Surface gloss and color perception of 3D objects

BEI XIAO[1] AND DAVID H. BRAINARD[2]

[1]Department of Neuroscience, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania
[2]Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania

Abstract

Two experiments explore the color perception of objects in complex scenes. The first experiment examines the color perception of objects across variation in surface gloss. Observers adjusted the color appearance of a matte sphere to match that of a test sphere. Across conditions we varied the body color and glossiness of the test sphere. The data indicate that observers do not simply match the average light reflected from the test. Indeed, the visual system compensates for the physical effect of varying the gloss, so that appearance is stabilized relative to what is predicted by the spatial average. The second experiment examines how people perceive color across locations on an object. We replaced the test sphere with a soccer ball that had one of its hexagonal faces colored. Observers were asked to adjust the match sphere have the same color appearance as this test patch. The test patch could be located at either an upper or lower location on the soccer ball. In addition, we varied the surface gloss of the entire soccer ball (including the test patch). The data show that there is an effect of test patch location on observers' color matching, but this effect is small compared to the physical change in the average light reflected from the test patch across the two locations. In addition, the effect of glossy highlights on the color appearance of the test patch was consistent with the results from Experiment 1.

Keywords: Color vision, Color appearance, Visual psychophysics, Surface gloss

1. Introduction

In daily life, we regularly judge the color appearance of three-dimensional (3D) objects. For example, it is not difficult to report that the mugs shown in Fig. 1 appear blue. The introspective ease of such judgments, however, belies the fact that the luminance and chromaticity of light reflected from an object can vary considerably across its surface. This variation arises because of geometrical variation in the incident illumination, coupled with the fact that the light reflected from objects depends on the direction of both the incident and reflected rays. Thus even when an object is made from a single homogeneous material, its shape can interact with the pattern of illumination to produce variation in the reflected light. For this reason, it is not trivial to perceive objects as having a unified color appearance: to do so requires aggregating information across the object's surface.

The degree to which reflected light varies across an object's surface depends on its material properties. The two rendered mugs depicted in Fig. 1 are both made of a homogeneous material. The mug on the left is matte, however, while the mug on the right is glossy. There is more variation across the surface of the glossy

mug than the matte mug, even though all other properties of the scene are held constant.

There is a large literature on the color appearance of flat matte surfaces. Some of this comes under the rubric of color appearance (Wyszecki, 1986; Shevell, 2003) and of color order systems (Derefeldt, 1991; Brainard, 2003). A second literature emphasizes the effects of changing illumination (Helson, 1940; Burnham et al., 1957; McCann, 1976; Arend & Reeves, 1986; Brainard & Wandell, 1992; Foster & Nascimento, 1994; Brainard, 1998; Bauml, 1999; Delahunt & Brainard, 2004; Hansen et al., 2007) or effects of scene geometry on the appearance of flat test surfaces (Bloj et al., 1999; Yang & Maloney, 2001; Khang & Zaidi, 2002; Boyaci et al., 2003, 2004; Doerschner et al., 2004; Ripamonti et al., 2004). The stimulus conditions of these studies eliminate variation in reflected light across the surface being judged.

A few studies examine how observers perceive the material from which an object is made, based on properties such as surface gloss (Hunter & Harold, 1987; Pellacini et al., 2000; Fleming et al., 2003; Obein et al., 2004; Fleming & Büthoff, 2005) and whether the visual system extracts information carried by specular highlights (Beck, 1964; Hurlbert et al., 1989; Yang & Maloney, 2001). Not much is known about the interaction of color appearance and the material from which an object is made; only a few studies examine this question in the more restricted domain of lightness perception (Nishida & Shinya, 1998; Todd et al., 2004; Motoyoshi et al., 2007). We return to the relation between our work and this literature in the discussion.

Address correspondence and reprint requests to: Bei Xiao, Department of Neuroscience, School of Medicine, University of Pennsylvania, Suite 330C, 3401 Walnut Street, Philadelphia, PA 19104. E-mail: beixiao@mail.med.upenn.edu
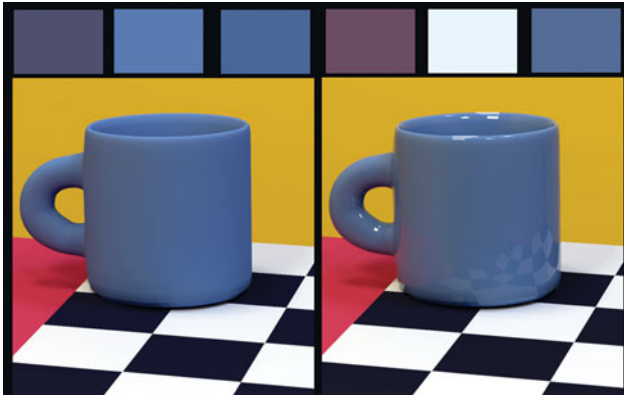
Fig. 1. Variation in reflected light across object surfaces. The left panel shows a graphics simulation of a matte mug in a synthetic scene, while the right panel shows a rendering of a glossy mug in the same scene. Both mugs have the same diffuse component, but differ in their glossiness. The inset at the top of each panel shows an expanded view of the image at three individual pixels. The pixels for both panels are taken from corresponding mug locations. For both mugs the inset shows that the reflected light varies across the mug. This effect is larger for the glossy mug. The figure is reproduced from Xiao and Brainard (2006) and is copyright ACM 2006.

Here we report exploratory experiments designed to clarify how people perceive the color of 3D objects. In Experiment 1, we studied how varying surface gloss affects color appearance. Observers adjusted the color of a matte sphere to match that of a test sphere. Across conditions, we varied the body color and glossiness of the test sphere.

One simple hypothesis about how observers integrate luminance and chromaticity across an object is that they take the spatial average of the reflected light. The data falsify this simple hypothesis. Indeed, the visual system compensates for the physical effect of varying the gloss, so that appearance is stabilized relative to what would be predicted under the averaging hypothesis.

In Experiment 2, we modified the stimuli to allow us to control what part of an object the observers judged. We replaced the test sphere with a soccer ball that had one of its hexagonal faces colored. This was the test patch. Observers were asked to adjust the color of a match sphere to be the same as that of the test patch. The match sphere was always matte. Across conditions, the test patch was presented at either an upper or lower location on the soccer ball, and the glossiness of the whole soccer ball was varied. The data show that there is an effect of test patch location on observers' color matching, but this effect was smaller than that predicted by the physical shift in average light reflected from the test patch. The effect of surface gloss on the color appearance of the test patch was similar to that found in Experiment 1.

2. General methods

2.1. Stimulus configuration

2.1.1. Surface reflectance and scene content

Observers viewed stereo image pairs of a rendered room. Left-eye images of the rendered room are shown in Fig. 2 (Experiment 1) and Fig. 8 (Experiment 2). In Experiment 1, the test object was a sphere, while in Experiment 2, the test object was a soccer ball. Observers in Experiment 1 judged the color of the test
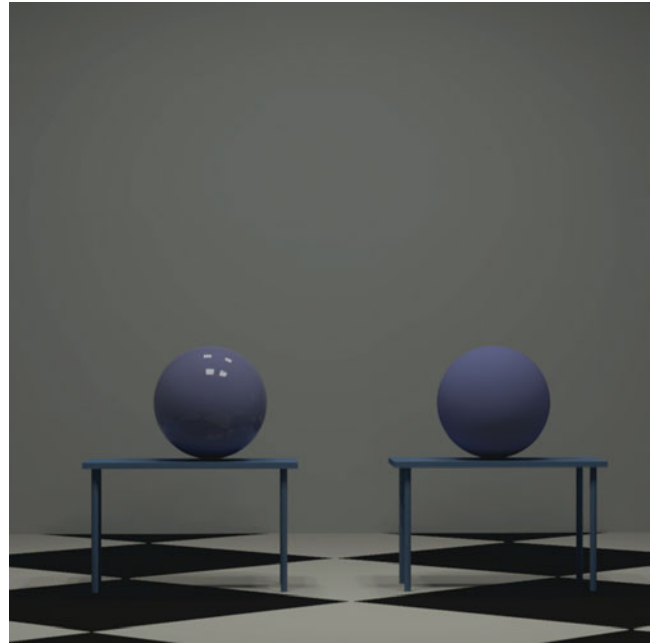
Fig. 2. Example of experimental image for Experiment 1. The test sphere is on the left and the match sphere is on the right. The surface reflectance properties of the match sphere varied from trial-to-trial. The match sphere was always match, and its diffuse reflectance was adjusted by observers. The displayed images in this and the other pictorial figures were rendered using our wavelength-by-wavelength rendering process and then converted to the sRGB color space (sRGB Standard, 1999) for publication (online version). The sRGB version was then converted to a CMYK representation (print version).

sphere as a whole, while those in Experiment 2 judged the color of a test patch consisting of one face of the soccer ball. Additional details about the scenes and images are provided for each experiment below.

The surface reflectance of all objects in the scenes conformed to the isotropic version of Ward light reflection model (Ward, 1992). This model represents surface reflectance as the sum of two components, diffuse and specular:

$$\rho(\theta_i, \phi_i, \theta_o, \phi_o, \lambda) = \frac{\rho_d(\lambda)}{\pi} + \rho_s \cdot \frac{\exp[-\tan^2 \delta/\alpha^2]}{4\pi\alpha^2 \sqrt{\cos \theta_i \cos \theta_o}}, \quad (1)$$

where $\rho(\theta_i, \phi_i, \theta_o, \phi_o, \lambda)$ is the surface BRDF, $\rho_d(\lambda)$ is the diffuse reflectance, which depends on wavelength, ρ_s governs the strength of the specular component, and α is a roughness parameter that describes the spread of the specular highlight. Angles θ_i and ϕ_i describe the direction of a light ray incident to the object, while θ_o and ϕ_o describe the direction of the reflected light. The quantity δ is computed from the four angles as described in Ward (1992). We varied $\rho_d(\lambda)$ to control the body color of objects, and varied ρ_s and α to control surface gloss.

2.2. Apparatus

Stimuli were presented to the observers stereoscopically, using two CRT monitors (Experiment 1 except Observer CGE, Hewlett-Packard, Palo Alto, CA, Model p1130; Experiment 2 and Observer CGE in Experiment 1, Viewsonic, Walnut, CA, Model G220fb). The monitors were driven at a 75 Hz refresh rate at 1152 by 870

spatial resolution, and with a color lookup table providing 14-bit (psuedo-color) resolution for each color channel (BITS++, Cambridge Research Systems, Rochester, England). Images delivered to the left and right eye were rendered separately using viewpoints separated horizontally by 6.3 cm.

We chose to use a stereo apparatus for two reasons. First, some previous studies (Yang & Shevell, 2002) have shown that adding binocular disparity improves color constancy in scenes containing specular highlights. Second, interpretation of the pattern of light reflected from simulated 3D objects could reasonably be expected to depend on the ability to see the 3D structure. We felt adding binocular disparity to the stimuli would maximize the veridicality of the perceived 3D structure. We did not systematically manipulate cues to depth or shape, however, nor did we make measurements to probe how depth and shape were perceived.

2.2.1. Image generation and rendering for display

All of the scenes used in the experiments were modeled with the Maya (Alias, Inc., San Rafael, CA) software tools. Model scenes were then exported to custom Matlab (Mathworks, Inc., Natick, MA) software. This software associated full spectra with each illuminant in the modeled scene, and parameters $\rho_d(\lambda)$, ρ_s, and α with each object. Spectra were represented using 10 nm sampling between 400 and 700 nm (31 wavelength samples). The Matlab software converted the scene representation to a format appropriate for the Radiance renderer (Ward, 1994) and invoked Radiance 31 times (once for each wavelength). This process resulted in a 31-plane hyperspectral image of the scene. This was converted to a three-plane LMS representation by computing at each pixel the excitations that would be produced in the human L-, M-, and S-cones. The LMS image was tone-mapped by truncating the pixel luminances at five times the mean luminance. It was then converted for display on the computer monitors using standard methods (Brainard et al., 2002) together with measurements of the monitors' phosphor emission spectra and gamma functions (PR-650 spectral radiometer, Photo Research, Inc., Chatsworth, CA). As part of the display conversion, tone-mapped LMS images were scaled to occupy 40% of the total luminance range available on the monitors.[1] The factor used to scale the the tone-mapped LMS images into the monitor luminance range varied very slightly from image to image because of differences in image mean, which corresponded to differences in the tone-mapping truncation level. This variation in scale factor was small, less than 1.1% across all images in Experiment 1 and less than 0.6% across all images in Experiment 2. The scale factor also varied across observers because different observers were run at different times, and the maximum luminance of the monitors changed as the monitors aged and with the switch in monitor models used. This change in maximum luminance was captured by periodic recalibration. Each individual observer was run using a single monitor model and calibration, and thus there was no within observer variation of scale factor because of monitor drift. The supplemental material[2] contains an archive of the scene models, representative stimulus images used in the experiments, maximum luminance of the monitors for each observer, and information to allow close re-rendering of the stimulus images from the scene specification.

---

[1] The use of only 40% of the available range, rather than a larger percentage, was due to a misspecification in the software that was discovered after the experiments were run.

[2] ⟨http://color.psych.upenn.edu/supplements/material_color/SupMaterial.html⟩

2.3. Procedure

The observer's task was to adjust the diffuse component of the match sphere until its color appearance was the same as that of the test. In Experiment 1, the observer controlled the match sphere using a game pad, which provided adjustments along the CIELAB L* (lightness), a* (red/green), and b* (blue/yellow) coordinates of the rendered sphere. In Experiment 2, we changed the coordinates of adjustment from CIELAB L*a*b* to the CIELAB HSV (hue, saturation, and value) space. In both cases, the coordinates were computed from the average LMS value of all pixels on the match sphere, and software transformed between LMS, CIELAB L*a*b*, and CIELAB HSV.

Because it was not feasible to re-render the entire displayed images in real time, an approximation was used. This was a variant of a method introduced by Griffin (1999). In the absence of secondary reflections, the light reflected from a surface is a linear function of the surface's spectral reflectance function. Thus in the absence of secondary reflections, we can synthesize the image reflected from a matte object whose surface reflectance is any weighted combination of N fixed reflectance functions by adding the image reflected from images of an object with the same shape and each of these N functions. We used this principle, and pre-rendered four images for each eye, containing red, green, blue, and white match spheres. These images were linearly combined using weights computed from the desired LMS coordinates of the match sphere and those of the precomputed images. We verified the validity of the approximation by comparison with a directly rendered image: the mean difference at each pixel on a directly rendered match sphere and its approximation was small relative to the DAC quantization of the video hardware.

As they performed the match, observers had several adjustment step sizes available. When the observer was satisfied with a match, he or she indicated this with a button press. Each match began with a randomly chosen color for the match sphere. The CIELAB L*a*b* starting value was randomly sampled from the following range: L* 30 to 75, a* −25 to 25, b* −25 to 25. Observers made three separate matches for each choice of test (10 conditions: two diffuse reflectances crossed with five material properties). After the observer completed each match, he or she was also asked to indicate whether the match was satisfactory or not. In practice, however, observers never indicated that a match was unsatisfactory.

3. Experiment 1

3.1. Stimuli

In Experiment 1, observers were asked to match the color appearance of a match sphere to that of a test sphere. The scene was a rendered room containing two spheres, each on its own table. The dimensions of the room specified in the scene space were approximately 35 cm (height) by 42 cm (width) by 56 cm (depth). The front wall of the scene contained an aperture through which objects in the room could be seen. An example of the stimulus image delivered to the left eye is shown in Fig. 2. The test sphere is on the left while the match sphere is on the right. The aperture into the room was 18.3 degrees (width) by 16.5 degrees (height) on the stimulus display (24.6 cm by 22.2 cm on the display). The displayed diameter of the test and match spheres was 2.9 degrees (3.8 cm). The stimulus images were viewed at a distance of 76.4 cm from the observer, and this was the the distance specified from the rendering viewpoint to the plane containing the test and match spheres.
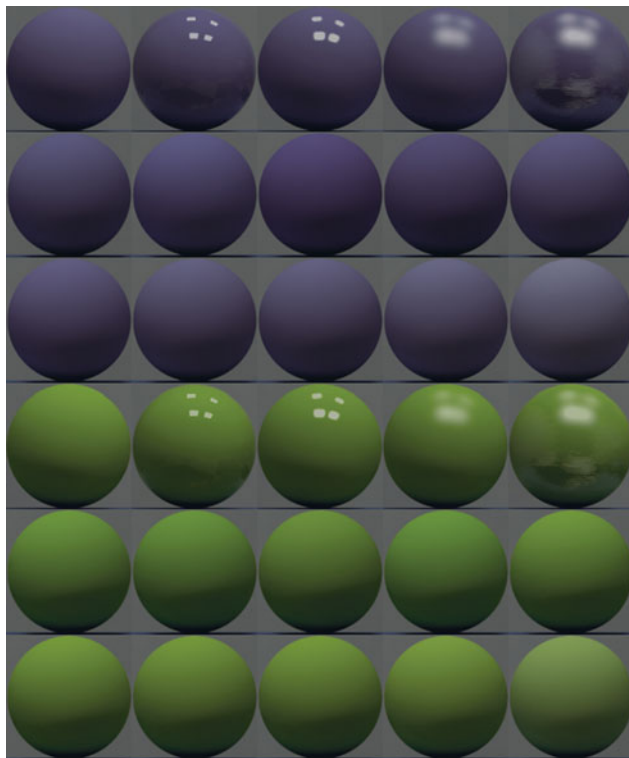
Fig. 3. Pictorial representation of the data for ACD. **Upper panel:** data for the purple body color. The test spheres are shown in the top row. These all have the same diffuse reflectance but differ in their surface gloss (from left to right: Matte, Conditions A-D). Observer ACD's mean matched spheres for the corresponding test spheres are shown in the second row. Matte spheres with the same average LMS values as the corresponding test spheres are shown in the third row. **Lower panel:** data for yellow-green body color, same format as the upper panel. The matte spheres shown in the first and third rows (left column) of each panel differ slightly from each other even though they have the same average LMS values. This is because the spatial distribution of the illumination is not exactly the same at the test and match locations.

The scene was not exactly left-right symmetric, so the illumination incident on the test and match spheres was not identical. An image of the scene with the test and match spheres rendered from mirror-like material is available as part of the supplemental material,[2] and provides a sense of the illumination incident on the two spheres.

Two different body colors and five different glossinesses were used for the test sphere. The body colors were chosen to have the diffuse reflectance spectra of two of the squares on the Macbeth Color Checker. The top row of the upper panel of Fig. 3 shows the five purple test spheres used. The material parameters are listed in Table 1, with conditions listed in order of increasing ρ_s and then α. The same materials were used for the yellow-green tests, which are shown in the top row of the lower panel of Fig. 3. The match sphere was always matte and its body color was adjusted by the observer. The rest of the objects in the scene were held fixed throughout the experiment and were as shown in Fig. 2.

The simulated scene was illuminated by four area lights and a diffuse illuminant. The four area lights were located near the ceiling (approximately 31 cm above the floor) of the rendered room. The simulated area lights had the spectrum of CIE D65

Table 1. *Test sphere material parameters for Experiment 1. This table provides the BRDF parameters for each of the test spheres. These parameters were given incorrectly in a preliminary report of this experiment (Xiao & Brainard, 2006)*

Condition	ρ_s	α
Matte	0.00	0.00
Condition A	0.02	0.00
Condition B	0.12	0.03
Condition C	0.12	0.16
Condition D	0.18	0.16

(CIE, 1986) and the diffuse ambient illuminant was spectrally flat. The contribution of the diffuse illuminant was small: approximately 2% of that from the four area lights.

3.2. Observers

Seven observers completed the experiment. All had normal color vision, 20/20 corrected acuity, and normal stereopsis. One observer (BX) was the first author, one (JMK) was a member of the lab familiar with the experiment, and five were naive paid volunteers. Before the data were collected, observers were asked to complete practice trials. The test spheres used in the practice trials had diffuse reflectances different from those used in the actual experiment. Observers repeated each of the six practice conditions (two diffuse reflectances, three material properties) three times.

3.3. Results

3.3.1. Pictorial representation of the data
Fig. 3 represents the data from Experiment 1 in pictorial format. The top row of the figure shows the test spheres for the purple body color. The next row shows the corresponding match spheres for one observer. The match spheres do not appear identical to each other, indicating that there is an effect, perhaps small, of material on color appearance. The third row shows matte spheres with the same average LMS values as the test spheres in the top row. These are what the match spheres would have looked like had observers matched average LMS coordinates. The matches deviate from the predictions based on the average: from left to right the match spheres become darker and more saturated while the corresponding predictions become lighter and less saturated. The fourth through sixth rows of the figure represent the data for the yellow-green body color for the same observer. The effect of material for this body color appears smaller; the deviation between the matches and the predictions based on the average is still apparent.

3.3.2. Quantitative representation of the data
Fig. 4 provides a quantitative representation of the data. We calculated the average LMS coordinates of the test and matte spheres and then converted these to the CIELAB L*a*b* coordinate system. The white point for CIE conversions for Observer ACD was specified as $XYZ = [43.4, 46.5, 44.7]$ where the units of luminance are cd/m^2. For the other observers, the white point was proportional to these values, but scaled according to the ratio of maximum monitor luminance of the monitors for that observer relative to that for Observer ACD. The conversion white points for
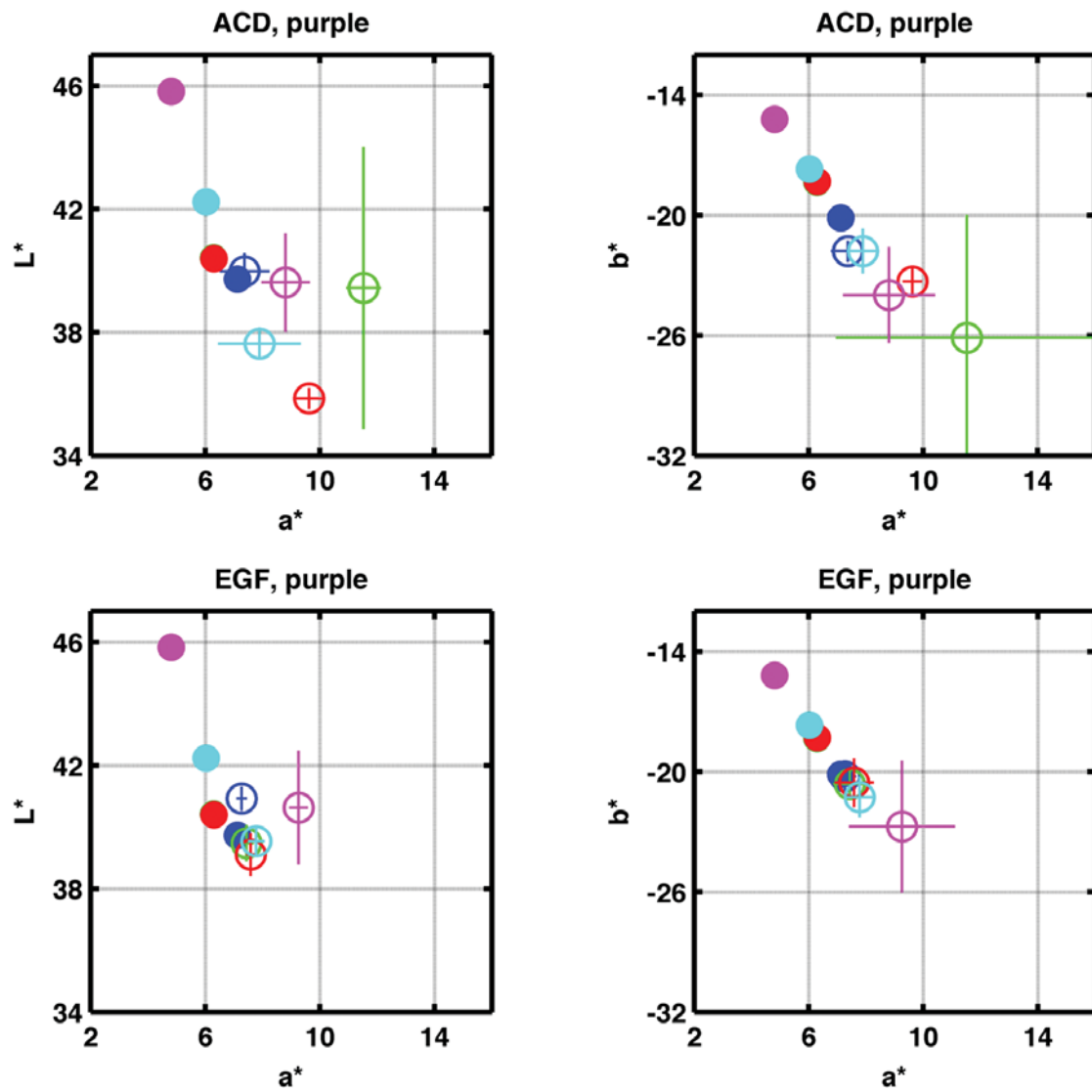
Fig. 4. Experiment 1 tests and matches in CIELAB L*a*b*. **Top panels:** data for ACD for the purple tests. **Lower panels:** data for EGF for the purple tests. The spatial average of the test spheres is shown by the filled symbols, while the spatial average of the matched spheres is shown by the open symbols. Left panels plot L* vs. a* while right panels plot b* vs. a*. Colors indicate the five surface material conditions. Blue: matte; Green: Condition A; Red: Condition B; Cyan: Condition C; Magenta: Condition D. This color key refers to the online version. In the print version, the blue may reproduce as dark purple-blue, the cyan as light blue, and the magenta as purple. The same reproduction differences apply for the plots in the other color figures below. The solid green circle is not explicitly visible in the figure because it is occluded by the solid red circle.

each observer in each experiment are provided in the supplemental material.[2]

The figure shows data for two observers for the purple body color. As we saw pictorially in Fig. 3, the data reveal both variation in the matches with the change of test sphere gloss, and that the variation in the spatial average of the tests does not predict the variation in the matches. The figure also shows that the effect of material variation on color appearance differs between the observers: ACD's data are more spread than EGF's. Data from all of the observers in the same format as Fig. 4 are available in the supplemental material.[2]

We computed the standard error of the L*, a*, and b* coordinates of each set of three matches and took the mean of this over all conditions. A histogram of these errors is provided in the supplemental material.[2] Generally, the standard errors were on the order of 1 CIELAB unit for each coordinate. Brainard (2003) computed that this magnitude is roughly the size of color difference ellipses measured by (MacAdam, 1942). We take this graph to indicate that observers are able to perform our object-to-object matching task with reasonable precision.

3.3.3. Dimensionality reduction

A disadvantage of the CIELAB L*a*b* representation used in Fig. 4 is that the 3D structure of the data must be inferred from two two-dimensional (2D) plots. It turns out, however, that the data for each body color cluster near a plane in the original LMS space. In this section, we develop a data representation that takes advantage of this observation and allows us to visualize the data in just 2D.
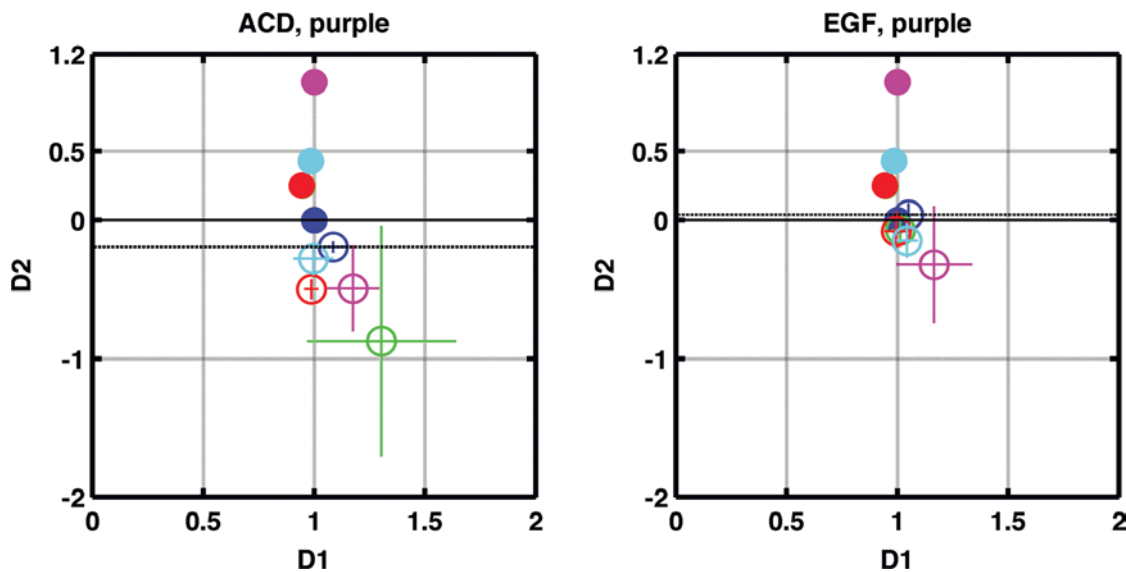
Fig. 5. Tests and matches for two observers for Experiment 1 in the 2D transformed space, purple tests. **Left panel:** ACD. **Right panel:** EGF. Both panels plot specular direction (D2) against diffuse direction (D1). Plot symbols have the same format as in Fig. 4. The value of the match to the matte test along the specular direction is shown by the horizontal dashed line. The solid green circle is not visible in the figure because it is occluded by the solid red circle.

This approach allows a detailed presentation of the data without the requirement of visualizing points in a 3D color space. We also use the 2D representation to derive summary indices of performance. These are the Material Influence Index (MII, Experiments 1 and 2) and the Location Effect Index (LEI, Experiment 2). Summaries of performance in terms of these indices are available in Fig. 7, Fig. 10, and Fig. 11. Readers primarily interested in the broad features of the data may choose to skip ahead to those figures.

The 2D coordinate system is constructed by using the data to define two axes in the LMS color space, which we denote by D1 and D2. The unit vector for the D1 axis (the diffuse direction) is the LMS coordinates of the matte test sphere. Moving along this axis corresponds to changing the magnitude of the diffuse component of the sphere. The unit vector for the D2 axis (the specular direction) is the difference in LMS coordinates between the Condition D test sphere and the matte test sphere. Changes along this dimension correspond closely to adding the LMS coordinates of the illuminant to those of the matte test sphere. Note that the test spheres differ from each other primarily by their value on this axis. A third axis (D3) is defined as orthogonal to D1 and D2, with its vector length set equal to the average of the vector lengths of the D1 and D2 axes. The fact that the data lie within the plane defined by D1 and D2 is indicated by the fact that the test and match spheres have very small values on this third axis (see supplemental material[2]). Thus providing the coordinates on the D1 and D2 axes characterizes the tests and matches.

We refer to the D1/D2 coordinates as the transformed representation. This representation is specific to each body color and Condition D material. Because the maximum luminance of the monitor changed over time (between observers), we computed the transformed representation separately for each observer, using the LMS coordinates of the matte and Condition D test spheres for that observer. Across observers, these differed only by an overall scale factor. Within a particular body color and observer, the matte test sphere always plots to transformed coordinates (1, 0) and the

Condition D test sphere plots to transformed coordinates (1, 1). The fact that the data are described by the diffuse and specular directions may be interpreted as indicating that observers' matches preserve the hue of the test stimuli but vary in lightness and saturation. Fig. 5 replots the data for ACD and EGF in the transformed space.

3.3.4. Effect of material properties

The shifts between match and test spheres can arise from two sources. One is the effect of material properties *per se*. The other is the effect of comparing spheres across two locations within the scene. This latter effect is characterized by the shift between test and match for the matte test sphere. We can remove this effect from the data by shifting the data for all of the match spheres by a constant amount, so that the coordinates of the matte test and match spheres are the same. The use of an additive shift for recentering has no deep theoretical significance and was chosen for convenience; we view it as providing a first order correction sufficient for our present purposes. Fig. 6 shows the data for all observers, recentered in this way. The data indicate that the spread in the match spheres is different from that for the test spheres. This confirms the observation made above for a subset of observers that the matches do not correspond to the average LMS coordinates of the tests.

To characterize the effect of material properties on color matching, we calculated the vector between matte condition coordinates and the glossy condition coordinates, for both tests and matches. We call these material influence vectors (MIVs). We used the MIVs to quantify the stability of color appearance across changes in surface gloss by computing a material influence index (MII). For each material and body color, this is the ratio of the length of the MIV for the match to the length of the MIV for the corresponding test. MII values near 0 indicate matches that are stable with respect to the change in surface gloss. Large MII values indicate matches that are highly influenced by material properties. An MII value of 1 corresponds to a change of match equal in
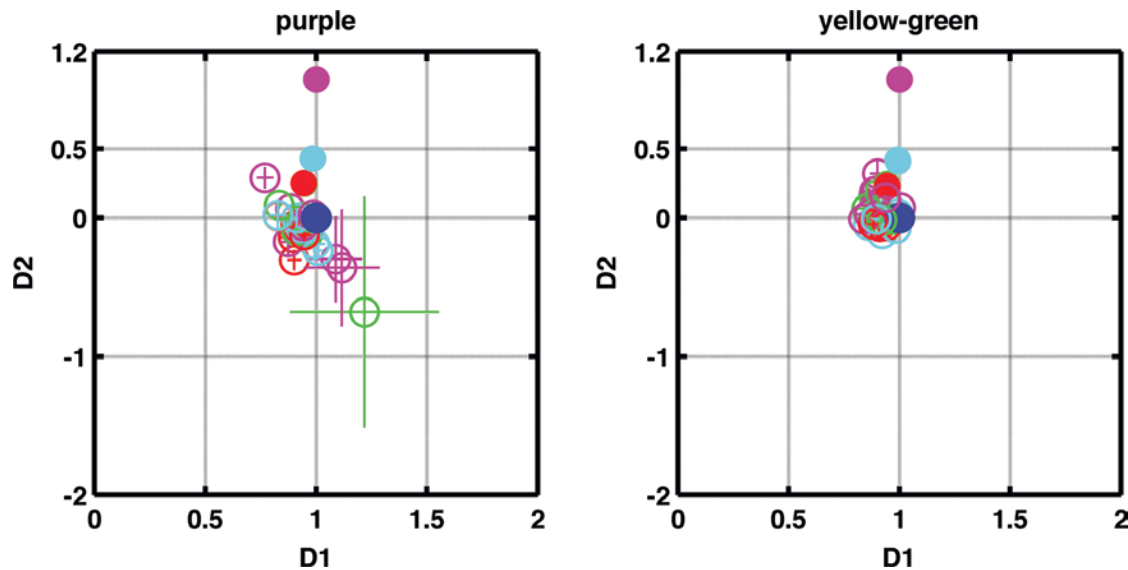
Fig. 6. Tests and re-centered matches of all observers for Experiment 1 in the 2D transformed space. **Left panel:** tests and all seven observers' matches for the purple condition. Each observer's matches have been recentered as described in the text. **Right panel:** tests and all seven observers' matches for the yellow-green condition. The symbols follow the same format as in Fig. 4. The solid green circle is not visible in the figure because it is occluded by the solid red circle.

magnitude to the corresponding change in test. Histograms of the MIIs for individual (not mean) matches from all observers are shown in Fig. 7. The mean MIIs are less than one (0.64 for the purple body color and 0.47 for the yellow-green body color), indicating that the matches are not perturbed as much as the tests. The fact that these indices are smaller than one means that the visual system has stabilized the color appearance of the spheres with respect to the change in average light produced by the changes in material.

It is important to note that the mean value of the MII would be positive even if there were no effect of material. The reason for this is that variability of the matches alone will produce match MIVs with non-zero length. The bottom panels of Fig. 7 show the magnitude of this variability-alone effect. Here we recomputed the index values but substituted for the match MIV the vector length of the deviation of each individual match from its own (same material and observer) mean match. Thus the bottom panel illustrates the magnitude that the MIIs would take on from matching variability alone. Here the mean values are 0.43 and 0.31 for the two body colors respectively. Given these values, we can correct the mean MII values by taking the difference, leading to mean variability-corrected MIIs of 0.21 and 0.16. The variability correction should be regarded as approximate, since vector lengths are not additive with respect to added measurement noise. Rather, the variability-corrected values provide a rough sense of the magnitude of the systematic effect of material relative to the physical shift of the tests.

Note also that the MII tells us about the magnitude of the perceptual shift produced by a change in material (relative to matte), but is insensitive to whether the perceptual shift is in the same direction as the physical shift and also insensitive to whether the shifts are consistent across observers.

To test for the significance of the effect of material, we performed two-way ANOVAS on the data. We treated observer as a random effect and material as a fixed effect and performed ANOVAS separately for each dimension of the transformed space and each body color. The results are provided in Table 2. There is a main effect of observer for each body color, and a main effect of material for the the yellow-green body color. The fact that the significance of the material effect depends on body color, together with the small magnitude of the variability-corrected MIIs, yields the overall conclusion that the effect of material on color appearance is small. The individual observer differences revealed by the ANOVAS are also apparent in an examination of the individual observer data plots provided in the supplemental material.[2]

4. Experiment 2

4.1. Purpose

Experiment 2 was similar to Experiment 1, but we modified the design to take more experimental control over what location on an object observers judged. We were motivated to do this for two reasons. First, we were interested both in the effect of material on color appearance and on how the judged location affects appearance. Second, we wondered whether the individual differences in Experiment 1 might have arisen because different observers judged different locations of the test sphere: some might have chosen to make an integrated judgment of the entire sphere, while others might have based their judgment on a location of the sphere that was distant from visible specular highlights.

4.2. Specific methods

4.2.1. Stimuli

The main difference between Experiment 1 and Experiment 2 is that we replaced the test sphere with a test patch. The test patch was one hexagonal face of a rendered soccer ball (Fig. 8). The scene geometry was otherwise the same as in Experiment 1, except that the front wall aperture differed in size from that used in Experiment 1 and the test sphere was slightly smaller. The aperture was 18.3 degrees (width) by 14.9 degrees (height) on the stimulus
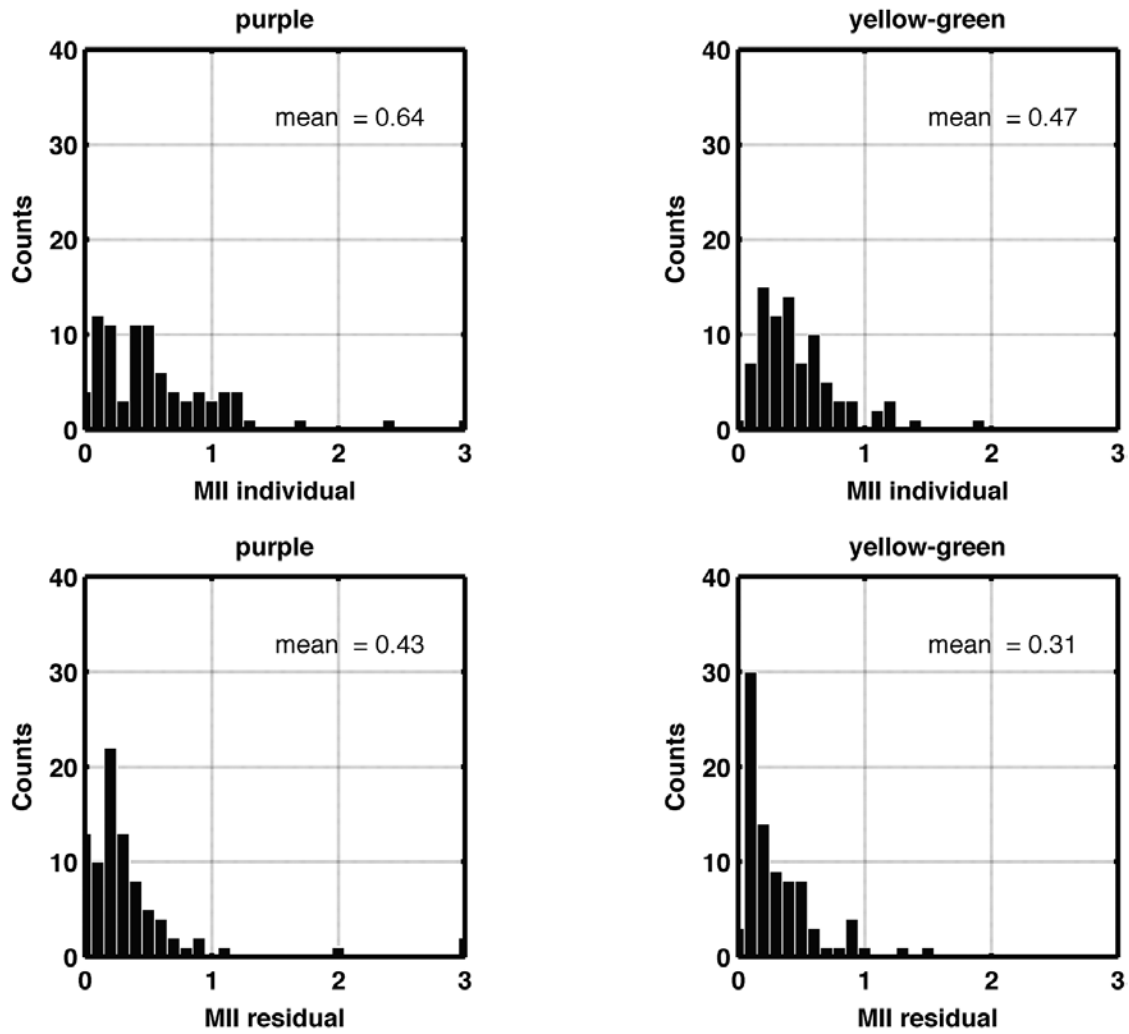
Fig. 7. Top panels: histograms of material influence indices (MIIs) for individual matches from Experiment 1 (purple on left, yellow-green on right). **Bottom panels:** histograms of the MII values we would expect based on measurement variability alone (see description in text).

Table 2. *Anovas for Experiment 1*

Purple					
D1			D2		
Effect	F	P value	Effect	F	P value
Observer	6.19	<0.001	Observer	12.72	<0.001
Material	1.63	0.20	Material	1.06	0.40
Observer * Material	0.85	0.66	Observer * Material	0.52	0.96

Yellow-green					
D1			D2		
Effect	F	P value	Effect	F	P value
Observer	11.49	<0.001	Observer	5.27	<0.005
Material	13.85	<0.001	Material	9.51	<0.001
Observer * Material	0.51	0.97	Observer * Material	0.55	0.95
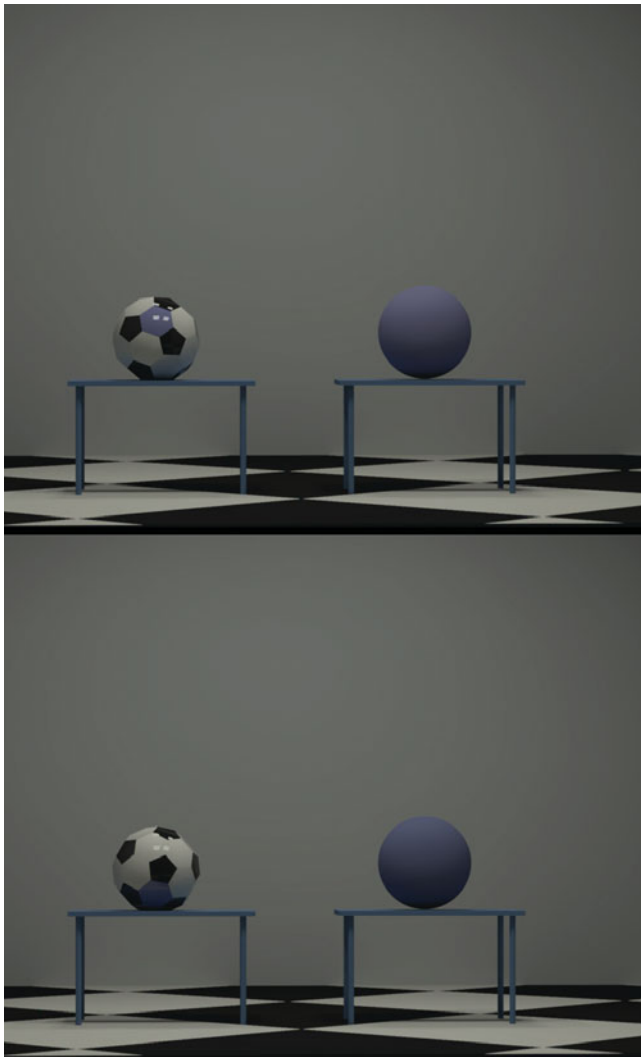
Fig. 8. Examples of the stimulus images used in Experiment 2. The scene contains one soccer ball and one sphere. The colored patch on the soccer ball was the test patch. Its surface reflectance properties were varied from trial to trial. On each trial, the test patch was at one of two locations on the soccer ball, as shown in the upper and lower panels. As in Experiment 1, the sphere on the right was matte and observers adjusted it to match the color appearance of the test patch.

display (24.6 cm by 19.9 cm on the display). The displayed diameter of the soccer ball was 2.5 degrees (3.2 cm) and that of the match sphere was 2.7 degrees (3.5 cm). The size of the test patch (measured as one of the diagonals at the lower test patch location) was 1 degree (1.4 cm).

Across conditions, we varied the body color of the test patch, the material properties of the entire soccer ball (including the test patch), and the location of the test patch on the soccer ball. This design allowed us to ask the same questions we addressed in Experiment 1, but under circumstances where the observer only judged a small portion of the test object. In addition, the design of Experiment 2 allowed us to examine the effect of test patch location. Fig. 8 illustrates the two possible locations of the test patch, which we refer to as the upper and lower locations. The upper location contained a visible specular highlight for the glossy material conditions, while no explicitly visible highlight was present at the lower location for any of the materials.

Table 3. *Test object material parameters for Experiment 2. BRDF parameters that were used to render each of the soccer balls in Experiment 2 are provided*

Conditions	ρ_s	α
Matte	0.00	0.00
Condition A	0.02	0.00
Condition B	0.08	0.02
Condition C	0.12	0.18
Condition D	0.18	0.12

The image generation, surface reflectance model, and scene dimensions were the same as those used in Experiment 1. The material parameters of the test patch are listed in Table 3. The non-matte conditions are ordered from A to D as the strength of the specular component ρ_s becomes larger. The material properties for the entire soccer ball always matched those of the test patch. There were thus a total of 20 conditions in Experiment 2. These resulted from crossing five surface materials, two diffuse reflectances (same as in Experiment 1), and two test patch locations. The simulated illumination was the same as for Experiment 1, and the scene itself was similar. As in Experiment 1, images were scaled to occupy the luminance range available on the monitor.

4.2.2. Procedure

In Experiment 2, the observer's task was to adjust a match sphere until it had the same color appearance as the test patch. As for Experiment 1, the match sphere was always matte. In this experiment, instead of adjusting the match sphere by controlling CIELAB L*a*b* coordinates directly, the observer was given control over CIELAB HSV (hue, saturation, and value) coordinates. This change was made on the basis of our introspection that using HSV coordinates facilitated homing in on any desired color for the match sphere.

4.2.3. Observers

Seven new observers participated in Experiment 2—none of these participated in Experiment 1. Observer ARS was a member of the lab and familiar with the design of the experiment. The other observers were naive paid volunteers. Before setting matches under the experimental conditions, each observer completed two sets of practice matches. In practice Condition 1, observers were asked to match a flat rectangular patch to another flat rectangular patch. Both patches were presented against a simple gray background. In this condition, observers learned the perceptual effect of the controls under simple symmetric matching conditions. In practice Condition 2, the stimuli were similar to those employed in Experiment 1 and observers were asked to match the color appearance of two matte spheres. The data from one naive observer were excluded from further analysis due to the high variability of his matches; data were analyzed for the six remaining observers.

4.3. Results

4.3.1. Match precision in CIE lab space

We calculated the spatial average of the LMS values across both the test patches and the matched spheres and transformed them to CIELAB coordinates. The matching precision was similar to that found in Experiment 1. The white point for CIE conversions

was specified as $XYZ = [37.6, 40.3, 38.7]$ where the units of luminance are cd/m^2.

4.3.2. 2D representation of the data

As with Experiment 1, it is convenient to transform the data to a 2D representation. We used the same procedure as for Experiment 1 to do so, and based the transformation on the coordinates of the test patch in the upper location. Thus the direction of the first axis is given by the LMS coordinates of the upper matte test patch and the direction of the second axis is given by the difference in LMS coordinates between the Condition D test patch and the matte test patch at the upper location. As in Experiment 1, we found that all of the test patches and matched spheres had a value of essentially zero along the third axis in the transformed space.

4.3.3. Effect of test patch location

Fig. 9 plots the data for Observers JHI and QW for the two test body colors. Similar plots for the other observers, as well as plots of the data in the CIELAB L*a*b* coordinate system are available in the supplemental material.[2] Test patches for the upper location (filled circles) plot along a vertical line, as we expect from the way the space was constructed. The test patches for the lower location (filled triangles) are shifted from those for the upper patches along the D1 axis, which is the diffuse direction. The distance between the two clusters represents the physical difference in the average light coming from the test patches at the two locations. Changes in material have a much smaller physical effect at the lower location.

The matches to the test patches at the upper location (open circles) and lower location (open triangles) are separated from one
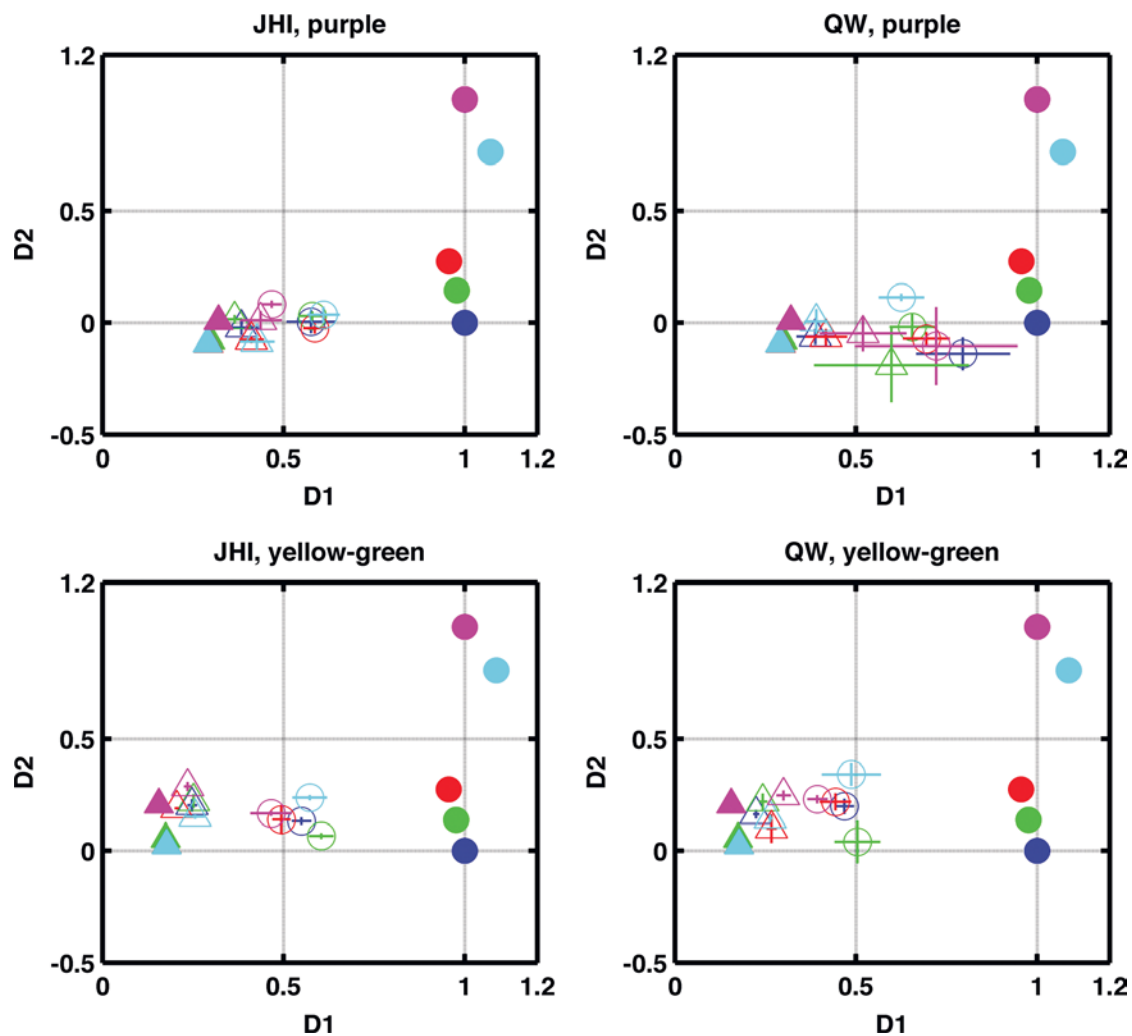
Fig. 9. Tests and matches of two observers for Experiment 2 in the 2D transformed space. The left panels show data for observer JHI for both test body colors, the right panels show data for observer QW. Top panels are for the purple body color, bottom panels are for the yellow-green body color. Filled symbols show the spatial averages of the test patches, while open symbols show the spatial averages of the matched spheres. Circles show the matches and tests for the upper patch, while triangles show the matches and tests for the lower patch. Colored symbols indicate five surface material conditions. Blue: matte condition; Green: Condition A; Red: Condition B; Cyan: Condition C; Magenta: Condition D. The surface material parameters are specified in Table 3. For the tests at the lower location, only the coordinates for Condition D vary substantially from those of the matte condition. This leads to overlap in the filled triangles, so that not all are visible in the plot. In particular, the filled red, green, and blue triangles are located underneath the filled cyan triangle.
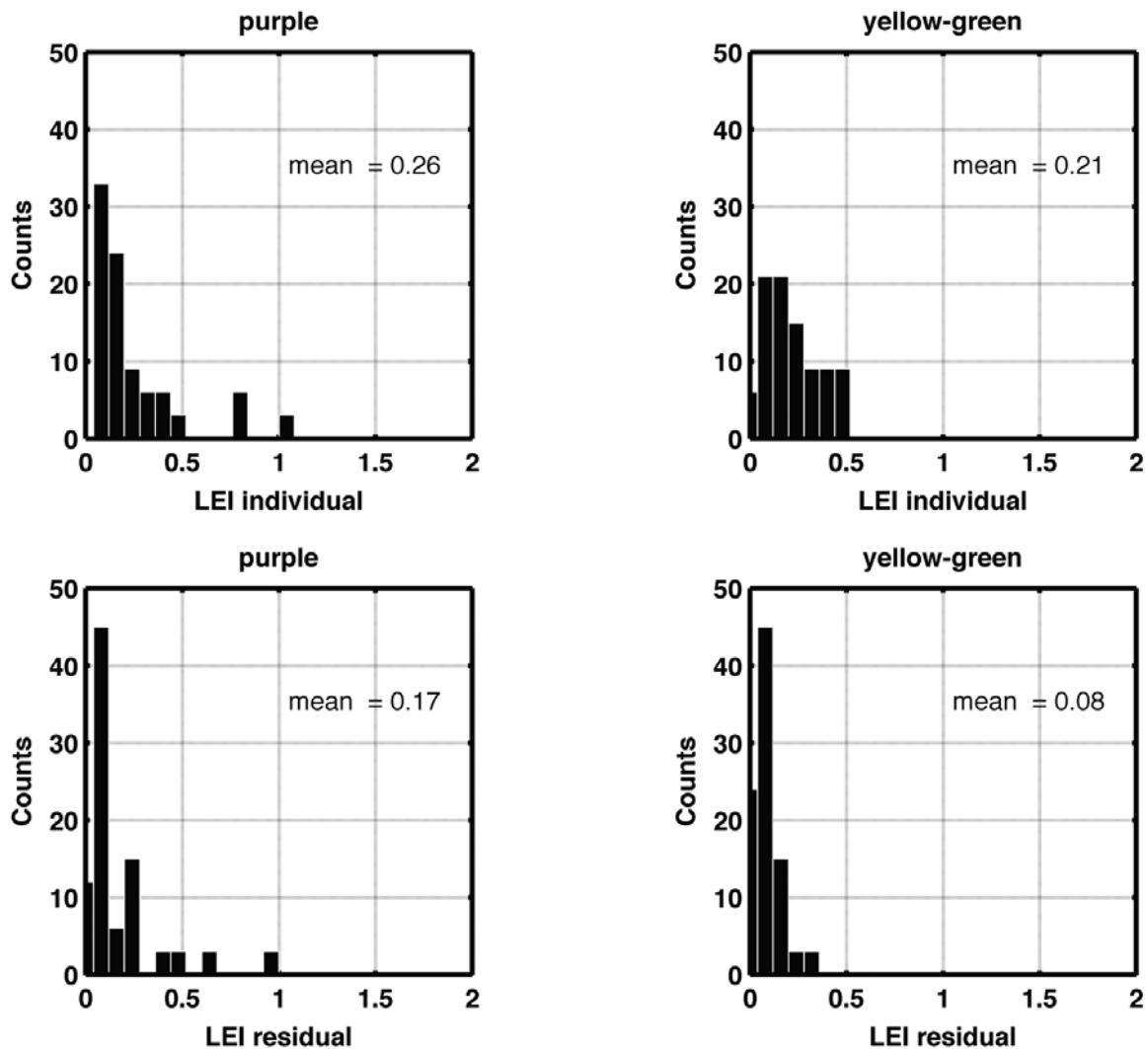
Fig. 10. Top panels: Histograms of Location Effect Indices (LEIs) obtained in Experiment 2 for the two body colors. **Bottom panels:** comparison histograms based on matching variability alone. To compute these, we shifted the individual matches for the top location so that their mean value (for each observer and material) was the same as the corresponding mean of the bottom location, then recomputed the LEIs.

another for both observers and body colors. This indicates that there is an effect of test patch location on observers' matches. To clarify the effect of location, we computed location effect vectors (LEVs) for tests and individual matches. For the tests, the LEVs were computed as the vector difference between the upper and lower test patches. For the matches, the LEVs were computed as the vector difference between individual matches to the corresponding patches. Since matches to the two locations were made separately, we arbitrarily paired these in the order they were set. For any material and body color, we then summarized the effect of test patch location by a location effect index (LEI). For each material, body color, and observer we computed this as the ratio of the LEV for the match to that for the corresponding test. The LEI is zero if the matches are invariant with respect to the change in test patch location. The index is one if the shift in the match is the same as the shift in the test. The top panels of Fig. 10 show a histogram of the LEIs for each body color. The mean value for the purple body color is 0.26 and for the yellow-green body color it is

0.21. As with the MII histogram introduced in Fig. 7, we also computed how the LEIs would behave based on noise alone. These histograms are shown in the bottom panels of the figure. The variability-corrected LEIs are 0.09 and 0.13. The effect of location on the matches is small compared to the physical effect of location on the tests.

4.3.4. Influence of material properties

Experiment 2 also allows us to assess the influence of object material properties on color perception. As shown in Fig. 9, changing material only had a substantial physical effect on the average light reflected from the test patch for the upper location. For this reason, we restricted our analysis of the effect of material on appearance to the data from that location. Fig. 11 provides histograms of the material influence indices from Experiment 2 for the upper location in the same format as Fig. 7. The variability corrected MIIs were 0.17 and 0.13 for the two body colors respectively, again small.
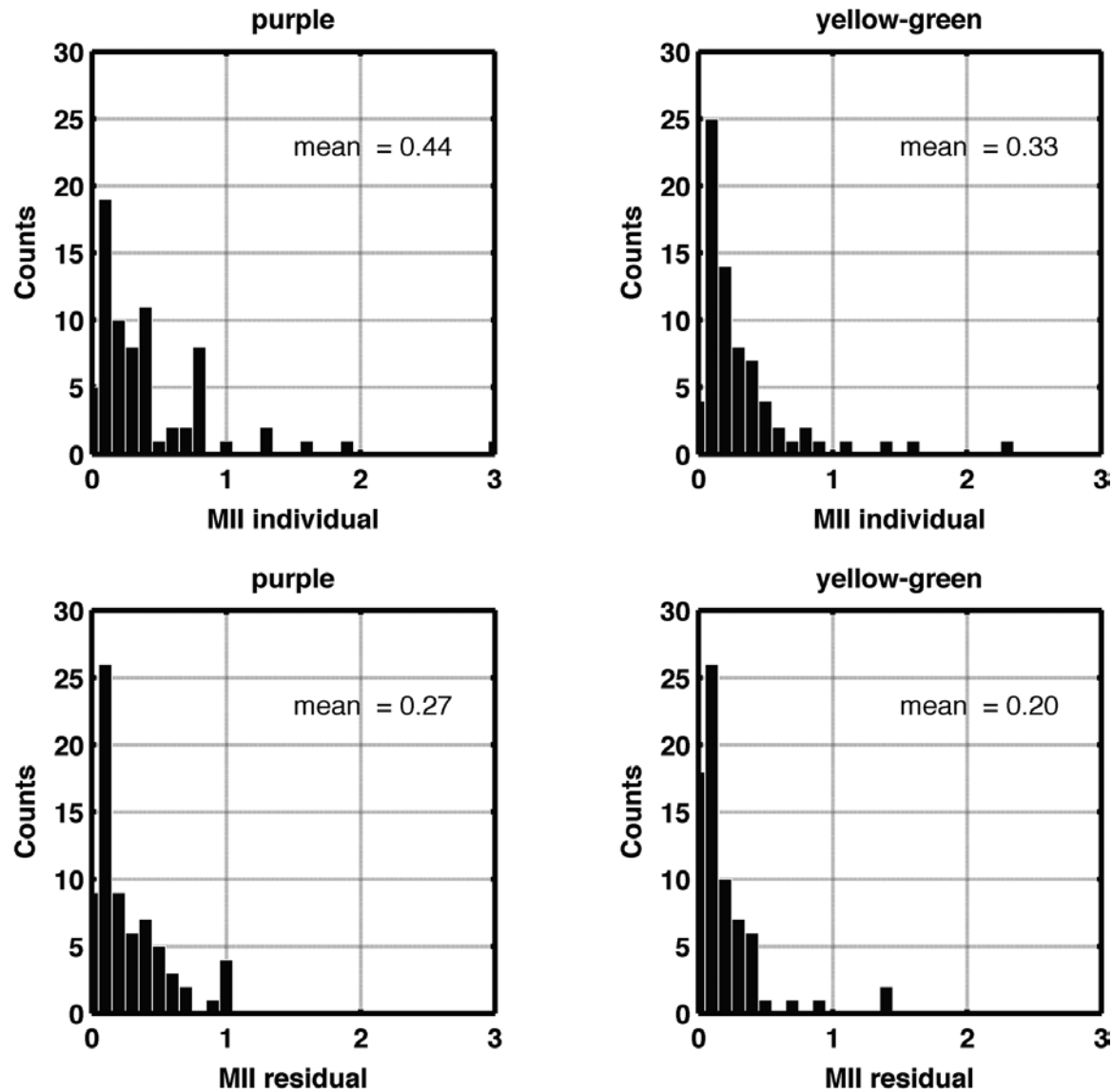
Fig. 11. Histogram of Material Influence Indices from Experiment 2. Same format as Fig. 7.

4.3.5. ANOVAS

We performed three-way ANOVAS on the data. We treated observer as a random effect with material and location as fixed effects, and performed ANOVAS for each transformed dimension and body color. The results are shown in Table 4. There is a significant main effect of location of on dimension D1 for the yellow-green body color, and a marginally significant effect for the purple body color. As in Experiment 1, there is a significant main effect of material for the yellow-green body color but not for the purple body color. And as in Experiment 1, the ANOVAS reveal individual differences, both in the form of main effects (D2) and in the form of interactions. Apparently the restriction from test sphere to test patch does not eliminate individual differences.

5. Summary and discussion

This paper provides data that explore how luminance and chromatic information is integrated to determine the appearance of 3D objects, and in particular how the color appearance of such objects is affected by changes in object material properties.

5.1. Effect of surface gloss

Experiment 1 studied the effect of varying surface gloss on the color appearance of spheres. Changing surface gloss has only a small effect on color appearance, so that matches set by observers vary little relative to the change in the average of the light reflected from the corresponding tests. This result falsifies the simple hypothesis that color appearance is determined by the spatial average of the light reflected from the object's surface. In Experiment 2, observers judged one part of a 3D object, rather than the entire object. For this configuration, the same general results were obtained.

5.2. Effect of test location

In Experiment 2, we found that test patch location affects the patch's color appearance. This effect was in the direction of the

Table 4. *Anovas for Experiment 2*

Purple					
D1			D2		
Effect	F	P value	Effect	F	P value
Observer	2.80	0.14	Observer	4.55	<0.05
Material	1.07	0.40	Material	0.34	0.85
Location	5.85	0.06	Location	3.70	0.11
Observer * Material	0.96	0.53	Observer * Material	1.58	0.16
Observer * Location	6.33	<0.005	Observer * Location	2.64	0.05
Material * Location	1.59	0.22	Material * Location	4.00	<0.05
Observer * Material * Location	0.54	0.94	Observer * Material * Location	0.62	0.89
Yellow-green					
D1			D2		
Effect	F	P value	Effect	F	P value
Observer	1.17	0.44	Observer	8.00	0.05
Material	6.32	<0.005	Material	15.13	<0.001
Location	24.02	<0.005	Location	3.20	0.13
Observer * Material	0.64	0.84	Observer * Material	0.63	0.84
Observer * Location	13.04	<0.001	Observer * Location	2.18	0.10
Material * Location	3.30	<0.05	Material * Location	12.54	<0.001
Observer * Material * Location	0.94	0.54	Observer * Material * Location	1.45	0.11

physical shift (as assessed by the average reflected light) caused by the change in location, but was considerably smaller than the physical shift. Thus, as with changes in material, the visual system appears to compensate for physical effects that occur with test patch location. On the other hand, the fact that appearance varies with test patch location raises interesting questions about how the color appearance of whole objects comes about. One possibility, which could be tested in experiments along the lines of those presented here, is that the overall appearance of an object may be predicted in some straightforward fashion by aggregating the appearance of its individual parts.

5.3. Individual observers

In both experiments, there was considerable variation in the data between observers. Thus our analysis and conclusions are based on aggregated data. The observer-to-observer variation reduced the overall power of the data, and we were unable to draw fine-grained conclusions about the separate effects of individual materials. Individual variation is perhaps not surprising given that the matched stimuli, while similar in color, still differed in perceptually salient ways (e.g., they looked like they were made of different materials). The use of a partial matching task presents ample opportunity for different observers to choose different criteria of accepting a match or to employ different matching strategies. Developing additional insight into the source of individual variation and using this insight to refine the experimental methods is an important goal for this general research direction.

5.4. Relation to other work

The literature on the interaction of material properties and color appearance is small, as is the literature on the color appearance of 3D objects. The studies most closely related to ours generally employ grayscale stimuli, rather than those that vary in both chromaticity and luminance.

Pessoa et al. (1996) studied how perceived lightness varies across locations of a matte 3D ellipsoid. Their task and stimuli differed from ours in important ways, but their broad conclusion is similar to the one we drew from Experiment 2: lightness can vary from one location of an object to another. Pessoa et al. (1996) did not vary the material properties of their objects, but did explore effects of object shape, and illumination direction. Todd et al. (2004) studied how the location of specular hightlights on a rendered grayscale ellipsoid affected lightness judgments between different locations on the ellipsoid. They also found location effects, and as with our results these effects were small compared to the change in the spatial averages of the stimuli.

Nishida and Shinya (1998) asked how well observers could match lightness and glossiness across variation in the 3D shape of objects. In their experiments, observers could adjust both the diffuse reflectance and glossiness of the match object. Nishida and Shinya (1998) found that there were systematic biases in the matches as object shape varied between test and match. They were able to model the matches on the assumption that a perceptual match occurred when the luminance histograms of the images of test and match objects were similar.

Motoyoshi et al. (2007) asked observers to rate the lightness and glossiness of images of grayscale stucco-like materials whose space-averaged luminance was held fixed. For these stimuli, they found considerable decrease in rated lightness as the glossiness of the materials was increased. This result is consistent with ours, in the sense that if we had decreased the average luminance of our glossy tests to match that of our matte tests, we would expect that observers matches would decrease in luminance. Motoyoshi et al. (2007) took a similar modeling approach to that of Nishida and Shinya (1998), in that they explained their results in terms of changes in the luminance histogram of the images of their objects.

In particular, they found that when the mean luminance was held fixed, the skewness of the histogram was negatively correlated with perceived lightness and positively correlated with perceived glossiness.

In another line of related work, Fleming et al. (2003, 2005) have examined how well humans can perceive material properties of objects. Fleming et al. (2003) found that observers' ability to accurately judge the material properties of spheres depends critically on the complexity of the simulated illuminant used to render the images, with more realistic illuminants leading to better performance. Our illumination geometry was produced by a realistic simulation of a 3D scene, but may not have been as complex as many naturally occurring illumination fields (Dror et al., 2004). Fleming and Büthoff (2005) studied the perception of translucent materials, and argued that important aspects of human performance for judging material properties may be understood in terms of statistics computed from the luminance histogram of the object's image. In this regard, their approach is consistent with that of Nishida and Shinya (1998) and Motoyoshi et al. (2007) as discussed above. These authors acknowledge that other factors, such as the spatial structure of the image, can play an important role in how material properties and lightness are perceived (see e.g., Fleming and Büthoff (2005), Fig. 21).

5.5. Do 3D objects have a well-defined color appearance?

In Experiment 1, we asked observers to compare the appearance of entire 3D objects with each other. This instruction is based on the assumption that observers are able to associate a color percept with whole objects. This assumption is not secure *a priori*, however, because examination of our stimuli (see Fig. 2) makes clear that it is possible to observe and attend to variation in color appearance across the surfaces of the test and match spheres.

A few observations lend support to the idea that observers are in fact able to associate a color appearance with the test and match spheres. First, our own introspection indicates that it is natural to think of most objects as having a color. Second, all of our observers were willing to perform the matching task without complaint that it was ill-defined. In addition, our procedure allowed observers to indicate that the best match they could obtain was not satisfactory, but none did so. Third, matches obtained in Experiment 2 were stable across object locations, relative to the physical shift in stimulus across the same locations.

That said, we note that none of these observations force the conclusion that the visual system extracts a single well-defined color for all 3D objects, or even for our stimuli. For example, introspection also suggests that it is difficult to produce a satisfactory color description for objects with strong specular components, and we did not include these in our stimulus set. Moreover, our introspections could be driven by the ease with which we categorize color names rather than by the availability of a more finely grained color representation. Finally, observers could have accomplished our matching task by focusing on specific locations on the test and matte objects. Experimental methods to determine the degree to which any given object elicits a unified color sensation would represent an important advance.

5.6. Implications for models

Much of the work reviewed above emphasizes using simple image statistics to predict how object appearance varies. The spatial average of the LMS values is one such statistic, and our data show that predictions based on this statistic alone do not account for observers' matches. It is possible that some other reasonably simple statistical measure of the properties of the stimuli can predict the matches. The work of Motoyoshi et al. (2007) suggests that examining the moments of the color histograms would provide a point of departure for developing a model of this sort.

Another approach is to consider inverse-optics models. In these models, observers would be postulated to estimate the geometrical properties of the light sources and the reflectance properties, including the BRDF, of the objects and base color appearance on these estimates. This class of models has been helpful in understanding color and lightness under simpler stimulus conditions (Brainard et al., 1997; Maloney, 1999; Bloj et al., 1999, 2004; Boyaci et al., 2003; Doerschner et al., 2004). To develop this class of models, we would look to techniques developed in computer vision (Lee, 1986; Ramamoorthi & Hanrahan, 2001; Johnson & Farid, 2007).

References

AREND, L.E. & REEVES, A. (1986). Simultaneous color constancy. *Journal of Optical Society of America A* **3**, 1743–1751.

BAUML, K.H. (1999). Simultaneous color constancy: How surface color perception varies with the illuminant. *Vision Research* **39**, 1531–1550.

BECK, J. (1964). The effect of gloss on perceived lightness. *The American Journal of Psychology* **77**, 54–63.

BLOJ, M., KERSTEN, D. & 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. & BRAINARD, D.H. (2004). An equivalent illuminant model for the effect of surface slant on perceived lightness. *Journal of Vision* **4**, 735–746.

BOYACI, H., DOERSCHNER, K. & MALONEY, L.T. (2004). Perceived surface color in binocularly viewed scenes with two light sources differing in chromaticity. *Journal of Vision* **4**, 664–679.

BOYACI, H., MALONEY, L.T. & HERSH, S. (2003). The effect of perceived surface orientation on perceived surface albedo in binocularly viewed scenes. *Journal of Vision* **3**, 541–553.

BRAINARD, D.H. (1998). Color constancy in the nearly natural image. 2. Achromatic loci. *Journal of the Optical Society of America A* **15**, 307–325.

BRAINARD, D.H. (2003). Color appearance and color difference specification. In *The Science of Color*, ed. Shevell, S.K., pp. 191–216. Washington, DC: Optical Society of America.

BRAINARD, D.H., BRUNT, W.A. & SPEIGLE, J.M. (1997). Color constancy in the nearly natural image. 1. Asymmetric matches. *Journal of the Optical Society of America A* **14**, 2091–2110.

BRAINARD, D.H., PEILI, D.G. & ROBSON, T. (2002). Display characterization. In *Encyclopedia of Imaging Science and Technology*, ed. Hornak, J., pp. 172–188. New York: John Wiley and Sons.

BRAINARD, D.H. & WANDELL, B.A. (1992). Asymmetric color-matching: How color appearance depends on the illuminant. *Journal of the Optical Society of America A* **9**, 1433–1448.

Burnham, R.W., Evans, R.M. & Newhall, S.M. (1957). Prediction of color appearance with different adaptation illuminations. *Journal of the Optical Society of America* **47**, 35–42.

CIE. (1986). Colorimetry. 2nd edition. Vienna: Bureau Central de la CIE.

Delahunt, P.B. & Brainard, D.H. (2004). Does human color constancy incorporate the statistical regularity of natural daylight? *Journal of Vision* **4**, 57–81.

Derefeldt, G. (1991). Colour appearance systems. In *The Perception of Colour*, ed. Gouras, P., pp. 218–261. Boca Raton, FL: CRC Press, Inc.

Doerschner, K., Boyaci, H. & Maloney, L.T. (2004). Human observers compensate for secondary illumination originating in nearby chromatic surfaces. *Journal of Vision* **4**, 92–105.

Dror, R.O., Willsky, A.S. & Adelson, E.H. (2004). Statistical characterization of real-world illumination. *Journal of Vision* **4**, 821–837.

Fleming, R.W. & Büthoff, H.H. (2005). Low-level image cues in the perception of translucent materials. *ACM Transaction on Applied Perception* **2**, 346–382.

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

Foster, D.H. & Nascimento, S.M.C. (1994). Relational colour constancy from invariant cone-excitation ratios. *Proceedings of the Royal Society of London B* **257**, 115–121.

Griffin, L.D. (1999). Partitive mixing of images: A tool for investigating pictorial perception. *Journal of the Optical Society of America* **16**, 2825–2835.

Hansen, T., Sebastian, W. & Gegenfurtner, K.R. (2007). Effects of spatial and temporal context on color categories and color constancy. *Journal of Vision* **7**, 1–15.

Helson, H. (1940). Fundamental problems in color vision. II. Hue, lightness, and saturation of selective samples in chromatic illumination. *Journal of Experimental Psychology* **26**, 1–27.

Hunter, R.S. & Harold, R.W. (1987). The Measurement of Appearance. 2nd edition. New York: John Wiley and Sons.

Hurlbert, A.C., Lee, H.C. & Bülthoff, H.H. (1989). Cues to the color of the illuminant. *Investigative Opthalmology and Visual Science* **30**, 221.

Johnson, M.K. & Farid, H. (2007). Exposing digital forgeries in complex lighting environments. *IEEE Transactions on Information Forensics and Security* **2**, 450–461.

Khang, B.G. & Zaidi, Q. (2002). Cues and strategies for color constancy: Perceptual scission, image junctions and transformational color matching. *Vision Research* **42**, 211–226.

Lee, H.C. (1986). Method for computing the scene-illuminant chromaticity from specular highlights. *Journal of Optical Society of America A* **3**, 1694–1699.

MacAdam, D.L. (1942). Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America* **32**, 247–274.

Maloney, L.T. (1999). Physics-based approaches to modeling surface color perception. In *Color Vision: From Genes to Perception*, ed. Gegenfurtner, K.R. & Sharpe, L.T., pp. 387–416. Cambridge University Press.

McCann, J.J. (1976). Quantitative studies in retinex theory: A comparison between theoretical predictions and observer responses to the 'color mondrian' experiments. *Vision Research* **16**, 445–458.

Motoyoshi, I., Nishida, S., Sharan, L. & Adelson, E.H. (2007). Image statistics and the perception of surface qualities. *Nature* **447**, 206–209.

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

Obein, G., Knoblauch, K. & Viénot, F. (2004). Difference scaling of gloss: Nonlinearity, binocularity, and constancy. *Journal of Vision* **4**, 711–720.

Pellacini, F., Ferwerda, J.A. & Greenberg, D.P. (2000). Toward a psychophysically-based light reflection model for image synthesis. In *SIGGRAPH '00: Proceedings of the 27th annual conference on computer graphics and interactive techniques*, pp. 55–64. Toronto, Canada: ACM Press.

Pessoa, L., Mingolla, E. & Arend, L.E. (1996). The perception of lightness in 3-d curved objects. *Perception and Psychophysics* **58**, 1293–1305.

Ramamoorthi, R. & Hanrahan, P. (2001). A signal-processing framework for inverse rendering. In *SIGGRAPH '01: Proceedings of the 28th annual conference on computer graphics and interactive techniques*, pp. 117–128. Toronto, Canada: ACM Press.

Ripamonti, C., Bloj, M., Hauck, R., Mitha, K., Greenwald, S., Maloney, S.I. & Brainard, D.H. (2004). Measurements of the effect of surface slant on perceived lightness. *Journal of Vision* **4**, 747–763.

Shevell, S.K. (2003). Color appearance. In *The Science of Color*, ed. Shevell, S.K., pp. 149–190. Washington, DC: Optical Society of America.

sRGB standard (1999). International Electrotechnical Commission Standard 61966-2-1. Geneva: International Electrotechnical Commission.

Todd, J.T., Norman, J.F. & Mingolla, E. (2004). Lightness constancy in the presence of specular highlights. *Psychological Science* **15**, 33–39.

Ward, G.J. (1992). Measuring and modeling anisotropic reflection. In *SIGGRAPH '92: Proceedings of the 19th annual conference on computer graphics and interactive techniques*, pp. 265–272. Toronto, Canada: ACM Press.

Ward, G.J. (1994). The radiance lighting simulation and rendering system. In *SIGGRAPH '94: Proceedings of the 21st annual conference on computer graphics and interactive techniques*, pp. 459–472. Toronto, Canada: ACM Press.

Wyszecki, G. (1986). Color appearance. In *Handbook of Perception and Human Performance: Sensory Processes and Perception*, ed. Boff, K.R., Kaufman, L. & Thomas, J.P., pp. 9.1–9.56. New York: John Wiley and Sons.

Xiao, B. & Brainard, D.H. (2006). Color perception of 3D objects: Constancy with respect to variation of surface gloss. In *APGV '06: Proceedings of the 3rd symposium on applied perception in graphics and visualization*, pp. 63–68. Toronto, Canada: ACM Press.

Yang, J.N. & Maloney, L.T. (2001). Illuminant cues in surface color perception: Tests of three candidate cues. *Vision Research* **41**, 2581–2600.

Yang, J.N. & Shevell, S.K. (2002). Stereo disparity improves color constancy. *Vision Research* **42**, 1979–1989.