13

Surface color perception and light field estimation in 3D scenes

LAURENCE T. MALONEY,
HOLLY E. GERHARD, HUSEYIN BOYACI,
AND KATJA DOERSCHNER

13.1   The light field

The spectral power distribution of the light emitted by the Sun is almost constant. The variations in daylight (Figure 13.1) that we experience over the course of a day and with changes in seasons are due to the interaction of sunlight with the Earth's atmosphere (Henderson, 1977). The resulting spectral distribution of daylight across the sky is typically spatially inhomogeneous and constantly changing (Lee and Hernández-Andrés, 2005a,b). The light arriving at each small patch of surface in the scene depends in general on the patch's location and orientation in the scene. Furthermore, objects in the scene create shadows or act as secondary light sources, adding further complexity to the light field (Gershun, 1936/1939; see also Adelson and Bergen, 1991) that describes the spectral power distribution of the light arriving from every direction at every point in the scene. The light field captures what a radiospectrophotometer placed at each point in the scene, pointing in all possible directions, would record (Figure 13.2).

When the light field is inhomogeneous, the light absorbed and reradiated by a matte[1] smooth surface patch can vary markedly with the orientation or location of the patch in the scene.

In Figure 13.3, for example, we illustrate the wide range of the light emitted by identical rectangular achromatic matte surfaces at many orientations,

---

[1] We use the term "matte" as a synonym for "Lambertian," a mathematical idealization of a matte surface (Haralick and Shapiro, 1993).
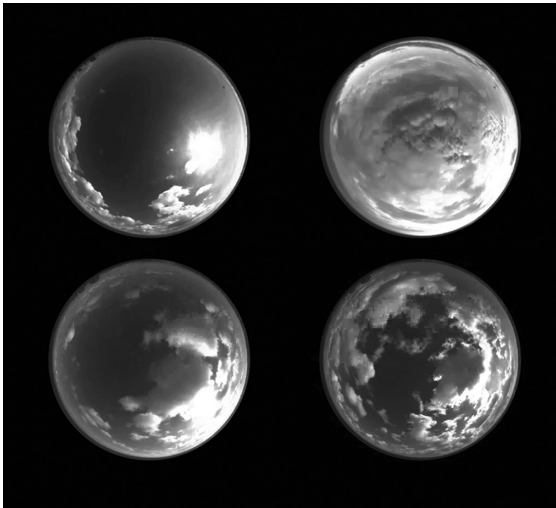
Figure 13.1 Terrestrial daylight. Four views of the sky over Los Angeles. Courtesy of Paul Debevec. A color version of this figure can be found on the publisher's website (www.cambridge.org/ 9781107001756).
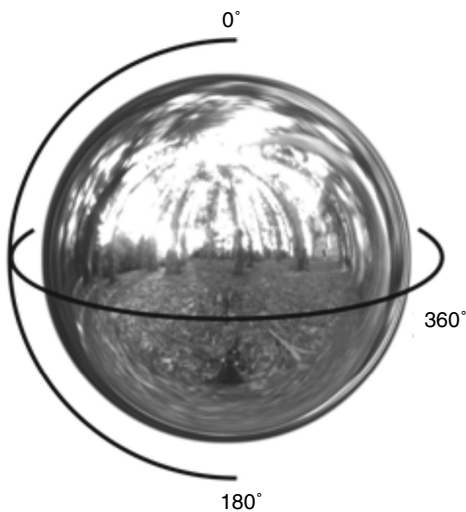

Figure 13.2 The light field. The light field is the spectral power distribution of light arriving from every possible direction at every point in the scene. For one wavelength and one location, the light field can be represented as a sphere as shown. Courtesy of Paul Debevec.

Figure 13.3 The effect of orientation. Identical rectangular matte patches at different locations and orientations, rendered under a distant neutral, collimated source placed along the line of sight. The luminance of each patch is proportional to the cosine of the angle between the surface normal and the direction towards the collimated source (Lambert's law) (Haralick and Shapiro, 1993).

illuminated by a distant neutral, collimated light source. Any visual system designed to estimate surface color properties (including lightness) is confronted with a new problem to solve with each change of surface orientation or location.

To arrive at a stable estimate of surface reflectance, the visual system has to discount the effects of the light field on a patch. In many scenes, discounting spatial variation in the light field is an underconstrained problem (Adelson and Pentland, 1996; Belhumeur *et al.*, 1999; Dror *et al.*, 2004). Moreover, detecting changes in the current light field and distinguishing them from changes in objects within the scene is itself a potentially difficult problem for the visual system. In this review, we describe recent research concerning surface color estimation, light field estimation, and discrimination of changes in the light field from other changes in the scene.

13.1.1   *The Mondrian singularity*

Previous research in color vision has typically avoided addressing the problems introduced by spatial and spectral variation in the light field, by the choice of scenes. These scenes, consisting of flat, coplanar matte surfaces,

have been referred to as Mondrians (Land and McCann, 1971). In such scenes, observers can accurately make a variety of judgments concerning surface color and lightness. Arend and Spehar (1993a,b) showed that observers were close to constant in estimating the lightness of a matte surface embedded in a two-dimensional Mondrian despite changes in illumination. Foster and Nascimento (1994) and Nascimento and Foster (2000) showed that observers can reliably distinguish whether the change in appearance of a Mondrian is due to an "illumination" change or a reflectance change, and that this could be explained by a model based on cone-excitation ratios. Bäuml (1999) showed that observers are capable of constant estimation of the color of surfaces in Mondrians following changes in illumination, and that his results could be well accounted by using the von Kries principle, which is a simple linear transformation of cone responses. However, these studies need to be extended for two reasons.

First, there is no obvious way to generalize these results to the normal viewing conditions of our ever-changing, three-dimensional world. In the flat, two-dimensional world of Mondrians, no matter how complex the light field is, the light emitted from a surface contains essentially no information about the spatio-spectral distribution of the light incident upon the surface (Maloney, 1999). A matte surface absorbs light from all directions in a hemisphere centered on its surface normal and then reemits uniformly in all directions a fraction of the total light absorbed: a matte surface "forgets" where the light came from. In previous work, we have called this phenomenon the Mondrian singularity (Boyaci *et al.*, 2006a).

A second reason why it is important to consider a wider range of stimuli in evaluating human color perception is that three-dimensional scenes can convey considerable information about the light field in a scene. Maloney (1999) noted that there are potential "cues to the illuminant" in three-dimensional scenes that signal illuminant chromaticity. Here we consider recent work directed towards determining what cues signal how the intensity and chromaticity of the illumination incident on a matte surface vary with surface orientation and location in three-dimensional scenes. We also consider recent studies of human ability to estimate the light field and to discriminate changes in the light field from changes in the actual contents of a scene, including the surface colors of objects in the scene.

13.1.2   *Illuminant cues*

The perception of surface color in Mondrian scenes is an intrinsically difficult problem. In order to estimate surface color accurately, the visual system must estimate the net intensity and chromaticity of the light incident on the Mondrian. The typical approach taken is to develop simple measures of the

central tendency, variance, and covariance of the photoreceptor excitations and use them as a basis for estimating the light intensity and chromaticity. "Gray world" algorithms (for reviews, see Hurlbert (1998) and Maloney (1999)), for example, use the mean chromaticity of the scene as an estimate of the chromaticity of the light. Mausfeld and Andres (2002) have conjectured that means, variances, and covariances contain all of the information used by the visual system in estimating surface color. Golz and Macleod (2002) and Macleod and Golz (2003) concluded that correlations between the chromaticities and luminance values of surfaces contained useful information about the chromaticity of the effectively uniform illumination of a Mondrian scene, but this conclusion has been challenged by recent work (Ciurea and Funt, 2004; Granzier *et al.*, 2005). These measurements, based on simple moments (mean, variance, and covariance) of distributions, eliminate what little spatial structure is present in the Mondrian. It is not clear that simple moments[2] derived from Mondrian scenes convey any information about the chromaticity of the illuminant or its intensity (Maloney, 1999), and they convey no information about spatial and spectral inhomogeneities in the light field.

When a scene is not restricted to flat, coplanar, matte surfaces arranged in a Mondrian, more information about the chromaticity of the illuminant (Maloney, 1999; Yang and Maloney, 2001) and the spatial and spectral distribution of the light field may be available to the observer. Researchers have shown that human observers are able to judge lower-order estimates of the light field such as diffuseness and mean illumination direction (te Pas and Pont, 2005; Pont and Koenderink 2004). We emphasize that any deviation from "flat" or "matte" in an otherwise Mondrian stimulus could disclose information about the light field, and we refer to these sources of information as illuminant cues (Kaiser and Boynton, 1996; Maloney, 1999; Yang and Maloney, 2001; Pont and Koenderink, 2003, 2004; Koenderink and van Doorn, 1996; Koenderink *et al.*, 2004). Illuminant cues, by definition, carry information about the illuminant. It is possible to develop algorithms that estimate the light field from such cues (see, e.g., Hara *et al.*, 2005) and Ramamoorthi and Hanrahan (2001a), but currently such algorithms depend upon restrictive assumptions about the scene and its illumination. Such algorithms are based on the physics of image formation, and, when they succeed, we can be sure they carry the desired information. The relevant question concerning an illuminant cue is whether it is used in human vision.

---

[2] Technically, any function of the retinal image is a statistic, and it is likely that the claim is vacuously true. In practice, researchers confine their attention to the moments of lowest degree of the photoreceptor excitations in the retinal image (e.g., Mausfeld and Andres, 2002), and we use the term "scene statistics" as a synonym for these moments.

In this chapter, we review recent work from a small number of research groups concerning how biological visual systems extract information about surfaces (albedo and color) and the flow of light in scenes outside the Mondrian singularity. We describe in more detail three sets of experiments, the first testing whether human observers can compensate for changes in surface orientation and examining what illuminant cues they may be using, and the second examining whether human observers can compensate for changes in surface location in scenes with a strong illuminant gradient in depth. In the third set of experiments, we assess human ability to estimate the light field and discriminate changes in the light field from other changes within a scene. The experimental results indicate that human color vision is well equipped to solve these apparently more complex problems in surface color perception outside the Mondrian singularity (for a review, see Maloney *et al.*, 2005).

13.2   Lightness and color perception with changes in orientation

Boyaci *et al.* (2003) investigated how human observers compensate for changes in surface orientation in binocularly viewed, computer-rendered scenes illuminated by a combination of neutral collimated[3] and diffuse light sources. The simulated source was sufficiently far from the rendered scene that it could be treated as a collimated source. The observer's task was to match the lightness (perceived albedo) of a test surface within the scene to a nearby lightness scale. The orientation of the test patch with respect to the collimated source was varied, and the question of interest was whether observers would compensate for test patch orientation, partially or fully. Previous work had found little or no compensation (Hochberg and Beck, 1954; Epstein, 1961; Flock and Freedberg, 1970; Redding and Lester, 1980) (see Boyaci *et al.* (2003) for details).

The methods and stimuli employed were similar to those of Boyaci *et al.* (2004), which we present in more detail below. In contrast to previous researchers, Boyaci *et al.* (2003) found that observers compensated substantially for changes in test patch orientation. Ripamonti *et al.* (2004) reached the same conclusions using scenes of similar design, composed of actual surfaces (not computer-rendered) viewed under a combination of collimated and diffuse light sources. The conclusion of both studies was that the visual system partially compensates for changes in surface orientation in scenes whose lighting model consists of a combination of a diffuse and a collimated source.

---

[3] For simplicity in rendering, collimated sources were approximated by point sources that were distant from the rendered scene. Elsewhere, we refer to these sources as "punctate." The difference is only terminological.

Boyaci *et al.* (2004) examined judgments of surface color in a similar experiment. The lighting model consisted of a distant collimated yellow light source ("sun") and a diffuse blue light source ("sky"). The test surface was illuminated by a mixture of the two that depended on the orientation of the test surface and the lighting model. The observer's task was to set the test patch to be achromatic (an achromatic-setting task). To do so, the observer first needed to estimate the blue–yellow balance of the light incident on the test patch, which was itself part of the spatial organization of a scene. Next, the observer needed to set the chromaticity of the light emitted by the surface to be consistent with that of an achromatic surface.

The collimated light source was simulated to be behind the observer at elevation $30°$ and azimuth $-15°$ (on the observer's left) or $15°$ (on the observer's right). The location of the light source remained constant during an experimental block. In every trial, each of four naive observers was presented with a rendered scene and asked to adjust a test surface to be achromatic. Scenes were rendered as a stereo image pair and viewed binocularly. A typical scene is shown in Figure 13.4. The test patch was always in the center and at the front
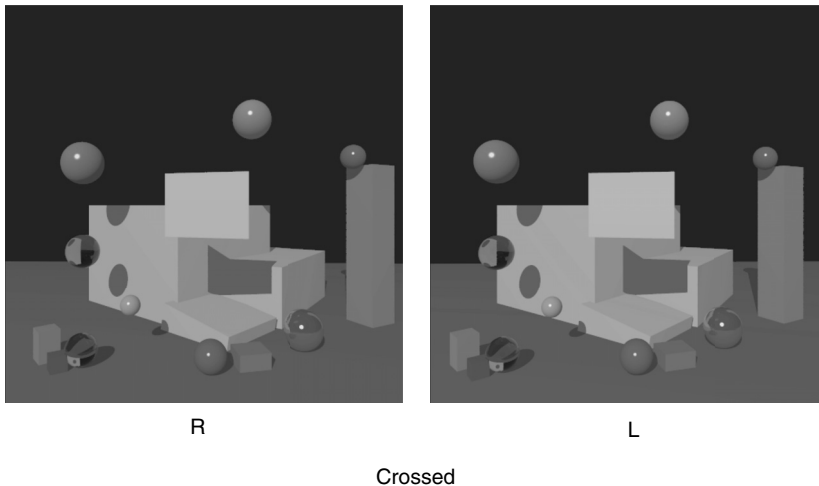

R   L

Crossed

Figure 13.4 A scene from Boyaci *et al.* (2004). Observers viewed rendered scenes binocularly. The two images permit crossed binocular fusion. The scenes were rendered with a combination of yellow collimated and blue diffuse light sources. The collimated source was always behind the observer, to the observer's left in half the trials and to the right in the remainder. The orientation of the test surface varied in azimuth and elevation from trial to trial. The observer's task was to set the test surface in the center of the scene to be achromatic. A color version of this figure can be found on the publisher's website (www.cambridge.org/9781107001756).

of the scene, and additional objects were provided as possible cues to the light field.

From trial to trial, the orientation of the test patch was varied in either azimuth or elevation, but not both. The dependent measure of interest was the relative blue intensity Λ^B in the observer's achromatic setting (for details, see Boyaci *et al.*, 2004). In theory, as the angle between the normal to the test surface and the direction to the yellow collimated light source increases, the observer should make achromatic settings that are "bluer." Boyaci *et al.* (2004) derived setting predictions for an ideal observer who chose achromatic settings that were color-constant, always picking the setting consistent with a test surface that was achromatic. These setting predictions are plotted in Figure 13.5a. There are two plots, one for the collimated light on the observer's left and one for the light on the observer's right. In each plot, the relative blue intensity Λ^B is plotted versus the azimuth of the test surface (solid curve) and versus the elevation (dashed curve). It is important to realize that each curve reaches a minimum when the test patch's orientation matches the direction of the yellow collimated light source.

The results are shown in Figure 13.5b for the subject closest to the ideal observer. All four subjects substantially discounted the effect of changes in orientation. Boyaci *et al.* (2004) were able to recover crude estimates of the light source direction from each observer's achromatic settings by estimating and plotting the minima of the four curves (Figure 13.6). There are four estimates of the azimuth (one for each observer) for the light source on the left (Figure 13.6a), and four for the light source on the right (Figure 13.6b). There are eight corresponding estimates of the elevation. The eight estimates of the elevation (Figure 13.6c) are within $10°$ of the true values.

The outcome of the experiment of Boyaci *et al.* (2004), together with the results of Boyaci *et al.* (2003) and Ripamonti *et al.* (2004), implies that the observer's visual system effectively develops an equivalent illumination model (Boyaci *et al.*, 2003) for a scene and uses this model to estimate the albedo and surface color of surfaces at different orientations. In order to do so, the visual system must use the cues present within the scene itself.

In a more recent experiment (Boyaci *et al.*, 2006a), we examined three possible "cues to the lighting model" that were present in the scenes described above: cast shadows, surface shading, and specular highlights. We asked the observers to judge the lightness of a rotating central test patch embedded in scenes that contained various cues to the lighting model. The methods and stimuli were similar to those in Boyaci *et al.* (2003) and the first experiment described above. We compared four conditions: the all-cues-present condition, where all three
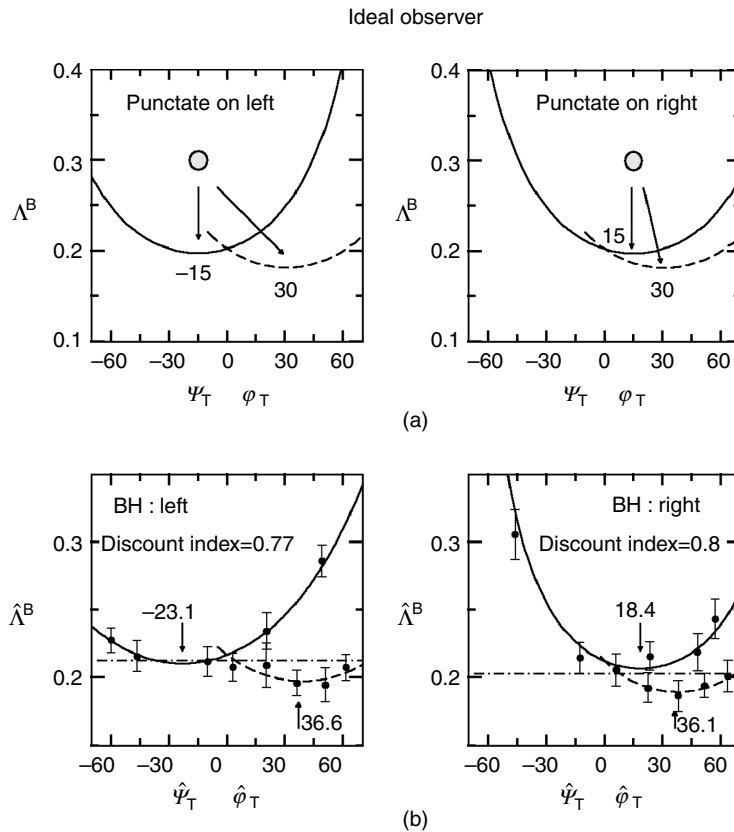
Figure 13.5 Achromatic-setting results from Boyaci *et al.* (2004). The dependent variable was the amount of blue (blue/total, or relative blue) in the observer's achromatic setting. (a) The settings for an ideal observer who perfectly compensated for changes in test patch orientation and collimated-source position. The left graph contains a plot of settings for trials where the collimated source was at 30° elevation and −15° azimuth (above and behind the observer, to his/her left). The right graph contains a plot of settings for trials where the collimated source was at 30° elevation and 15° azimuth (above and behind the observer, to his/her right). In both graphs, the horizontal axis is used to plot either the azimuth or the elevation of the test surface. The vertical axis is the relative blue intensity in the observer's settings. The solid curve in each graph signifies the settings of the ideal observer that compensate for changes in the test surface azimuth. The dashed curve in each graph signifies the settings of the ideal observer that compensate for changes in the test surface elevation. Note that both curves reach a minimum when the test surface is closest in azimuth or elevation to the "yellow" collimated source. (b) Settings of one observer, from Boyaci *et al.* (2004). The format is identical to that of (a). The lines through the data are based on an equivalent illumination model not described here (see Boyaci *et al.*, 2004).
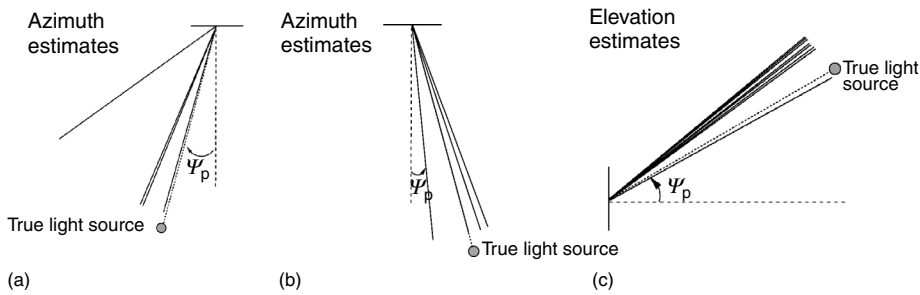
Figure 13.6 Estimates of collimated-source direction from Boyaci *et al.* (2004). For each observer, Boyaci *et al.* (2004) estimated the minimum of the achromatic-setting curves for each subject (Figure 13.4) and interpreted these as estimates of the collimated-light-source direction. The true value is plotted as a dashed line, and the observer's estimates as solid lines. (a) Azimuth estimates, collimated source at $-15°$ azimuth. (b) Azimuth estimates, collimated source at $15°$ azimuth. (c) Elevation estimates.

cues were present in the scene; the cast-shadows only condition; the shading-only condition; and the specular-highlights-only condition.

Boyaci *et al.* found that observers corrected for the test patch orientation in all four cue conditions, suggesting that they used each candidate cue in isolation. We also performed a reliability analysis to address to what extent observers combined the cues when all three cues were present (all-cues condition). The results of that analysis indicated that the reliability of the observers' settings in the all-cues condition was higher than for the best individual cue ("effective cue combination" (Oruc *et al.* (2003)); nevertheless, it was smaller than the reliability predicted by optimal cue combination rules.

In the next section, we describe two experiments where the orientation of the test surface never changed. Instead, there was a strong gradient of illumination in depth within the scene and the test surface was moved from a dimly lit region to a brightly lit region. The only difference between the two experiments was the presence of specular surfaces that served as candidate illumination cues. A comparison of performance in the two experiments revealed whether observers used these specular cues to estimate the light field.

13.3   Lightness perception with changes in location

In indoor scenes, the light field can vary markedly with location as walls serve to block or reflect incident light. The celebrated experiments of Gilchrist (1977, 1980) (see also Kardos, 1934) demonstrated that the visual system partially compensates for light fields that vary across space. Gilchrist *et al.*

(1999) proposed that observers segment complex scenes into illumination *frameworks* (Katz, 1935; Gilchrist and Annan, 2002) and discount the illumination (light field) within each framework. The rules for organizing frameworks and assigning surfaces to them are complex and not fully understood. The three-dimensional structure of a scene could also guide the segmentation of scenes into frameworks, and it is likely that there are analogous effects of three-dimensional organization on color perception (e.g., Bloj *et al.*, 1999; Doerschner *et al.*, 2004).

Ikeda *et al.* (1998) examined lightness perception in scenes comprising two small rooms arranged in depth with a doorway between them, patterned after Gilchrist (1977). The lighting of the rooms was complex, consisting of multiple fluorescent tubes placed above both rooms, and the observer could not see these light sources. The intensity of the light incident on a test surface placed along the line of sight through the doorway varied with depth. Ikeda *et al.* (1998) measured the apparent lightness for surfaces at different depths. Their observers viewed a test square placed at several different depths along the line of sight and passing in depth through the centers of both rooms. The observers' task was to match the test square to a lightness scale. Ikeda *et al.* found that observers substantially discounted the actual illumination profiles at different depths in the scene.

We next describe two experiments by Snyder *et al.* (2005) using rendered scenes similar in design to those of Gilchrist (1977) and Ikeda *et al.* (1998). All scenes were presented binocularly and consisted of two rooms arranged along the line of sight with walls composed of random, achromatic Mondrians. A top view of the simulated scenes is shown in Figure 13.7a. The far room was lit by two light sources not visible to the observer. The near room was lit by diffuse light only. The test surface (called a standard surface from this point onwards) varied in depth from trial to trial as shown. The observer adjusted an adjustable surface in the near room until the standard and adjustable surfaces seemed to "be cut from the same piece of paper." In the second experiment, a candidate cue was also added to the spatial distribution of the illumination: 11 specular spheres placed at random in the scene (but never in front of either the standard or the adjustable surface). The relative luminance of the light (with respect to the back wall of the far room) is plotted in Figure 13.7b. It varied by roughly a factor of five from the far room to the near room. An example of a scene with spheres (for Experiment 2) is shown in Figure 13.8. The scenes for Experiment 1 were similar but lacked specular spheres.

The results of Snyder *et al.* (2005) for five subjects, four naive and one (JLS) an author of the study, are shown in Figure 13.9. In both experiments, Snyder *et al.* estimated the ratio of the luminance of the standard surface to that of the adjustable surface (the relative luminance) at each location in the room. If
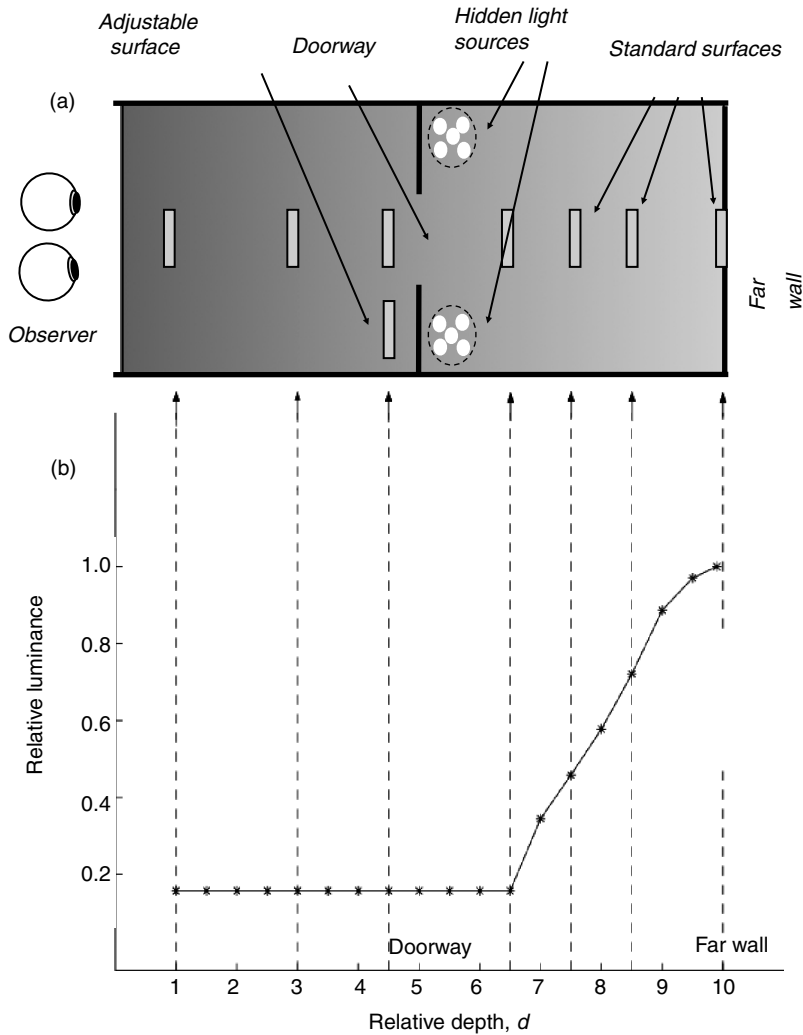
Figure 13.7 Schematic illustration of the scenes used by Snyder *et al.* (2005). (a) schematic top view of the scenes. (b) The actual relative illumination profile: the intensity of light incident on a matte surface perpendicular to the observer's line of sight as a function of depth.

the observers were lightness-constant, these settings would follow the relative-illumination profile in Figure 13.7, which is replotted with the results for each subject in Figure 13.9. The horizontal dashed line corresponds to the settings that would be chosen if the observers were simply matching luminance. The results for Experiment 1 are plotted with hollow circles, and those for Experiment 2 with filled circles.
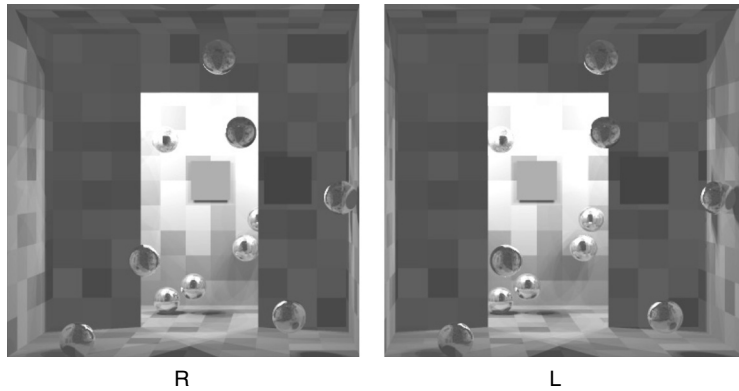
R   L

Figure 13.8 A scene from Snyder *et al.* (2005). Observers viewed rendered scenes binocularly. The two images permit crossed binocular fusion. The scenes were rendered with a combination of collimated and diffuse light sources. The collimated sources were in the far room behind the wall containing the doorway. The standard patch was in the center of the scene and moved from trial to trial along the observer's line of sight in depth. The adjustable patch was next to the doorway on the right. The observer adjusted the luminance of this patch until its lightness (perceived albedo) matched the lightness of the test patch.

Snyder *et al.* concluded that all observers significantly discounted the gradient of illumination in depth in both experiments and that their degree of constancy significantly improved with the addition of the specular spheres.

13.4   Representing the light field

Boyaci *et al.* (2004) found that observers could compensate for changes in surface orientation in scenes illuminated by a single yellow collimated light and a blue diffuse light. Boyaci *et al.* (2003) and Ripamonti *et al.* (2004) came to similar conclusions for asymmetric lightness matches in scenes illuminated by a single collimated source and a diffuse source. These results indicate that the observers' visual systems effectively estimated a representation of the spatial and chromatic distribution of the illuminant at the points in the scene probed. But what if there is more than one collimated light source? Doerschner *et al.* (2007) investigated whether the visual system can represent and discount more complex spatial and chromatic light fields.

The argument of Doerschner *et al.* (2007) is based on mathematical results due to Basri and Jacobs (2001) and Ramamoorthi and Hanrahan (2001a,b) that we will not reproduce here. The key idea is that a Lambertian surface effectively blurs the light field incident on it so that, for example, multiple collimated
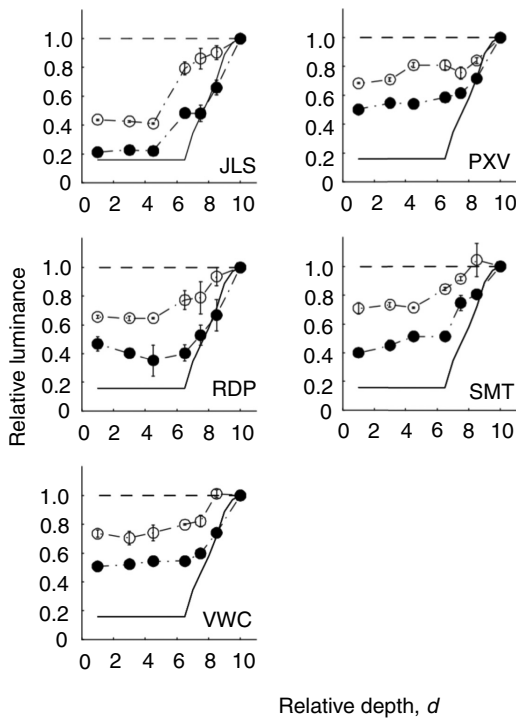
Figure 13.9 Results from Snyder *et al.* (2005). The relative luminance of the observers' lightness matches is plotted as a function of depth, with specular spheres (solid circles) and without (open circles). The actual relative-illumination profile is also included as a solid curve. An observer with perfect lightness constancy would have settings on this line. The horizontal dashed line signifies settings for an observer with no lightness constancy (luminance matching). The observers partially discounted the actual gradient of light intensity, with and without the specular spheres. With the specular spheres, their performance was markedly closer to that of a lightness-constant observer. The results suggest that the spheres act as cues to spatial variations in the light field.

sources that arrive from almost the same direction are equivalent to a single extended source with a single maximum of intensity. Surprisingly, even when the angle between the collimated sources is as great as 90°, they effectively merge into a single extended source. When the separation is as great as 160°, the effective light field has two distinct maxima. The goal of Doerschner *et al.* was to determine whether the human visual system could discriminate between these two configurations.

The stimuli were computer-rendered 3D scenes, containing a rectangular test patch at the center. Observers viewed the stimuli in a stereoscope. The scenes
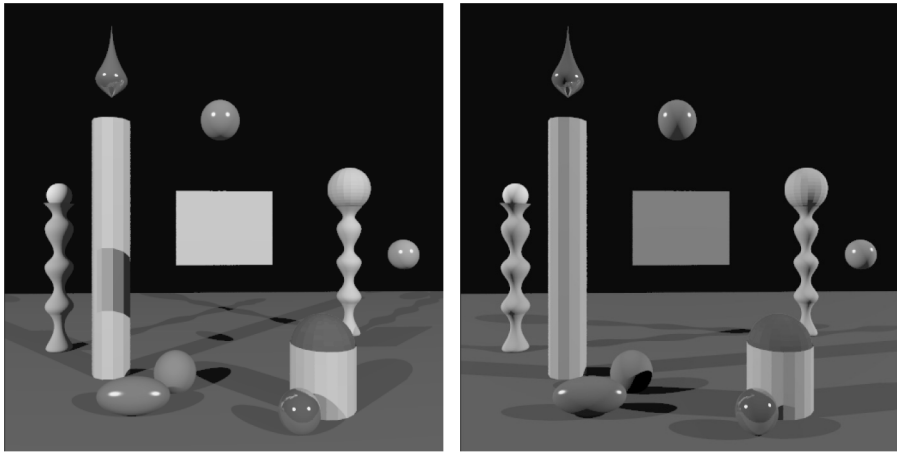
Figure 13.10 Example of scene illumination conditions from Doerschner *et al.* (2007). Only the left image of each stereo pair is shown. The scenes were illuminated by a combination of a diffuse blue source and two yellow collimated sources either 90° apart (left) or 160° apart (right).

were illuminated by a combination of a diffuse blue source and two yellow collimated sources placed symmetrically about the observer's line of sight and either 90° apart or 160° apart. Examples of the stimuli are shown in Figure 13.10.

A condition with a blue diffuse source and a single yellow collimated source was included as a control. The orientation of the test patch was randomly varied among nine orientations from −60° to 60°. In each trial, the observer adjusted the color of the test patch until it was perceived to be achromatic. We analyzed the amount of relative blue in the observers' achromatic settings as a function of test patch orientation (just as in Boyaci *et al.*, 2004). Six naive observers repeated each orientation-and-light condition 20 times. We fitted a generalization of the equivalent illumination model developed by Boyaci *et al.* (2003, 2004) (the model of Bloj *et al.* (2004) is essentially identical to that of Boyaci *et al.* (2003)) to predict the settings at each test patch orientation for an ideal observer with imperfect knowledge of the spatial and chromatic distribution of the illuminants.

The observers systematically discounted the relative contributions of diffuse and collimated light sources at the various test patch orientations for all illuminants. We conclude that the visual system effectively represents complex lighting models that include multiple collimated sources. Doerschner *et al.* (2007) argued further that the ability of the human visual system to discriminate the presence of multiple light sources is well matched to the task of estimating the surface color and lightness of Lambertian surfaces (Figure 13.11). We will not reproduce their argument here.
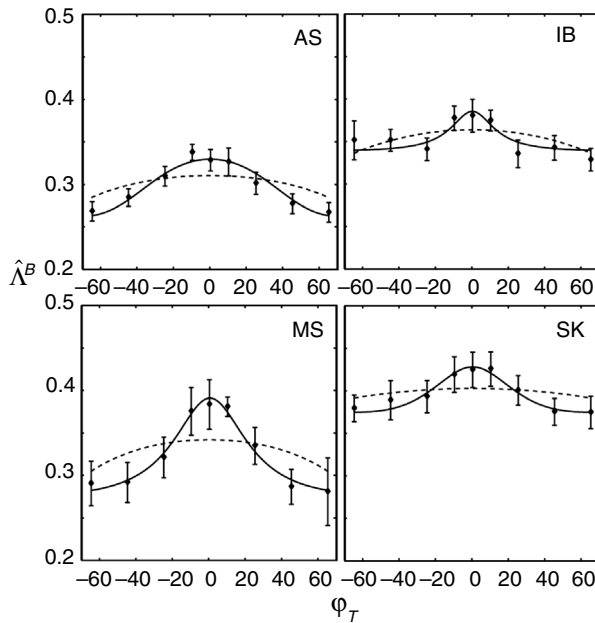
Figure 13.11 Data and model fits for the 160° condition from Doerschner *et al.* (2007). Λ^B is plotted as a function of the test patch orientation φ_T. The figure shows observers' data (diamond symbols). The error bars are plus or minus twice the standard error of the mean, which corresponds approximately to the 95% confidence interval. The figure illustrates clearly that the data are fitted better when the observer's equivalent illumination model is approximated with a 9D spherical harmonic subspace (solid line) than with a 4D harmonic subspace (dashed line), indicating that the visual system can resolve directional variation in the illumination up to at least a 9D subspace. All fits were obtained by means of maximum likelihood estimation, as described in Doerschner *et al.* (2007).

13.5   The psychophysics of the light field

The work discussed above demonstrates that observers discount the effects of the light field when interpreting the color or albedo of a surface, suggesting that the observers construct an internal representation of the light field, as described by Boyaci's *et al.*'s (2003) equivalent lighting model. In this section, we review work from two laboratories which have directly studied the psychophysics of the light field. First, we describe the work of Koenderink *et al.* (2007), who directly evaluated the human visual system's ability to estimate the light field in a novel experiment. Their results suggest that observers estimate accurately a light field filling the entire visual space. Second, we review work by Gerhard *et al.* (2007) on the temporal dynamics of light field inference, which
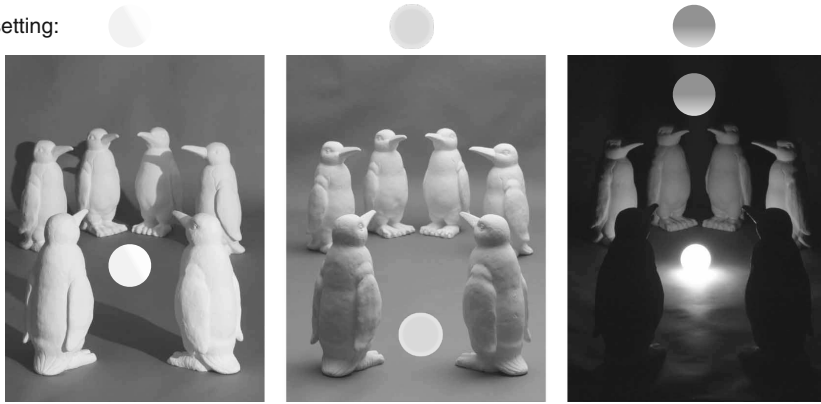
Figure 13.12 Light field estimation. Stimuli from Koenderink *et al.* (2007). Observers set the central sphere to agree with the local light field as inferred from the scene.

revealed that observers detect changes in the light field rapidly and accurately. Last, we show that observers effectively use this estimated light field to improve sensitivity to detecting concomitant changes in surface color.

Koenderink *et al.* (2007) took stereo photographs of real scenes containing matte-white-painted penguins facing each other, standing on a matte gray ground (Figure 13.12). Three lighting conditions were used: one simulating daylight with a distant collimated light source, one simulating an overcast day, and one simulating a "nativity scene" painting, in which the sole light source was at the feet of the group in the middle of the circle. A white matte sphere was also photographed at various positions, either floating or resting on the floor of the scene. During the experiment, observers viewed stereo photographs of the scene where the matte sphere was replaced by a virtual probe sphere with adjustable shading. The probe's shading started at a random setting, and the observer's task was to adjust four light field parameters until the sphere appeared to fit well into the scene. There were two position parameters, tilt and slant, and two quality parameters, directedness and intensity. Using four sliders, the observers spent typically one minute adjusting the lighting online until the sphere's shading looked correct. In order to produce these settings, it was necessary for the observers to infer the properties of the light field using only the penguins and the ground plane as cues to the spatial variance of the light field's intensity and then to infer how that light field would shade a novel object placed in the scene.
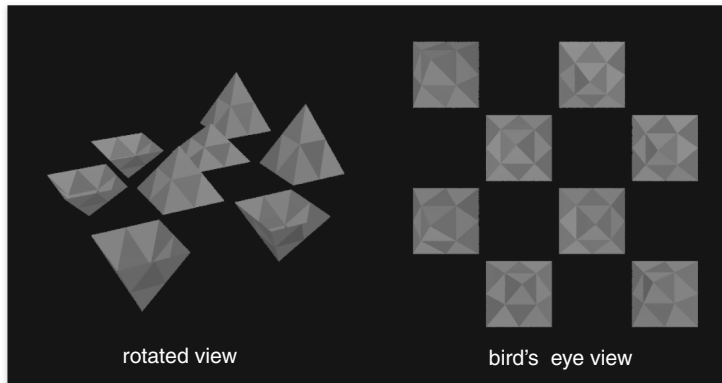
Figure 13.13 Stimuli from Gerhard *et al.* (2007) and Gerhard and Maloney (2008). Observers viewed rendered scenes of achromatic concave and convex pyramids floating in a black space. On the left is a rotated view illustrating the three-dimensional structure of the scene, and on the right is the bird's-eye view which the observer had through a stereoscope.

The observers' light field settings varied monotonically with the veridical values, with the observers being particularly keen at setting the light source tilt correctly (within 10° of the veridical tilt). The correlations between the images produced by the observers settings and the predicted probe images computed from the true lighting settings were quite high, with R^2 values for the regressions in the range 0.7–0.9. However, the comparison of observers' settings with the veridical light field parameters was not the most important part of the analysis of these results; the observers' settings need not reproduce the true lighting conditions, as the internal representations of the light field might be subject to systematic errors as other domains of visual processing are (Boyaci *et al.*, 2003; Ripamonti *et al.*, 2004).

The important result is that the observers were remarkably consistent both across sessions and with each other. The light source direction settings were both highly precise and reproducible across sessions. The light quality settings, i.e., directedness and intensity, were similar between observers and fairly reproducible. Koenderink *et al.*'s (2007) novel method confirmed that human observers can estimate the spatial variation of the light field across the entire visual space.

Gerhard *et al.* (2007) constructed rendered three-dimensional scenes which allowed precise control over image luminances in order to evaluate dynamic light field perception. The scenes were grayscale pyramids floating in a black space, viewed stereoscopically from above. Example stimuli are seen in Figure 13.13.

The sides of the pyramids were various shades of gray and were illuminated according to Lambert's cosine law by a distant collimated source and a diffuse source one-quarter as intense. Rendered trials were created where the collimated source rotated angularly in one of four directions by 2, 4, 6, or 8°, over two 1 s frames.

Importantly, three-dimensional shape was necessary to disambiguate the changes, as some pyramids were concave and others convex, determining the flow of shading as the light source moved. This "yoked" stimulus design was used in order to address the only previous work on light field change, in which changes in luminance ratios between adjacent surfaces were predictive of surface color changes and not light changes, which preserve edge ratios in flat-world scenes (Foster and Nascimento, 1994). However, in three-dimensional scenes, edge ratios vary as a light source rotates about the scene; see Figure 13.14 for an illustration.
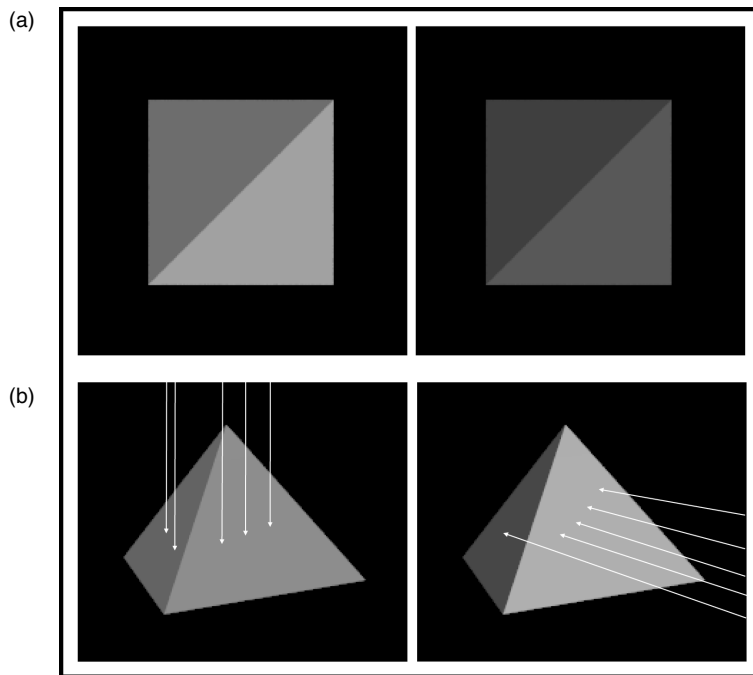

Figure 13.14 The effect of light field transformations on luminance edge ratios. (a) In a flat world with a homogeneous light field, an intensity or color change in the light field does not affect the ratio of luminance between two adjacent surfaces. (b) In a three-dimensional world in which the position of a point light source changes, the ratio of the luminances of adjacent surfaces changes.

For each light-transformation trial, we created a "yoked" nonlight trial, in which each pyramid's luminance values were retained in both frames, including the luminance ratios between adjacent edges, but the pyramids were rotated at random within the scene so that the resulting change in scene was not consistent with any possible change in the direction of the collimated light source.

In each trial, observers were first asked to judge whether the scene changes were consistent with a change in location of the collimated source or not. If they reported that the scene changes were consistent with a change in location of the collimated source, they were then asked to report the direction in which the light source had moved.

At the lowest magnitude of light source movement, 2°, two of the four observers were above chance in discriminating the globally consistent light change from the inconsistent version, and at 4°, all four observers were well above chance in discriminating the changes, with discriminability increasing with the magnitude of the angular rotation. Importantly, all observers were above chance in reporting the direction of the lighting change even in trials with the smallest magnitude of the change in lighting direction. Discriminability measures are plotted in Figure 13.15, and results for the ability to defect the movement direction of the light are plotted in Figure 13.16. The observers excelled at discriminating the changes even at low magnitudes, indicating that the human visual system is sensitive to small light-field-induced changes that disturb luminance edge ratios. The nature of the discrimination in this experiment required processing of the three-dimensional structure of the scene. Given only one of the two images in a stereo pair, it was not possible to accurately predict the direction of movement of the light source. This result demonstrates that light field perception cannot be modeled by image-based computations on a single image and cannot be modeled by algorithms making use only of changes in edge ratios.

In a second experiment, Gerhard and Maloney (2008) investigated whether an observer's ability to discriminate changes in the light field would aid the observer in detecting simultaneous changes in surface color. If, for example, the visual system can accurately estimate the changes in luminance due to a light field change, then it could potentially detect a further change in surface albedo more accurately.

In half of the trials chosen at random, Gerhard and Maloney added a surface albedo perturbation on top of the two global scene changes presented in their earlier work. Instead of detecting globally consistent or inconsistent light field changes, observers were asked to detect whether one facet of one pyramid at
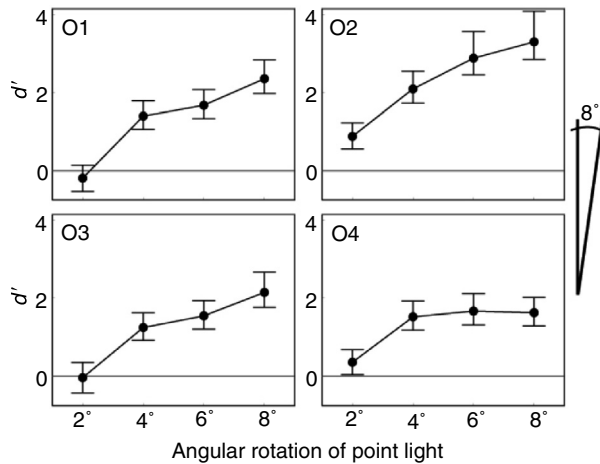
Figure 13.15 Light field discriminability results from Gerhard *et al.* (2007). The discriminability of light field changes from matched nonlight changes, measured as d' (Green and Swets, 1966), is plotted for each observer as a function of the degree of angular rotation that the collimated source underwent. The d' for each level of change magnitude was calculated from response data for a set of trials, half of which contained globally consistent light changes and the other half of which were statistically matched trials that did not contain a global light field change. Nonparametric bootstrapping was used to obtain 95% confidence intervals, and the results indicate performance above chance for all observers at all levels except the lowest level, at which observers 1 and 3 were at chance. On the right is an 8° angle for reference; it is the largest magnitude of light source rotation tested. A color version of this figure can be found on the publisher's website (www.cambridge.org/9781107001756).

a random location in the scene had changed albedo. In half of the trials, one facet did increase or decrease in luminance at random.

Trials were prepared in the same yoking fashion as before to preserve luminance edge ratio information, as well as all other low-order image statistics. If the visual system can more accurately detect surface changes simultaneously with a change in the position of the collimated light source than simultaneously with matched changes not consistent with a change in the position of the collimated source, then the d' measures should be higher than in the previous experiment.

The observers demonstrated an average benefit of 1.65 times higher albedo perturbation sensitivity when the global change could be perceived as light-field-induced. The improvement was significant ($t = 3.27$; $p < 0.001$). See Figure 13.17 for the perturbation sensitivity for each observer. These results
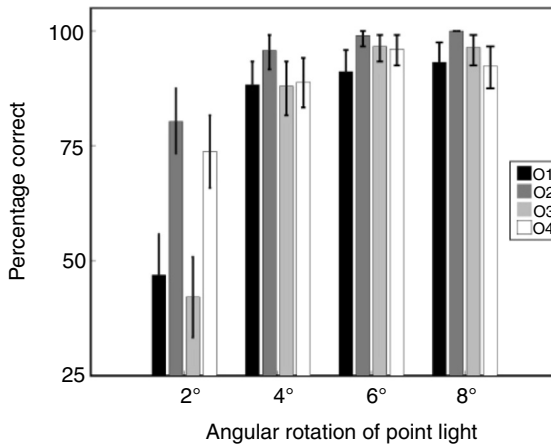
Figure 13.16 Light movement direction results from Gerhard *et al.* (2007). The percentage correct for each observer is plotted as a function of the degree of angular rotation of the light source. All observers were above chance, which was 25%, at discriminating the direction in which the light source had moved in trials in which they detected a true light source rotation.

suggest that the variability in the image luminances was effectively reduced when the observers perceived the global change as driven by a dynamic light field, and that the observers discounted the component of change due to the changing light field.

13.6   Conclusions

The world in which we live is three-dimensional. Claims about the usefulness of visual information should be based on performance in three-dimensional environments. Many researchers in color vision have limited their choice of experimental conditions to conditions that are very different from the world in which we live. Such studies have, consequently, yielded limited results. A fruitful alternative is to examine human color perception in three-dimensional scenes that contain veridical cues to the light field.

In this chapter, we first reviewed recent work by researchers on the evaluation of surface color, lightness perception, and constancy in three-dimensional scenes with moderately complex lighting models, and we presented two recent studies in detail. The implication of this research is that the human visual system can compensate for spatially and spectrally inhomogeneous light fields. In the discussion, we found that performance is affected by the availability of
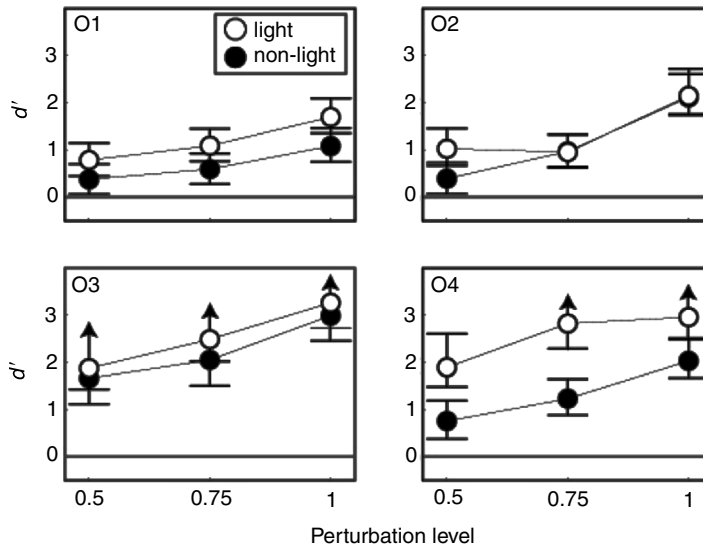
Figure 13.17 Detection of albedo change, from Gerhard and Maloney (2008). The detectability of an albedo perturbation, measured as d' (Green and Swets, 1966), is plotted for each observer as a function of the level of albedo perturbation, expressed as a factor multiplying the original surface reflectance. Nonparametric bootstrapping was used to obtain 95% confidence intervals; arrows denote confidence intervals that include infinity, indicating that the task was trivial for some observers at some perturbation magnitudes. The open circles show the detectability of albedo perturbations in global light field change trials, and the filled circles are for image-statistic-matched nonlight trials. The average benefit for albedo perturbation detectability under a light field change was a 1.65 increase in d'. A color version of this figure can be found on the publisher's website (www.cambridge.org/9781107001756).

specular illuminant cues that signal the light field (Snyder *et al.*, 2005; Boyaci *et al.*, 2006b).

The results of Snyder *et al.* are particularly interesting. The stimuli were presented binocularly and, if we view either of the binocular images in isolation, we see that the only change in the stimulus from trial to trial is small shifts to the left and right against an otherwise constant background (Figure 13.8). If we were to attempt to explain the perceived lightness of the test surface in terms of its immediate surroundings, then we would only predict that there would be little or no trial-to-trial variation. Yet we find large changes in perceived lightness (Figure 13.9) as a function of depth. The binocular-disparity cues that lead to an altered perception of lightness are not present in either image alone. These results are consistent with those of Ikeda *et al.* (1998). Color perception

in three-dimensional scenes cannot be readily predicted given only the results of experiments on Mondrian scenes.

We next described work by Doerschner *et al.* (2007) testing whether human observers can compensate for light fields generated by more than one collimated light source and found that they could do so. Last, we looked at human ability to estimate the light field and changes in the light field and whether this ability benefited human ability to detect changes in surface albedo in scenes where the lighting is changing. It did so. We emphasize that the human ability demonstrated in these last experiments cannot be explained by algorithms limited to consideration of edge ratios and other local computations or even to single images taken from a stereo pair.

In the last part of the chapter, we described recent work directly assessing human ability to estimate the light field (Koenderink *et al.*, 2007) and whether humans can use estimates of changes in the light field to enhance their ability to detect concurrent changes in surface albedo (Gerhard *et al.*, 2007; Gerhard and Maloney, 2008).

Reframing the problem of illuminant estimation in terms of combination of veridical cues to the light field opens up new and potentially valuable directions for research (Maloney, 1999).

In this review, we have focused on surface color perception. It would also be of interest to see how human observers judge color relations (Foster and Nascimento, 1994; Nascimento and Foster, 2000) between surfaces free to differ in orientation and location in the kinds of scenes used in the experiments presented here. Equally, it would be of interest to assess how judgments of material properties (Fleming *et al.*, 2003) vary in such scenes.

From the broadest perspective, a full description of human color perception requires that we examine how the human visual system operates in fully three-dimensional scenes with adequate cues to the illuminant. Understanding human color perception in the case of the Mondrian singularity remains an extremely important research topic, and work in this area is contributing to our understanding of visual perception. The work described here serves to complement and extend the large existing body of literature on surface color perception.

Acknowledgments

References

Adelson, E. H. and Bergen, J. R. (1991). The plenoptic function and the elements of early vision. In M. S. Landy and J. A. Movshon (eds.), *Computational Models of Visual Processing*, pp. 3–20. Cambridge, MA: MIT Press.

Adelson, E. H. and Pentland, A. P. (1996). The perception of shading and reflectance. In D. Knill and W. Richards (eds.), *Perception as Bayesian Inference*, pp. 409–423. New York: Cambridge University Press.

Arend, L. E. and Spehar, B. (1993a), Lightness, brightness, and brightness contrast: 1. Illuminance variation. *Percept. Psychophys.*, 54: 446–456.

Arend, L. E. and Spehar, B. (1993b), Lightness, brightness, and brightness contrast: 2. Reflectance variation. *Percept. Psychophys.*, 54: 457–468.

Basri, R. and Jacobs, D. (2001). Lambertian reflectance and linear subspaces. In *Proceedings of the IEEE International Conference on Computer Vision*, Vancouver, pp. 383–390.

Bäuml, K.-H. (1999). Simultaneous color constancy: how surface color perception varies with the illuminant. *Vis. Res.*, 39: 1531–1550.

Belhumeur, P. N., Kriegman, D., and Yuille, A. (1999). The bas-relief ambiguity. *Int. J. Comput. Vis.*, 35: 33–44.

Bloj, M. G., Kersten D., and Hurlbert, A. C. (1999). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature*, 402: 877–879.

Bloj, M., Ripamonti, C., Mitha, K., Hauck, R., Greenwald, S., and Brainard, D. H. (2004). An equivalent illuminant model for the effect of surface slant on perceived lightness. *J. Vis.* 4(9): 735–746.

Boyaci, H., Maloney, L. T., and Hersh, S. (2003). The effect of perceived surface orientation on perceived surface albedo in three-dimensional scenes. *J. Vis.*, 3: 541–553.

Boyaci, H., Doerschner, K., and Maloney, L. T. (2004). Perceived surface color in binocularly-viewed scenes with two light sources differing in chromaticity. *J. Vis.*, 4: 664–679.

Boyaci, H., Doerschner, K., and Maloney, L. T. (2006a). Cues to an equivalent lighting model. *J. Vis.*, 6: 106–118.

Boyaci, H., Doerschner, K., Snyder, J. L., and Maloney, L. T. (2006b). Surface color perception in three-dimensional scenes. *Vis. Neurosci.*, 23: 311–321.

Ciurea, F. and Funt, B. (2004). Failure of luminance-redness correlation for illuminant estimation. In *Proceedings of the Twelfth Color Imaging Conference*, Scottsdale, AZ, pp. 42–46.

Doerschner, K., Boyaci, H., and Maloney, L. T. (2004). Human observers compensate for secondary illumination originating in nearby chromatic surfaces. *J. Vis.*, 4: 92–105.

Doerschner, K., Boyaci, H., and Maloney, L. T. (2007). Testing limits on matte surface color perception in three-dimensional scenes with complex light fields. *Vis. Res.*, 47: 3409–3423.

Dror, R. O., Willsky, A., and Adelson, E. H. (2004). Statistical characterization of real-world illumination. *J. Vis.*, 4: 821–837.

Epstein, W. (1961). Phenomenal orientation and perceived achromatic color. *J. Psychol.*, 52: 51–53.

Fleming, R. W., Dror, R. O., and Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *J. Vis.*, 3: 347–368.

Flock, H. R. and Freedberg, E. (1970), Perceived angle of incidence and achromatic surface color. *Percept. Psychophys.*, 8: 251–256.

Foster, D. H. and Nascimento, S. M. C. (1994). Relational colour constancy from invariant cone-excitation ratios. *Proc. R. Soc. Lond. B*, 257: 115–121.

Gerhard, H. E. and Maloney, L. T. (2008). Albedo perturbation detection under illumination transformations: a dynamic analogue of lightness constancy. [Abstract.] *J. Vis.*, 8: 289.

Gerhard, H. E. and Maloney, L. T. (2010). Detection of light transformations and concomitant changes in surface albedo. *J. Vis.* 10: 1–14.

Gerhard, H. E., Khan, R., and Maloney, L. T. (2007). Relational color constancy in the absence of ratio constancy. [Abstract.] *J. Vis.*, 7: 459.

Gershun, A. (1936/1939). *Svetovoe Pole* [The Light Field]. Moscow, 1936. Translated by P. Moon and G. Timoshenko (1939) in *J. Math. Phys.*, 18: 51–151.

Gilchrist, A. L. (1977). Perceived lightness depends on spatial arrangement. *Science*, 195: 185–187.

Gilchrist, A. L. (1980). When does perceived lightness depend on perceived spatial arrangement? *Percept. Psychophys.*, 28: 527–538.

Gilchrist, A. L. and Annan, A., Jr. (2002). Articulation effects in lightness: historical background and theoretical implications. *Perception*, 31: 141–150.

Gilchrist, A. L., Kossyfidis, C., Bonato F., Agostini, T., Cataliotti, J., Li, X. J., Spehar, B., Annan, V., and Economou, E. (1999). An anchoring theory of lightness perception. *Psychol. Rev.*, 106: 795–834.

Golz, J. and MacLeod, D. I. A. (2002). Influence of scene statistics on colour constancy. *Nature*, 415: 637–640.

Granzier, J. J. M., Brenner, E., Cornelissen, F. W., and Smeets, J. B. J. (2005). Luminance–color correlation is not used to estimate the color of the illumination. *J. Vis.*, 5: 20–27.

Green, D. M. and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.

Hara, K., Nishino, K., and Ikeuchi, K. (2005). Light source position and reflectance estimation from a single view without the distant illumination assumption. *IEEE Trans. Pattern Anal. Machine Intell.*, 27: 493–505.

Haralick, R. M. and Shapiro, L. G. (1993). *Computer and Robot Vision*, Vol. 2. Reading, MA: Addison-Wesley.

Henderson, S. T. (1977). *Daylight and Its Spectrum*, 2nd edn. Bristol: Adam Hilger.

Hochberg, J. E. and Beck, J. (1954). Apparent spatial arrangements and perceived brightness. *J. Exp. Psychol.*, 47: 263–266.

Hurlbert, A. C. (1998). Computational models of colour constancy. In V. Walsh and J. Kulikowski (eds.), *Perceptual Constancy: Why Things Look as They Do*, pp. 283–322, Cambridge: Cambridge University Press.

Ikeda, M., Shinoda, H., and Mizokami, Y. (1998). Three dimensionality of the recognized visual space of illumination proved by hidden illumination, *Opt. Rev.*, 5: 200–205.

Kaiser, P. K. and Boynton. R. M. (1996). *Human Color Vision*, 2nd edn. Washington, DC: Optical Society of America.

Kardos, L. (1934). Ding und Schatten; Eine experimentelle Untersuchung über die Grundlagen des Farbsehens. Z. *Psychol. Physiol. Sinnesorgane, Ergänzungsband*, 23. Leipzig, Germany: Verlag von J. A. Barth.

Katz, D. (1935). *The World of Colour*. London: Kegan, Paul, Trench, Trubner and Co.

Koenderink, J. J. and van Doorn, A. J. (1996). Illuminance texture due to surface mesostructure. *J. Opt. Soc. Am. A*, 13: 452–463.

Koenderink, J. J., van Doorn, A. J., and Pont S. C (2004). Light direction from shad(ow)ed random Gaussian surfaces. *Perception*, 33: 1403–1404. Special issue, Shadows and Illumination II.

Koenderink, J. J., Pont S. C., van Doorn, A. J., Kappers, A. M. L., and Todd J. T. (2007). The visual light field. *Perception*, 36: 1595–1610.

Land, E. H. and McCann, J. J. (1971). Lightness and retinex theory. *J. Opt. Soc. Am.*, 61: 1–11.

Lee, R. L., Jr. and Hernández-Andrés, J. (2005a). Short-term variability of overcast brightness. *Appl. Opt.*, 44: 5704–5711.

Lee, R. L., Jr. and Hernández-Andrés, J. (2005b). Colors of the daytime overcast sky. *Appl. Opt.*, 44: 5712–5722.

MacLeod, D. I. A. and Golz, J. (2003). A computational analysis of colour constancy. In R. Mausfeld and D. Heyer (eds.), *Colour Perception: Mind and the Physical World*, pp. 205–242. Oxford: Oxford University Press.

Maloney, L. T. (1999). Physics-based approaches to modeling surface color perception. In K. R. Gegenfurtner and L. T. Sharpe (eds.), *Color Vision: From Genes to Perception*, pp. 387–422. Cambridge: Cambridge University Press.

Maloney, L. T., Boyaci, H., and Doerschner, K. (2005). Surface color perception as an inverse problem in biological vision. *Proc. SPIE*, 5674: 15–26.

Mausfeld, R. and Andres, J. (2002). Second-order statistics of colour codes modulate transformations that effectuate varying degrees of scene invariance and illumination invariance. *Perception*, 31: 209–224.

Nascimento, S. M. C. and Foster, D. H. (2000). Relational color constancy in achromatic and isoluminant images. *J. Opt. Soc. Am. A*, 17: 225–231.

Oruc, I., Maloney, L. T., and Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error. *Vis. Res.*, 43: 2451–2458.

Pont, S. C. and Koenderink, J. J. (2003). Illuminance flow. In N. Petkov and M. A. Wetsenberg (eds.), *Computer Analysis of Images and Patterns*, pp. 90–97. Berlin: Springer.

Pont, S. C. and Koenderink, J. J. (2004). Surface illuminance flow. In Y. Aloimonos and G. Taubin (eds.), *Proceedings of the second International symposium on SD Data Processing Visualization and Transmission*, Thessaloniki, Greece. Piscataway, NJ: IEEE.

Ramamoorthi, R. and Hanrahan, P. (2001a). On the relationship between radiance and irradiance: determining the illumination from images of a convex Lambertian object. *J. Opt. Soc. Am. A*, 18: 2448–2458.

Ramamoorthi, R. and Hanrahan, P. (2001b). An efficient representation for irradiance environment maps. In *Proceedings of SIGGRAPH 2001*, Los Angeles, CA, pp. 497–500. New York: ACM Press.

Redding, G. M. and Lester, C. F. (1980). Achromatic color matching as a function of apparent test orientation, test and background luminance, and lightness or brightness instructions. *Percept. Psychophys.*, 27: 557–563.

Ripamonti, C., Bloj, M., Hauck, R., Kiran, K., Greenwald, S., Maloney, S. I., and Brainard, D. H. (2004). Measurements of the effect of surface slant on perceived lightness. *J. Vis.*, 4: 747–763.

Snyder, J. L., Doerschner, K., and Maloney, L. T. (2005), Illumination estimation in three-dimensional scenes with and without specular cues. *J. Vis.*, 5: 863–877.

te Pas, S. F. and Pont S. C. (2005). Comparison of material and illumination discrimination performance for real rough, real smooth and computer generated smooth spheres. In *Proceedings of the 2nd Symposium on Applied Perception in Graphics and Visualization*, A Coruña, Spain, Vol. 95, pp. 75–81. New York: ACM Press.

Yang, J. N. and Maloney, L. T. (2001). Illuminant cues in surface color perception: tests of three candidate cues. *Vis. Res.*, 41: 2581–2600.