International Journal of Computer Vision 34(2/3), 193–204 (1999)

When Shadows Become Interreflections

M.S. LANGER
NEC Research Institute, 4 Independence Way, Princeton, NJ 08540

Received August 7, 1997; Revised June 12, 1998

Abstract. Shadows and interreflections are present in all real scenes and provide a rich set of photometric cues for vision. In this paper, we show how shadows and interreflections are intrinsically related. Shadows tend to occur in those parts of a scene in which interreflections have the largest gain. We provide several basic results concerning this relationship in terms of the interreflection modes of a scene. We show that for a given scene, the interreflection mode having the largest gain is a physically realizable radiance function. We derive bounds on the gain of this mode and discuss how this mode is related to shadows. We analyze how well an n-bounce model of interreflections approximates an infinite-bounce model and how shadows affect this approximation. Finally, we introduce a novel method for inferring surface color in a uni-chromatic scene. The method is based on the relative contrast of the scene in different color channels.

Keywords: shading, shadows, interreflection, illumination, color

1. Introduction

Shadows and interreflections are present in all real scenes and provide a rich set of cues for vision. These cues greatly enhance the realism of synthetic imagery. This has been demonstrated convincingly in computer graphics (Foley et al., 1990), and indicates that the human visual system makes use of these cues. In computer vision, however, shadows and interreflections have been poorly exploited. Most photometric methods do not consider shadows or interreflections, and those that do have considered one of the two cues only. That is, methods based on shadows have ignored interreflections (Waltz, 1975; Shafer, 1985; Belhumeur et al., 1997) and methods based on interreflections have ignored shadows (Nayar et al., 1991; Funt and Drew, 1993). In real scenes, both shadow *and* interreflections are present (Forsyth and Zisserman, 1991; Haddon and Forsyth, 1998) and thus both types of cues should be used.

In this paper, we take several steps toward the goal of using both shadows and interreflections, by showing how the two cues are intrinsically related. The main observation that motivates our analysis is that shadows tend to occur in those parts of a scene in which interreflections have the largest gain. To illustrate, consider the two examples in Fig. 1. On the left is a spherical concavity embedded in a ground plane. The parameter ϕ is the angle between the vertical direction and the direction marking the boundary of the concavity, measured from the center of the sphere. Observe that the smaller the angle ϕ, the less of the hemispheric sky above the ground plane will be visible from within the concavity and so the more shadowing will tend to occur within the concavity (Langer and Zucker, 1994). Similarly, the smaller the angle ϕ, the greater will be the number of interreflection bounces within the concavity since each bounce will be more likely to remain within the concavity than to exit from it. Together these two observations imply that the smaller the angle ϕ, the more shadows *and* interreflections tend to occur within the concavity.

The same line of reasoning holds for the wedge concavity on the right, and it holds even when the wedge surface has mirror-like reflectance. In this case, any ray of light that enters the concavity will bounce from

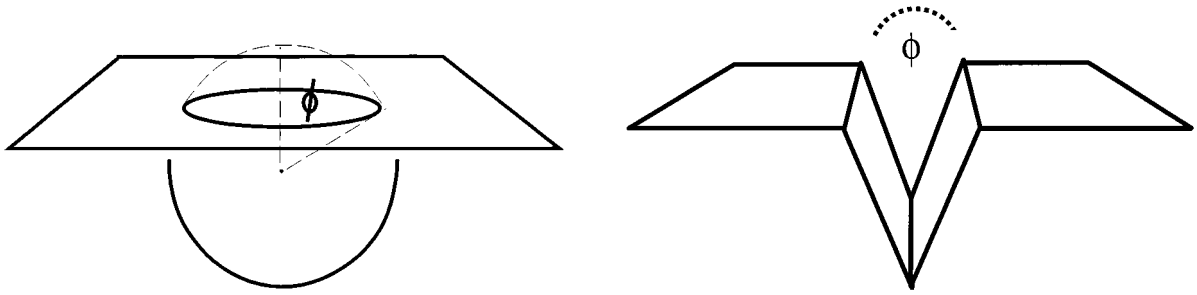

Figure 1. (Left) A spherical concavity. $\phi = 0$ is a spherical enclosure; $\phi = 90$ degrees is a hemispheric concavity; $\phi = 180$ degrees is a ground plane. (Right) A wedge concavity.

one side of the concavity to the other a finite number of times before exiting. On each bounce, the horizontal component of the ray's direction will reverse, and the vertical component will increase towards the exit. The smaller is ϕ, the smaller will be the change in the vertical component, thus the greater will be the number of interreflections.

In this paper, we provide several basic results concerning this general relationship of shadows and interreflections. We focus on the case of a Lambertian scene of uniform albedo.

2. Outline

In Section 3, we provide the background. We review the radiosity equation which is the classical infinite-bounce model of interreflections (Siegel and Howell, 1981; Foley et al., 1990). We also review the theory of interreflection modes, introduced by Moon (1940) and developed further by Kœnderink and van Doorn (1983). In Sections 4 and 5, we consider two example scenes and calculate their interreflection modes. In Section 6, we discuss contrast reversal and how it can cause shadows to appear as interreflections. In Section 7, we introduce the principal interreflection mode of a scene. This is the interreflection mode having the largest gain. We show that this mode is a physically realizable radiance function and we derive bounds on the gain of this mode. We also show how this mode is related to classical "ambient light" term for rendering interreflections in computer graphics. In Section 8, we examine the errors that occur when using an n-bounce model of interreflections to approximate the infinite-bounce model. We also examine how shadows affect these errors. Finally, in Section 9 we consider how color appearance is affected by shadows and interreflections. We present a novel method for inferring surface color in a uni-chromatic scene based on relative contrast in different RGB channels.

3. Background

Interreflections are complicated because they involve an infinite recursion. If one surface reflects light to a second, then the second must reflect light back to the first, and so on. The classical model of interreflections is the radiosity equation (Siegel and Howell, 1981; Kœnderink and van Doorn, 1983; Forsyth and Zisserman, 1991; Nayar et al., 1991; Foley et al., 1990). Consider a scene composed of Lambertian surfaces of uniform albedo ρ. Let $I(\mathbf{x})$ be the radiance from surface point $\mathbf{x}$ and let $I_{\text{direct}}(\mathbf{x})$ be the component of radiance that is due to direct illumination from the source only. The *radiosity equation* is

$$I(\mathbf{x}) = I_{\text{direct}}(\mathbf{x}) + \rho \int_{\mathcal{S}} I(\mathbf{y})\, K(\mathbf{x}, \mathbf{y})\, d\mathbf{y}. \qquad (1)$$

The integral is over the set of all surface points $\mathcal{S}$ in the scene. $K(\mathbf{x}, \mathbf{y})$ is a symmetric weighting function that depends on the relative orientation and distances of $\mathbf{x}$ and $\mathbf{y}$, and on whether the two points are visible from each other.

Moon (1940) observed that interreflections could be analyzed in terms of the eigenvectors of the kernel $K(\mathbf{x}, \mathbf{y})$. (This model was developed further by Kœnderink and van Doorn (1983).) These eigenvectors have a natural physical interpretation. Each defines a linear function space that is invariant to interreflections. As such, the eigenvectors are called *interreflection modes* (Kœnderink and van Doorn, 1983). The

modes $I_m(\mathbf{x})$ are defined by the equation:

$$\mu_m\, I_m(\mathbf{x}) = \int_{\mathcal{S}} I_m(\mathbf{y})\, K(\mathbf{x}, \mathbf{y})\, d\mathbf{y}, \tag{2}$$

where μ_m is the mth eigenvalue and $I_m(\mathbf{x})$ is the mth eigenvector. Each eigenvalue has absolute value less than or equal to one. The set of eigenvectors forms an orthonormal basis for the space of well-behaved functions defined on the surface $\mathcal{S}$.

For a given lighting condition, the direct radiance may be decomposed as a sum of these eigenvectors:

$$I_{\text{direct}}(\mathbf{x}) = \rho \sum_{m=0}^{\infty} h_m I_m(\mathbf{x}).$$

The coefficients of direct radiance, h_m, are typically non-zero for all m. One can show (Kœnderink and van Doorn, 1983) that the net radiance (direct + indirect) is

$$I(\mathbf{x}) = \rho \sum_{m=0}^{\infty} \frac{h_m}{1 - \rho\, \mu_m} I_m(\mathbf{x}). \tag{3}$$

We see that the effect of interreflections is to multiply the mth component of direct radiance by a *gain*,

$$\text{gain}_m = \frac{1}{1 - \rho\, \mu_m}.$$

The gain is an increasing function of the eigenvalue μ_m and of the albedo ρ. Note that when either the eigenvalue or albedo is near zero, the gain is near unity and so interreflections have little effect. When both the eigenvalue and the albedo are near one, the gain is large and interreflections have a large effect.

Interreflection modes for a number of simple scene geometries are calculated in (Moon, 1940; Kœnderink and van Doorn, 1983). No attempt was made, however, to relate the modes to the direct component of radiance. That is our goal in this paper.

4. Example 1: The Spherical Concavity Family

The spherical concavity family shown in Fig. 1 is defined by a single parameter ϕ. When ϕ is zero, the scene is a hollow sphere. When ϕ is 90 degrees, the scene is a hemispherical concavity. When ϕ is 180 degrees, the scene is a ground plane. The interreflection modes of this family of surfaces are quite simple.

Proposition 1. *A spherical concavity of unit radius has only one non-zero eigenvalue*,

$$\mu_0 = \frac{1}{2}(1 + \cos\,\phi). \tag{4}$$

The corresponding eigenvector, $I_0(\mathbf{x})$, *is constant within the concavity and zero outside the concavity.*

Proof: See (Moon, 1940) and (Moon and Spencer, 1981), p. 165. □

It follows that, for a given spherical concavity and for a given lighting condition, interreflections add a constant radiance to points within the concavity but add no radiance to points on the ground plane. For example, consider a uniform hemispheric source above the ground plane, having unit radiance. This source approximates the diffuse lighting from the sky on an overcast day. Under such lighting, the direct radiance is constant within the concavity (Moon and Spencer, 1981),

$$I_{\text{direct}} = \frac{\rho}{2}(1 - \cos\,\phi). \tag{5}$$

Observe that I_{direct} is an increasing function of ϕ while the eigenvalue μ_0 (and hence the gain) is a decreasing function of ϕ. This is the inverse relationship between shadows and interreflections that we discussed intuitively in Section 1.

5. Example 2: A Sphere on a Disk

A second example that illustrates the relation between shadows and interreflections is a convex object, such as a sphere, resting on the ground. This example is generic in the sense that it consists of two smooth surfaces making first order contact. This occurs often in natural scenes. Since no analytic solution is known for the interreflection modes of this scene, we solve for the modes numerically. We do so by considering a unit sphere centered on a disk of radius two. The two interreflection modes illustrated in Fig. 2 are those for which the eigenvalues are largest in absolute value. Note that we use a different notation (**I** vs. I) for the discrete case than the continuous case (see Appendix A for more details on the discretization.)

A few observations may be made immediately. First, the region of support for these two modes is spatially concentrated near the point of contact between

Figure 2. Two interreflection modes $\mathbf{I}_0$ and $\mathbf{I}_1$ for a sphere resting on a disk. (Black indicates negative values, white indicates positive, and grey indicates zero.) The mode on the left has the largest eigenvalue of all modes ($\mu_0 = 0.83$) and the mode on the right has the smallest eigenvalue ($\mu_1 = -0.83$). The functions $\mathbf{I}_0$ and $\mathbf{I}_1$ are identical in absolute value. $\mathbf{I}_0$ is positive everywhere, and $\mathbf{I}_1$ is positive on the sphere but negative on the disk. The next few interreflection modes (not shown) have similar eigenvalues. These modes differ from the first two by having a sinusoidal variation (of period 2π radians) as a function of longitude angle.

the sphere and the disk, since any illumination that reaches these regions is likely to undergo multiple reflections. Second, the mode having the largest eigenvalue is non-negative everywhere and thus corresponds to a physically realizable radiance function. We will see in Section 7 that this property holds for a general scene.

Next consider Fig. 3 which shows a scatter plot of $\mathbf{I}_0$ versus $\mathbf{I}_{\text{direct}}$ under uniform diffuse lighting. An inverse relationship is clearly present, which is consistent with the idea that shadows tend to occur in the parts of a scene where interreflections have the largest gain. $\mathbf{I}_{\text{direct}}$ and $\mathbf{I}_0$ are not perfectly anti-correlated, however. That is, the photographic negative of the direct radiance is not identical to the interreflection mode of largest gain. Taking the negative does yield an interesting perceptual effect, however, and we consider this effect in the following section.

6. Contrast Reversal and the Perception of Glow

The shading in a photographic negative does not correspond to natural shading. Yet the visual system often

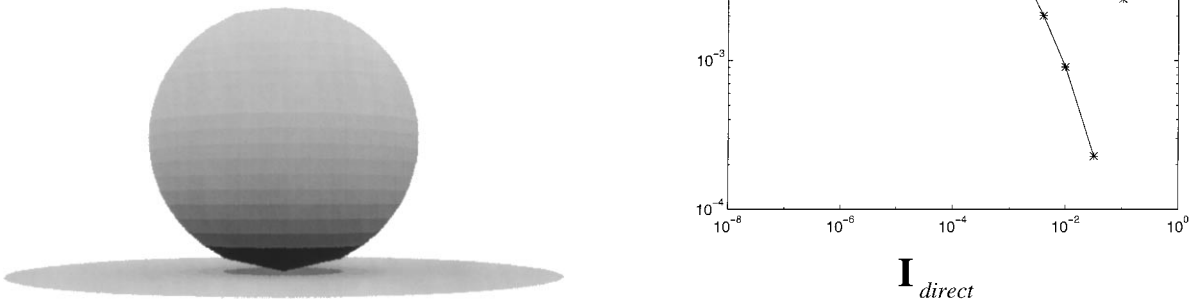

Figure 3. (Left) The direct component of radiance $I_{\text{direct}}(\mathbf{x})$ under uniform diffuse lighting. (Right) A scatter plot of $\mathbf{I}_{\text{direct}}$ vs. $\mathbf{I}_0$. The two clusters correspond to the sphere and disk. Note that there is an inverse relationship between $\mathbf{I}_{\text{direct}}$ and $\mathbf{I}_0$.

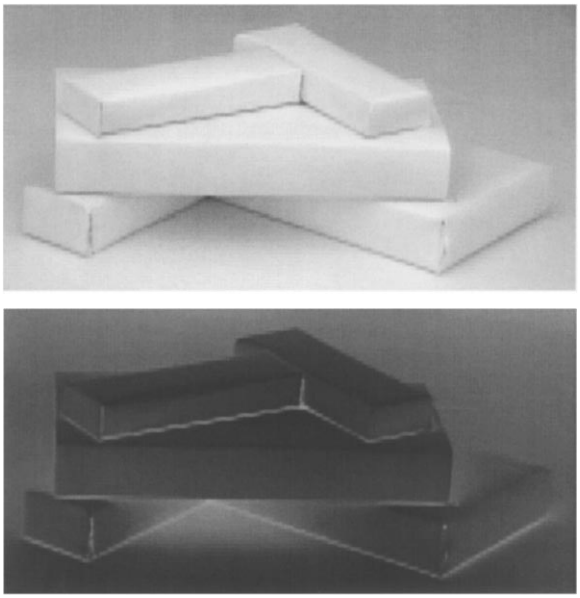

Figure 4. A blocks world scene under diffuse lighting and its photographic negative.

imposes an interpretation of this shading nonetheless (Anstis, 1992). A common illusion is that surfaces in a negative appear to glow. This occurs, for example, in the negative image of a face, where the nostrils, the mouth, and the pupils of the eye appear to glow. Another example is the blocks world image and its negative in Fig. 4.

How can we account for this illusion of glow in negatives? First, observe that contrast reversal does not alter the positions and orientations of lines and edges. These are geometric features, and inform the visual system about object boundaries and scene layout. By why the appearance of glow? We believe that the answer may be related to interreflection modes. A basic fact about interreflections is that when a light source is placed inside a concavity, the concavity will become bright because the modes of a concavity have high gain. When an image contains lines and edges that indicate a certain region of a scene is a concavity and when this concavity is observed to be bright relative to its surroundings, the visual system may infer that a light source is present within the concavity, i.e. that the interreflection modes of the concavity have been stimulated by a source. Note that the inferred source need not be visible in the image for the inference to be valid since a concavity may be partially occluded.

7. The Principal Interreflection Mode

In the two examples of Sections 4 and 5, the interreflection mode having the largest eigenvalue was a non-negative function. Thus it is physically realizable. One can show that this property holds for a general scene.

Proposition 2. *For a given scene, the interreflection mode having the largest eigenvalue is a non-negative function.* (*See Appendix B for proof.*)

We refer to this mode, $I_0(\mathbf{x})$, as the *principal mode* of the scene, and its eigenvalue, μ_0, as the *principal eigenvalue*. In this section, we study the basic properties of this mode.

First, the principal mode need not be unique. For example, a ground plane containing a number of disjoint spherical concavities of identical ϕ will have several principal modes, one for each concavity. Any linear combination of these principal modes will also be a principal mode. Typically though, the principal mode for a scene will be unique.

We may derive bounds on the principal eigenvalue as follows. These bounds help us to understand interreflections in extreme scene geometries such as a low-relief surface or an enclosure.

Proposition 3.

$$\min_{\mathbf{x}} \int_{\mathcal{S}} K(\mathbf{x}, \mathbf{y})\, d\mathbf{y} \leq \mu_0 \leq \max_{\mathbf{x}} \int_{\mathcal{S}} K(\mathbf{x}, \mathbf{y})\, d\mathbf{y}.$$

To interpret these bounds, note that for a given surface point $\mathbf{x} \in \mathcal{S}$, we have

$$\int_{\mathcal{S}} K(\mathbf{x}, \mathbf{y})\, d\mathbf{y} \equiv \int_{\mathcal{S}} \mathbf{N}(\mathbf{x}) \cdot \mathbf{L}\, d\Omega,$$

where the unit direction $\mathbf{L}$ points from $\mathbf{x}$ to some other surface point $\mathbf{y} \in \mathcal{S}$, and $\mathbf{L}$ is contained in the infinitesimal solid angle $d\Omega$.

An example of the upper bound on μ_0 is a scene having low relief. For each point on a low relief surface, the fraction of the hemisphere of incident directions in which other surfaces are visible is near zero. The upper bound tells us that μ_0 is near zero for a low relief surface. Hence the gain from interreflections will be near unity and interreflections will have little effect.

An example of the lower bound is an indoor room having only a small window. For all surface points

within the room, the angle subtended by the window is near zero and hence the fraction of the hemisphere of incident directions in which other surfaces are visible is near one. The lower bound tells us that μ_0 is near one. If the albedo is also near one then the gain will large and interreflections will have a large effect.

Corollary 1 (Ambient light). *Let a scene be completely enclosed, that is,* $\int_S K(\mathbf{x}, \mathbf{y})\, d\mathbf{y} = 1$ *for all* $\mathbf{x}$. *Then* $\mu_0 = 1$ *and* $I_0(\mathbf{x})$ *is constant over* $\mathbf{x}$.

This corollary (also stated in (Moon, 1940)) provides insight into a classical method used in computer graphics for rendering interreflections, in which a constant "ambient" intensity is added to each pixel. For an enclosure, this method amounts to considering the principal interreflection mode only. The method is insufficient for accurate renderings since modes other than the principal mode may have large gains. This is the underlying reason why radiosity methods (Foley et al., 1990), which implicitly consider all modes, have yielded more realistic renderings than those based on the ambient term alone.

8. The Accuracy of *n*-bounce Models

While it is widely accepted that interreflections affect scene appearance significantly, little is known about how significant the effects actually are. For example, many techniques in computer vision such as shape from shading or photometric stereo do not consider interreflections at all, and others that do consider interreflections account for one interreflection bounce only (Funt and Drew, 1993; Oren and Nayar, 1994). The accuracy of these approximations is rarely discussed.

In this section we address this question directly by asking how well an n-bounce model of interreflections approximates the radiosity equation. By an n-bounce model we mean the following. A 1-bounce model accounts for direct illumination only. A 2-bounce model accounts for direct illumination plus one interreflection. A 3-bounce model accounts for direct illumination plus two interreflections, and so on. Formally, an n-bounce model is obtained by truncating a Taylor series expansion of the gain to n terms, so that Eq. (3) is approximated by

$$I^{(n)}(\mathbf{x}) \equiv \sum_{m=0}^{\infty} \rho\, h_m (1 + \mu_m \rho + \cdots (\mu_m \rho)^{n-1})\, I_m(\mathbf{x}). \tag{6}$$

It is clear that for many scenes, the geometric series will converge for small n, for example, when the albedo is small or when the scene has low relief so that all eigenvalues are near zero. When both the albedo and relief are large, however, the high order terms in the Taylor series may be large as well. Truncating these terms could lead to large approximation errors.

We examine the truncation errors using two measures:

$$\text{absolute error}\,(\mathbf{x}) \equiv \frac{I(\mathbf{x}) - I^{(n)}(\mathbf{x})}{\max_{\mathbf{x}} I(\mathbf{x})}$$

$$\text{relative error}\,(\mathbf{x}) \equiv \frac{I(\mathbf{x}) - I^{(n)}(\mathbf{x})}{I(\mathbf{x})}.$$

The absolute error is the percent error at $\mathbf{x}$, measured relative to the maximum radiance in the scene. The relative error is the percent error at $\mathbf{x}$, measured relative to the radiance at $\mathbf{x}$.

To simplify the analysis, we consider the spherical concavity family of Section 4. We also assume that the light source is a sum of point sources at infinity. In this case, the absolute error does not depend on $\mathbf{x}$ since the concavity has only one non-zero eigenvalue and the corresponding eigenvector is constant. The relative error will depend on $\mathbf{x}$, however. For example, suppose an oblique point source at infinity were to produce a shadow within a concavity. The relative error for a 1-bounce model would be 100 percent within the shadow, but less than 100 percent within the non-shadowed region of the concavity. To avoid this dependence on $\mathbf{x}$, we consider the special case of uniform diffuse source since in this case the direct component of radiance is constant within a given spherical concavity (recall Eq. (5)) and so the relative error is also constant.

The errors for the spherical concavity family and the uniform diffuse source are shown in Fig. 5. The left column shows iso-contour plots of the absolute errors for a n-bounce model for $n = 1, 2, 3$, as a function of the albedo ρ and the angle ϕ. The right column shows iso–contour plots of the relative errors. Both types of error clearly decrease with n, as one expects. However, the absolute errors are typically quite small. For a hemispheric concavity ($\phi = 90$ degrees), a 1-bounce model and an albedo of 0.5 yield an absolute error of only 15%, and this error falls to 8% for a 2-bounce model. For deep concavities (small ϕ) and large albedos, absolute errors are near zero. For a given albedo ρ, the absolute error is maximized at an angle ϕ that is

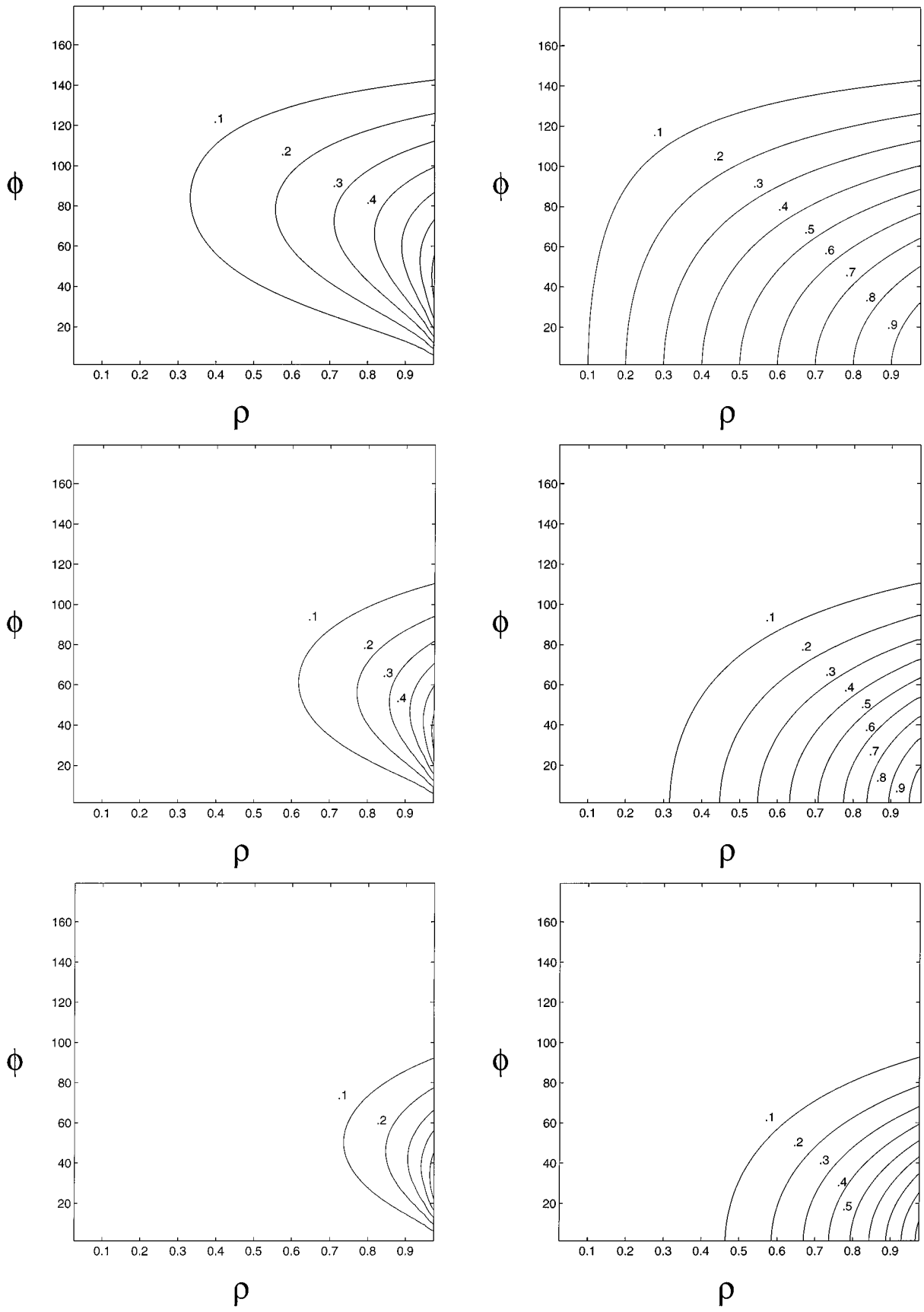

Figure 5. A iso-contour plot of the *absolute errors* (left) and *relative errors* (right) of an n-bounce model, where $n = 1$ (top row), 2 (middle row), 3 (bottom row), for the spherical concavity family. ρ is the albedo and ϕ is the angle shown in Fig. 1 (left).

strictly greater than zero degrees. The result is surprising since it contradicts a common intuition that interreflection effects will be greatest within the deepest concavities. We find rather that as ϕ tends to zero, the decrease in direct radiance outweighs the gain from interreflections, and so the indirect radiance decreases as a concavity becomes very deep. Note that this effect does not hold for the relative error which increases monotonically as ϕ tends to zero.

A few remarks should be made concerning the generality of the results. First, as mentioned above, the absolute errors for the spherical concavity family do not require a uniform diffuse source. Any sum of point sources at infinity will produce identical error curves. Indeed the errors will be similar for a point source that is not at infinity but only a few multiples of the concavity radius away. If the point source is within the concavity, however, then the errors curves no longer apply. For example, consider a spherical concavity that is illuminated by an isotropic point source at the center of the sphere and suppose I_{direct} is unity. The net radiance within the concavity would be

$$1 + \mu_0\rho + (\mu_0\rho)^2 + \cdots = \frac{1}{1 - \mu_0\rho}.$$

A 1-bounce interreflection model would yield an absolute error of $\rho\mu_0$, and a 2-bounce model would yield an absolute error of $\rho^2\mu_0^2$. For a concavity that is enclosed or nearly enclosed ($\mu_0 \approx 1$) and that has a high albedo, these absolute errors are quite large.

Do these errors generalize to other concavity shapes? For example, recall the family of finite-depth wedge concavities shown in Fig. 1, i.e. wedges that have been dug out of a ground plane. We compare these finite-depth wedges to infinite-depth wedges, which are defined by two intersecting infinite planes. One can show that, for any n and for any ϕ, the n-bounce models for the finite- and infinite-depth wedges are identical in the limit as $\mathbf{x}$ goes to the deepest point(s) of the wedge. Hence the n-bounce errors are also identical. This is useful, since the eigenvalues and eigenvectors for the infinite-depth wedge are easily computed. The principal eigenvalue is identical to that of the spherical concavity family for ϕ's defined as in Fig. 1, and so Eq. (4) applies exactly (for proof, see Moon, (1940), p. 202)). Moreover the principal eigenvector is constant as in the case of the spherical concavity. Thus for uniform diffuse lighting and for a given ϕ, the n-bounce errors are identical to those for a spherical concavity, and so the error curves of Fig. 5 are identical as well. As mentioned above, this identity holds only in the limit of points at the bottom of the concavity.

To summarize, we have two surprising results about the accuracy of n-bounce models. First, low-bounce models produce only small absolute errors, even for surfaces of high albedo. The reason is that interreflections tend to occur within shadowed regions and so the gain from interreflections acts on a small direct radiance only. One qualifier is that the result holds only when the source is well-outside the concavity. It does not hold for indoor lighting. Second, the indirect radiance (the absolute error of a 1-bounce model) vanishes as the concavity becomes very deep. This contradicts a common intuition that interreflection effects are largest in deep concavities such as cracks. That intuition is correct only for the relative error, not for the absolute error. Results were presented for the spherical concavity and wedge concavity families only, but we expect similar error curves for other concavity shapes.

9. Color From Relative Contrast

Until now we have presented results on how shadows and interreflections are related, but we have not yet specified how to use these results to solve vision problems. In this section, we use the results to motivate a novel method for inferring surface color.

To introduce the problem, we begin with an observation commonly made by skiers that, under diffuse lighting, images of snow-covered hills have low visual contrast. This low-contrast effect is evident mathematically for the spherical concavities of Section 4. If a surface has an albedo near unity and the source is diffuse, then the radiance of the concavity will be nearly equal to that of the ground plane. A similar contrast effect occurs for more general surfaces as well, as we will see below.

Interestingly, there is evidence that the human visual system uses this contrast cue to infer surface albedo. In one experiment, (Gilchrist and Jacobsen, 1984) observers were asked to judge a pair of achromatic miniature rooms containing multiple objects such that the two rooms were identical geometrically but in one room all surfaces were painted black and in the other room the surfaces were painted white. Observers correctly judged the white room to be lighter shade of paint than the black room, even though the mean radiance in the two rooms was controlled using sources of different power.

A followup experiment extended this idea to colored rooms. Observers were shown one room that was painted red and illuminated by white light, and a second room that was painted grey and illuminated by a red light. Again the majority of observers could distinguish which room was which (Gilchrist and Ramachandran, 1993). For the colored rooms, it was argued that observers might have used the color saturation within the shadows as a cue. The shadows in the red painted room were a more saturated red than shadows in the white painted room because of multiple red interreflections in the red painted room. (The fact that interreflections can affect room color is well-known in lighting design. See for example (Spencer and Sanborn, 1961).)

Strictly speaking, the saturation cue would only hold under white light. A more general cue could be based on RMS contrast. For example, consider the following (for a proof, see Appendix B).

Proposition 4. *Given an enclosure of constant albedo and an arbitrary lighting condition, the root mean square (RMS) contrast of the scene radiance decreases monotonically with the albedo.*

To see how to apply this fact, consider a camera with three color channels (RGB) having non-overlapping spectral sensitivity. A single RGB image would then be identical to three grey-level images of three scenes having identical geometry but different albedos. (In this sense, the black and white experiment described above is formally equivalent to a color experiment with two channels.) Proposition 4 implies that the RMS contrast within a given color channel is inversely related to the surface spectral reflectance over wavelengths of that channel. (We assume for simplicity that the surface spectral reflectance is constant over the wavelengths defined by a given color channel.) The method for inferring surface spectral reflectance is simply to examine the contrast within each of the channels, and infer that the order of the reflectances in the three channels is opposite to the order of the contrasts.

9.1. Experiment

To test this method, we performed the following experiment. We measured the RMS contrast of RGB images of various uni-chromatic surfaces, made of matte construction paper. Each surface was crumpled and illuminated by a grid of fluorescent panel lights in the ceiling. The images were taken using an NEC-NX18 color video CCD camera and the camera's gamma response was set to unity. We verified that the response was approximately linear by comparing images taken with and without a neutral density filter.

Table 1 shows the albedos $\vec{\rho}$ and RMS contrasts $\vec{\kappa}$ for various crumpled surfaces and various illuminants.

Table 1. RGB albedo $\vec{\rho}$ and contrast $\vec{\kappa}$ for three different illuminants $\vec{\mathbf{e}}$.

		R	G	B	R	G	B	R	G	B
Illuminant	$\vec{\mathbf{e}}$	150	130	190	95	170	124	180	125	130
White (standard)	$\vec{\rho}$	100	100	100	100	100	100	100	100	100
	$\vec{\kappa}$	15	15	13	17	14	15	17	20	19
Pink	$\vec{\rho}$	100	41	53	111	50	58	102	31	43
	$\vec{\kappa}$	29	48	43	32	39	41	28	60	59
Green	$\vec{\rho}$	**27**	**63**	**41**	27	54	38	29	67	43
	$\vec{\kappa}$	**39**	**32**	**34**	40	29	31	49	38	41
Yellow	$\vec{\rho}$	92	95	12	91	89	18	91	93	−9
	$\vec{\kappa}$	29	32	119	25	27	59	22	24	89
Off-white	$\vec{\rho}$	88	91	73	90	88	77	91	90	68
	$\vec{\kappa}$	32	33	36	23	23	25	25	27	32
Red	$\vec{\rho}$	57	14	15	59	22	22	58	12	5
	$\vec{\kappa}$	38	63	71	29	38	42	20	42	67
Blue	$\vec{\rho}$	19	23	56	23	25	48	19	26	62
	$\vec{\kappa}$	50	50	41	50	46	39	52	50	43

(The vectors $\vec{\rho}$ and $\vec{\kappa}$ will be defined shortly.) The illuminants were varied by adjusting the white balance of the camera, rather than adjusting the illuminant spectra. For a given setting of the white balance, the illuminant $\vec{\mathbf{e}}$ was defined as the within-channel mean intensity of an image of a flat piece of white paper. (Thus the illuminant was a 3-vector rather than an entire spectral distribution.) This white paper served as a standard.

For a given surface material and a given illuminant, the albedo $\vec{\rho}$ was defined by taking the within-channel ratios of the mean intensity of flattened surface (i.e. uncrumpled) image to the mean intensity of the flattened white standard image. By this definition, the albedo of the white standard was 100 per cent in all three channels. Moreover, the albedo of another surface could be greater than 100 per cent in a channel since the white standard was not totally reflective. Indeed, the pink paper had higher reflectance than the white standard within the R channel. Finally, the albedo $\vec{\rho}$ could (and generally) did vary with different illuminants $\vec{\mathbf{e}}$. This is a result of the sensors having broad spectral sensitivity.

The RMS contrast $\vec{\kappa}$ was defined for each channel and each image as the ratio of the standard deviation of the intensities to the mean intensity. We subtracted the dark noise from each pixel before computing the albedo and the RMS contrast. The dark noise was the intensity with lens cap on.

The data are shown in Table 1 and are roughly as predicted. For a given illuminant and a given surface, the color channel in which the albedo was highest was typically the channel in which the RMS contrast was lowest, and vice-versa. For example, consider the data (shown in boldface) for source $\vec{\mathbf{e}} = (150, 130, 190)$ and for the green paper. The albedo $\vec{\rho}$ is largest in the G channel (63 per cent) and smallest in the R channel (27 per cent). The contrast has the opposite trend. It is smallest in the G channel (32 per cent) and largest in the R channel (39 per cent).

A few caveats about the method and data should be mentioned. First, the method is only valid for light colored surfaces, that is, surfaces whose reflectance is above say 20 percent (middle gray on the Munsell scale) in at least one of the channels. For dark surfaces, interreflections will have little effect on surface radiance and so the contrast will be roughly the same in all three color channels. For example, the method cannot distinguish a dark red surface under white light from a dark grey surface under reddish light. In our experiments, we only used light colored surfaces.

Further caveats concern the experimental apparatus. The camera response was not perfectly linear, nor was the dark noise simply an additive constant. For example, when the yellow paper was illuminated by the source $\mathbf{e} = (180, 125, 130)$, the mean intensity in the B channel was lower than the dark noise! (This is why the albedo is shown as negative in the table for that case.) Third, the sensors had a broad spectral sensitivity. This could cause problems since the method assumes that the spectral reflectance is constant over the wavelengths to which the sensor is sensitive. The contrast effects that we obtained were presumably affected by all three of these factors.

9.2. Previous Work

The idea of using shadows and interrreflections to infer surface color is not entirely new. For example, Ruben and Richards (1982) posed the problem of detecting whether an image edge was due to an illumination or to a material discontinuity. To solve the problem, they made a grey world assumption that the spectrum of the "ambient light" was the same as that of the direct light. Gershon et al. (1990) improved on that method by allowing for chromatic ambient light. Both of these methods assume that ambient light is an additive constant, of course, and we have discussed limitations of this assumption in Section 7.

Funt and Drew considered a more realistic model of interreflections and focussed on a specific geometry: a wedge concavity (Funt et al., 1991; Funt and Drew, 1993). They assumed that the two sides of a given wedge are of different material and that a 2-bounce interreflection model is valid. With these assumptions they were able to provide constraints on the spectral reflectance of the surfaces using interreflection cues. A related psychophysics experiment by Kersten and Hurlbert showed that human observers can partially discount the interreflections within a wedge concavity, at least for the case that one of the sides of the wedge is white (Kersten and Hurlbert, 1996). Finally, Nayar and Gong (1992) extended the shape-from-interreflections method of Nayar et al. (1991) to color images. Nayar's method assumes the lighting conditions are given, however, and that no shadows are present within the concavity. It is unclear how to generalize that method to a situation in which shadows are present, e.g. under diffuse lighting.

10. Conclusion

Broadly speaking, the visual world contains three types of illumination phenomena. First, there is classical

shading which depends on the surface orientation with respect to a collimated illuminant and on the relative distance between a surface patch and the illuminant. Second, there is shadowing which depends on the visibility of the light source from different surface points. Third, there are interreflections between surfaces. These illumination phenomena arise canonically under most lighting conditions and scene geometries, and each is present to some extent in every scene.

Over the past two decades there has been great progress both in computer graphics and computer vision in going beyond classical shading and in addressing shadows and interreflections as well. Our contribution in this paper is to articulate several basic properties of shadows and interreflections, in particular, how they are related. We have focussed on the case of a Lambertian surface of uniform albedo and we have shown that a wide range of appearance phenomena emerge for this case. In future work, we hope to broaden our analysis and to study richer scene domains in which surfaces have non-uniform and non-Lambertian reflectance. In these domains, the relationships between the three types of illumination phenomena remain largely unexplored.

Appendix A

In this appendix we discretize the radiosity equation and compute the interreflection modes for the ball on a disk example of Section 5.

We approximate the radiosity equation by discretizing the surfaces of the scene into a set of N polygons (Siegel and Howell, 1981; Foley et al., 1990), and assume that the radiance is constant over each polygon. Let $\mathbf{F}_{ij}$ denote the fraction of the radiant flux from polygon j that is received by polygon i. Equation (1) is approximated by the *discrete radiosity equation*:

$$\mathbf{I} = \mathbf{I}_{\text{direct}} + \rho\,\mathbf{F}\,\mathbf{I}. \tag{7}$$

The vectors $\mathbf{I}_{\text{direct}}$ and $\mathbf{I}$ are the direct radiance and net radiance, respectively. The jth element of $\mathbf{I}$ is the radiance from polygon j.

For simplicity, we require all N polygons to have the same area. In that case, $\mathbf{F}$ is a symmetric $N \times N$ matrix and has a complete, orthogonal set of eigenvectors, defined by

$$\mu_m\,\mathbf{I}_m = \mathbf{F}\,\mathbf{I}_m. \tag{8}$$

For a given lighting condition, the direct component of radiance is a sum of the eigenvectors:

$$\mathbf{I}_{\text{direct}} = \rho \sum_{m=1}^{N} h_m\,\mathbf{I}_m,$$

and the net radiance is

$$\mathbf{I} = \rho \sum_{m=1}^{N} \frac{h_m}{1-\rho\,\mu_m}\,\mathbf{I}_m. \tag{9}$$

For the sphere on a disk example, we tile the sphere and disk using 800 polygons each, so that $N = 1600$. To accurately compute the form factors, $\mathbf{F}_{ji}$, we use a recursive subdivision algorithm. For a given pair of polygons, we consider whether the distance between the polygons is large relative to the size of the polygons (Baum et al., 1989). If so, then the form factor can be easily computed using a "point mass" approximation. If not, then the polygons are subdivided, and the test is repeated recursively.

Appendix B

Proof of Proposition 2: We consider the discrete case only (see Appendix A). Let $\mathbf{I}_0$ be the mode having the largest eigenvalue, and suppose without loss of generality that $\mathbf{I}_0$ is strictly positive at polygon j. (We assume there is a unique mode having the largest eigenvalue, though the proof can easily be extended to the case of multiple modes having the largest eigenvalue. In such a case, the modes would have disjoint regions of support.) Let $\mathbf{p}_j$ be the $N \times 1$ vector that is 1 at polygon j and 0 elsewhere. Then $\mathbf{I}_0 \cdot \mathbf{p}_j > 0$. Since μ_0 is the largest eigenvalue, for any $\mathbf{I}$ satisfying $\mathbf{I} \cdot \mathbf{I}_0 > 0$ we must have

$$\lim_{n\to\infty} \frac{1}{\mu_0^n}\,\mathbf{F}^n\,\mathbf{I} = \mathbf{I}_0.$$

This limit depends on $\mathbf{F}$ being symmetric and hence having a complete set of eigenvectors. (Recall from Appendix A that we can ensured the symmetry of $\mathbf{F}$ by choosing the polygons to all have the same area.) Take $\mathbf{I} \equiv \mathbf{p}_j$. Since all elements of $\mathbf{F}$ are non-negative, we must have $\mathbf{F}^n\,\mathbf{p}_j \geq 0$ for all j and for any n. This is true in the limit as $n \to \infty$ as well, and hence $\mathbf{I}_0 \geq 0$. □

Proof of Proposition 3: To derive the upper bound, observe that

$$\mu_0\,I_0(\mathbf{x}) \leq \left\{ \max_{\mathbf{y}}\; I_0(\mathbf{y}) \right\} \int_{\mathcal{S}} K(\mathbf{x},\mathbf{y})\,d\mathbf{y}.$$

Suppose that $\max_{\mathbf{y}} I_0(\mathbf{y})$ is attained at $\mathbf{x}^*$. Then

$$\mu_0 \, I_0(\mathbf{x}^*) \leq I_0(\mathbf{x}^*) \int_{\mathcal{S}} K(\mathbf{x}^*, \mathbf{y}) \, d\mathbf{y}.$$

Dividing both sides by $I_0(\mathbf{x}^*)$ completes the proof. □

To derive the lower bound, let $\alpha = \min_{\mathbf{x}} \int K(\mathbf{x}, \mathbf{y}) \, d\mathbf{y}$. Define a lighting condition such that $I_{\text{direct}}(\mathbf{x}) = 1$ for all $\mathbf{x}$. In this case, one can show that $I(\mathbf{x}) \geq (1 - \rho\alpha)^{-1}$ for all $\mathbf{x}$. However, since μ_0 is the largest eigenvalue, one can also show that $I(\mathbf{x}) \leq (1 - \rho\mu_0)^{-1}$ for all $\mathbf{x}$. Putting these two inequalities together yields the lower bound.

Proof of Proposition 4: Recall Corollary 1 which states that, for an enclosure, the principal mode $I_0(\mathbf{x})$ is constant over $\mathbf{x}$ and the principal eigenvalue μ_0 is unity. The mean radiance of $I(\mathbf{x})$ over the scene is thus $\frac{h_0}{1-\rho}$. By definition, the RMS contrast is the standard deviation of the scene radiance divided by the mean radiance, and so

$$\text{RMS contrast} = \frac{\sqrt{\sum_{i>0} \left(\frac{h_i}{1-\rho\mu_i} \right)^2}}{\frac{h_0}{1-\rho}}.$$

It is easy to show that expression decreases with the albedo ρ. □

Acknowledgments

Thanks to John Haddon, David Jacobs, John Oliensis and the reviewers for many helpful comments and suggestions.

References

Anstis, S. 1992. Visual adaptation to a negative brightness reversed world: Some preliminary observations. *Neural Networks for Vision and Image Processing*. The MIT Press.

Baum, D.R., Rushmeier, H.E., and Winget, J.M. 1989. Improving radiosity solutions through the use of analytically determined form factors. In *Proc. SIGGRAPH*.

Belhumeur, P.N., Kriegman, D.J., and Yuille, A. 1997. The bas-relief ambiguity. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, San Juan, PR.

Foley, J.D., van Dam, A., Feiner, S.K., and Hughes, J.F. 1990. *Computer Graphics: Principles and Practice*. Addison-Wesley: Reading, MA, 2nd edition.

Forsyth, D. and Zisserman, A. 1991. Reflections on shading. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:671–679.

Funt, B.V. and Drew, M.S. 1993. Color space analysis of mutual illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(12):1319–1326.

Funt, B.V., Drew, M.S., and Ho, J. 1991. Color constancy from mutual reflection. *International Journal of Computer Vision*, 6(1):5–24.

Gershon, R., Jepson, A.D., and Tsotsos, J.K. 1990. Ambient illumination and the determination of material changes. *Journal of the Optical Society of America*, 7(10):2041–2047.

Gilchrist, A. and Jacobsen, A. 1984. Perception of lightness and illumination in a world of one reflectance. *Perception*, 13:5–19.

Gilchrist, A.L. and Ramachandran, V.S. 1993. Red rooms in white light look different than white rooms in red light. In *Invest. Oph. and Vis. Sci. (ARVO abstract)*.

Haddon, J. and Forsyth, D. 1998. Shading primitives: Finding folds and shallow grooves. In *Proceedings of the Sixth International Conference on Computer Vision*, Bombay, India.

Kersten, D.K. and Hurlbert, A.C. 1996. Discounting the color of mutual illumination: A 3-d-shape-induced color phenomenon. In *Invest. Oph. and Vis. Sci. (abstract)*.

Kœnderink, J.J. and van Doorn, A.J. 1983. Geometrical modes as a general method to treat diffuse interreflections in radiometry. *J. Opt. Soc. Am.*, 73(6):843–850.

Langer, M.S. and Zucker, S.W. 1994. Shape from shading on a cloudy day. *Journal of the Optical Society of America A*, 11(2):467–478.

Moon, P. 1940. On interreflections. *Journal of the Optical Society of America*, 30:195–205.

Moon, P.H. and Spencer, D.E. 1981. *The Photic Field*. MIT Press: Cambridge, MA.

Nayar, S.K. and Gong, Y.G. 1992. Colored interreflections and shape recovery. In *Image Understanding Workshop*, San Diego, DARPA, pp. 333–343.

Nayar, S.K., Ikeuchi, K., and Kanade, T. 1991. Shape from interreflections. *International Journal of Computer Vision*, 6:173–195.

Oren, M. and Nayar, S.K. 1994. Seeing beyond lambert's law. In *ECCV '94*. Lecture Notes in Computer Science 801, Springer-Verlag, pp. 269–280.

Rubin, J.M. and Richards, W.A. 1982. Color vision and image intensities: When are the changes material? *Biological Cybernetics*, (45):215–226.

Shafer, S.A. 1985. *Shadows and Silhouettes in Computer Vision*. Kluwer Academic Publishers.

Siegel, R. and Howell, J.R. 1981. *Thermal Radiation Heat Transfer*. Hemisphere Publ. Corp.

Spencer, D.E. and Sanborn, S.E. 1961. Interreflections and color. *Journal of the Franklin Institute*, 252:413–426.

Waltz, D. 1975. Understanding line drawings of scenes with shadows. *The Psychology of Computer Vision*, McGraw-Hill: New York, pp. 19–91.