the separability of the harmonics in elevation and azimuth to use an FFT algorithm to transform variations in azimuth.

I show two transforms in figure 14.11. The spherical harmonic $Y_{6,4}$ is depicted at the top middle in the space

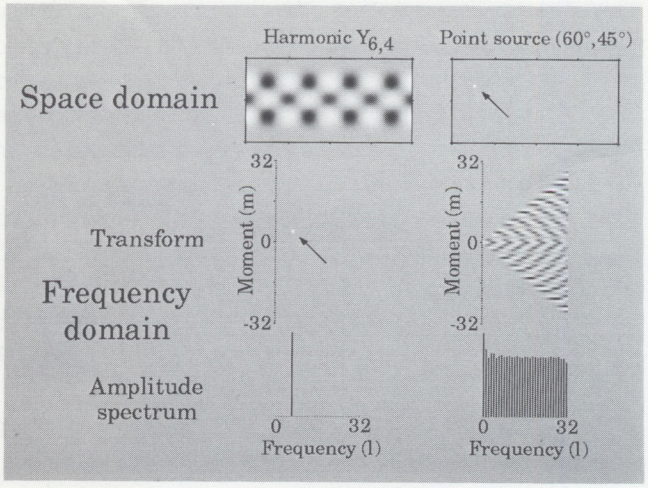

Fig. 14.11
Spherical harmonic transform. *Left*, Transform of a harmonic. At top is depicted the harmonic $Y_{6,4}$ in the space domain. Immediately below is its transform, shown as a function of frequency (increasing from 0 at left to 32 at right) and moment (increasing from −32 at bottom to 32 at top). Positive-valued coefficients are shown using pixels brighter than the background, while negative-valued coefficients are shown using pixels darker than the background. The harmonic has energy only at $(\ell\, m) = (6\ 4)$. At the bottom is shown the amplitude as a function of frequency (increasing from 0 at left to 32 at right). *Right*, Transform of a point source. Shown are the point source in the space domain (*top*), its full transform (*middle*), and its amplitude spectrum (*bottom*).

domain. Below are its full transform and its amplitude spectrum, both of which are discrete; frequency is represented along the abscissa in these latter plots. The ordinate for the plot of the full transform is the moment and, as expected, there is only one nonzero term (represented by the white dot at $\ell = 6, m = 4$) in the linear combination of harmonics that sums to the harmonic at the top. The amplitude spectrum (calculated in a way that I describe just below) shows that there is energy only at a frequency ℓ of six. At the top right is a point source in the space domain. The figure shows that nonzero terms at all frequencies and at many moments are needed to represent the point source in the frequency domain.

To define the "amplitude" of a function $f(\theta, \phi)$ at a particular frequency ℓ', we must take into account the number of moments at that frequency. Let $f_{\ell' m}$ be the

1. Note that a differential area on the surface of the sphere is $\sin\theta \, d\theta \, d\phi$, not simply $d\theta \, d\phi$. This dooms the attempt to generate a basis for analysis on the sphere by replacing the associated Legendre functions in equation 5 with sinusoidal functions of elevation: The resulting functions are not orthogonal to one another.

set of transform coefficients at frequency ℓ'. There are $2\ell' + 1$ of these coefficients, so that the amplitude $A_{\ell'}$ is defined as follows:

$$A_{\ell'} = \frac{1}{\sqrt{2\ell' + 1}} \sqrt{\sum_{m=-\ell'}^{\ell'} f_{\ell' m}^2}. \tag{7}$$

We can represent the "phase" of the function at frequency ℓ' by the vector of coefficients $f_{\ell' m} / \|f_{\ell' m}\|$, which is normalized to unit length.

Shift-Invariant Filters and Point-Spread Functions

On the line we most easily filter a function by passing it through a shift-invariant filter. The filter is represented either by a point-spread function (PSF) in the space domain (more generally an impulse response) with which an input function is convolved, or by a transfer function in the Fourier domain, namely the transform of the PSF, which acts multiplicatively on the transform of the input function (Bracewell, 1978). Shift-invariant filtering on the sphere is very similar. Figure 14.12 shows a noisy pattern in the left column; on the right is shown the result of filtering the noise with a shift-invariant filter whose multiplicative attenuation at each frequency is described by a (discrete) Gaussian function. The spectral properties of the filter are shown as an amplitude at the bottom of the middle column and as a (rotationally symmetric) PSF centered on the north pole at the top of the middle column. While the amplitude at each frequency is attenuated according to the Gaussian, the phase at each frequency is unchanged, so that the filtering blurs the input but preserves its position.

Shift-Invariant Reflectances

The simplest sort of reflectance, typified by Phong's (1975) model, closely resembles a shift-invariant filter: In response to a point-source irradiance, the reflectance produces an exitant radiance pattern (impulse response) that does not depend on incident direction. However, the reflectance does not blur the incident irradiance but blurs the result of using a perfect mirror to reflect the incident irradiance. To calculate the exitant radiance, we rotate the incident irradiance pattern about the surface normal by π radians and apply a shift-invariant filter to the result.

We can readily compute such an exitant radiance in the Fourier domain. A first step is to make the incident irradiance transform coefficients $a'_{\ell m}$ into a (column) vector $\mathbf{a}'$. A natural way to order the coefficients is to order by frequency ℓ: $(a'_{0,0}\ a'_{1,-1}\ a'_{1,0}\ a'_{1,1}\ a'_{2,-2}\ a'_{2,-1}\ a'_{2,0}\ a'_{2,1}\ a'_{2,2}\ \ldots)^{\mathrm{T}}$. The second step is to represent with matrices the linear operations to be performed on $\mathbf{a}'$. To construct the "perfect reflection" matrix, note that (see figures 14.8 to 14.10): (1) rotating a spherical harmonic with an even moment m by π radians about the z-axis returns unscathed the same harmonic, and (2) rotating a harmonic with an odd moment changes its sign. The matrix $\mathbf{P}$ that represents perfect reflection is thus a diagonal matrix (entries off the diagonal are equal to 0) with entries along the diagonal $+1$ or -1 according to whether the harmonic has even or odd moment, respectively. The matrix $\mathbf{P}$ is its own inverse (idempotent).

To construct the matrix $\mathbf{R}$ that represents the shift-invariant filtering by the reflectance, recall that such filters attenuate each spherical harmonic in a way that depends on its frequency ℓ alone. Such a reflectance is completely characterized by a set of numbers $\rho(\ell)$ that describe the

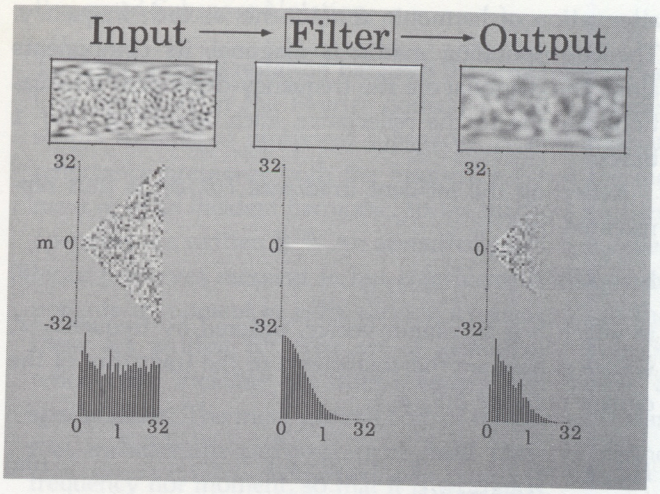

Fig. 14.12
Shift-invariant filtering. A noisy pattern (*left column*) is filtered by a shift-invariant filter (*middle column*) to produce a blurred noisy pattern (*right column*). The patterns at left and right are shown as space-varying functions (*top*), as spherical harmonic transforms (*middle*), and as amplitude spectra (*bottom*). The shift-invariant filter (*middle column*) is described by a (discrete) Gaussian function in the transform domain. To the spectrum representing the shift-invariant filter (*bottom middle*) corresponds a spherical harmonic transform (*middle middle*) with no energy at nonzero moments. The corresponding point-spread function in the space domain (*top middle*) is a linear combination of zonal harmonics and is symmetric about the z-axis (surface normal).

attenuation of harmonic components at each frequency. The corresponding matrix is diagonal; its components along the diagonal are the frequency-dependent attenuations specific to the reflectance, each appearing $2\ell + 1$ times.

Reflecting the incident irradiance $I'(\theta_i, \phi_i)$ is thus represented by

$$\mathbf{b} = \mathbf{RPa}', \tag{8}$$

in which $\mathbf{b}$ is a column vector, ordered by frequency ℓ, whose entries are the coefficients of the transform of the exitant radiance $E(\theta_e, \phi_e)$:

$$E(\theta_e, \phi_e) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} b_{\ell m} Y_{\ell m}(\theta_e, \phi_e), \tag{9a}$$

where

$$b_{\ell m} = \int_0^{2\pi} \int_0^{\pi} E(\theta_e, \phi_e) Y_{\ell m}(\theta_e, \phi_e) \sin \theta_e \, d\theta_e d\phi_e. \tag{9b}$$

One sets a frequency limit when doing numerical calculations based on these equations; the vectors and matrices in equation 8 are then finite-dimensional. The minimal frequency limit (which allows the most efficient computation) is set by the particular cutoff frequency $\ell_{max}(\mathbf{R})$ of a reflectance $\mathbf{R}$; beyond its cutoff frequency, the reflectance transfers numerically insignificant amounts of incident energy.

Azimuthally Isotropic Reflectances

Reflectance functions that are azimuthally isotropic are a more general class than that discussed above; reflectances modelled by Cook and Torrance (1982/1987) fall into this category. It is possible to represent their action on an irradiance in a form like that of equation 8.

To determine the form that an azimuthally isotropic reflectance takes in the Fourier domain, note that rotating harmonics about the surface normal does not alter the frequencies of the sinusoids that describe variation in azimuth: the absolute value of the moment, $|m|$, is preserved under rotations about the z-axis (see equation 5a and figures 14.8 to 14.10). Azimuthally isotropic reflectances must preserve the absolute value $|m|$ of the moment of an incident harmonic in the reflected pattern of light. Reflectances modeled by Cook and Torrance also preserve the *sign* of the moment: Factoring out the perfect reflection $\mathbf{P}$, these reflectances preserve the azimuthal phase of an incident harmonic. The reflectances preserve subspaces labeled by moment m.

Because these reflectances may vary with shift in elevation, they are not generally represented by diagonal matrices, rather they transform linearly terms with varying ℓ but fixed m. In addition, these linear transformations labelled by m must be symmetric in incident and exitant indices because the symmetry of reflectances (equation 4) is unaltered by a linear transformation such as the spherical harmonic transform. Finally, note that the reflectance is insensitive to the phase of the sinusoid that describes a harmonic's variation in azimuth: The linear transformation at some non-negative moment m (cosine phase) must be identical to the transformation at $-m$ (sine phase).

The matrix that represents an azimuthally isotropic reflectance is, therefore, a symmetric block-diagonal matrix in which the blocks are labelled by m with identical blocks at $\pm m$ (figure 14.13). One can order by moment m the components $a'_{\ell m}$ within the vector $\mathbf{a}'$ that represents the incident irradiance to maintain consistency with this block structure. The most general type of reflectance, a directional reflectance typical of scored surfaces, need be isotropic in neither elevation nor azimuth and so is represented in the Fourier domain by a symmetric but otherwise unstructured matrix.

Intensity to Irradiance Revisited

A final ingredient is needed to round out the Fourier-domain description of light reflected by some differential area, namely a representation for the cosine-windowing operator $\max(\cos(\theta), 0)x$ of equation 1 that turns an intensity I into an irradiance I'. An intensity I and an irradiance I' related in this fashion may be transformed into the Fourier domain, in which they are described by a vector $\mathbf{a}$ and a vector $\mathbf{a}'$, respectively. To determine the Fourier-domain structure of the cosine-windowing operator taking $\mathbf{a}$ to $\mathbf{a}'$, note that (1) the operator is linear and symmetric, and that (2) while the operator varies with elevation, it is azimuthally isotropic. The transformation of intensity to irradiance is thus represented by a symmetric block-diagonal matrix that I will call $\mathbf{C}$ with blocks labelled by moment $|m|$ (figure 14.14, right):

$$\mathbf{a}' = \mathbf{Ca}. \tag{10}$$

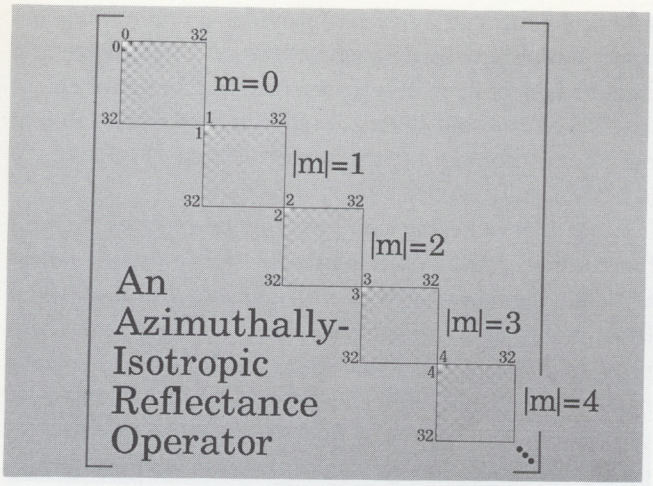

Fig. 14.13
Matrix representation of an azimuthally isotropic reflectance $\mathbf{R}^m$. The reflectance is calculated from Cook and Torrance's model (1982/1987) for a specular reflectance ($s = 1.0$) with coefficient m set to 0.4. Shown are terms in the block-diagonal matrix for blocks labeled by the first several moments. Blocks with the same absolute moment $|m|$ are identical. Positive-valued matrix entries are shown using pixels brighter than the background, while negative-valued coefficients are shown using pixels darker than the background.

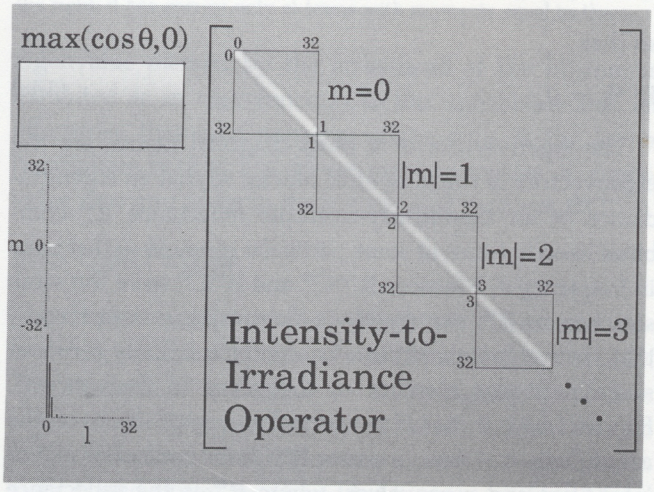

Fig. 14.14
Matrix representation of the intensity-to-irradiance operator. *Left*, The function $\max(\cos\theta, 0)$ in the space (*top*) and transform (*middle* and *bottom*) domains. *Right*, Terms in the block-diagonal matrix representing the corresponding multiplicative operator $\mathbf{C}^m$ for blocks labelled by the first several moments. Blocks with the same absolute moment $|m|$ are identical.

Preserved-Subspace Notation

To distinguish in notation between the structures of various operators in the Fourier domain I shall use superscripts to refer to preserved subspaces. A shift-invariant reflectance preserves both the frequency ℓ and the moment m of an incident harmonic, so the matrices for these operators are written $\mathbf{R}^{\ell m}$. An azimuthally isotropic reflectance preserves moment m but transforms linearly components at moment m with varying frequency ℓ, so that these are written $\mathbf{R}^m$. The set of azimuthally isotropic reflectances includes the shift-invariant reflectances as special cases. The most general sort of reflectance (a class that includes the previous two) need preserve neither frequency nor moment, so that it is written $\mathbf{R}$.

Following this convention, the perfect reflection matrix is $\mathbf{P}^{\ell m}$, while the cosine-windowing operator of equation 10 is written $\mathbf{C}^m$. Note also that reflectances of the form $\mathbf{R}^{\ell m}$ or $\mathbf{R}^m$ commute with the perfect reflection matrix $\mathbf{P}^{\ell m}$:

$$\mathbf{R}^{\ell m}\mathbf{P}^{\ell m} = \mathbf{P}^{\ell m}\mathbf{R}^{\ell m}, \tag{11a}$$

and

$$\mathbf{R}^{m}\mathbf{P}^{\ell m} = \mathbf{P}^{\ell m}\mathbf{R}^{m}. \tag{11b}$$

This is not necessarily true for some directional reflectance $\mathbf{R}$; in this case one can always determine the matrix $\mathbf{R}'$ that satisfies

$$\mathbf{R}'\mathbf{P}^{\ell m} = \mathbf{P}^{\ell m}\mathbf{R}. \tag{11c}$$

To formulate the reflection of a pattern of intensity that is expressed as a vector $\mathbf{a}$ in the frequency domain, we combine equation 8 (exitant radiance from incident irradiance) and equation 10 (incident irradiance from incident intensity). This combination takes three forms:

(a) For shift-invariant reflectances:

$$\mathbf{b} = \mathbf{R}^{\ell m}\mathbf{P}^{\ell m}\mathbf{C}^{m}\mathbf{a} = \mathbf{P}^{\ell m}\mathbf{R}^{\ell m}\mathbf{C}^{m}\mathbf{a}, \tag{12a}$$

(b) For azimuthally isotropic reflectances:

$$\mathbf{b} = \mathbf{R}^{m}\mathbf{P}^{\ell m}\mathbf{C}^{m}\mathbf{a} = \mathbf{P}^{\ell m}\mathbf{R}^{m}\mathbf{C}^{m}\mathbf{a}, \tag{12b}$$

(c) For directional reflectances:

$$\mathbf{b} = \mathbf{R}'\mathbf{P}^{\ell m}\mathbf{C}^{m}\mathbf{a} = \mathbf{P}^{\ell m}\mathbf{R}\mathbf{C}^{m}\mathbf{a}. \tag{12c}$$

The decomposition of surface reflection separates non-varying,[2] geometrical components of reflection, represented by $\mathbf{C}^m$ (which takes into account surface orientation) and $\mathbf{P}^{\ell m}$ (which represents perfect reflection), from those components represented by $\mathbf{R}$ that vary from surface to surface in a way that depends on surface roughness and material properties.

Ambiguity

Arbitrary Views of Flat Surfaces

I turn now to the question of how reflectances and illuminants may be traded off against one another to produce identical reflected light patterns. The trade-offs described below involve determining the inverses of the reflectances used in equation 12 to find exitant radiances; the properties of these inverses are found by studying eigenilluminants.

Eigenilluminants and Kernels

An eigenfunction of a linear operator $\boldsymbol{W}$ is a function G that passes unscathed through the operation except for multiplication by its corresponding eigenvalue ω, which is some scalar:

$$\boldsymbol{W}G = \omega G. \tag{13}$$

Because reflectances are real symmetric linear operators (equation 4), the eigenvalues of a reflectance are real-valued; furthermore, the (normalized) eigenfunctions of a reflectance form a complete orthonormal basis for the space of all possible lights on which the reflectance acts. I call the eigenfunctions of a reflectance eigenilluminants and distinguish, when the need arises, between eigenirradiances, which are the eigenfunctions of a reflectance represented by a matrix $\mathbf{R}$, and eigenintensities, which I take to be the eigenfunctions of an operator represented by a matrix $\mathbf{RC}^m$ (see equation 12).

The eigenirradiances of a shift-invariant reflectance $\mathbf{R}^{\ell m}$ are the spherical harmonics themselves; the corresponding eigenvalues are the frequency-dependent attenuations $\rho(\ell)$ particular to that reflectance. Such an eigenirradiance can be represented by a vector on which $\mathbf{R}^{\ell m}$ acts: The spherical harmonic transform provides a one-to-one correspondence between eigenfunctions of a reflectance and eigenvectors of its matrix representation. The set of eigenirradiances of a shift-invariant reflectance is represented in the Fourier domain by the set of vectors $\{(1\,0\,0\,0\ldots)^T, (0\,1\,0\,0\ldots)^T, (0\,0\,1\,0\ldots)^T, \ldots\}$.

Note that spherical harmonics at frequencies beyond the cutoff $\ell_{max}(\mathbf{R}^{\ell m})$ of a particular shift-invariant reflectance are eigenirradiances with eigenvalue zero: $\rho(\ell) = 0$ for $\ell > \ell_{max}(\mathbf{R}^{\ell m})$. Eigenirradiances of eigenvalue zero provide a basis for a subspace of incident irradiance patterns characterized by having no impact whatsoever on exitant light patterns. Such a subspace is called an operator's kernel and can be written $\ker(\boldsymbol{W})$ for arbitrary operator $\boldsymbol{W}$. As applied to reflectances, we have, for example: if

$$K(\theta,\phi) = \sum_{\ell=\ell_{max}(\mathbf{R}^{\ell m})+1}^{\infty} \sum_{m=-\ell}^{\ell} k_{\ell m} Y_{\ell m}(\theta,\phi), \tag{14a}$$

for some arbitrary real coefficients $k_{\ell m}$ and some shift-invariant reflectance $\mathbf{R}^{\ell m}$, then the transform $\mathbf{k}$ of $K(\theta,\phi)$ satisfies

$$\mathbf{k} \in \ker(\mathbf{R}^{\ell m}), \tag{14b}$$

so that

$$\mathbf{R}^{\ell m}\mathbf{k} = 0. \tag{14c}$$

The eigenvectors of a block-diagonal matrix are the eigenvectors of the individual blocks, so that an eigenirradiance of an azimuthally-isotropic reflectance $\mathbf{R}^m$ combines linearly terms at some particular moment m that vary in frequency ℓ. Matrices $\mathbf{R}^{\ell m}\mathbf{C}^m$ and $\mathbf{R}^m\mathbf{C}^m$ have the same structure as $\mathbf{R}^m$, namely block-diagonal in moment m, so that their eigenintensities also combine linearly terms at some particular moment m that vary in frequency ℓ. Figure 14.15 shows the spectral decomposition of the azimuthally isotropic operator $\mathbf{R}^m\mathbf{C}^m$ in terms of its eigenintensities and eigenvalues, where $\mathbf{R}^m$ is the reflectance depicted in figure 14.13. The eigenintensities at low values of m are exhibited as (column) eigenvectors at the top left of figure 14.15; corresponding eigenvalues are exhibited in the spectrum at the bottom left. The particular eigenintensity marked by the vertical dotted line with

2. This decomposition is suitable for most opaque surfaces. By suitably altering the operators it can be extended to handle thin translucent surfaces.

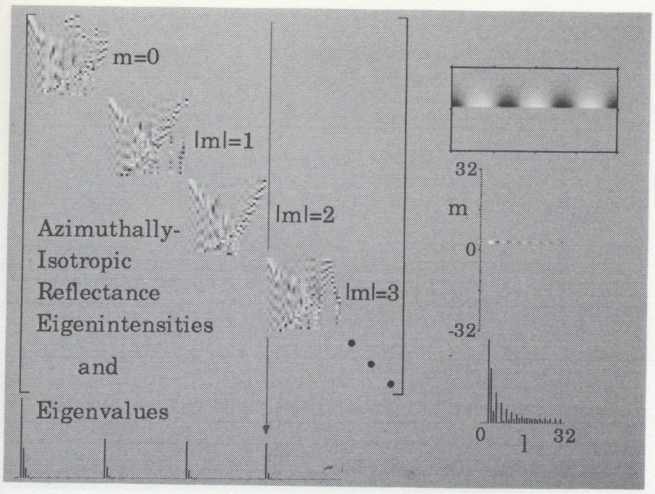

Fig. 14.15
Eigenintensities of the azimuthally isotropic reflectance $\mathbf{R}^m$ depicted in figure 14.13. At left are blocks labelled by moment that contain column vectors, each of which is an eigenvector of matrix $\mathbf{R}^m\mathbf{C}^m$. Beneath these eigenvectors in corresponding columns are the corresponding eigenvalues, shown as a spectrum. The eigenvectors in the right halves of each block correspond to eigenfunctions (of eigenvalue zero) that are nonzero only behind the surface. The eigenvector marked by the vertical dotted line (with eigenvalue marked by the arrow at the bottom) is shown as an eigenfunction at the right; it has energy only at terms with moments equal to three.

eigenvalue marked by the arrowhead at the bottom is exhibited as an eigenfunction to the right. Note that all intensities that take nonzero values only in the lower hemisphere ($\pi/2 \leqslant \theta \leqslant \pi$) are eigenintensities of eigenvalue zero and so lie in the kernel of all matrices $\mathbf{RC}^m$.

Black Light Patterns

The simplest example of shading ambiguity due to the confounding of illumination and reflectance follows readily. The example is analogous to an ambiguity studied in color vision (Wyszecki & Stiles, 1982): Lights that lie in the kernel of some chromatic operator (e.g., trichromatic transduction) may be added to any transferred (e.g., visible) light to produce a metameric light. In analogy, eigenilluminants of a particular reflectance that have an eigenvalue of zero may be added to an incident light pattern to produce a set of "metameric" illuminants for that reflectance; lights reflected *toward all viewpoints* are identical. Following terminology in color research on metamerism (Brainard, Wandell & Cowan, 1989), I call these eigenilluminants, within the kernel of a reflectance, black light patterns.

This notion applies in a straightforward way to flat, extended surfaces of uniform reflectance if one supposes that light sources are sufficiently distant so that intensities expressed about local normals to the surface are the same. More formally, if $\mathbf{k} \in \ker(\mathbf{RC}^m)$, then for any incident intensity $\mathbf{a}$ we may rely on linearity and equation 12 to write

$$\mathbf{b} = \mathbf{P}'^m\mathbf{RC}^m\mathbf{a} = \mathbf{P}'^m\mathbf{RC}^m(\mathbf{a} + \mathbf{k}).^3 \qquad (15)$$

A viewer of such a surface who has prior knowledge of its reflectance properties can thus recover the pattern of illumination from the reflected light only up to an equivalence class of reflectance-dependent metamers.

Adequate Illumination

Can a viewer of a flat, extended surface with prior knowlege of its position and pattern of illumination (due to sufficiently distant sources) determine the spatial properties of its reflectance function? In the simplest case the reflectance is represented in the transform domain by a diagonal matrix with frequency-dependent entries; evidently the only requirement for complete recovery is that the illuminant have nonzero energy at all frequencies passed by the reflectance. The more generally stated requirement—that an illuminant have nonzero projections along all of a reflectance's significant eigenilluminants—is one of *adequate illumination*. By significant eigenilluminants are meant those with nonzero eigenvalues. This requirement may be weakened in the simplest case if the viewer has prior knowledge of the general form of the reflectance[4]; in this case, interpolation and extrapolation can be used to provide an accurate estimate of reflectance despite inadequate illumination. Point sources, environments that present edges in the illumination, and white, brown, etc. noise are prototypical adequate illuminants. The prototypical inadequate illuminant is spatially uniform lighting.

3. In practice we must, of course, meet the constraint that incident intensities are non-negative.

4. The specular component in the Phong model, for instance, is drawn from the family of point-spread functions $[\cos\theta]^n$.

Inadequate Illumination

It is through inadequate illumination that further instances of confounding are most easily devised. Intuitively, one makes the light reflected from a shiny surface like that from a dull surface by lighting the former with patterns limited to low frequencies. Conversely, one can attempt to make the light reflected from a dull surface like that from a shiny surface by lighting the former with patterns with greater energy at high frequencies.

To make identical exitant lights from two flat surfaces with different reflectances, one need but choose an illuminant for one of the surfaces that lies in the intersection of the subspaces spanned by the two surfaces' significant eigenilluminants; the illuminant for the other surface is calculated by matrix inversion. A wide variety of possibilities corresponding to this choice of the first surface's illuminant are then generated by adding patterns to each surface's illuminant that are black light patterns for the respective reflectances.

I shall work this through in detail for two flat surfaces with shift-invariant reflectances $\mathbf{R}_1^{\ell m}$ and $\mathbf{R}_2^{\ell m}$. Supposing that the first surface has incident intensity $\mathbf{a}_1$, the task is to find, if possible, an incident intensity $\mathbf{a}_2$ for the second surface so that the exitant radiance patterns from the two surfaces are identical towards all viewpoints. Using equation 12a, the desired state of affairs can be written

$$\mathbf{b} = \mathbf{P}^{\ell m}\mathbf{R}_2^{\ell m}\mathbf{C}^m\mathbf{a}_2 = \mathbf{P}^{\ell m}\mathbf{R}_1^{\ell m}\mathbf{C}^m\mathbf{a}_1. \qquad (16a)$$

By multiplying the three parts of this equation on the left by the inverse $[\mathbf{P}^{\ell m}]^{-1}$ of $\mathbf{P}^{\ell m}$ (namely $\mathbf{P}^{\ell m}$ itself) and by converting intensities $\mathbf{a}_i$ into irradiances $\mathbf{a}_i'$, we find the following condition for equality:

$$\mathbf{R}_2^{\ell m}\mathbf{a}_2' = \mathbf{R}_1^{\ell m}\mathbf{a}_1'. \qquad (16b)$$

The two reflectances share the same eigenfunctions, namely the spherical harmonics. $\mathbf{R}_1^{\ell m}$ and $\mathbf{R}_2^{\ell m}$ differ in their transfer functions, which are described by the two sets of eigenvalues $\rho_1(\ell)$ and $\rho_2(\ell)$, respectively (see under Shift-Invariant Reflectances). We can define the inverse $[\mathbf{R}_2^{\ell m}]^{-1}$ of $\mathbf{R}_2^{\ell m}$ to be the transfer function $1/\rho_2(\ell)$ for those frequencies ℓ at which $\rho_2(\ell) > 0$. With this restricted inverse we can solve equation 16b for $\mathbf{a}_2'$ using

$$\mathbf{a}_2' = [\mathbf{R}_2^{\ell m}]^{-1}\mathbf{R}_1^{\ell m}\mathbf{a}_1', \qquad (16c)$$

if we simultaneously impose the restriction that the exitant radiance $\mathbf{R}_1^{\ell m}\mathbf{a}_1'$ have no energy at frequencies ℓ at which $\rho_2(\ell) = 0$. The incident intensity $\mathbf{a}_2$ on the upper hemisphere is then found by multiplying the space-domain function represented by $\mathbf{a}_2'$ by the inverse $1/\cos(\theta)$ of the cosine-windowing operator over the restricted domain $\theta \in [0, \pi/2)$.

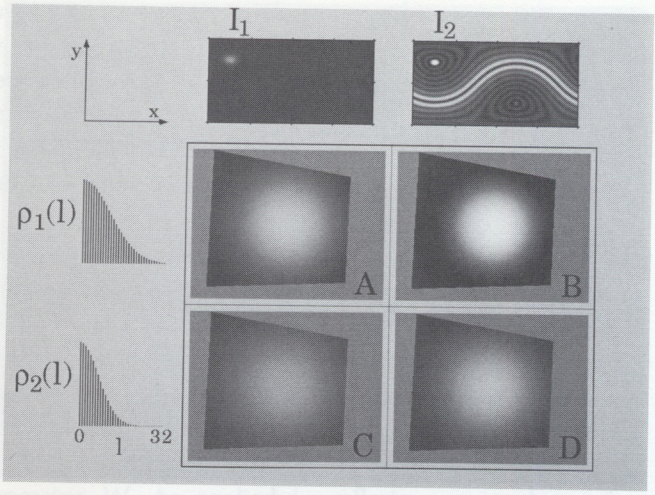

Fig. 14.16
Images of two surfaces with distinct reflectances. Flat surfaces with distinct shift-invariant reflectances shown at the left are lit by two distinct illuminants with intensities shown at the top. The 2 by 2 matrix of simulated reflected lights shows that distinct combinations of reflectance and illumination can produce identical reflected lights (A and D). *Left*, two distinct shift-invariant reflectances represented by their (discrete) Gaussian eigenvalues $\rho_1(\ell)$ and $\rho_2(\ell)$, respectively. *Top*, patterns of intensity I_1 and I_2. I_1 is a Gaussian-blurred point source at $(\theta\,\phi) = (45°\ 45°)$ to which has been added some spatially uniform illumination. The x- and y-axes lie in the page (see *top left*); the z-axis points out towards the viewer. The pattern I_2 was determined by (1) transforming I_1 into an irradiance; (2) convolving the result with $\rho_1(\ell)/\rho_2(\ell)$ (boosting the high frequencies), and (3) transforming the resulting irradiance into the intensity I_2. The shinier surface lit by the diffuse source (A) produces a highlight identical to that produced by the duller surface lit by the "sharp" source (D).

Figure 14.16 shows two simulated flat surfaces distinguished by their (shift-invariant) reflectance functions; these functions are represented to the left by their eigenvalues $\rho_1(\ell)$ and $\rho_2(\ell)$ given, in this example, by (discrete) Gaussian functions of frequency. The surfaces are lit by two distant light source patterns shown at the top; these are chosen so that the surface with reflectance described by $\rho_1(\ell)$ under the first light provides an exitant radiance pattern identical to that of the surface with reflectance $\rho_2(\ell)$ under the second light found using equation 16c. Adding black light patterns to these incident lights will not alter the exitant radiances.

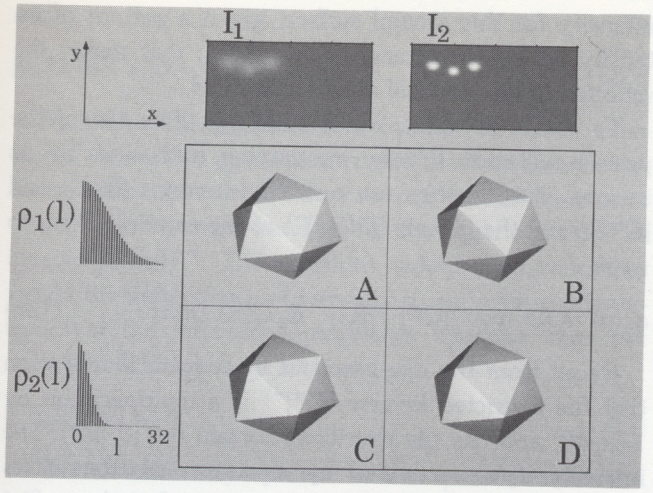

Fig. 14.17
Similar images of two icosahedral surfaces with distinct reflectances. Icosahedra with distinct shift-invariant reflectances shown at the left are lit by two distinct illuminants with intensities shown at the top. Intensity I_1 is found from intensity I_2 by convolving I_2 with $\rho_2(\ell)/\rho_1(\ell)$ (lowpass filtering); the shading patterns in (A) and (D) are indistinguishable.

If one supposes that the amplitudes of natural patterns of incident illumination fall off with frequency, on average, then patterns with a high-frequency boost needed to make a dull surface resemble a shiny surface are, on the whole, unlikely. Surfaces with exitant radiance patterns limited to low frequencies are far more ambiguous under this natural lighting supposition.

Arbitrary View of Convex Surfaces

In the attempt to extend from flat surfaces to identically shaped convex surfaces this way to generate physically distinct pairs of surfaces that appear identical *from all viewpoints*, we find two results: (1) almost all such matches can be shown to be physically imperfect, but (2) one can often find patterns of illumination that can make two convex[5] objects with distinct reflectances differ negligibly in their appearance. The evidence for the latter assertion is best seen; I show in figure 14.17 two icosahedra with the same shape but with distinct (shift-invariant) reflectance functions that nevertheless look very much the same when lights are chosen in accordance with equation 16.

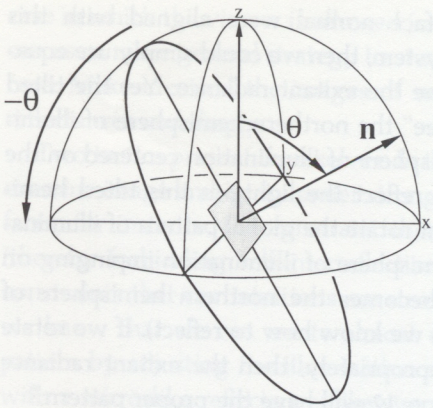

Fig. 14.18
Rotating the pattern of intensity to match surface orientation. The light reaching a piece of surface with unit normal **n** comes from directions lying in the hemisphere aligned about the normal. To reflect this pattern of light using equation 12, we must rotate the pattern of light intensity so that the hemisphere of lights incident on the oriented surface is aligned with the (northern) hemisphere in which we have represented the reflectance operators. If, as shown in this figure, the surface normal is described by an elevation θ in some global coordinate system, then we must rotate the pattern of light intensity through $-\theta$ to perform this alignment.

To see why the reflected light patterns (at all viewpoints) are not completely identical, we can construct an equation similar to equation 16a that takes into account the additional fact that incident light patterns, drawn from some "global" pattern of distant light sources, vary with surface orientation. With this equation we find that the pattern of intensity that compensates for a change in reflectance depends on surface orientation, so that there is no one global pattern of light sources that compensates correctly at all surface orientations for all viewpoints.

Rotations

To describe the light exiting a convex surface at a location with some particular surface orientation, we must first describe the relationship between incident irradiance, a local quantity, and the global pattern of illumination. Figure 14.18 shows a tilted piece of surface within the global coordinate system in which are described the distant light sources that form the illumination pattern

5. By limiting ourselves to convex objects we can ignore problems that arise with self-interreflection. Kajiya (1986) reviews methods for treating reflected light under circumstances in which interreflection plays a role.

$I(\theta_g, \phi_g)$. If the surface normal were aligned with this global coordinate system, then we could simply use equation 12 to determine the exitant radiance. Yet the tilted surface does not "see" the northern hemisphere of illumination but the hemisphere of illumination centered on the surface normal. To reflect the lights in this tilted hemisphere, we must first rotate the global pattern of illumination so that the hemisphere of illumination impinging on the tilted surface becomes the northern hemisphere of illumination (which we know how to reflect). If we rotate the illumination appropriately, then the exitant radiance found using equation 12 will have the proper pattern.[6]

Let us express this use of rotations more formally by letting $\mathbf{D}(\mathbf{n})$ stand for the rotation that takes the hemisphere of illumination impinging on the piece of surface with unit normal $\mathbf{n}$ into the northern hemisphere. The exitant radiance $\mathbf{b}(\mathbf{n})$, which depends on the surface normal, is then given by

$$\mathbf{b}(\mathbf{n}) = \mathbf{P}^{\ell m}\mathbf{R}\mathbf{C}^{m}\mathbf{D}(\mathbf{n})\mathbf{a}. \tag{17}$$

The form of the rotation matrix $\mathbf{D}$ used in equation 17 has been discussed above (see especially under Spherical Harmonics and Azimuthally Isotropic Reflectances). Rotations leave invariant subspaces labelled by frequency ℓ but transform linearly terms at each frequency that vary in moment. Rotations are thus represented by block-diagonal matrices $\mathbf{D}^{\ell}$. The entries of these matrices depend on the particular rotation under consideration; these are most often parameterized by the three Euler angles.[7] I shall continue, however, to parameterize the rotations by the unit surface normal as in equation 17.

In analogy to equation 16a, the equation that expresses the desired identity between the lights reflected by two identically shaped convex surfaces with distinct reflectances is

$$\mathbf{b}(\mathbf{n}) = \mathbf{P}^{\ell m}\mathbf{R}_2\mathbf{C}^{m}\mathbf{D}^{\ell}(\mathbf{n})\mathbf{a}_2(\mathbf{n}) = \mathbf{P}^{\ell m}\mathbf{R}_1\mathbf{C}^{m}\mathbf{D}^{\ell}(\mathbf{n})\mathbf{a}_1. \tag{18a}$$

This equation expresses the sought pattern of intensity $\mathbf{a}_2(\mathbf{n})$ for the second surface as a function of surface orientation: It is not clear that there exists *a* pattern of intensity for this second surface, given a pattern of intensity $\mathbf{a}_1$ and reflectances $\mathbf{R}_1$ and $\mathbf{R}_2$, that makes the reflected lights identical at all viewpoints.

To solve for $\mathbf{a}_2(\mathbf{n})$, we first rid equation 18a of the perfect reflection $\mathbf{P}^{\ell m}$ by multiplying both sides by its inverse. We can then use restricted inverses $[\mathbf{R}_2]^{-1}$ and $[\mathbf{C}^{m}]^{-1}$ and the inverse $[\mathbf{D}^{\ell}(\mathbf{n})]^{-1}$ of the rotation $\mathbf{D}^{\ell}(\mathbf{n})$ to write

$$\mathbf{a}_2(\mathbf{n}) = [\mathbf{D}^{\ell}(\mathbf{n})]^{-1}[\mathbf{C}^{m}]^{-1}[\mathbf{R}_2]^{-1}\mathbf{R}_1\mathbf{C}^{m}\mathbf{D}^{\ell}(\mathbf{n})\mathbf{a}_1. \tag{18b}$$

Recall from the discussion in Inadequate Illumination that the restricted inverse $[\mathbf{C}^{m}]^{-1}$ is a multiplication by $1/\cos(\theta)$ and has the restricted domain $\theta \in [0, \pi/2)$. The intensity $\mathbf{a}_2(\mathbf{n})$ recovered by equation 18b thus takes values only within the hemisphere centered on the surface normal. Two distinct surface normals $\mathbf{n}_1$ and $\mathbf{n}_2$ will provide two hemispheric intensities $\mathbf{a}_2(\mathbf{n}_1)$ and $\mathbf{a}_2(\mathbf{n}_2)$, and the question is whether it is possible for these intensities to agree at those directions where the hemispheres overlap.

Commutativity

The question can be answered by considering the commutativity of the various operators used in equation 18b. We know, for instance, that a shift-invariant reflectance commutes with an arbitrary rotation: the order in which we apply reflectance $\mathbf{R}^{\ell m}$ and rotation $\mathbf{D}^{\ell}(\mathbf{n})$ does not matter. More formally,

$$\mathbf{R}^{\ell m}\mathbf{D}^{\ell}(\mathbf{n}) = \mathbf{D}^{\ell}(\mathbf{n})\mathbf{R}^{\ell m}. \tag{19}$$

This commutativity is handily expressed in terms of the commutator $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$. With this notation, the commutativity expressed by equation 19 is equivalent to saying that the commutator $[\mathbf{R}^{\ell m}, \mathbf{D}^{\ell}(\mathbf{n})]$ is equal to $\mathbf{0}$. Note that neither azimuthally isotropic reflectances nor general reflectances commute with all rotations (particularly those rotations *not* about the surface normal), so that *in general*

$$[\mathbf{R}^{\ell m}, \mathbf{D}^{\ell}(\mathbf{n})] = \mathbf{0}, \tag{20a}$$

$$[\mathbf{R}^{m}, \mathbf{D}^{\ell}(\mathbf{n})] \neq \mathbf{0}, \tag{20b}$$

6. The exitant radiance found this way (equation 17) will not generally have the correct position: It is aligned with the northern hemisphere. If we want the exitant light to have both the proper pattern and the proper position in global coordinates, then we must rotate it back to the position determined by the surface normal using the rotation that is the inverse of the initial rotation.

7. While the formula for these entries (e.g., Talman (1968) equations 9.19 and 9.20) is somewhat forbidding (and must be altered to apply to real-valued rather than to complex-valued harmonics), the entries are readily computed.

and

$$[\mathbf{R}, \mathbf{D}^{\ell}(\mathbf{n})] \neq \mathbf{0}. \tag{20c}$$

By examining equation 18b for shift-invariant reflectances, we see that *if* rotations commute with the intensity-to-irradiance operator $\mathbf{C}^{m}$ then we can shift $\mathbf{D}^{\ell}(\mathbf{n})$ over to a position in the right-hand side of the equation beside its inverse $[\mathbf{D}^{\ell}(\mathbf{n})]^{-1}$. The rotations would then cancel, as would the dependence of intensity $\mathbf{a}_2$ on surface orientation. Yet the intensity-to-irradiance operator does not commute with arbitrary rotations:

$$[\mathbf{C}^{m}, \mathbf{D}^{\ell}(\mathbf{n})] \neq \mathbf{0}. \tag{21}$$

This is intuitively evident in the space domain and follows in the Fourier domain from the distinct structures of the matrices that represent the operators.

We could still succeed in cancelling the rotations (and the dependence on surface orientation) if reflectances commuted with the intensity-to-irradiance operator. If, in equation 18b, we could pass $\mathbf{C}^{m}$ through the reflectances so that it came to rest beside its inverse, then we could cancel the operators with which rotations do not commute. We could then rid the equation of its dependence on surface normal if the reflectances were shift-invariant (equation 20a). Shift-invariant reflectances do *not* commute with the intensity-to-irradiance operator, however:

$$[\mathbf{R}^{\ell m}, \mathbf{C}^{m}] \neq \mathbf{0}. \tag{22}$$

Finally, while it is theoretically possible for an azimuthally isotropic reflectance $\mathbf{R}^{m}$ to commute with $\mathbf{C}^{m}$, we still cannot remove the dependence on surface normal from equation 18b because such reflectances do not commute with arbitrary rotations (equation 20b).

It is generally impossible to make the lights from identically shaped convex surfaces with distinct reflectances identical towards all viewpoints. For shift-invariant reflectances, the trade-off between reflectance and illumination can be accomplished in an approximate fashion if the known illumination pattern $\mathbf{a}_1$ and the reflectances $\mathbf{R}_1$ and $\mathbf{R}_2$ are chosen to minimize the result of applying the commutator $[[\mathbf{R}_2^{\ell m}]^{-1}\mathbf{R}_1^{\ell m}, \mathbf{C}^{m}]$ to the local intensities $\mathbf{D}^{\ell}(\mathbf{n})\mathbf{a}_1$.

Single Views

Ambiguity increases as the number of views decreases. Let us now turn from the extreme case of infinitely many views considered above to the other extreme, namely the case in which the viewer is allowed just a single view. We can use a more general strategy to study ambiguity in this case: Make identical the images of surfaces with distinct reflectances *and* shapes.

Two tactics present themselves. In the first we carefully examine image ambiguity under strong constraints on surfaces, for instance, the constraint that shapes be convex (thereby forbidding self-interreflection). It is under such constraints that our intuitions are strongest. For instance, we know that we can create almost any image shading pattern by adjusting the illumination of a convex surface with a mirrorlike reflectance; likewise, the set of shading patterns that we can create using a matte surface is far more limited (for results of this sort with point sources see Horn, Szeliski & Yuille, 1989).

The second tactic is to rely solely on the constraints that we have used throughout, namely that light sources be distant from the viewed surfaces and that observation by the viewer does not interfere with the system. In tandem with the linearity of reflectance, we can use these constraints to provide a general way to choose the lighting of a particular surface or collection of surfaces so that the pattern of shading matches, as well as possible, some given pattern of shading.

Fourier-Domain Representation of the Linear Rendering Operator

Let us suppose that our lights are distant and that we are viewing an arbitrarily complex set of surfaces with known shapes, positions, and reflectances. We can render such a scene (i.e., calculate the light returned towards the viewer from each image location) either by (1) reflecting some sum of point sources at each position and then using this point-wise information to make a spatially extended image, or (2) calculating the spatially extended image due to each point source and then making the final image by summing over images from each point source. Either method can be employed if one uses a ray tracing procedure to render the scene. The latter method works also with more sophisticated rendering procedures such as radiosity (Cohen et al., 1988; Goral et al., 1984; Kajiya, 1986), in which interreflections among surfaces are taken into account.

Rather than compute the images that result from individual point sources and then sum over these to find the final image, we can form the images that arise from each spherical harmonic component in the illumination and

then sum over these. Let us call $\boldsymbol{S}$ the linear operator that, for some given scene and choice of viewpoint, transforms a pattern of illumination $I(\theta_g, \phi_g)$ into its corresponding continuous image $J(x, y)$:

$$J(x, y) = \boldsymbol{S}(I(\theta_g, \phi_g)). \tag{23}$$

The images $J_{\ell m}(x, y)$ that arise from spherical harmonic patterns of illumination are then given by

$$J_{\ell m}(x, y) = \boldsymbol{S}(Y_{\ell m}(\theta_g, \phi_g)), \tag{24}$$

so that the image $J(x, y)$ that arises from illumination pattern $I(\theta_g, \phi_g)$ with vector representation $\mathbf{a}$ is

$$\begin{aligned} J(x, y) &= \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} J_{\ell m}(x, y) \\ &= \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} \boldsymbol{S}(Y_{\ell m}(\theta_g, \phi_g)) \\ &= \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \boldsymbol{S}(a_{\ell m} Y_{\ell m}(\theta_g, \phi_g)) \\ &= \boldsymbol{S}\left(\sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\theta_g, \phi_g) \right) \\ &= \boldsymbol{S}(I(\theta_g, \phi_g)). \end{aligned} \tag{25}$$

We must discretize not only the illumination but also the spatially continuous image to put this decomposition to use. The latter is done most simply using the discrete Fourier transform (Bracewell, 1978), which assigns to a continuous image $J(x, y)$ with rectangular domain a vector $\mathbf{j}$ with complex components j_{uv}. The (space-domain) operator $\boldsymbol{S}$ is then represented in the Fourier domain by a matrix $\mathbf{S}$ that takes as input an illumination vector $\mathbf{a}$ and provides as output the transform $\mathbf{j}$ of the resulting image:

$$\mathbf{j} = \mathbf{S}\mathbf{a}. \tag{26a}$$

In terms of the discrete Fourier transform indices u and v for the image and the spherical harmonic transform indices ℓ and m for the illumination, we have

$$j_{uv} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} S_{uv,\ell m} a_{\ell m}. \tag{26b}$$

Least-Squares Lighting

Determining the entries $S_{uv,\ell m}$ in the matrix $\mathbf{S}$ is clearly a laborious procedure. One must first render a scene for each individual spherical harmonic $Y_{\ell m}(\theta_g, \phi_g)$ to find the corresponding continuous image $J_{\ell m}(x, y)$. Each of these images must then be Fourier transformed to provide, for each choice of ℓ and m, the entries $S_{uv,\ell m}$.[8] Having computed these entries, however, we can rapidly find the discrete Fourier transform of an image arising from an arbitrary pattern of illumination. With the entries we can, furthermore, invert the rendering process.

The matrix $\mathbf{S}$ has both a range, namely the subspace of image transforms that can arise from the scene, and a kernel, namely the transforms of the set of illumination patterns that are mapped by $\boldsymbol{S}$ to the zero image $J(x, y) = 0$. If some particular image transform lies within the range of the matrix $\mathbf{S}$, then $\mathbf{S}$ may be inverted to determine the spherical harmonic transform of the pattern of illumination that gives rise to that image. If, on the other hand, the image lies outside of the range of $\mathbf{S}$, then there is no pattern of illumination for the scene and viewpoint represented by $\mathbf{S}$ that will produce the image. Under these circumstances it is nevertheless possible to find the pattern of illumination that produces an image that is the best possible fit in the least-squares sense to the given image. This approximation is found by performing a singular value decomposition of $\mathbf{S}$, which is a standard numerical procedure (Press et al., 1988) for "inverting" singular matrices. The procedure provides not only a restricted inverse $\mathbf{S}^{-1}$ but also a basis for the range of images that could arise from the scene and a basis for the illumination patterns that lie in the scene's kernel.

With the restricted inverse $\mathbf{S}^{-1}$ we can find the pattern of illumination $\mathbf{a}$ for the scene[9] that provides an image that is the best possible fit to an arbitrary image $\mathbf{j}$:

$$\mathbf{a} = \mathbf{S}^{-1}\mathbf{j}. \tag{27}$$

This is a very general result, as befits the tactic of placing constraints on neither surface shapes nor reflectances. It rests on the fact that the map between illumination and image is linear; the particular representation

8. If the image in the space domain is already discrete, then this step may obviously be skipped; the transformed images would be preferred, in this case, only if the image has negligible energy at high spatial frequencies, allowing one to truncate the representation.

9. As noted earlier, the inverse transform $I(\theta_g, \phi_g)$ of $\mathbf{a}$ must be a non-negative function in practice.

for the map used here relies on the constraint that the illumination be distant. It is more difficult to find such a representation if we loosen the constraints by allowing arbitrary light source positions.

Discussion

By representing lights and reflectances in the Fourier domain, we have found several results on shading ambiguity. The first is that a reflectance has an infinite number of black light patterns, so that a viewer who knows a surface's reflectance properties can recover the pattern of illumination from the reflected light only up to an equivalence class of reflectance-dependent metamers. The second is that a pattern of illumination must be adequate if the viewer is to recover the spatial properties of reflectance from the shading of a surface. Third, we can use inadequate illumination to make surfaces with distinct reflectances look alike, although such matches are almost always imperfect in the case of closed surfaces potentially viewed from all directions. Fourth, we have found a way to choose the lighting of an arbitrary collection of surfaces, seen from a particular viewpoint, so that the resulting visual image matches some given image as well as possible. Finally, this approach makes evident the linear relationship between illumination and image for complex scenes.

The constraint that illumination be distant and unperturbed by the viewer disallows translation-varying intensity, shadows cast by the viewer, and interreflections between surfaces and the viewer. While these features of more realistic scenes are readily synthesized in computer graphics applications, their analysis appears far more difficult. With the constraints, we can analyze for *extended surfaces* the intuition that surface reflectances blur incident light patterns and study lights reflected towards all viewpoints as well as towards single viewpoints.

The analysis shows that the *spatial* trade-offs between reflectance and illumination are formally similar to the *chromatic* trade-offs between reflectance and illumination so important in studies of surface color appearance. The chromatic spectrum of the light from a small area of (optically inactive) surface is given by the wavelength-by-wavelength product of the surface's reflectance function and the illuminant's spectral power distribution: $E(\lambda) = R(\lambda)I(\lambda)$. The action of a reflectance can be represented by a simple multiplication in the wavelength domain because light incident at a particular wavelength gives rise only to exitant light at the same wavelength: Monochromatic lights are eigenfunctions for the chromatic component of reflectances.

While the "blurring" action of a reflectance on spatial patterns of incident light seems far more complex than the simple multiplication in the wavelength domain, this is so only because point sources of light are not eigenfunctions for the spatial component of physically realizable reflectances. The simplest type of reflectance has spherical harmonic patterns of light as its eigenfunctions; such a reflectance acts on incident light patterns in the space domain by convolution and in the Fourier domain by simple multiplication. The present results on black light patterns, on the adequacy of illumination and on ambiguity in cases of inadequate illumination are wholly analogous to corresponding results in the chromatic domain.

Extending this work on shading ambiguity to the simultaneous consideration of spatial and chromatic properties of reflectance is straightforward, particularly so for reflectances with "specular" components that have separable spatial and chromatic properties. While this separability is known not to hold in some cases (e.g., for materials like red metals with refractive indices that vary substantially with wavelength; see Cook & Torrance, 1982/1987), Healey (1989) has recently shown that many of these inseparable specular reflectances are approximated well by separable functions. Under the condition that the specular component of a reflectance is spatiochromatically separable, the standard models of reflectance used in the vision literature (see, e.g., D'Zmura & Lennie, 1986; Healey, 1989; Klinker, Shafer & Kanade, 1988; Shafer, 1984; Tominaga & Wandell, 1989) are translated readily from the space domain into the Fourier domain. The reflection of spatiochromatically inhomogeneous illuminants by colored surfaces may be analyzed in a way similar to that presented here.

Patterns of shading, highlights in particular, are very sensitive to the precise details of surface material and roughness and to the spatial pattern of illumination. While these details can be controlled in a manufacturing inspection setting in which, for instance, specularity can be used as a cue to shape (e.g., Healey & Binford, 1988; Sanderson, Weiss & Nayar, 1988), this control is not present in the normal, everyday viewing of scenes. If we

are to infer precise shape from shading patterns, we need accurate estimates of both lighting and reflectance.

One thus cannot find fault with the human visual system for using shading information to estimate shape in such a half-hearted fashion. For instance, Barrow and Tenenbaum (1981) showed that contour generally overrides shading information in determining perceived shape (see also Witkin & Tenenbaum, 1983). Mingolla and Todd (1986; see also Todd & Mingolla, 1983) have shown that shading provides a weak input to the perception of variations in depth, in that observers consistently underestimate depth variation from shading information alone. Bülthoff and Mallot (1988) have shown that human shape-from-shading mechanisms perform better in stereo, where shading disparity can cue degree of convexity or concavity (Blake, 1985), but that edges nevertheless override shading disparity.

We must likewise know surface shape and pattern of illumination if we are to infer correctly surface reflectance properties. While the pattern of inference from shading to the spatial properties of reflectance is incompletely understood, there is one well-known inference illustrated by Beck (1972). Beck made two pictures of a vase under known illumination that are identical except for the presence or the (artificial) absence of a highlight. The vase with the highlight appears shiny across its entire surface, while the picture of the same vase with highlight removed appears completely matte. The visual system seems to infer automatically that a surface is glossy from the presence of a highlight. We know that this is a robust inference because a highlight is seen by the viewer only if several conditions are met: (1) the surface must be glossy; (2) the illumination must be adequate, and (3) the viewer must occupy a viewpoint at which the highlight is visible. The failure to see a highlight, on the other hand, can be caused by a matte reflectance, by inadequate illumination, or by poor choice of viewpoint, so that the viewer must have further knowledge of these factors if the correct inference about surface reflectance is to be drawn.

I have invoked a lot of machinery to make a simple point, namely that variation in the spatial properties of reflectances and illuminants leads to ambiguity in shaded images. Yet the machinery and the intuitions behind its use are similar to those belonging to linear systems analysis on the line or on the plane. Furthermore, spherical harmonic transforms, rotation matrices, and the like are readily implemented numerically on a personal computer,[10] albeit at a low resolution. I hope that the methods presented here will prove useful in psychophysical studies; taking shading ambiguity into account is an important ingredient in conducting controlled psychophysical experiments on the perception of shape, reflectance, and illumination.

Acknowledgments

I thank Al Ahumada, Dave Brainard, Geoff Iverson, Mike Landy, and Per Møller for helpful discussions and comments. This work was supported by NEI grants EY04440, EY01319, and RR03825 and by funds from the School of Social Sciences at UC Irvine.

References

Barrow, H. G. & Tenenbann, J. M. (1981). Computational vision. *Proceedings of the IEEE, 69,* 572–595.

Beck, J. (1972). *Surface color perception.* Ithaca, NY: Cornell.

Bennett, J. M. & Mattsson, L. (1989). *Introduction to surface roughness and scattering.* Washington, D.C.: Optical Society of America.

Blake, A. (1985). Specular stereo. *Proceedings of the 9th Intentional Joint Conference on Artificial Intelligence,* Los Angeles, 973–976.

Blinn, J. F. (1977/1988). Models of light reflection for computer synthesized pictures. In W. Richards (Ed.), *Natural computation* (pp. 214–223). Cambridge, MA: MIT.

Bracewell, R. N. (1978). *The Fourier transform and its applications* (2nd ed.). New York: McGraw-Hill.

Brainard, D. H., Wandell, B. A. & Cowan, W. B. (1989). Black light: How sensors filter spectral variation of the illuminant. *IEEE Transactions in Biomedical Engineering, 36,* 140–149.

Bülthoff, H. H. & Mallot, H. A. (1988). Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America A, 5,* 1749–1758.

10. Rendering complex surfaces using equation 17 on a computer such as an Apple Macintosh II can readily tax one's patience, however; even more demanding are equations 26 and 27, which should be avoided completely on such a platform.

Cabral, B., Max, N. & Springmeyer, R. (1987). Bidirectional reflection functions from surface bump maps. *ACM Computer Graphics, 21,* 273–281.

Cohen, M. F., Chen, S. E., Wallace, J. R. & Greenberg, D. P. (1988). A progressive refinement approach to fast radiosity image generation. *ACM Computer Graphics, 22,* 75–84.

Cook, R. L. & Torrance, K. E. (1982/1987). A reflectance model for computer graphics. In W. Richards & S. Ullman (Eds.), *Image Understanding 1985–86* (pp. 1–19). Norwood, New Jersey: Ablex.

D'Zmura, M. & Lennie, P. (1986). Mechanisms of color constancy. *Journal of the Optical Society of America A, 3,* 1662–1672.

Gel'Fand, I. M., Minlos, R. A. & Shapiro, Z. (1963). *Representations of the rotation and Lorentz groups and their applications.* Oxford: Pergamon.

Goral, C. M., Torrance, K. E., Greenberg, D. P. & Battaile, B. (1984). Modeling the interaction of light between diffuse surfaces. *ACM Computer Graphics, 18,* 213–222.

Healey, G. (1989). Using color for geometry-insensitive segmentation. *Journal of the Optical Society of America A, 6,* 920–937.

Healey, G. & Binford, T. O. (1988). Local shape from specularity. *Computer Vision, Graphics, and Image Processing, 42,* 62–88.

Hobson, E. W. (1965). *The theory of spherical and ellipsoidal harmonics.* New York: Chelsea.

Hoffman, K. & Kunze, R. (1961). *Linear algebra.* Englewood Cliffs, NJ: Prentice-Hall.

Horn, B. K. P. (1975). Obtaining shape from shading information. In P. H. Winston (Ed.), *The Psychology of machine vision* (pp. 115–155). New York: McGraw-Hill.

Horn, B. K. P. (1986). *Robot vision.* New York: McGraw-Hill.

Horn, B. K. P.& Sjoberg, R. W. (1979). Calculating the reflectance map. *Applied Optics, 18,* 1770–1779.

Horn, B. K. P., Szeliski, R. S. & Yuille, A. L. (1989). Impossible shaded images. *MIT AI Lab TR.*

Hunter, R. S. (1975). *The measurement of appearance,* New York: Wiley.

Kajiya, J. T. (1986). The rendering equation. *ACM Computer Graphics, 20,* 143–150.

Klinker, G. J., Shafer, S. A. & Kanade, T. (1988). The measurement of highlights in color images. *International Journal of Computer Vision, 2,* 7–32.

MacRobert, T. M. (1967). *Spherical harmonics. An elementary treatise on harmonic functions with applications* (3rd ed.). New York: Pergamon.

Mingolla, E. & Todd, J. T. (1986). Perception of solid shape from shading. *Biological Cybernetics, 53,* 137–151.

Nicodemus, F. Σ., Richmond, J. C., Hsia, J. J., Ginsberg, I. W. & Limperis, T. (1977). Geometrical considerations and nomenclature for reflectance. *NBS Monograph 160.* Washington, D.C.: National Bureau of Standards.

Pentland, A. P. (1982). Finding the illuminant direction. *Journal of the Optical Society of America, 72,* 448–455.

Pentland, A. P. (1984). Local shading analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6,* 170–187.

Pentland, A. P. (1988). Shape information from shading: a theory about human perception, *MIT Media Lab Visual Science TR,* 103.

Phong, B. T. (1975). Illumination for computer generated pictures. *Communications of the ACM, 18,* 311–317.

Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1988). *Numerical recipes in C. The art of scientific computing.* New York: Cambridge.

Sanderson, A. C., Weiss, L. E. & Nayar, S. K. (1988). Structured highlight inspection of specular surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 10,* 44–55.

Shafer, S. A. (1984). Using color to separate reflection components. *Department of Computer Science TR, 136.* New York: University of Rochester.

Swarztrauber, P. N. (1979). On the spectral approximation of discrete scalar and vector functions on the sphere. *SIAM Journal of Numerical Analysis, 16,* 934–949.

Talman, J. D. (1968). *Special functions: A group theoretic approach.* New York: W. A. Benjamin.

Terras, A. (1985). *Harmonic analysis on symmetric spaces and applications I.* New York: Springer.

Todd, J. T. & Mingolla, E. (1983). Perception of surface curvature and direction of illumination from patterns of shading. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 583–595.

Tominaga, S. & Wandell, B. A. (1989). The standard surface reflectance model and illuminant estimation. *Journal of the Optical Society of America A, 6,* 576–584.

Torrance, K. E. & Sparrow, E. M. (1967). Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America, 57,* 1105–1114.

Walker, J. S. (1988). *Fourier analysis.* New York: Oxford.

Witkin, A. P. & Tenenbaum, J. M. (1983). On the role of structure in vision. In J. Beck, B. Hope & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 481–543). New York: Academic.

Wyszecki, G. & Stiles, W. S. (1982). *Color science. Concepts and methods, quantitative data and formulae* (2nd ed.). New York: Wiley.