Real-world illumination and the perception of surface reflectance properties

Roland W. Fleming Massachusetts Institute of Technology, Cambridge, MA, USA

Ron O. Dror Massachusetts Institute of Technology, Cambridge, MA, USA

Edward H. Adelson Massachusetts Institute of Technology, Cambridge, MA, USA

Under typical viewing conditions, we find it easy to distinguish between different materials, such as metal, plastic, and paper. Recognizing materials from their surface reflectance properties (such as lightness and gloss) is a nontrivial accomplishment because of confounding effects of illumination. However, if subjects have tacit knowledge of the statistics of illumination encountered in the real world, then it is possible to reject unlikely image interpretations, and thus to estimate surface reflectance even when the precise illumination is unknown. A surface reflectance matching task was used to measure the accuracy of human surface reflectance estimation. The results of the matching task demonstrate that subjects can match surface reflectance properties reliably and accurately in the absence of context, as long as the illumination is realistic. Matching performance declines when the illumination statistics are not representative of the real world. Together these findings suggest that subjects do use stored assumptions about the statistics of real-world illumination to estimate surface reflectance. Systematic manipulations of pixel and wavelet properties of illuminations reveal that the visual system's assumptions about illumination are of intermediate complexity (e.g., presence of edges and bright light sources), rather than of high complexity (e.g., presence of recognizable objects in the environment).

Keywords: reflectance estimation, gloss, specularity, lightness constancy, illumination, natural image statistics, material perception, texture recognition

1. Introduction

All objects in the world are made of some material or another, and we usually have a good idea what, just by looking. Under typical viewing conditions, we find it trivial to distinguish between different materials, such as metal, plastic and paper, irrespective of the form of the object or the conditions of illumination. Given this observation, and given the enormous variety of substances to be found in the environment, it seems reasonable to presume that our capacity for recognizing different *materials* rivals our ability to recognize different *objects*. And yet very little research has been carried out to determine how (although see Nishida & Shinya, 1998; Adelson, 2001; and a number of ongoing projects of Koenderink and colleagues). Key questions include the following: What are the necessary and sufficient conditions to recognize different materials? What sources of information are available to an observer as a result of the different ways that materials interact with light? What are the principle dimensions underlying the representation of materials in the observer's visual system?

One very important source of information about material identity results from the wide range of optical properties that different materials exhibit. Different materials reflect, transmit, refract, disperse, and polarize light to different extents and in different ways; this provides a rich set of optical cues for distinguishing materials. For most materials, the majority of the light that is not absorbed is reflected from the surface, and thus a material's surface reflectance properties are surely some of its most important optical attributes. When light is reflected from a surface, it is generally scattered in many directions, producing a pattern that is characteristic of the material. Variation in the distribution of scatter gives rise to such varied visual appearances as bronze, plaster-of-Paris, gloss paint, and gold. In this work, we present a number of theoretical and empirical observations on the conditions under which humans are good at estimating surface reflectance properties. We also discuss a number of cues that appear to underlie this aptitude.

1.1 Surface Reflectance Estimation

Estimating surface reflectance is difficult because the image presented by a material depends not only on the reflectance properties but also on the conditions of illumination. The image of a chrome sphere, for example, is simply a distorted reflection of the world around it, and thus the image of the sphere depends solely on the context in which it is viewed (see Figure 1).

DOI 10:1167/3.5.3 Received October 12, 2002; published July 1, 2003 ISSN 1534-7362

And yet something about the appearance of the sphere remains the same across all of these contexts: it still looks like chrome. This variability prevents the brain from recognizing materials by simply matching the raw image to a stored template. In this respect, the task of recognizing materials with *uniform* surface reflectance properties (i.e., untextured materials, which have the same reflectance at each location on the surface) resembles the task of recognizing *textures* (i.e., materials whose reflectance properties vary across the surface in distinctive statistical patterns). In both cases there is some characteristic pattern of features that are common to all samples within a class, and yet the specific image varies from sample to sample. In the following arguments, we draw close parallels between texture recognition and surface reflectance estimation, and also

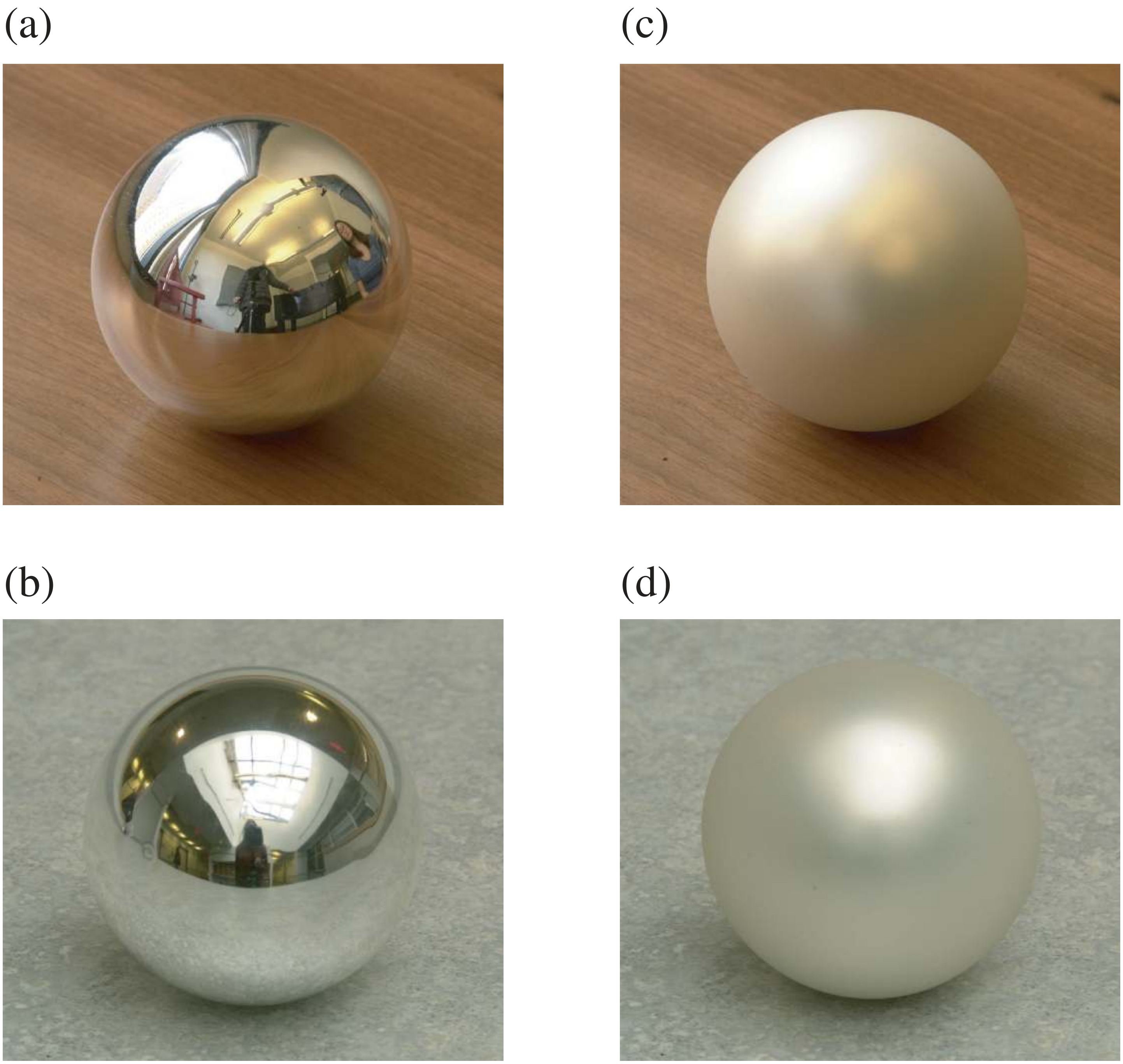

Figure 1. Surface reflectance estimation is difficult because of confounding effects of illumination. The same sphere is shown in two different scenes in (a) and (b). Because of the change of environment, the images of the spheres are quite different, although the material appears the same. Images (c) and (d) are photographs of a different sphere in the same scenes as (a) and (b). On a pixel-by-pixel basis, (c) is more similar to (a) than (b) is, despite the difference in material composition.

discuss a few critical differences.

A second problem for surface reflectance estimation is posed by the conditions of viewing. Under carefully contrived viewing conditions, a chrome sphere, for example, can be made to produce the same image as a sphere of any other material. This could be achieved by painting the world in such a way that its distorted reflection in the sphere perfectly reproduced the pattern of light that *would* be reflected from a matte red sphere, for example, viewed under more normal conditions. If the precise conditions of viewing are not known to the observer, then surface reflectance estimation is under-constrained because many different combinations of material and scene are consistent with a given image.

To summarize, identical materials can lead to different images, whereas different materials can lead to identical images. These examples serve to demonstrate the deep relationship between illumination and surface reflectance estimation. The characteristics of everyday illumination play a major role in the arguments that follow.

1.2 Real-World Illumination

As discussed above, the image of a surface depends not only on the material from which the surface is made, but also on the pattern of light that impinges on it from the environment. Therefore, if we want to understand surface reflectance estimation, we must also understand the patterns of light that typically illuminate surfaces in the real world. Figure 2 demonstrates the importance of the pattern of incoming light in determining the appearance of a material. Three spheres were computer rendered under different illuminations. In (a) the sphere was rendered under an isolated point-light source floating in space, while the spheres in (b) and (c) were rendered in environments that are more typical of the real world. The impression of the material properties is clearer in (b) and (c) than in (a). This observation motivates the arguments that follow, and is corroborated by our experiments.

What is illumination, and what determines its structure? In the real world, light is typically incident on a surface from nearly every direction. Some of this light comes directly from luminous sources, such as the sun, and some comes indirectly, reflected from other surfaces. However, all of the light is treated equally by the surface, regardless of its origin, and thus all of this light is "illumination."

Each point in space receives different illumination, as a different set of rays converge on that point. In order to characterize the illumination at a given point in space, we would have to measure the light arriving at that point from every direction. This would create a spherical image or "illumination map" for that point in space. The value at each location on the spherical image represents the amount of light arriving from that direction, as depicted in Figure 3. Such spherical images or "illumination maps" have been captured photographically from locations in the world and used to render objects for the purposes of realistic computer graphics (Debevec, 1998; Debevec, Hawkins, Tchou, Duiker, Sarokin, & Sagar, 2000). Indeed, the spheres in (b) and (c) of Figure 2 were rendered using two of these illuminations.

Why might real-world illumination facilitate surface reflectance estimation? Our argument is that illumination maps derived from very different scenes in the world nevertheless share certain statistical regularities in structure and these regularities allow the visual system to make certain assumptions in interpreting images of surfaces. Recent work has shown that the spatial

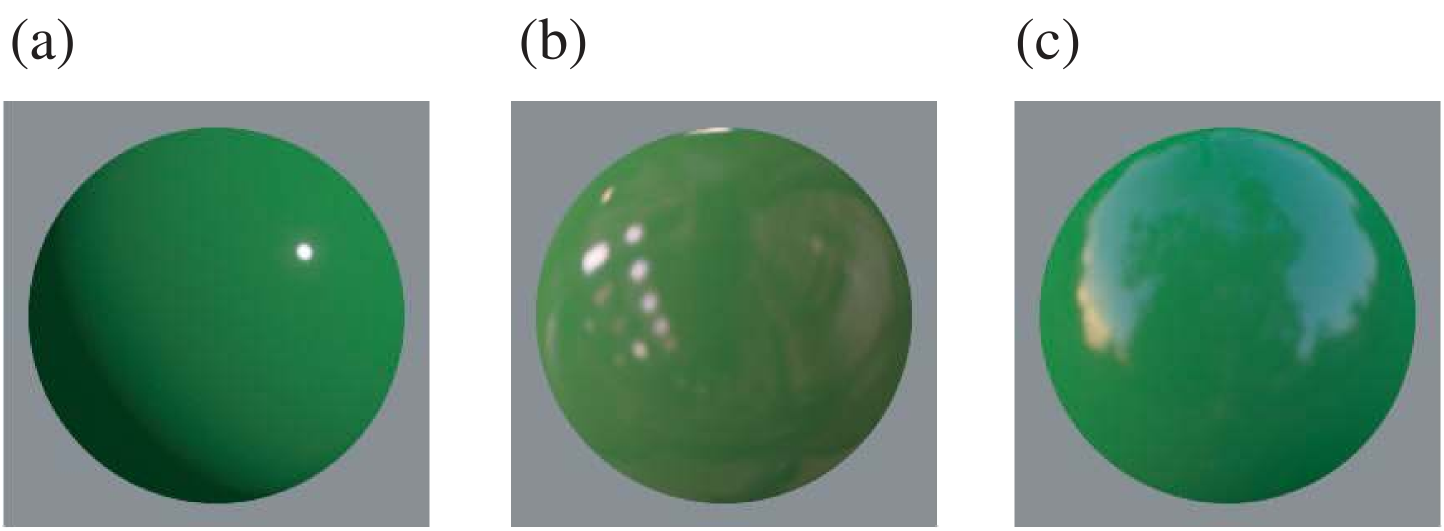

Figure 2. The sphere in (a) was rendered under point-source illumination, while the spheres in (b) and (c) were rendered under photographically captured real-world illuminations. Most observers agree that the impression of material qualities is clearer for (b) and (c) than for (a). This demonstrates the important role of real-world statistics in the perception of surface reflectance.

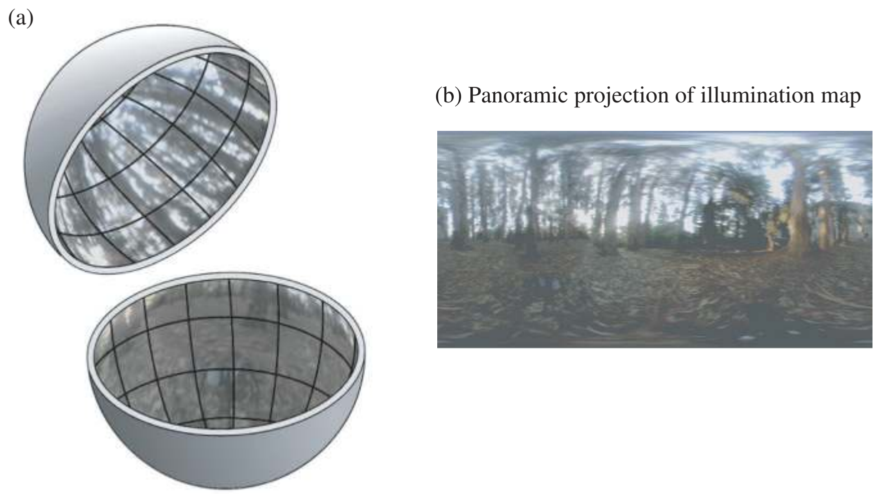

Figure 3. Illumination at a point in space is defined as the set of rays that converge on that point from every direction. The set of all rays forms a spherical image of incoming light such that each location on the sphere represents the amount of light arriving from the corresponding direction. Such spherical images can be acquired photographically for points in the real world. (a) shows one such map, acquired by Debevec et al. (2000). (b) shows the same illumination map projected onto a two-dimensional plane. Real-world illumination maps exhibit statistical regularities similar to those of conventional real-world images.

structure of real-world illumination maps possesses statistical regularity similar to that of natural images (Dror, Adelson, & Willsky, 2001; Dror, Leung, Willsky, & Adelson, 2001); this is not surprising because the structure of the maps is derived from the layout of objects of the environment. The visual system could in principle rely on these statistical regularities to eliminate unlikely image interpretations. We discuss some key statistical properties of illumination below.

1.3 Exploiting the Statistical Regularities Of Real-World Illumination to Estimate Surface Reflectance

Although many combinations of illumination and material are consistent with a given image, some combinations are more likely than others if we take into account the statistics of the real world. We reason that humans exploit tacit knowledge of the statistics of real-world illuminations to reject interpretations that are unlikely to occur under normal viewing conditions. This makes it possible to recognize materials even when the precise illumination is unknown, without performing "inverse optics."

Figure 4 is a photograph of a pearlescent sphere, which has been cut out of its original context and placed against a neutral background. The only information that observers can use to determine the material is the pattern of light within the sphere itself, and yet our impression of the surface reflectance is unambiguous and fairly accurate (e.g., we do not mistake the sphere for bronze or chalk). We argue that this is possible because the image is full of blurred features, such as the one highlighted in Figure 4. When confronted with a blurred feature, two of the possible interpretations are (a) the feature could be a blurred reflection of an inherently sharp world, or (b) it could be a sharp reflection of an inherently blurry world. We argue that the visual system can reject the latter

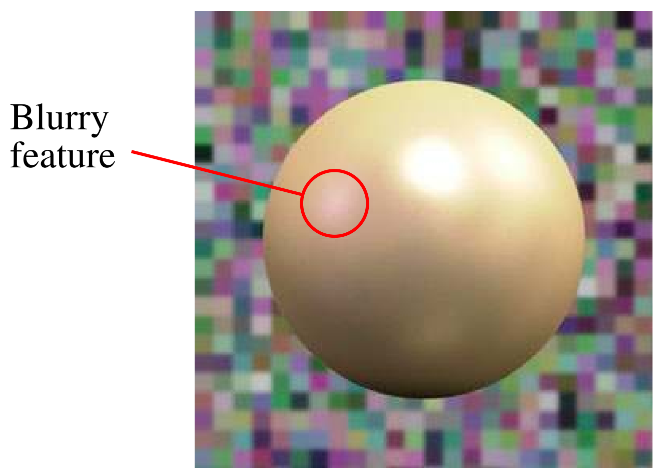

Figure 4. A photograph of a pearlescent sphere that has been cropped and placed on an arbitrary synthetic background. We have a fairly clear impression of the material qualities of the sphere, even though there is no context to specify the illumination. When real-world illumination is reflected in a surface, it reliably leads to image features, such as the one highlighted in red, that are characteristic of the surface's reflectance properties. Despite the inherent ambiguity of interpreting the feature, the regularities of the real world allow the visual system to reject interpretations that are improbable given the statistics of the world.

interpretation because most of the time the world is not blurry, and thus it is much more likely that it is the reflection that is blurred.

This logic effectively converts the problem of surface reflectance estimation into a problem analogous to texture recognition. Different textures can be recognized because they contain some characteristic set of statistical image features. Likewise, different materials have characteristic appearances because the reflection of the world in their surfaces reliably leads to some set of statistical image features, such as the blurred feature in the pearlescent sphere. Thus surface reflectance properties can be estimated directly from the image, without performing inverse optics. Note that this approach is only available to the observer because certain statistical properties are highly conserved across real-world illuminations (i.e., in the real world, illumination is not arbitrary). This is what we mean when we say that the visual system exploits the statistical regularities of real-world illuminations to eliminate improbable image interpretations.

One consequence of this "image-based" approach is that subjects can recognize materials across variations in illumination, as demonstrated in Figure 1. Even though on a pixel-by-pixel basis the image of a surface varies dramatically from illumination to illumination, subjects can nevertheless "ignore" the variations that are due to illumination and reliably recognize the material. Our argument is that subjects do this by tracking diagnostic features that are well conserved across illuminations. In the following experiment, we find that subjects can reliably match surface reflectance properties across variations in illumination.

A second consequence of the image-based approach to surface reflectance estimation is that it should be more difficult to recognize materials under illuminations with statistics that are *not* typical of the real world. Image features that are reliable cues for surface reflectance under typical illuminations may lead to spurious estimates of surface reflectance when the illumination statistics are not typical of the real world. In the following experiment, we measure the accuracy of human surface reflectance estimation under illuminations with typical and atypical statistics.

1.3.1 The role of context in surface reflectance estimation

A third consequence of the image-based approach to surface reflectance estimation is that subjects can estimate certain surface reflectance properties (e.g., gloss) for isolated surfaces; that is, in the absence of context. This is consistent with the example of the pearlescent sphere above, and our previous report (Fleming, Dror, & Adelson, 2001). Figure 5 further demonstrates the effects of changing context. The image in (a) shows a sphere rendered under an illumination that was photographically captured from the real world by Debevec et al. (2000). In this image, the sphere is shown against its true background.[1] In (b), the image of the sphere has been cropped out of its original background and pasted onto a different real-world background. Although the image somehow looks "wrong," or internally inconsistent, this has remarkably little effect on the perceived surface reflectance properties of the sphere itself; specifically, the background has practically no effect on perceived gloss. Image (c) shows the same sphere, this time against a third real-world background. Again the background looks inappropriate, but the surface reflectance properties of the sphere remain largely unaffected. These observations are consistent with the finding of Hartung and Kersten

(a) (b) (c)

Figure 5. The negligible effects of context on perceived gloss. Sphere (a) is shown against its true background, acquired photographically by Debevec et al. (2000). Images (b) and (c) were created by cropping the sphere out of image (a) and placing it against other backgrounds. This has relatively little effect on our perception of the surface reflectance properties of the sphere.

(2002) that when the object under scrutiny and its background do not belong together, this has little effect on the perception of gloss.

It is worth noting that this observation is seemingly at odds with a couple of well-known phenomena in lightness perception, namely that (a) it is impossible to estimate the albedo of an isolated patch of Lambertian material (Gelb, 1929); and (b) the perceived lightness of a patch of uniform intensity can be dramatically altered by the context in which it is placed (Gelb, 1929; Katz, 1935; Gilchrist, 1977, 1979, 1994; Adelson, 1999; for a review, see Gilchrist et al., 1999). Why does context seem to play so much less of a role for our stimuli? We argue that it is because of the structured specular reflections present in our stimuli. The complex patterns of reflection supply the visual system with sufficient diagnostic image features to estimate the specular reflectance properties directly from the image of the object, without having to derive any estimate of the prevailing illumination from the context. We discuss the role of context further in the "Appendix."

In the experiments that follow, images of spheres were removed from their original contexts. This is justified because of the apparently small effects of context on surface reflectance estimation for these stimuli. If performance is good in the absence of context, it supports our suggestion that subjects can estimate certain reflectance properties directly from the image, without performing inverse optics.

2. Methods

In order to measure the accuracy of human surface reflectance estimation, we asked subjects to perform a surface reflectance-matching task. Subjects were presented with the images of two spheres that had been computer rendered under different illuminations (see Figure 6). Their task was to adjust the surface reflectance of one sphere (the "Match") until it appeared to be made of the same material as the other sphere (the "Test"), despite the difference in illumination.

2.1 Observers

Four subjects with normal or corrected-to-normal vision participated in the experiments. One was an author (R.F.), two were experienced observers (J.M. and M.S.), who were naïve to the purpose of the study, and one was a novice observer (R.A.), who was paid for participating.

2.2 Stimuli

2.2.1 Reflectance properties

The spheres were all spatially uniform in surface reflectance (i.e., untextured). Reflectance was represented using the isotropic Ward model (Ward, 1992), which is a parametric model of reflectance like the Phong shading model. Unlike the Phong model, the Ward model is constrained to obey fundamental physical laws, such as conservation of energy and reciprocity. The Ward model represents surface reflectance as the sum of two components: diffuse and specular reflection. Diffuse (or "Lambertian") reflection occurs when light is scattered equally in all directions as it reflects from the surface. The proportion of incoming light reflected in this way determines the albedo (ρ_D) of the surface (see Figure 7). Small values of the albedo parameter lead to black and dark grey surfaces, while large values lead to light-grey and white surfaces. As lightness perception has been studied extensively, this parameter was held fixed at red = 0.1; green = 0.3; blue = 0.1 for all stimuli in the experiment. This yields a dark green color.

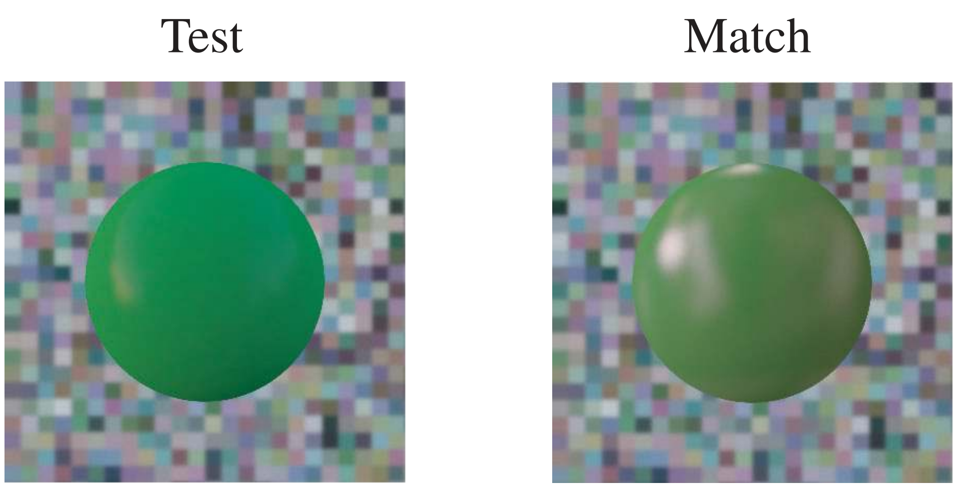

Figure 6. Example stimuli from the surface reflectance matching task. Subjects adjusted the reflectance properties of the Match sphere until it appeared to be made of the same material as the Test sphere, despite the difference in illumination. Note that in this image the spheres have different surface reflectance properties.

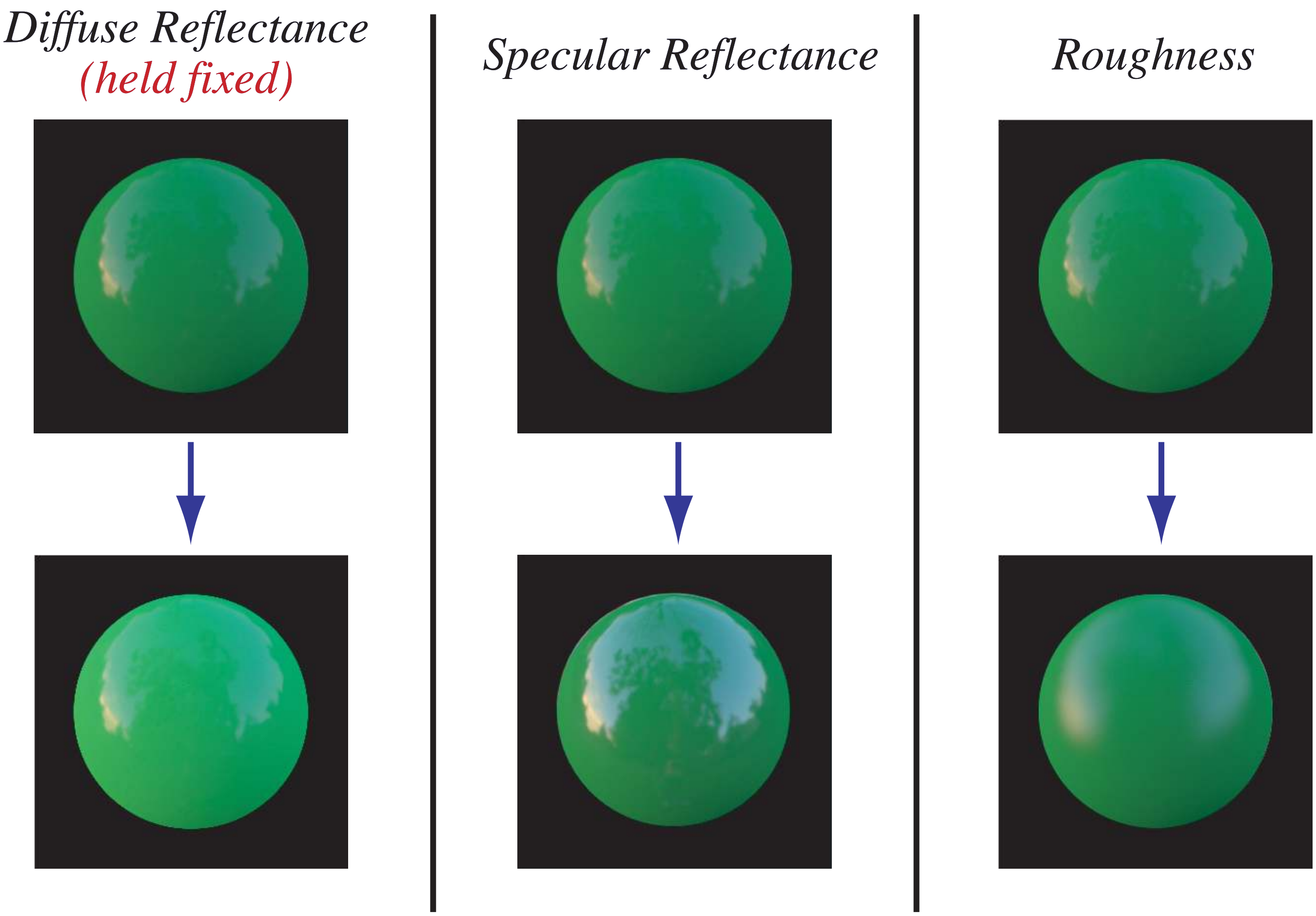

Figure 7. The three parameters of the Ward reflectance model. Diffuse reflectance specifies the proportion of incoming light reflected by the diffuse (Lambertian) component. In the matching experiments this was held constant for all stimuli. Specular reflectance controls the proportion of incoming light reflected by the specular component, surface roughness controls the spread or blur of the specular reflection. Subjects adjusted the latter two parameters to match surface reflectance.

The second component of the Ward model represents specular reflection. This reflectance component is characterized by the fact that the angle of reflectance is equal to the angle of incidence (or distributed thereabout). Specular reflection leads to a mirrorlike or glossy appearance. Unlike diffuse reflectance, there are two parameters associated with specular reflection in the Ward model. The specular reflectance (ρ_S) parameter controls the proportion of incoming light that is reflected in this way. Small values of this parameter yield matte surfaces such as soot and chalk; intermediate values yield glossy surfaces such as plastic and glass; and large values yield lustrous surfaces such as platinum (see Figure 7). A final parameter (α) controls the roughness of the surface at a microscopic scale. Changing this parameter leads to changes in the "spread" or blur of the specular reflection. Small values of the roughness parameter lead to smooth surfaces with crisp specular reflections, like polished chrome. Large values lead to rough surfaces with blurred reflections, like unpolished aluminum or sandblasted plastic (see Figure 7). A wide range of materials, such as metals, plastics, and paints, have been modeled with the aforementioned three parameters.[2] The parameter scales were stretched nonlinearly to make the step-sizes perceptually equal. This reparameterization was performed according to the psychophysically uniform space proposed by Pellacini, Ferwerda, and Greenberg (2000).[3]

Subjects simultaneously adjusted the specular reflectance and roughness parameters of the specular reflection to match the material. Ten values were used for the specular reflectance parameter and eleven for the roughness parameter, making a total of 110 possible surface reflectances. These values spanned a range greater than but including the range of reflectances exhibited by isotropic "plastics," such as gloss paint and sandblasted plastic in the real world (see Figure 8). Specifically, values for the specular reflectance parameter ran from c = 0.019 to 0.190 in 10 even steps in the Pellacini et al. parameterization, which is equivalent to a range of ρ_S = 0.0139 to 0.193 in the Ward model. Values for the surface roughness parameter ran from d = 0.900 to 1.00

in 11 even steps in the Pellacini et al. parameterization, which is equivalent to a range of α = 0.00 to 0.10 in the Ward model.

2.2.2 Illuminations

The spheres were rendered under nine real-world illuminations, and five artificial illuminations with various atypical statistics. The real-world illuminations that we used were taken from a database originally acquired by Debevec et al. (2000) from a variety of indoor and outdoor scenes, using high-dynamic range photography.[4] The overall brightness of the different illuminations was normalized such that a standard Lambertian patch oriented perpendicular to the observer yielded the same luminance under each of the illuminations. Figure 9 shows spheres viewed under each of the eight real-world illuminations used to render Test stimuli; all spheres in this figure have the same surface reflectance properties. The Match sphere that the subjects adjusted was viewed under the "Galileo" real-world illumination for all conditions (see Figure 10). This illumination was never used to render Test stimuli.

The artificial illuminations were designed to have specific atypical spatial or statistical properties; they consisted of (a) a single point source; (b) multiple point sources; (c) a single extended rectangular source; (d) Gaussian white noise; and (e) Gaussian noise with a 1/f amplitude spectrum (pink noise). Example spheres rendered under each of these illuminations are shown in Figure 11; the spheres all have the same reflectance as the spheres rendered under real-world illuminations in Figure 9. It is worth noting that the impression of the reflectance properties is generally less distinct for the spheres viewed under the artificial illuminations than for

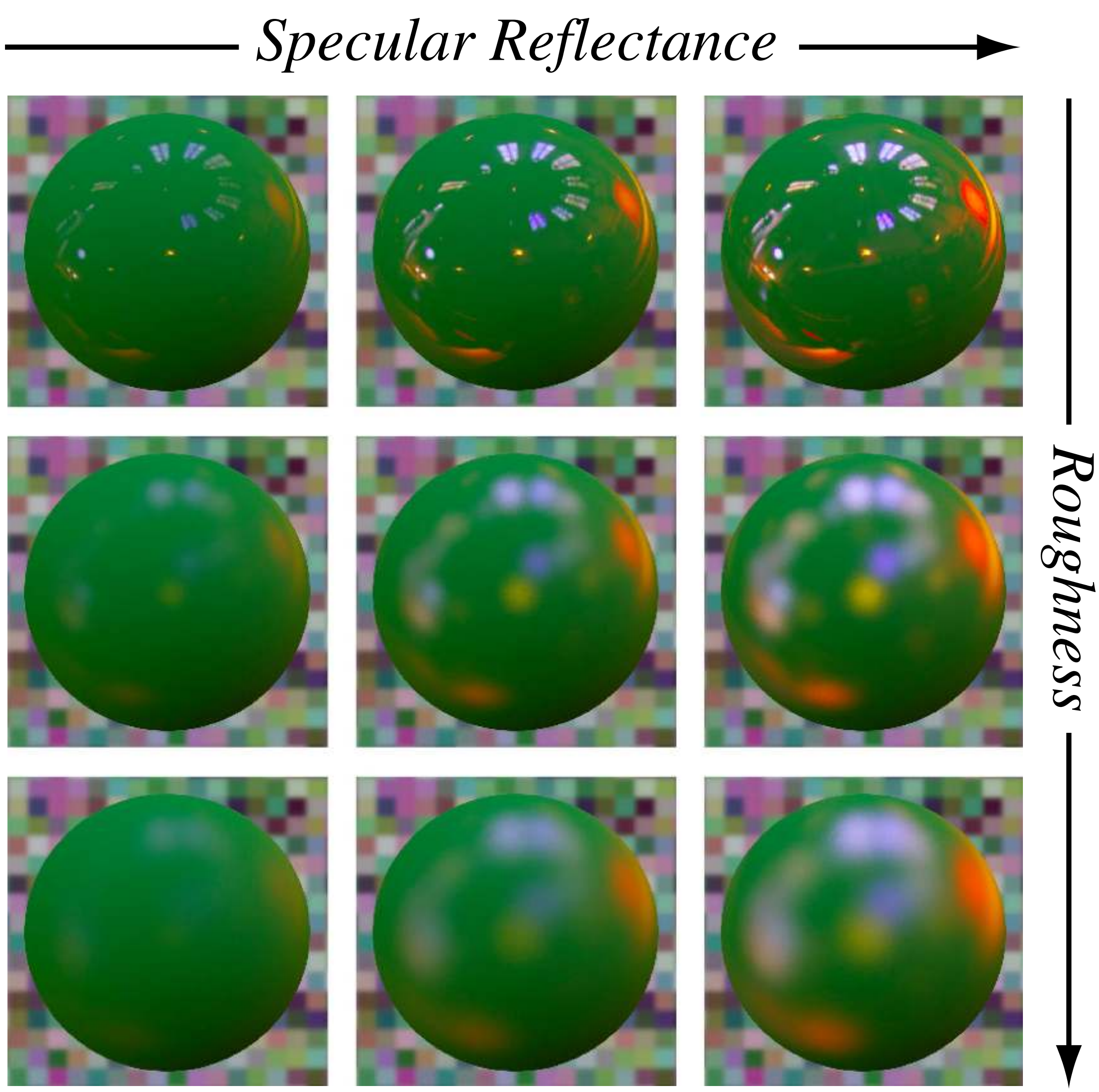

Figure 8. Subjects adjusted specular reflectance and surface roughness to match the appearance of the spheres. Ten values were used for specular reflectance and 11 for roughness yielding a total of 110 possible surface reflectances. The scales of these parameters were adjusted to form a perceptually uniform space, using the nonlinear scaling proposed by Pellacini et al. (2000).

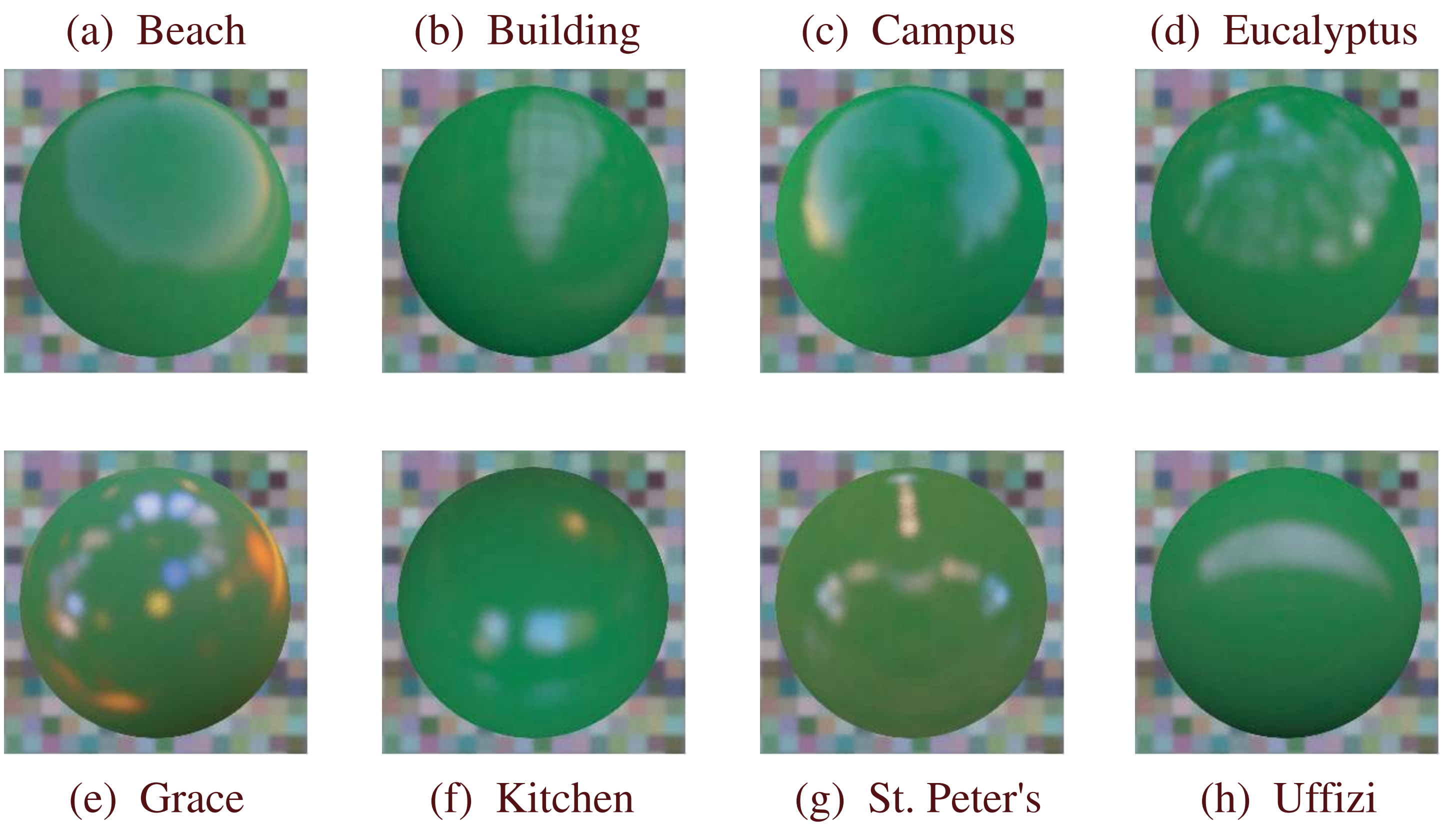

Figure 9. Spheres rendered under each of the real-world illuminations used in the matching experiments. All spheres shown here have the same surface reflectance properties. It should be noted that these spheres do not have the maximum specular reflectance or minimum roughness used in the experiments. Therefore additional detail was visible in some experimental conditions.

those rendered under real-world illuminations; the one exception is the illumination featuring the single rectangular source.

The white noise illumination map was generated by summing spherical harmonics whose coefficients up to a fixed order were chosen from independent Gaussian distributions of equal variance. For the pink noise, the spherical harmonic coefficients were again chosen from independent Gaussian distributions, but the standard deviation of the distributions was inversely proportional to the spherical harmonic order (this is the spherical analogue of frequency). This process yields a characteristic "cloudlike" pattern, whose power spectrum is similar to that of many real-world illuminations, but whose phase characteristics are not typical of the real world.

Figure 10. Sphere rendered under the illumination used for the match sphere in the experiments. As in Figure 9, this sphere has neither the sharpest nor brightest specular reflectance values used in the experiments.

2.2.3 Rendering

Rendering was performed using the RADIANCE rendering software (Ward, 1994; http://radsite.lbl.gov/radiance/HOME.html). Illuminations were stored and loaded using the RADIANCE native high-dynamic range format (.hdr or .pic). The illumination data were treated as illumination arriving from infinite distance and from all directions for the evaluation of the Ward reflectance model. This can be achieved by representing the data as a "glow source" in the RADIANCE scene description. Further details are given in Dror's (2002) doctoral thesis.

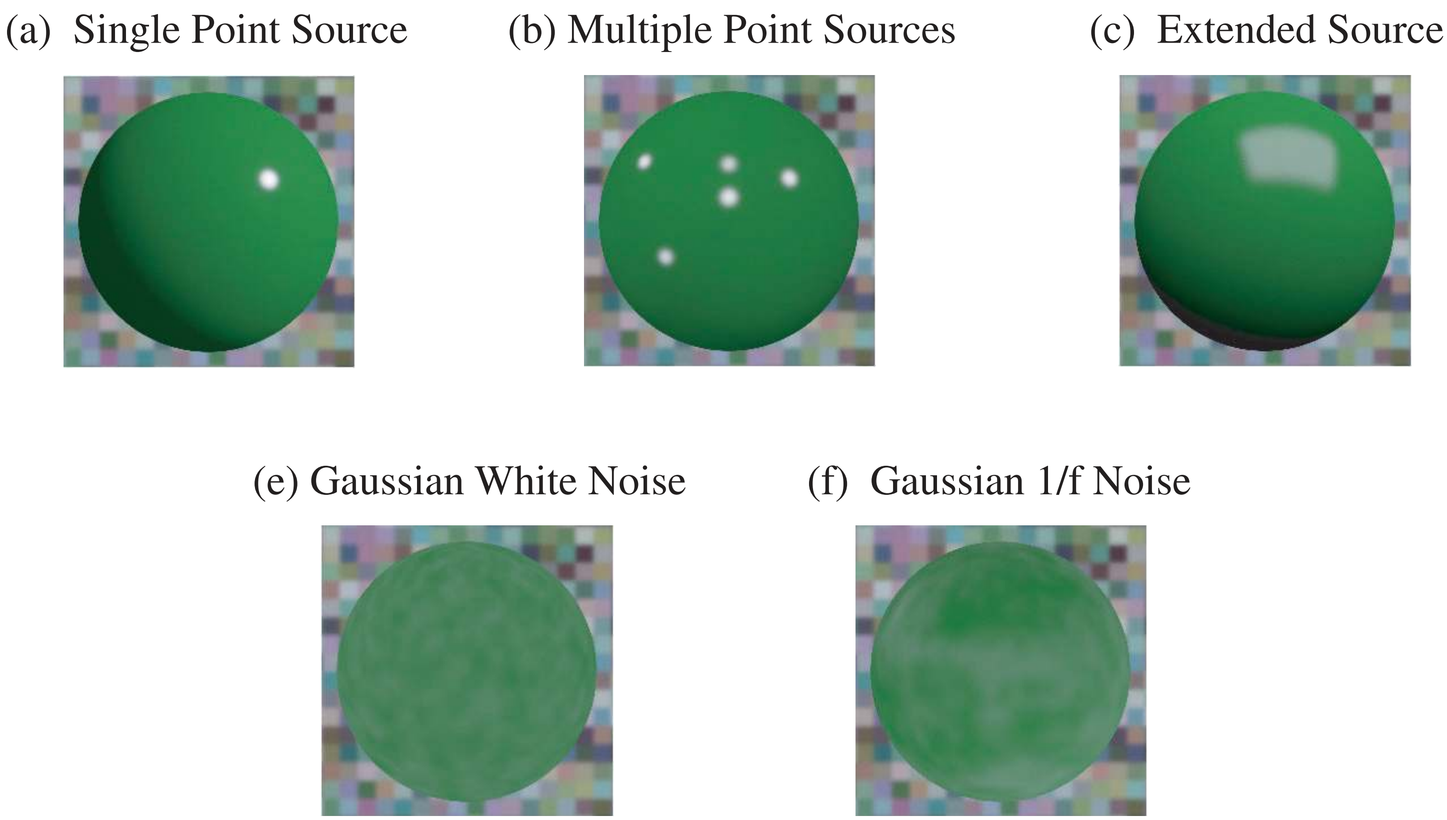

Figure 11. Spheres rendered under each of the synthetic illuminations used in the matching experiment. Each illumination was designed to have some key properties in common with real-world illuminations, but otherwise to have atypical statistics. If subjects' stored assumptions about illuminations are infringed, performance should be impaired. It should be noted that perceived surface reflectance is less clear for these spheres than for those in Figure 9, with the possible exception of (c), which was rendered in a world featuring a single extended rectangular source.

2.2.4 Display limitations on the CRT

The range of luminances that results from viewing a specular surface under ordinary viewing conditions can be several orders of magnitude larger than what is possible with a good monitor. It is possible that the sheer intensity of real highlights facilitates reflectance estimation, and this cannot be reproduced using current display technology. However, in an attempt to overcome this we used a number of presentation devices, to maximize the utility of the available range.

First, all images were presented in a black room with the lights off, to decrease the luminance of the darkest blacks in the image. We estimated that as a consequence of this we were able to achieve a dynamic range of about 30:1 for high spatial frequency information, and up to about 120:1 for larger regions.

Second, rather than allowing the image values to clip, the images were passed through a compressive nonlinearity of the type described by Tumblin, Hodgins, and Guenter (1999). This is a sigmoidal nonlinearity that is linear for intermediate luminances but compresses low and high values. The same tone-mapping function was used for every experimental condition. The monitor was calibrated to ensure linearity before every session of the experiment.

Third, we applied synthetic glare to the rendered images in order to mimic the optical effects of viewing high luminances with the human eye. This was done according to specifications derived by Ward Larson, Rushmeier, and Piatko (1997) from empirical measurements of the optical properties of the eye. This process simulates the glare that would be experienced had the brightest points in the images really been shown at full intensity. The process has little effect except for bright point sources.

2.3 Procedure

Each illumination condition was run in a separate block and the order of the blocks was randomized across subjects. Within a block, subjects made 110 observations, one for each of the possible reflectances of the Test sphere. Hence, for a given value of specular reflectance, subjects would perform 11 matches (each with a different roughness). Conversely, for a given value of roughness, subjects would perform 10 matches (each with a different specular reflectance). The reflectances within a block were shown in random order.

Subjects could adjust both parameters simultaneously using the keyboard, and were given unlimited time. Subjects were informed by a beep if they tried to exceed the range of Match reflectances.

3. Results

3.1 Can Subjects Match Surface Reflectance Without Knowing The Specific Illumination?

Figure 12 shows example matching data for three subjects; each subject was matching spheres under a different real-world illumination. For each subject, matches for the specular reflectance parameter are plotted on top, with matches for roughness underneath. The x-axes represent the value of the Test sphere, the y-axes represent the subject's match. The grey level in the graph indicates density of responses, such that if a subject always provided the same match value for a given test value, the square would be white; the rarer the response, the darker the grey. The diagonal line shows ideal performance.

Figure 13 summarizes the complete data set, pooled across all subjects and all real-world illuminations. Again, matches for specular reflectance and roughness are plotted on separate graphs.

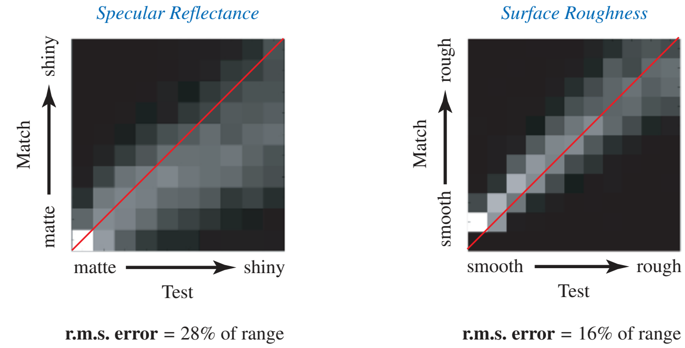

Figure 13. Matching data pooled across all subjects and all real-world illuminations. The two parameters are plotted separately. Veridical performance would fall along red lines. Grey level indicates density of subjects' responses. Root mean square error between subjects' matches and Test values can be expressed as a percentage of the range of Test values used.

The data show that subjects can match specular surface reflectance properties across variations in illumination fairly reliably and accurately. Specifically, subjects' matches are not independent of the Test value, as would be predicted if the subjects were incapable of estimating surface reflectance. This is important as it confirms our observation that the pattern of light within an object provides a cue to surface reflectance, despite the potential ambiguity of the image features. Such a strategy is available only because of the statistical regularities that

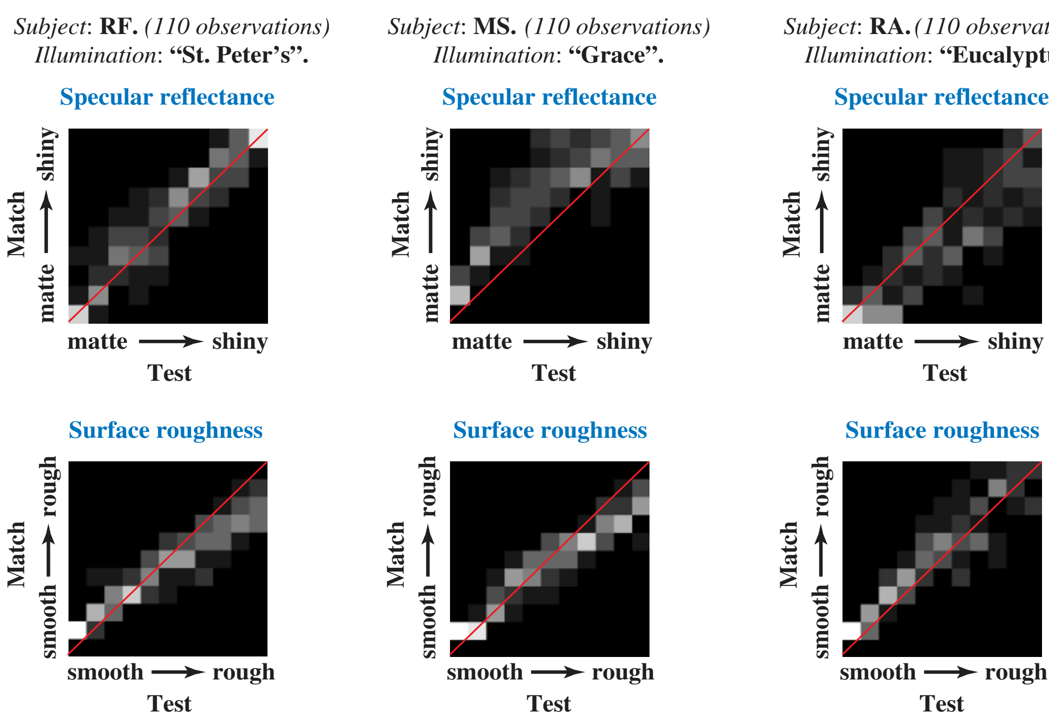

Figure 12. Examples of matching data from three subjects' viewing spheres under three real-world illuminations. Matches for the two parameters are plotted separately. Abscissa represents value of Test parameter, Match axis represents subject's estimate. Veridical performance would fall along red line. Grey level indicates density of subject's responses.

are conserved across real-world illuminations. Without the statistical regularities, the image features would be ambiguous and matching performance would be at chance across illuminations. That subjects can match surface reflectance properties accurately even though the images differ considerably on a pixel-by-pixel basis implies that they are using higher-level image features to perform the match. The finding also confirms our observation that gloss constancy does not require context, as long as the statistics of the illumination are typical of the real world.

3.2 Differences In Matching Performance Across Real-World Illuminations

Although matching performance is well above chance, there are statistically significant differences in matching performance across variations in illumination. Put another way, constancy is not perfect under our viewing conditions. For example, estimates of the specular reflectance parameter are systematically lower under the "Uffizi" illumination than under the "Galileo" illumination (see Figure 14). However, the fact that constancy is not perfect does not undermine our basic observations. That performance is better than chance (i) across illuminations and (ii) in the absence of context demonstrates that subjects can use higher-level image features to match surface reflectance properties. Furthermore, although certain statistical regularities are *well conserved* across real-world illuminations, we do not expect them to be *perfectly* conserved – residual differences in the statistics across illuminations ought to lead to biases in subjects' estimates of the surface reflectance properties. Thus, differences in matching performance are to be expected when subjects' assumptions are not perfectly satisfied.

3.3 How Accurate Are Subjects' Matches?

It is clear from Figures 12 and 13 that subjects are performing above chance. But exactly how well can subjects match surface reflectance in the absence of context? In order to quantify accuracy, we took the root mean squared (RMS) error between subjects' responses and the true values of the Test stimulus. This measure of accuracy can be expressed as percentage of the total range of values that we used in the experiments: the larger the percentage, the worse the performance.

The RMS error for the specular reflectance matches, pooled across all subjects and all real-world illuminations (see Figure 13) was 28% of the range of values we used.

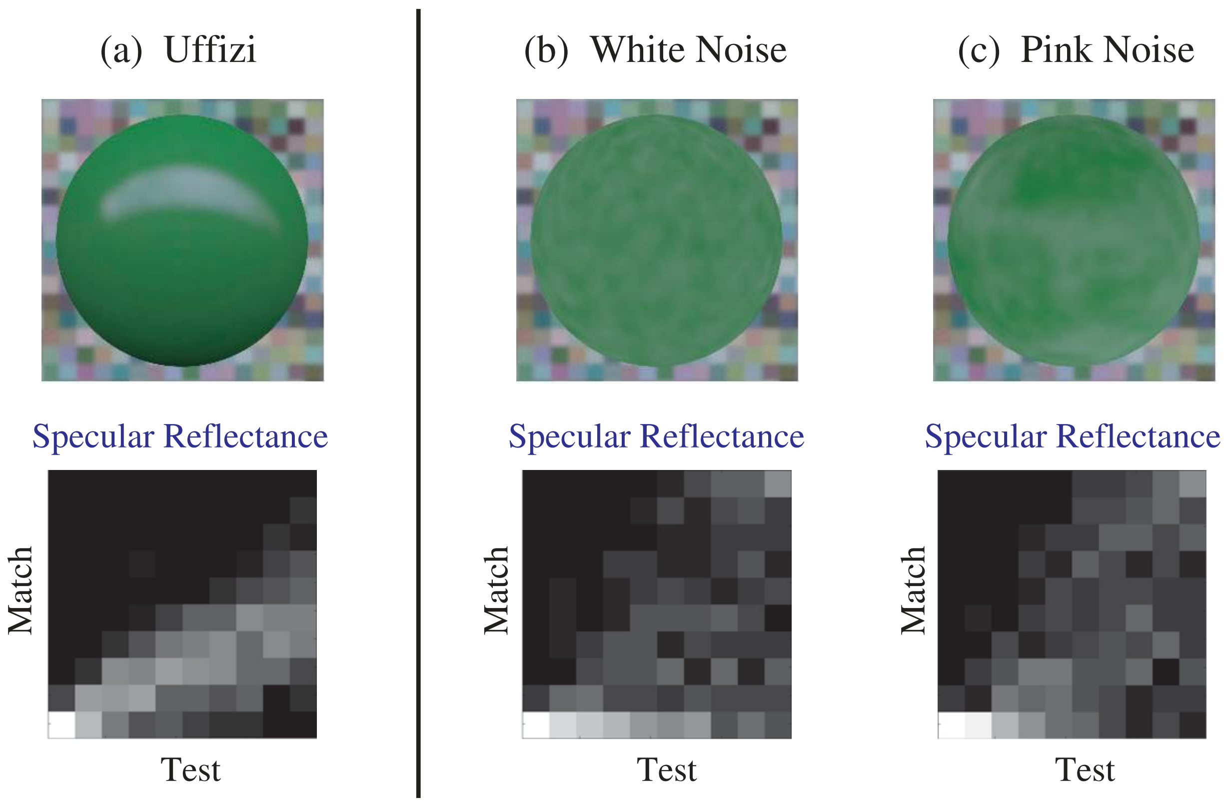

Figure 14. Matching performance under real-world and noise illuminations. (a) shows matches pooled across subjects for the Uffizi illumination, which yielded the least accurate performance of all the real-world illuminations. Poor performance reflects a systematic bias in matching. By contrast, performance for the noise stimuli shown in (b) and (c) is disorganized, presumably manifesting the difficulty subjects had in interpreting the patterns in the spheres as specular reflections.

This error represents the tendency for subjects to underestimate the specular reflectance of the Test surface relative to the Match surface seen in Figure 13 (i.e., the slope is less than 1). This tendency to underestimate specular reflectance appears to be partly due to a response bias that leads subjects to avoid the highest values on the scale. If there were no response bias, then swapping the illumination maps used for the Test and Match spheres should lead to a symmetrical change in the matching slope (i.e., slopes of less than 1 should become greater than 1). However, when subjects adjusted Match spheres viewed under the "Eucalyptus" illumination map to match Test spheres viewed under the "Galileo" illumination map, matching slopes were also less than 1, suggesting a response bias.

The RMS error for the roughness matches, pooled across all subjects and all real-world illuminations (see Figure 13), was 16% of the range of values we used.

3.4 Are the Parameters Perceptually Independent?

In Figures 12 and 13, matches for specular reflectance and roughness were plotted on separate graphs. This is only appropriate if the parameters are perceptually independent (i.e., if perceived specular reflectance is not a function of roughness and vice versa). When Pellacini et al. (2000) proposed their psychophysically uniform reparameterization of the Ward model, they reported that the two parameters are independent. Our data support this finding: there was no statistical dependence of perceived specular reflectance on surface roughness, nor of perceived roughness on surface specular reflectance, when the data were pooled across subjects and illuminations.

3.5 Comparison Between Real-World and Artificial Illuminations

Figure 15 shows matching error for each of the real-world and artificial illuminations, pooled across subjects. The red lines are the mean errors for the real-world illuminations. Subjects are generally less reliable and less accurate at matching surface reflectance properties under artificial illuminations (dark blue) than under real-world illuminations (light blue). One notable exception is for the illumination featuring an extended rectangular source (see Figure 11), for which matching performance is comparable to matches performed under real-world illumination for the roughness parameter.

Matching is especially disorganized for the white and pink noise stimuli. Figure 14 shows data pooled across subjects for the "Uffizi" illumination, which yielded the least accurate performance of the real-world illuminations. Although subjects' matches are inaccurate, their errors reflect a systematic bias, presumably resulting from some idiosyncratic statistics of that illumination. By contrast, matches for the noise illuminations are highly unreliable, as shown in Figure 14. It is likely that this unreliability reflects the difficulty that subjects experienced in interpreting these patterns as specular reflections. Subjects reported that the spheres viewed under noise illumination did not look glossy; some subjects also reported that the objects did not even look spherical, but rather flat and matte. Indeed, the example images shown in Figure 11 demonstrate that random patterns of illumination do not lead to distinct percepts of gloss.

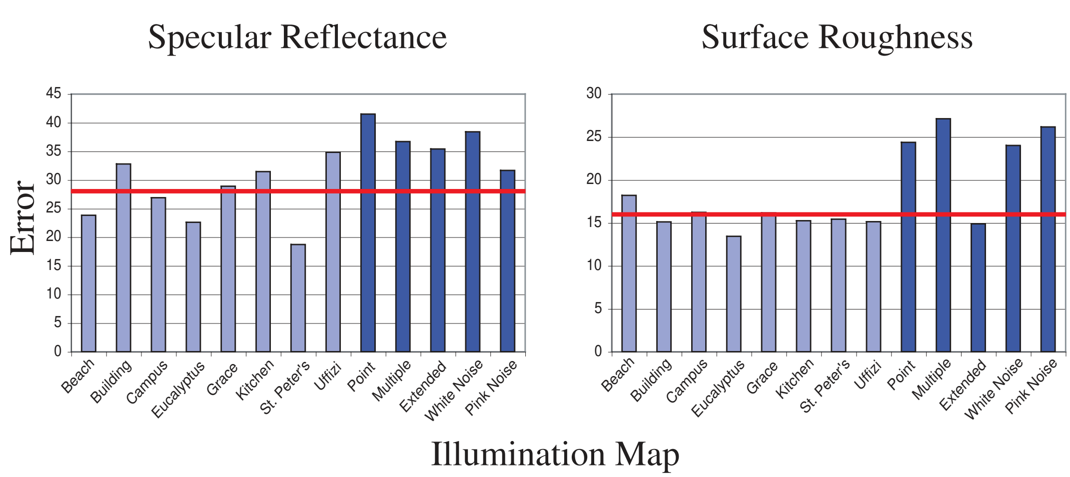

Figure 15. Comparison between matching performance for real-world and artificial illuminations. Error for two parameters is plotted separately. Real-world illuminations shown in light blue, artificial illuminations in dark blue. Error axis represents RMS error between subjects' match and Test value, expressed as a percentage of the total range of Test values used in the experiment. The red line indicates mean error for real-world illuminations.

4. Discussion

Taken together, these findings corroborate our initial observations. First, subjects can reliably match surface reflectance properties across variations in illumination, even though the images are quite different on a pixel-by-pixel basis. Matching performance across real-world illuminations is well above chance. This demonstrates that image features that are (i) more abstract than pixels and (ii) local to the image of the surface provide a reliable cue to surface reflectance properties across variations in illumination. If illumination were arbitrary, this would not be possible, as the origin of a given image feature would be ambiguous. Therefore, in order to perform the task, subjects must somehow exploit the statistical regularities of real-world illumination. We argue that subjects do this by tracking diagnostic image features that are well conserved across illuminations.

Second, subjects are better at estimating surface reflectance when the object under scrutiny is illuminated by a world with typical statistics (or at least by illuminations taken from the real world). When the illumination statistics are not representative of those found under ordinary viewing conditions, surface reflectance estimation is less accurate. This supports our hypothesis that subjects rely on stored assumptions about the statistics of the world, because performance deteriorates when the assumptions are infringed.

Third, as observed earlier, subjects can estimate surface reflectance directly from images of objects; they do not need to estimate the illumination precisely from the context. We know this because subjects can match surface reflectance reliably and accurately even when the precise conditions of illumination are unknown.

4.1 What Are The Stored Assumptions?

We have suggested that subjects should be able to use the distinctive patterns that are reflected from generic materials to estimate the surface reflectance of the materials. We argued that this is possible because of the statistical regularities of real-world illumination. We are left with the deeper question, however: *what* statistical properties do subjects exploit? What measurements does the visual system perform to estimate surface reflectance? In the "Introduction," we drew a parallel between surface reflectance estimation and texture recognition. Different samples of the same texture look similar even though on a pixel-by-pixel basis, the image changes from sample to sample. Likewise, the same material looks similar under different illuminations even though the pattern of reflection varies with the illumination. There must be some set of diagnostic features,[5] some set of statistical properties that is common across typical images of a given material or a given texture. The question is: what are the features?

The results of the matching experiment already provide some important clues. One obvious hypothesis is that the visual system looks for local highlights — the small very bright "first bounce"[6] specularities that result from the reflection of light arriving directly from luminous sources. Beck and Prazdny (1980) showed that a matte surface can be given a glossy appearance simply by adding a few local highlights. By contrast, our finding that point source illumination leads to poor surface reflectance estimates suggests that the visual system requires more varied or more extended features than local highlights in order to estimate surface reflectance. It is important to recall that the low dynamic range of the CRT limits the intensity of the highlights in our displays. It is possible that with higher dynamic range, performance would be somewhat improved under the point light sources. Nevertheless, no matter how intense a single highlight becomes, it will never possess the extended spatial structure that results from real-world illumination.

Recent work by Berzhanskaya, Swaminathan, Beck, and Mingolla (2002) suggests that perceived specular reflectance falls off with distance from the highlight. Could it be that local highlights *are* good cues to surface reflectance but that the impression fails to propagate across the whole surface? The result with the multiple point sources makes this seem unlikely, because matching was still poor even when more of the surface was "close to a highlight." We suggest that highlights are good for distinguishing glossy from matte surfaces (hence the Beck demonstration), but do not provide sufficient information to specify the *degree* of specular reflectance or the roughness of the surface.

The white noise illumination leads to detectable contrasts right across the surface of the object. However, we found that matches were also poor under white noise illumination, suggesting that the ubiquitous and varied contrasts are not sufficient features for estimating surface reflectance.

The fact that surface reflectance estimation is also poor under the pink noise illumination confirms this, but also rejects another hypothesis. The subject's stored assumptions about the statistics of real-world illuminations do not simply consist of information about spatial frequencies, as the pink noise illumination has a similar power spectrum to typical real-world illuminations and yet matching performance was poor. Clearly "structural" or "configurative" regularities are also important.

Of the spheres shown in Figure 11, the one illuminated under a single extended rectangular source looks more similar to the "real-world" spheres than the other "artificial" spheres. For comparison, Figure 16 shows example spheres illuminated under the Uffizi real-world illumination, under the extended artificial source, and under the pink noise illumination. The similarity in appearance between the sphere rendered under the

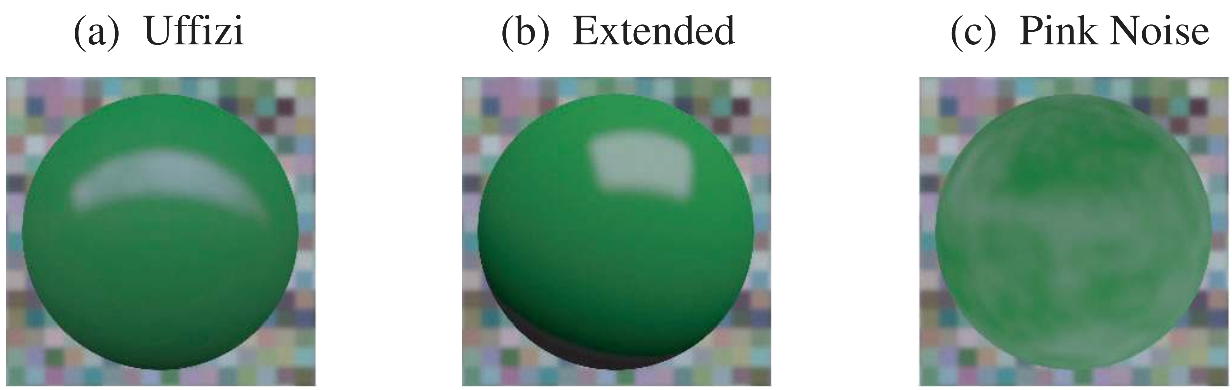

Figure 16. Example spheres illuminated under (a) real-world illumination, (b) artificial illumination featuring an extended rectangular source, and (c) pink noise illumination. Both (b) and (c) are synthetic illuminations, and yet the impression of surface reflectance properties is clearer for (a) and (b) than for (c); we suggest that this is because the extended source illumination shares important properties in common with real-world illuminations, while the pink noise illumination infringes many of the assumptions used by the visual system to estimate surface reflectance.

extended source and the Uffizi illumination is reflected in the matching results; accuracy under the extended source was comparable to the real-world illuminations, at least for the roughness parameter. There are three striking features of this illumination: (1) it has a dominant direction of illumination, unlike the noise illuminations; (2) it contains extended edges, unlike all the other artificial illuminations; (3) the edges are organized into a regular, meaningful shape. These are important candidate features that the visual system might require in order to estimate surface reflectance properties accurately. In the following section, we discuss these and other possible features.

4.2 Further Observations on the Features Underlying Surface Reflectance Estimation

We have discussed several image features that subjects may use to estimate surface reflectance properties. However, there are countless other features that may be important, ranging from the sharpness of the brightest edge to the presence of recognizable objects in the reflection. The most direct way to test the importance of a feature is to see if selectively manipulating that feature has a systematic effect on surface reflectance estimation.

4.2.1 Direct manipulations of the image

In Figure 17, we directly modify the brightest highlights in images of spheres rendered under one real-world and one artificial illumination. When we remove the brightest highlights from the image of a sphere rendered under the “St. Peter’s” illumination (Figure 17b), the result looks somewhat less glossy than the original (a). This is consistent with Beck and Prazdny’s (1980) observation, discussed above. However, it should be noted that the sphere does not appear uniformly matte. It is easy to interpret the remaining, lower-contrast patterns as reflections, suggesting that these features also play a role in the glossy appearance, as argued above. Likewise, when we blur the brightest highlights (Figure 17c), the resulting sphere appears somewhat rougher than the original. However, the effect does not extend uniformly across the entire surface, nor does the visual system attribute all of the blur to the environment. The sphere looks non-uniform in reflectance, but it still looks essentially like a glossy sphere, demonstrating that under real-world illumination many features act simultaneously to produce the impression of gloss.

With artificial illumination, manipulating the local features can have a much more pronounced effect. Figure 17d shows a sphere rendered under multiple, isolated point sources, as used in the matching experiment. When we remove the bright highlights, the sphere looks entirely matte (Figure 17e). This is, of course, because no other features are available to produce the impression of specularity. Likewise, when we blur the bright highlights, the entire sphere appears rougher (Figure 17f). Real-world illumination provides much richer specular reflections. In turn, the visual system has more features available with which to estimate the surface reflectance properties.

4.2.2 Manipulations of the Illumination

Doctoring the image directly allows us to test the role of specific local features in surface reflectance estimation. However, there are two advantages to manipulating the illuminations as opposed to adjusting the features of the rendered image directly. The first is that the results are guaranteed to be physically possible (within the limits of the display device). The second is that features of illumination are independent of the three-dimensional shape of the object being viewed, and thus we do not need to have a theory of shape perception to make statements about which properties of illumination are important for the perception of surface reflectance. To this end, we have rendered objects under a number of manipulated or fabricated illuminations to demonstrate the importance (or lack thereof) of various salient properties of illumination for surface reflectance estimation.

A priori, we would expect a given property of the illumination to be important for estimating surface reflectance if (a) the property is well conserved across illuminations so that it gives reliable information across instances, and (b) variations in surface reflectance systematically map that illumination property into detectable, reliable image features. Before showing surfaces rendered under manipulated and fabricated illuminations, it is instructive to review some of the most salient statistical regularities of illumination [see Dror, Leung, Willsky, & Adelson (2001) and Dror (2002) for a more thorough account]. Illuminations tend to have the following properties, which we have grouped by the complexity, and the extent to which they can be measured locally.

(1) Properties based on the raw luminance values

a. High dynamic range. The "Campus" illumination map (Figure 9c), for example, has a range of luminances spanning over three orders of magnitude (2000:1).
b. Pixel histograms that are heavily skewed toward low-intensity values.

(2) Quasi-local and nonlocal properties

a. Nearby pixels are correlated in intensity, such that power falls off at higher spatial frequencies. At higher spatial frequencies, amplitude typically falls off as 1/f, where f is the modulus of the spatial frequency.
b. Distributions of bandpass filter coefficients (e.g., wavelet coefficients) are highly kurtotic. In other words, large wavelet coefficients are sparse.
c. Approximate scale invariance. Distributions of wavelet coefficients are similar at different scales.
d. Although approximately decorrelated, wavelet coefficients exhibit higher-order dependencies across scale, orientation, and location. These dependencies reflect the presence of image features such as extended edges.

(3) Global / nonstationary properties

a. Dominant direction of illumination.
b. Presence of recognizable objects such as buildings and trees.
c. Cardinal axes corresponding to the ground plane and perpendicular structures erected thereupon.

We will now discuss the role of a number of these properties in surface reflectance estimation by rendering under systematically manipulated illuminations.

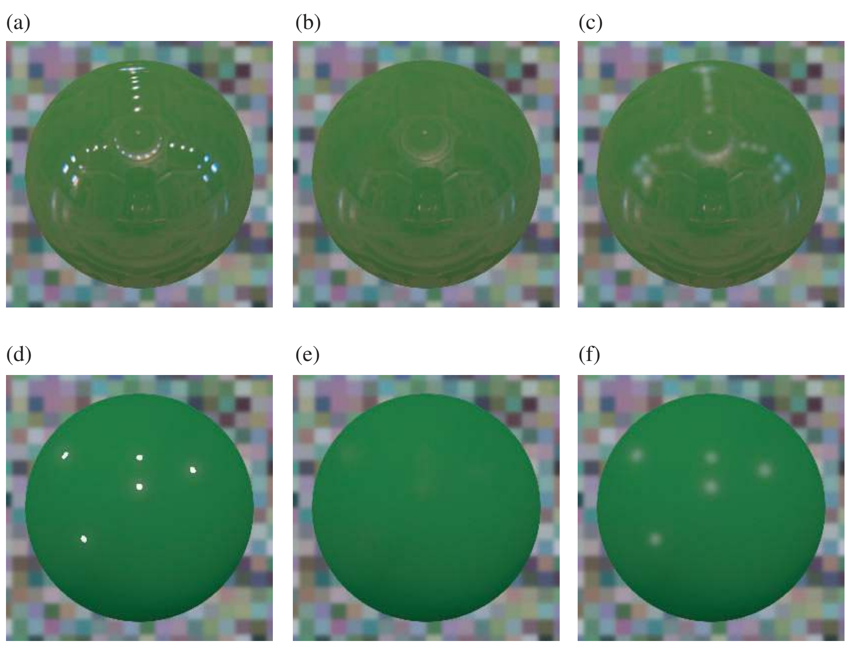

Figure 17. Direct manipulation of highlights. Original images are shown in (a) and (d). In (b) and (e), highlights have been removed; in (c) and (f), they have been blurred. The consequences are more pronounced for the artificial illumination than for the real-world illumination. Real-world illumination provides the visual system with many features with which to estimate surface reflectance, unlike the artificial illumination shown here.

4.2.3 Manipulations of the illumination histogram

Real-world illuminations tend to have pixel histograms with moderately well conserved higher-order statistics. One particularly salient feature of the distributions of illumination intensities found in the real world is that they are heavily skewed toward small values, such that the vast majority of pixels are many orders of magnitude darker than the few brightest. This reflects the fact that radiant sources in the real world are generally fairly compact. Is the distribution of intensities found in real-world illumination one of the stored assumptions humans use to estimate surface reflectance?

Consider the sphere shown in Figure 11(e), which was illuminated under the pink noise illumination. This sphere yields a poor impression of surface reflectance. The illumination map was synthesized to have a Gaussian pixel histogram, which is atypical of real-world illumination.[7] If the visual system assumes that the distribution of illumination intensities is generally heavily skewed, then the poor impression of surface reflectance associated with the Gaussian pink noise illumination may in part be due to the fact that it violates this assumption. We can test this hypothesis directly by enforcing a more realistic histogram on the illumination and rendering a new sphere. If the alteration improves the percept of surface reflectance, then this suggests that the visual system does expect objects to be illuminated by skewed distributions of light, as they tend to be in the real world.

Conversely, we can enforce a Gaussian (i.e., unrealistic) histogram on one of the real-world illuminations, and thus rob that illumination of one of its characteristic properties. If this property is important for surface reflectance estimation, then the modified illumination should yield poor impressions of surface reflectance, just as the Gaussian pink noise illumination does. The results of these two manipulations are shown in Figure 18.

Spheres rendered under the original versions of the illuminations are shown in (a) and (b), along with their pixel histograms.[8] By passing the intensities of the Gaussian noise illumination through a carefully chosen static nonlinearity, we can force the illumination to have a very similar pixel histogram to the Campus illumination shown in (a)[9]; this process is known as histogram matching. Rendering under this modified noise illumination yields the sphere shown in (d). On close inspection, it is clear to the observer that the world reflected in this sphere is made of meaningless clumps, rather than meaningful objects such as buildings and trees. However, at first glance, the sphere tends to give the impression of being spherical and glossy, as opposed to relatively flat and matte, as is the case in (b). This suggests that simply by skewing the illumination

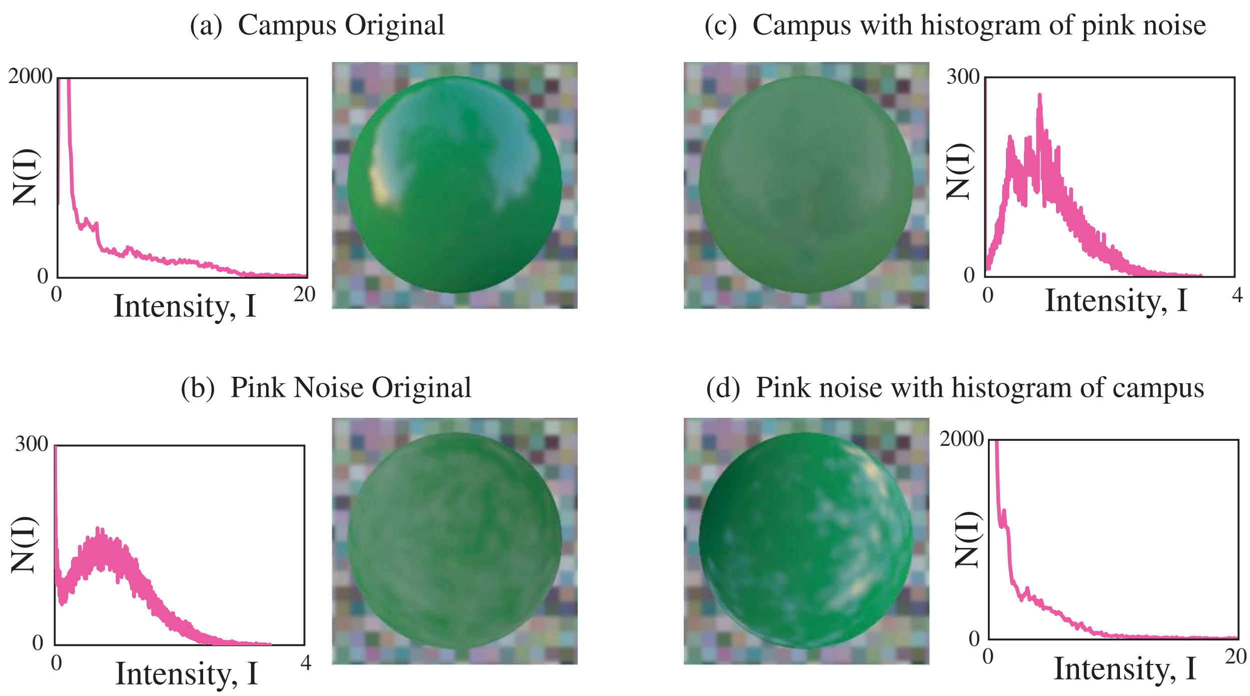

Figure 18. Spheres rendered under illuminations with modified pixel histograms. The sphere in (a) was rendered under the original Campus illumination, which has a heavily skewed pixel histogram. The sphere in (b) was rendered under pink noise illumination with a truncated Gaussian pixel histogram. These illuminations were then modified using histogram matching. The sphere in (c) was rendered under modified Campus illumination with an approximately Gaussian histogram derived from (b). The sphere in (d) was rendered under modified noise illumination with a histogram derived from (a). Note the difference in scale on the pixel histogram plots; the original Campus is considerably more skewed than the original Noise.

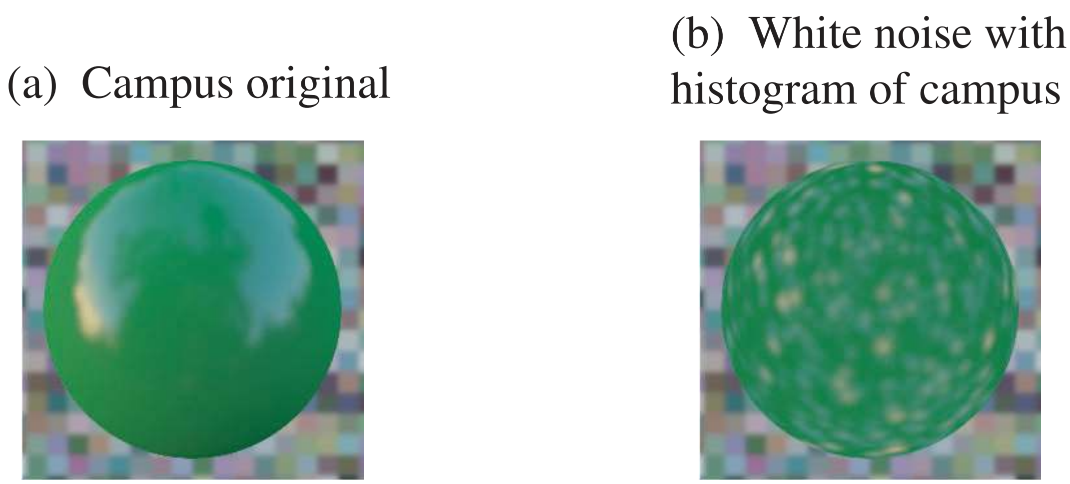

Figure 19. Sphere (a) was rendered under an unmodified real-world illumination. Sphere (b) was rendered under white noise illumination that had been modified using histogram matching to have the same pixel histogram as the illumination in (a). Unlike the modified pink noise shown in Figure 18, the modified white noise does not lead to compelling impressions of gloss.

histogram, we have satisfied one of the major assumptions held by the visual system about the statistics of real-world illumination.

Conversely, when we force the Campus illumination to have a Gaussian histogram, the corresponding sphere, shown in (c), appears dull and flat. Scrutiny reveals that the sphere contains the same pattern as in (a), but it lacks the depth and shading associated with a skewed pixel histogram. These two demonstrations suggest that a skewed distribution of illuminant intensities is a necessary condition for perceiving surface reflectance. It seems likely that this is one of the stored assumptions held by the visual system.

Why might a skewed illumination histogram lead to good impressions of surface reflectance? Although only a small proportion of the world is very bright, it is in fact those few bright sources that are responsible for the majority of the light that is reflected from a surface. A skewed illumination histogram allows the few brightest directions to dominate the image, leading to good shading information from the diffuse component of reflectance, and bright, localized highlights from the specular component. Moreover, the skew tends to increase the contrast between the darkest and brightest regions of the image. This could be important for distinguishing spatial variations in illumination (i.e., highlights) from spatial variations in the intrinsic reflectance of the material (i.e., surface texture). In the real world, variation in pigment reflectance spans a range of about 30:1, whereas first-bounce highlights can be many orders of magnitude brighter than their surroundings. We argue that under illuminations with skewed histograms, the visual system more readily interprets the pattern as reflections and thus better estimates the intrinsic properties of the surface.

Although the visual system seems to expect a skewed illumination histogram, it is by no means sufficient alone. This is demonstrated in Figure 19. This sphere was rendered under modified white noise. As before, the illumination was given the same pixel histogram as the Campus illumination, and yet this time the impression of a glossy surface is much less vivid. The reason for the difference between the modified pink noise and the modified white noise is due to the spatial distribution of the brightest pixels. In pink noise, the intensity of neighboring pixels is correlated, and thus there is a good chance that the brightest pixels will be clumped together in space. When these pixels are made much brighter than the rest by the histogram matching process, they form a directional source that leads to vivid shading and highlights, as discussed above. By contrast, the brightest pixels in white noise are randomly distributed about the illumination map. When these pixels are "boosted" by the histogram matching process, they do not aggregate into a predominant direction of illumination, but rather add bright light from many directions at once. This leads to poor shading information and lower contrast, and hence a weaker impression of a glossy surface.

4.2.4 The role of illumination wavelet statistics

If the reason that the modified white noise illumination yields poor impressions of surface reflectance is because it lacks some of the spatial structure of real-world illuminations, then synthesized illuminations that share such spatial structure should lead to compelling percepts of gloss. In order to test this, we need a method to describe the relevant properties of spatial structure. We suggest that histograms of wavelet coefficients at various scales and orientations may serve as a formal measure of this property for the following reasons. First, wavelet histogram properties are well conserved across real-world illuminations (Dror et al., 2001) and therefore lead to reliable cues. Second, wavelet histograms offer a means to capture some of the important spatial structure of illuminations, including that captured by power spectra. However, wavelets are more powerful than power spectra in that they capture the effects of local image features, such as edges. Third, because of their local, multi-scale nature, illuminations that are synthesized to have constrained wavelet histograms will exhibit structure at all scales and contrasts. Put another way, wavelet histograms can capture the relatively low-contrast structure that results from secondary sources in the environment (i.e., non-emitting surfaces such as walls, trees, and people) as well as structure resulting from bright light sources. Thus, wavelet statistics represent reliable features of intermediate complexity, which capture a number of important quasi-local properties of illumination. It is also worth noting that a computer vision system can perform reflectance estimation using the statistics of the wavelet and pixel histograms considered here (Dror, Adelson, & Willsky, 2001).

Heeger and Bergen (1995) provide an iterative algorithm for synthesizing textures with specified pixel histograms and with specified histograms of the wavelet coefficients at each scale and orientation. We can use their texture synthesis algorithm to generate new illuminations that are constrained to have the same pixel and wavelet coefficient distributions as real-world illuminations. If such histograms capture the types of features that the visual system expects from glossy surfaces, then these illuminations should yield compelling impressions of surface reflectance.

Figure 20 shows spheres rendered under eight such illuminations. Each illumination was generated from a different random initial state and was forced by the algorithm to have the same pixel and wavelet coefficient histograms as the real-world illumination with the corresponding name. Thus, the first sphere, for example, was rendered under a synthetic illumination with the same pixel and wavelet coefficient histograms as the Beach illumination (which was captured photographically from the real world).

On close inspection, it is clear that the world reflected in each of these spheres does not contain meaningful objects such as houses and people, and yet, at first glance, the impression of gloss is quite compelling in most of the cases. Nothing about the synthesis process forces the bright sources to organize themselves into regular or naturalistic configurations; none of the global properties of real-world illumination are captured, such as the fact that in the real-world light generally comes from above. In addition, the synthesis process does not capture features such as extended straight edges; such features

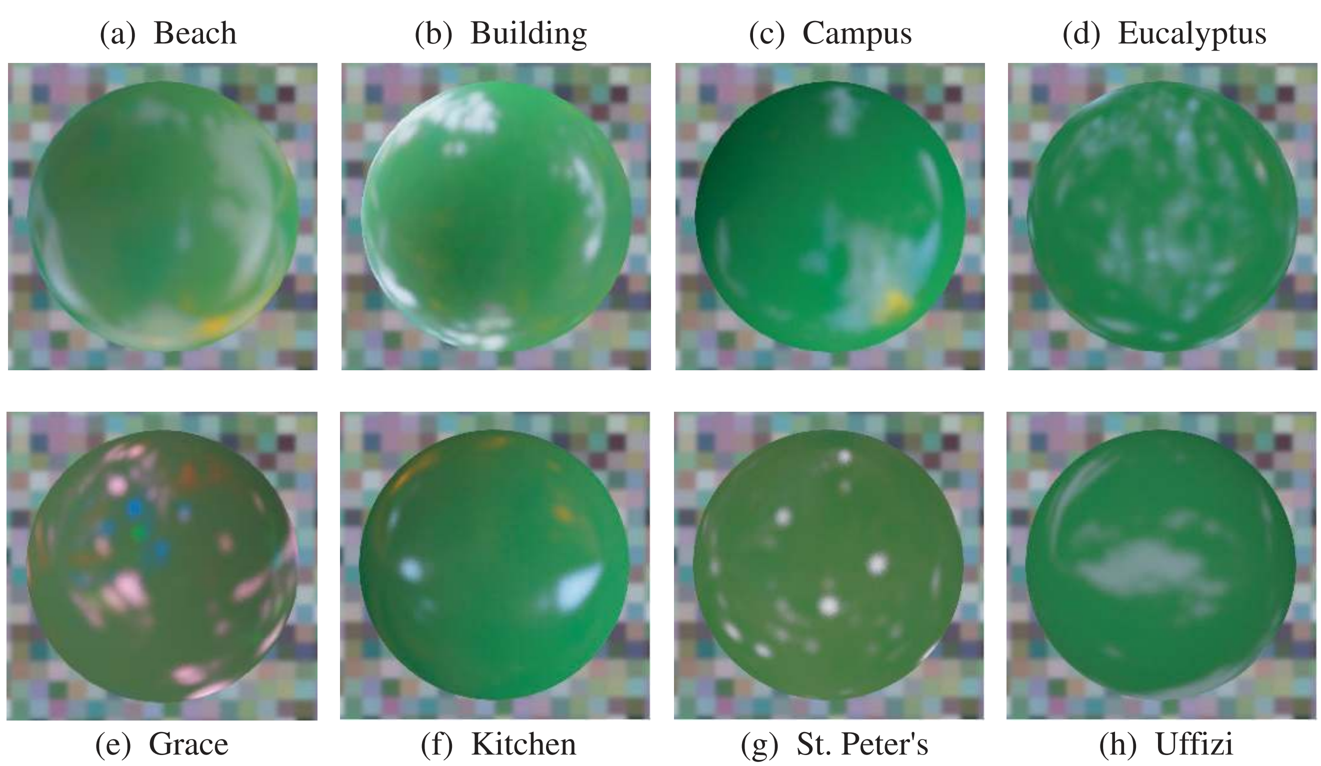

Figure 20. Spheres rendered under synthetic illuminations with same wavelet and pixel histograms as real-world illuminations. Each illumination was synthesized from a random initial state by an iterative procedure that constrains the wavelet and pixel histograms to be the same as a given real-world illumination. The statistics of each illumination were matched to the real-world illumination denoted in the title (cf. Figure 9).

result in interdependencies between specific wavelet coefficients at different scales, but are not described by the histograms we used. Despite this, observers agree that the spheres do generally lead to a compelling sense of surface reflectance.

This has two consequences for our theory of the stored assumptions about illumination used by the visual system to estimate surface reflectance. First, we can infer that wavelet properties capture some of the essential features of illumination expected by the visual system. Second, although real-world illumination certainly contains higher-order regularities (e.g., cardinal axes), the visual system does not *require* these to pertain in order for an object to yield a clear impression of gloss. Specifically, it does not seem to be important for the structures of the environment to be organized into *recognizable* forms. Enforcing additional, higher-order regularities would no doubt yield even better impressions of surface reflectance. For example, objects illuminated from above tend to look more "normal" or "realistic" than those lit from behind or below.[10] However, that the spheres in Figure 20 look quite compelling without enforcing additional constraints suggests that a number of important assumptions have already been captured, and that additional constraints would yield diminishing returns.

5. Conclusions

Recognizing materials by their reflectance properties is difficult because the image of a material depends not only on the material but also on the conditions of illumination. Many combinations of illumination and material are consistent with a given image, and yet we usually have a clear and unique impression of the material attributes of an object. The results of the matching experiment suggest that this aptitude does not require knowledge of the specific conditions of illumination, as subjects can accurately perceive surface reflectance in the absence of contextual information to specify the illumination. Indeed, we have demonstrated that it is possible to vary the context considerably with little effect on the apparent glossiness of the surface.

We have argued that subjects use tacit knowledge of the statistics of real-world illumination to eliminate improbable image interpretations. Our claim is that the statistical regularities of real-world illumination manifest themselves as diagnostic image features that can be reliably interpreted as resulting from a given surface reflectance. Thus, the recognition of glossy surfaces can be treated as a problem analogous to texture recognition. Our demonstrations suggest that surface reflectance properties are clearer when objects are viewed under real-world illuminations than when they are viewed under atypical illuminations such as a single point light source. This observation is supported by our finding that subjects are poorer at matching reflectance properties under illuminations with atypical statistics than under real-world illuminations. We have also identified some of the properties of illumination that lead to reliable image features. Localized point sources and random noise patterns yield poor estimates of surface reflectance. Mimicking the power spectrum of real-world illumination is insufficient to create a compelling impression of gloss. By contrast, extended edges and a predominant direction of illumination tend to lead to good impressions of gloss.

Direct manipulation of the highlights in the image suggests that under ordinary viewing conditions, many features play a role in the perception of gloss, and not just local highlights. By manipulating the conditions of illumination systematically, we have identified additional properties of illumination that are important for human surface reflectance estimation. We have demonstrated that some important properties of illumination can be captured by relatively simple measurements using the pixel histograms and wavelet coefficient histograms of illumination maps. This suggests that the visual system's stored assumptions include local spatial properties of intermediate complexity, as opposed to complex, global, nonstationary, or configurative properties, such as cardinal axes of orientation and the organization of environmental structures into recognizable forms. Although higher-order regularities found in the environment are likely to facilitate realism, they are not *required* for compelling impressions of surface reflectance.

Appendix

We observe that surfaces viewed under real-world illumination appear remarkably constant across changes in the background against which they are viewed (Figure 5). This can be contrasted with the dramatic role that context plays in lightness perception. Our explanation for this discrepancy is that observers estimate surface reflectance directly from the image of the object under scrutiny and thus do not require context to provide an estimate of the prevailing illumination. In this section, we specify more precisely how much surface reflectance constancy can be achieved without context.

Without context, the visual system can estimate the distribution of light scatter (e.g., the ratio of specular to Lambertian reflectance) directly from the image. However, without context, it cannot even in principle estimate the overall scaling factor for the reflectance distribution, because this is confounded by the overall intensity of the illumination.

This point is illustrated in Figure 21, where we consider only surfaces whose reflectance is a combination of a monochromatic Lambertian component and a monochromatic perfect specular (mirrored) component (additional dimensions are also possible). Surfaces along the Lambertian axis vary from matte black to matte white, while surfaces along the specular axis vary from matte black, through glossy black (like a black billiard ball) to perfectly mirrored (like chrome). Our claim is that the visual system can distinguish two surfaces without context, as long as they do not fall on a straight line that passes through the origin of this space. Put another way, it is possible to estimate the ratio of Lambertian to specular reflectance, but not the total proportion of reflected light. This is because the same image could be produced by simultaneously halving the absolute reflectance of the surface and doubling the illuminant intensity. Thus, in the absence of context, the visual system could solve the problem up to a one-dimensional scaling ambiguity. To resolve the remaining ambiguity requires context.[11]

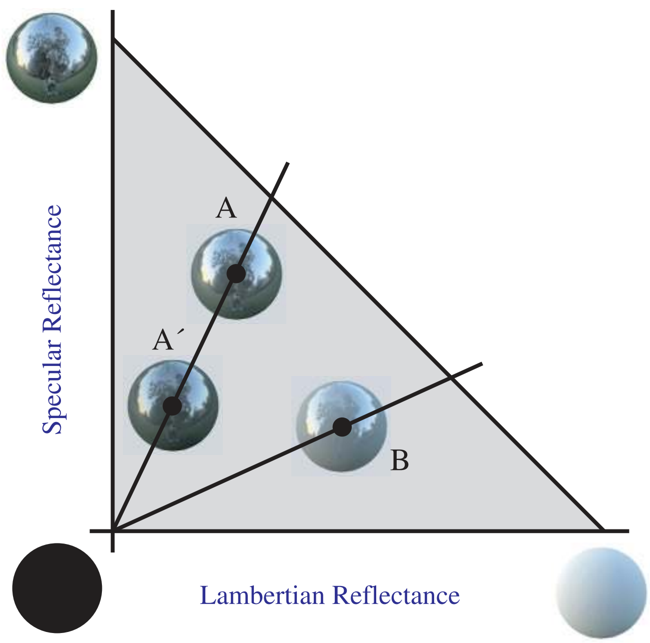

Figure 21. Ambiguity in surface reflectance estimation without context. Physical reflectances are constrained to fall within the grey triangle. Without context, surface reflectance can be estimated up to an unknown scale factor. Thus, two reflectances can be distinguished without context as long as they do not fall on a straight line that passes through the origin. Hence, the visual system could distinguish spheres A and B, but cannot tell A and A´ apart without context. This is a two-dimensional example, but the principle holds for arbitrary dimensions.

The reflectances of the Lambertian materials considered by Gelb (1929) and others fall along one axis of the space in Figure 21. Because these surfaces differ only by a scaling factor, the visual system requires context to resolve them. Thus our claim is not at odds with previous claims about the role of context in lightness perception. Rather, by considering a wider range of materials, we make explicit the possible reduction in ambiguity that the visual system could achieve without context.

Acknowledgments

This research was supported by National Institutes of Health Grant EY12690-02 to E.H.A., a Nippon Telegraph and Telephone Corporation grant to the MIT Artificial Intelligence Lab, a contract with Unilever Research, a National Defense Science and Engineering Graduate Fellowship and a Whitaker Fellowship to R.O.D., and a Hugh Hampton Young Fellowship to R.W.F.

Commercial relationships: none.

Footnotes

1. Specifically, what is visible is a portion of the spherical illumination map in which the object was placed in order to render it.
2. Additional parameters are required to represent reflection as a function of wavelength. For plastics and other dielectrics, only diffuse reflectance varies with wavelength: specular reflection varies little with the wavelength of the incident light. However, for metals such as gold, the specular reflectance can be wavelength sensitive.
3. Specifically, the reparameterization of specular reflectance was such that $\rho_S = (c + (\rho_D/2)^{1/3})^3 - \rho_D/2$. To compute ρ_D, we projected the red, green, and blue channels of ρ_D in the Ward model onto the CIE Y dimension using the values specified in the RADIANCE documentation. The reparameterization of surface roughness was such that $\alpha = 1\text{-}d$.
4. The high dynamic range light probe images are available online at http://www.debevec.org/Probes.
5. We intend the term “features” to refer to any measurable property of the image, and not simply local image tokens such as edges and junctions.
6. The term “first-bounce” refers to the fact that these reflections are the first time that the light bounces after leaving the luminous source.
7. Note that pixel values cannot be negative, so the distribution of intensities is truncated at zero, and therefore is not strictly Gaussian, although we refer to it as such for brevity. This is not the case for wavelet coefficients, however, which can be negative as well as positive.
8. It should be noted that the axes of the histograms are on different scales – the real world histogram has a much longer tail than the pink noise histogram, indicating that a few points in the map are much brighter than the majority.
9. No steps were taken to preserve the spatial frequency content of the modified noise stimulus, and thus there is no guarantee that it has a 1/f amplitude spectrum. Nevertheless, the modified illumination is still essentially “noise” because it was generated by a nondeterministic

process, and it contains none of the phase structure of a typical real-world illumination.

10. Conversely, objects illuminated from below or behind can be given an unnatural appearance, a fact exploited in film- and stage-lighting for dramatic effect.

11. The visual system could in principle learn priors on the scaling factor of particular materials (e.g., purely specular surfaces are more likely to be metals than black billiard balls, and therefore more likely to have a large scaling factor), although we know of no evidence so far that this plays a role in perception.

References

Adelson, E. H. (1999). Lightness perception and lightness illusions. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 339–351). Cambridge: MIT Press.

Adelson, E. H. (2001). On seeing stuff: The perception of materials by humans and machines. In *Proceedings of the SPIE*, *4299*, 1—12.

Beck, J., and Prazdny, S. (1980). Highlights and the perception of glossiness. *Perception & Psychophysics*, 30, 407—410. [PubMed]

Berzhanskaya, J., Swaminathan, G., Beck, J., & Mingolla, E. (2002). Highlights and surface gloss perception [Abstract]. *Journal of Vision*, 2(7), 93a, http://journalofvision.org/2/7/93/, DOI 10.1167/2.7.93. [Abstract]

Debevec, P. E. (1998). Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. *Proceedings of SIGGRAPH, 1998,* 189-198.

Debevec, P. E., Hawkins, T., Tchou, C., Duiker, H.-P., Sarokin, W., and Sagar, M. (2000). Acquiring the reflectance field of a human face. *Proceedings of SIGGRAPH, 2000,* 145-156.

Dror, R. O. (2002). *Surface reflectance recognition and real-world illumination statistics*. Doctoral Thesis, Massachusetts Institute of Technology. [Published as AI Lab Technical Report, AITR-2002-009: ftp://publications.ai.mit.edu/ai-publications/2002/AITR-2002-009.pdf] http://www.ai.mit.edu/research/publications/

Dror, R. O., Adelson, E. H., and Willsky, A. S. (2001). Surface reflectance estimation and natural illumination statistics. In *Proceedings of IEEE Workshop on Statistical and Computational Theories of Vision*, Vancouver, Canada.

Dror, R. O., Leung, T., Willsky, A. S., and Adelson, E. H. (2001). Statistics of real-world illumination. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Lihue, HI.

Fleming, R. W., Dror, R. O., and Adelson, E. H. (2001). How do humans determine reflectance properties under unknown illumination? In *Proceedings of IEEE Workshop on Identifying Objects Across Variations in Lighting*, Lihue, HI.

Gelb, A. (1929). Die "Farbenkonstanz" der Sehdinge [Colour constancy of visual objects]. In W. A. von Bethe (Ed.), *Handuch norm. und pathol. Psychologie* (pp. 594—678). Berlin: Springer.

Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. *Science*, *195*, 185—187. [PubMed]

Gilchrist, A. L. (1979). The perception of surface blacks and whites. *Scientific American*, *240*, 112—123. [PubMed]

Gilchrist, A. L. (Ed.). (1994). *Lightness, brightness and transparency*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Hartung, B., and Kersten, D. (2002). Distinguishing shiny from matte [Abstract]. *Journal of Vision, 2(7), 551a,* http://journalofvision.org/2/7/551/, *DOI 10.1167/2.7.551*. [Abstract]

Heeger, D. J., and Bergen, J. R. (1995). Pyramid-based texture analysis/synthesis. Proceedings of SIGGRAPH. *1995,* 229—238.

Katz, D. (1935). *The world of colour*. London: Kegan Paul, Trench, Trubner.

Land, E. H., and McCann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America A*, *61*, 1—11. [PubMed]

Nishida, S., and Shinya, M. (1998). Use of image-based information in judgments of surface-reflectance properties. *Journal of the Optical Society of America A*, *15*, 2951–2965. [PubMed]

Pellacini, F., Ferwerda, J. A., and Greenberg, D. P. (2000). Toward a psychophysically-based light reflection model for image synthesis. *Computer Graphics, 34*(2), 55-64.

Tumblin, J., Hodgins, J. K., and Guenter, B. K. (1999). Two methods for display of high-contrast images. *ACM Transactions on Graphics*, *18*, 56–94.

Ward, G. J. (1992). Measuring and modeling anisotropic reflection. *Computer Graphics*, *26*(2), 265—72.

Ward, G. J. (1994). The RADIANCE lighting simulation and rendering system. *Computer Graphics, 28*(2), 459-472.

Ward Larson, G., Rushmeier, H., and Piatko, C. (1997). A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics*, 3(4), 291-306.