Visual Neuroscience (2006), **23**, 351–356. Printed in the USA.
Copyright © 2006 Cambridge University Press 0952-5238/06 $16.00
DOI: 10.1017/S0952523806233285

Color constancy in natural scenes with and without an explicit illuminant cue

KINJIRO AMANO,[1] DAVID H. FOSTER,[1] AND SÉRGIO M.C. NASCIMENTO[2]

[1]Sensing, Imaging, and Signal Processing Group, School of Electrical and Electronic Engineering, University of Manchester, Manchester, United Kingdom
[2]Department of Physics, Gualtar Campus, University of Minho, Braga, Portugal

(RECEIVED March 8, 2006; ACCEPTED March 8, 2006)

Abstract

Observers can generally make reliable judgments of surface color in natural scenes despite changes in an illuminant that is out of view. This ability has sometimes been attributed to observers' estimating the spectral properties of the illuminant in order to compensate for its effects. To test this hypothesis, two surface-color-matching experiments were performed with images of natural scenes obtained from high-resolution hyperspectral images. In the first experiment, the sky illuminating the scene was directly visible to the observer, and its color was manipulated. In the second experiment, a large gray sphere was introduced into the scene so that its illumination by the sun and sky was also directly visible to the observer, and the color of that illumination was manipulated. Although the degree of color constancy varied across this and other variations of the images, there was no reliable effect of illuminant color. Even when the sky was eliminated from view, color constancy did not worsen. Judging surface color in natural scenes seems to be independent of an explicit illuminant cue.

Keywords: Color constancy, Sky, Natural scenes, Illuminant estimate, Spatial cone-excitation ratios, Specular highlights

Introduction

The spectrum of the light reaching the eye from a scene depends both on the reflecting properties of surfaces within it and on the spectrum of the light illuminating it. Consequently, to make an accurate judgment of the color of a particular surface, an observer must somehow discount the effects of the illuminant (von Helmholtz, 1867), even when the illuminant itself is not directly visible. To a variable extent, human observers are successful in this task, that is, color constancy holds, and there is a large literature, partially summarized elsewhere (Foster, 2003; Smithson, 2005), reporting estimates of the degree of color constancy with geometric stimuli, both planar (e.g. Arend & Reeves, 1986) and three-dimensional (e.g. Kraft & Brainard, 1999; de Almeida et al., 2004).

Some theories of color constancy assume that the observer estimates the spectral properties of the illuminant in order to compensate for its effects, an estimation that is theoretically possible in some circumstances (D'Zmura & Iverson, 1993*a*, 1993*b*, 1994; Maloney, 1999). Yet just because observers can estimate the illuminant spectrum does not necessarily imply that they do. Certainly, for matching colored papers in Mondrian-like patterns, it appears to be unnecessary (Amano et al., 2005). But is it possible that explicit information about an illuminant, rather than indirect inference, might influence judgments, particularly in those natural scenes where the illuminant forms part of the normal field of view?

To test this hypothesis, two surface-color-matching experiments were performed with images of natural scenes obtained from high-resolution hyperspectral images (Foster et al., 2004). The cues to the illuminant in the scenes were manipulated using an approach that was similar to the cue-perturbation method adopted by Yang and Maloney (2001) (see also Linnell & Foster, 1997). The images used here were of an urban and a rural scene, as shown in Fig. 1. In the first experiment, the sky (but not the sun) illuminating the scene was made clearly visible to the observer, and its color was varied. The rationale was that including the sky in the scene may or may not improve performance, since its spectrum was consistent with the radiance spectrum of the scene, but changing its spectrum, so that it was inconsistent with the radiance spectrum of the scene, ought to have a biasing effect. In fact, although the degree of color constancy varied across this and other changes to the images, there was no reliable effect of the color of the sky. When the sky was eliminated from view, performance was neither better nor worse. Despite the absence of any reliable effect of a visible sky, it might be argued that observers were able to separate the effects of the ambient illumination it provided from the direct illumination by the sun, and ignore the

---

Address correspondence and reprint requests to: Kinjiro Amano, Sensing, Imaging, and Signal Processing Group, School of Electrical and Electronic Engineering, University of Manchester, Manchester M60 1QD, UK. E-mail: k.amano@manchester.ac.uk

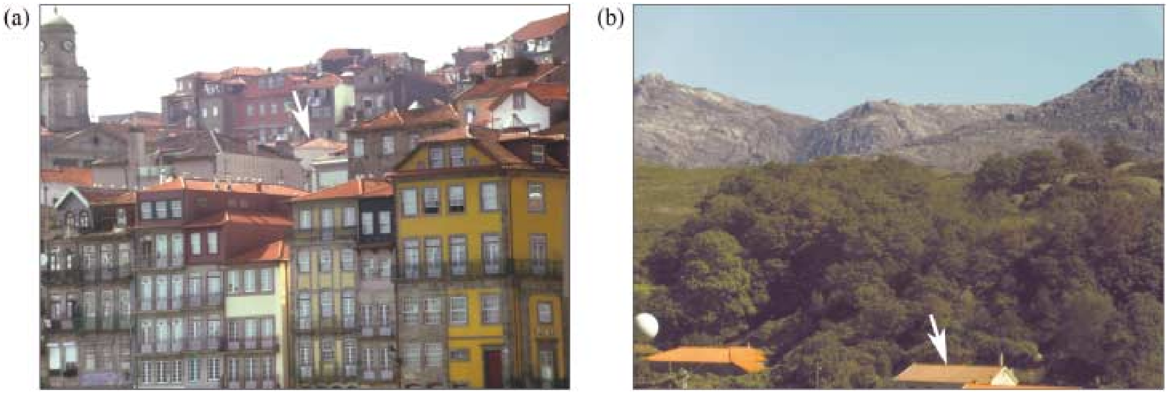

Fig. 1. Sample images of the (a) urban and (b) rural natural scenes used in Experiments 1 and 2. The test surface in each scene is indicated by an arrow.

former. Since there are practical problems in imaging a scene with a visible sun, in the second experiment, a large gray sphere was introduced into the scene so that its illumination by the sun (and sky) could be reliably inferred from its diffuse highlight, and the color of the illumination on the sphere varied. Again, there was no reliable effect of illuminant color.

Judging surface color in natural scenes seems to be independent of explicit knowledge of the color of the illumination.

Materials and methods

Stimuli

Two natural scenes were selected from a hyperspectral image database (Foster et al., 2004; Foster et al., 2006*a*). Hyperspectral data were used to enable the accurate and independent control of illuminant and reflectance spectra. The data were obtained with a high-resolution hyperspectral imaging system, based on a digital camera with a spatial resolution of 1344 × 1024 pixels (Hamamatsu, model C4742-95-12ER, Hamamatsu Photonics K.K., Japan) with a fast tuneable liquid-crystal filter (VariSpec, model VS-VIS2-10-HC-35-SQ, Cambridge Research & Instrumentation, Inc., MA) mounted in front of the lens, together with an infrared blocking filter. Peak-transmission wavelength of the filter was varied in 10-nm steps over 400–720 nm. Further details of the system and its calibration are given in Foster et al. (2004) and Foster et al. (2006*a*).

An urban scene was used for the first experiment and a rural scene for the second experiment, as shown in Fig. 1. The urban scene, from Porto, Portugal, consisted of a group of sixteenth–eighteenth-century buildings and sky. The test surface was one of the roofs in the middle of the image, which was illuminated by the sky. The rural scene, from the Brufe region, Portugal, consisted of fields and trees, some low farm buildings, distant mountains, and sky. A large (30-cm diam.) gray sphere (painted with Munsell N7 matt emulsion paint [VeriVide Ltd., Leicester, UK]) was introduced into the scene so that it reflected light from the sun, which was to the right of the observer. The sphere was the brightest object in the scene (the intensity of its image was clipped in the printed figure), and the diffuse highlight on it gave a reliable cue to the color of the direct illumination (Yang & Maloney, 2001). The test surface was the roof of one of the farm buildings in the lower center of the scene. In both scenes the sky occupied about a quarter of the image and was clearly visible to observers.

Display system and calibration

Stimuli were produced on the screen of a 21-inch RGB color monitor (Trinitron Color Graphic Display, model GDM-F500R, Sony Corp., Tokyo, Japan), with spatial resolution 1600 × 1200 pixels, controlled by a color-graphics workstation (Fuel V12, Silicon Graphics, Inc., Mountain View, CA) whose 10-bit digital-to-analog converters provided an intensity resolution of 1024 levels on each of the red, green, and blue guns. Each image was limited to 80–85% of the displayable area of the screen. The images of the two scenes subtended 1339 × 972 and 1339 × 1018 pixels. A calibrated telespectroradiometer (SpectraColorimeter, PR-650, Photo Research Inc., Chatsworth, CA) and photometer (LMT, L1003, Lichtmesstechnik GmbH, Berlin, Germany) were used to monitor and calibrate the display system. Calibration data included the phosphor coordinates and voltage-intensity look-up tables for the three guns. The monitor was allowed 1 h to warm up before use.

Routine monitoring of the display system tested whether errors in the displayed CIE (x, y, Y) coordinates of a white test patch were <0.005 in (x, y) and $<5\%$ in Y ($<10\%$ at low light levels). Tests of image fidelity used images from the experiments, as described in Foster et al. (2006*b*). Errors for patches of width >20 pixels were ≤ 0.002 in (u', v') coordinates, less than 15% of the 0.015 grid spacing in the (u', v') plane used to sample observers' responses. Since images were presented sequentially in the same position on the screen, position-dependent chromatic errors in each pair of images were the same. Other details of stimulus generation and display are given in Foster et al. (2006*a*).

Stimulus variation

Images were prepared off-line. From the hyperspectral acquisition, the color signal for each scene, that is, the original radiance image $C_0(\lambda; x, y)$ as a function of wavelength λ and position (x, y), was recovered, along with the spectrum $E_0(\lambda)$ of the global illumination on the scene, recorded at a particular reference point (scenes were recorded in direct sunlight under a cloudless sky or under a sky with uniform cloud; Foster et al., 2004). A new radiance image $C(\lambda; x, y)$ was generated for a new global illuminant $E(\lambda)$ by putting $C(\lambda; x, y) = C_0(\lambda; x, y)E(\lambda)/E_0(\lambda)$. In effect, for each surface in the scene, the new radiance at each point (x, y) is the same as that obtained by multiplying a spectral reflectance $R(\lambda; x, y)$ at that point by the new global illuminant $E(\lambda)$, where, by definition, $R(\lambda;x,y) = C_0(\lambda;x,y)/E_0(\lambda)$ (an analysis of effective reflectances with surfaces under direct and indirect illumination in terms of bidirectional reflectance functions is given in Foster et al. 2006*b*). For luminous and non-spectrally selective reflecting areas in the scene, such as the sky or specular highlights, this global illuminant change from $C_0(\lambda;x,y)$ to $C(\lambda;x,y)$ has the critical property that it preserves the spectral relationships between these areas and the surfaces of the scene; for example, if the illumination on the scene becomes more blue, then the sky and specular highlights also become more blue.

In Experiment 1, the global illuminant $E(\lambda)$ was first a daylight with correlated color temperature 25,000 K and then 6700 K; in Experiment 2, the global illuminant was first 15,000 K and then 5700 K. This difference between the pairs of illuminants was due to limits on the displayable color gamuts of the scenes, but the color difference in 1976 CIE (u',v') space between the members of each pair of illuminants was the same.

The reflectance of the test surface in the first image was manipulated independently of the global illuminant: five different initial test-surface colors ("yellowish," "greenish," etc.) were tested in five separate blocks. In each block, the spectral reflectance of the test surface in the second image varied randomly, from trial to trial, in one of 65 ways (all randomization was without replacement). This variation was achieved by a computational device, as follows. Suppose that the initial spectral reflectance was $R(\lambda;x,y)$ at each point (x,y) of the surface and the global illuminant spectrum was $E(\lambda)$, so that the color signal was $R(\lambda;x,y)E(\lambda)$. With a change in spectral reflectance to $R'(\lambda;x,y)$, say, the color signal becomes $R'(\lambda;x,y)E(\lambda)$. But the same color signal can be achieved with the original $R(\lambda;x,y)$ by replacing $E(\lambda)$ locally by a different daylight $E'(\lambda)$ such that $R'(\lambda;x,y)E(\lambda) = R(\lambda;x,y)$ $E'(\lambda)$; the change in reflectance $k(\lambda)$ is given by $k(\lambda) = R'(\lambda;x,y)/R(\lambda;x,y) = E'(\lambda)/E(\lambda)$. Varying the chromaticity of this local illuminant $E'(\lambda)$ is closely related to varying the chromaticity of the test surface, but the representation of changes in spectral reflectances $R'(\lambda;x,y)/R(\lambda;x,y)$ in terms of changes in local illuminants $E'(\lambda)/E(\lambda)$ has the advantage of a natural colorimetric parameterization and of a quantification that is independent of the initial spectral reflectance of the test surface, so that averages may be calculated over stimuli (see Foster et al., 2001*a*).

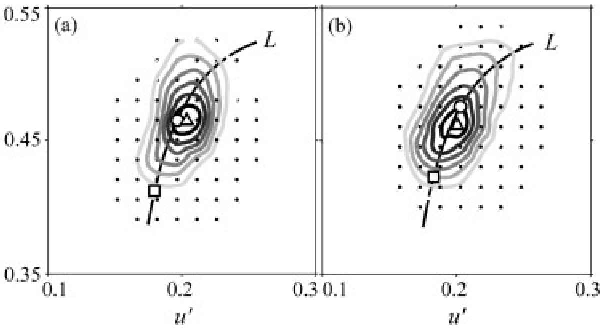

Fig. 2. Contour plots showing the relative frequency of "illuminant-change" responses by observers in the CIE 1976 (u',v') chromaticity diagram as a function of the chromaticity of the reflectance change of the test surface. The square symbols show the first illuminant, a daylight with correlated color temperature, 25,000 K in (a) and 15,000 K in (b); the circles the second illuminant, 6700 K in (a) and 5700 K in (b); and the triangles the mode, from which the color-constancy index was derived. The line marked *L* is the daylight locus.

These local illuminants were constructed from a linear combination of the daylight spectral basis functions (Judd et al., 1964) whose corresponding chromaticities were drawn from the gamut in the (u',v') diagram consisting of 65 locations, with spacing 0.015 in the u' and v' directions, shown by the small solid points in the graphs of Fig. 2. The same technique was used to produce the five different initial test-surface spectra, whose corresponding chromaticities were shifted from the original or Munsell N5 or N7 by (0.015, 0), (0, 0.015), (−0.015, 0), (0, −0.015), and (0, 0).

Changes to the spectrum of the sky in Experiment 1 and to the spectrum of the illumination on the gray reflecting sphere in Experiment 2 were made in the same way; that is, the radiance spectrum $C(\lambda;x,y)$ at each point (x,y) of the sky or sphere was replaced by the spectrum $C(\lambda;x,y)k(\lambda)$, making the sky or sphere appear, for example, more blue or more red, depending on $k(\lambda)$. This illuminant change $k(\lambda)$ was the same for the first and second images, for making inconsistent changes across the two images is known to worsen color-constancy judgments (Foster et al., 2006*a*).

Procedure

In each trial, two images of a particular scene were presented in sequence on the screen of the color monitor, each for 1 s, with no interval. The images differed in their global illuminants. As already mentioned, during the global illuminant change, the spectral reflectance of the test surface in the second image also changed, by a random amount. The observer's task was to decide whether the test surface in the successive images was the same or different, that is, whether an illuminant change or an illuminant change accompanied by a change in the spectral reflectance of the test surface had occurred (Craven & Foster, 1992). Responses were made with mouse buttons connected to a computer. Observers were allowed to move their eyes freely. At the beginning of the session, the experimenter indicated the identity of the test surface to the observer verbally and by pointing and gave a demonstration of illuminant and varying sizes of reflectance changes.

In each experimental session, there was just one color change to the sky or sphere and just one test-surface color. In all, 20 conditions were tested in Experiment 1 and 15 in Experiment 2. A further control condition was introduced in Experiment 2, in which the scene was cropped to remove both the sphere and sky.

The images on the screen of the monitor were viewed binocularly at 100 cm and subtended approximately 18 deg × 13 deg visual angle. The test surfaces subtended approximately 1 deg × 0.5 deg and 3 deg × 1 deg for Experiments 1 and 2, respectively. The reflecting sphere of Experiment 2 subtended approximately 1 deg. Images were presented in a dark surround, and the luminance at each pixel varied from 0 to 33 cd m^{-2}. Room luminance was approximately 0.5 cd m^{-2}. Observers each performed at least 1300 trials in all.

Observers

Twelve observers, aged 23–34 years, took part in the experiments: two male and four female for Experiment 1 and the same for Experiment 2. All observers had normal color vision verified with the Farnsworth-Munsell 100-Hue test; Ishihara pseudoisochromatic plates (24-plates edition, 1964); Rayleigh and Moreland anomaloscopy; and luminance matching (Interzeag Color Vision meter 712, Schlieren, Switzerland). All had normal or corrected-to-normal visual acuity. The experiments were conducted in accordance with principles embodied in the Declaration of Helsinki (Code of Ethics of the World Medical Association) and were approved by the Research Ethics Committee of the University of Manchester. All observers were unaware of the purpose of the experiment.

Analysis

Only one out of the 65 changes in the spectral reflectance of the test surface was a null change, that is, corresponded to a pure illuminant change on the scene. An observer with perfect color constancy would therefore give "illuminant-change" responses only to this stimulus combination and "reflectance-change" responses to all the others. The frequency of "illuminant-change" responses in each condition was therefore plotted as a function of the chromaticity of the local illuminant in the CIE 1976 (u',v') chromaticity diagram. This frequency plot was then smoothed by a two-dimensional nonparametric locally weighted quadratic regression ("loess"; Cleveland & Devlin, 1988) and contour plots derived as shown in Fig. 2 (cf. Bramwell & Hurlbert, 1996, who used a two-dimensional Gaussian model). Each contour represents a constant relative frequency, with differences between contours of approximately 0.10–0.15. The position of the maximum of each distribution was obtained numerically from the loess analysis (shown by the triangles in Fig. 2). If the observer had perfect color constancy, that position would coincide with the position of the second illuminant (circles).

To summarize the error in the surface-color judgment, that is, the bias, a standard color-constancy index (Arend et al., 1991) was then derived. Thus, if a is the distance between the positions of the maximum (triangle) and the 6700-K, or 5700-K illuminant (circle) (the bias in observers' responses) and b the distance between the positions of the 25,000-K or 15,000-K illuminant (square) and 6700-K or 5700-K, respectively, illuminant, then the constancy index is $1 - a/b$. The standard error (SE) of this index was estimated with a bootstrap procedure, based on 1000 replications, with resampling over observers (Efron & Tibshirani, 1993). Perfect constancy corresponds to an index of unity. Perfect inconstancy corresponds to an index of 0, which occurs when the response peak coincides with the first global illuminant.

Results

Recall that these experiments were designed to test for a systematic effect on surface-color judgments of initial shifts in color of a visible sky or of the color of illumination on a visible sphere. The distributions of observers' responses without these color shifts are shown in Figs. 2a and b. Fig. 3 shows mean color constancy indices for the six observers of Experiment 1 in which the color of the sky was shifted in increments along the daylight locus. Data are grouped according to the shift in sky color. Within groups, each data point corresponds to a different initial color of the test surface. Vertical bars show ±1 SE of the mean. The control condition is discussed later.

Color constancy indices ranged from 0.56 to 0.88, with an overall mean of 0.71. Neither the initial shift in sky color nor the initial test-surface color seemed to affect performance. This was confirmed by a repeated-measures analysis of variance (ANOVA). Thus, sky-color shift had no significant effect ($F(3,15) = 1.3$, $P = 0.3$); nor did test-surface color ($F(4,20) = 0.8$, $P = 0.5$); nor was there an interaction between the two ($F(12,60) = 1.0$, $P = 0.4$).

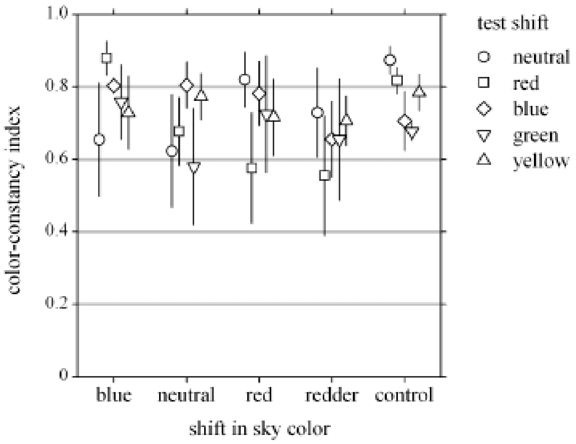

Fig. 3. Color-constancy indices from Experiment 1 in which the color of the sky in the scene (Fig. 1a) was shifted in increments along the daylight locus. Data are grouped according to the shift in initial sky color; within groups, each data point corresponds to a different shift in initial color of the test surface. Data are averaged over six observers. Vertical bars show ±1 SE of the mean.

Since data from the same scene location were also available from a separate experiment in which the same kinds of surface-color judgments were made without a visible illuminant (Foster et al., 2006*a*), the mean level of color constancy is also included in Fig. 3 for comparison (labeled "control"). Despite the absence of the sky, and a different group of observers, performance was similar. An ANOVA showed formally that the visible sky had no significant effect ($F(4,135) = 1.1$, $P = 0.4$) and neither did the initial test-surface color ($F(4,135) = 0.5$, $P = 0.8$).

Fig. 4 shows mean color constancy indices for the six observers of Experiment 2, in which the color of the light illuminating the sphere in the scene was shifted in increments along the daylight locus. As in Fig. 3, data are grouped according to the shift of the sphere illuminant color. Within groups, each data point corre-

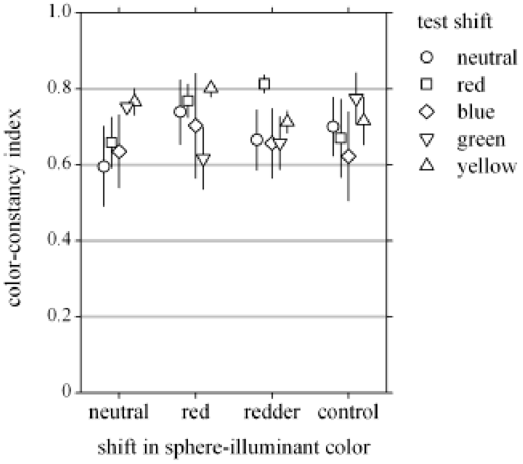

Fig. 4. Color-constancy indices from Experiment 2 in which the color of the light from the sun and sky illuminating the sphere in the scene (Fig. 1b) was shifted in increments along the daylight locus. Other details as for Fig. 3.

sponds to a different initial color of the test surface. Vertical bars show ±1 SE of the mean.

Color constancy indices ranged from 0.60 to 0.81, with an overall mean of 0.70. As with the main experiment, neither the initial shift in illuminant color nor the initial test-surface color seemed to affect performance. This was confirmed by a repeated-measures ANOVA. Thus, sphere-illuminant color shift had no significant effect ($F(2,10) = 0.6$, $P = 0.5$); nor did test-surface color ($F(4,20) = 1.7$, $P = 0.2$); nor was there an interaction between the two ($F(8,40) = 1.0$, $P = 0.4$).

Results for the control condition, in which the scene was cropped to remove both the sphere and sky, were similar. A repeated-measures ANOVA showed formally that the presence of the sphere had no significant effect ($F(1,5) = 0.3$, $P > 0.5$) and neither did the initial test-surface color ($F(4,20) = 1.0$, $P = 0.4$).

Discussion

If the gamut of surfaces in a scene is sufficiently large, then, in theory, it is possible to make a reliable estimate of the color of the illumination, for example, by assuming that it coincides with a spatial average color of the scene (Buchsbaum, 1980), or with the color of the highest-luminance surface (Land & McCann, 1971), or by appealing to other statistical properties of the image (Finlayson et al., 2001; Golz & MacLeod, 2002). In experiments with Mondrian-like patterns, observers are able to estimate the color of an illuminant on a scene (Linnell & Foster, 2002), and, as expected, the accuracy of estimation improves as the number of surfaces in the patterns increases. Space-average scene color may also be used as the cue in some surface-color matching experiments in which both the illuminant and test-surface position vary (Amano & Foster, 2004). Even so, there are clearly conditions where space-average scene color gives an unreliable cue to the illuminant, most notably where the color gamut is limited. Making information about the illuminant on a scene explicit, rather than inferential, ought therefore to influence judgments, particularly in natural scenes.

As was shown here, however, shifting the color of a directly visible sky along the daylight locus so that it differed from the true illuminant spectrum on the scene had no reliable effect on the accuracy of surface-color judgments, nor did removing the sky from the field of view. Shifting the color of the light from the sun and sky illuminating a clearly visible sphere in the scene so that it, too, differed from the true illuminant spectrum on the scene also had no reliable effect. It might be argued that in the images of Experiment 2 (Fig. 1b), the right-hand side of the house whose roof was the test surface could also have been used to infer the color of the direct illumination, but it was less bright than the sphere and had no diffuse highlight that allowed the illuminant color to be inferred.

These shifts in color of the sky and the light on the sphere were substantial, in the 1976 CIE (u',v') color diagram of the same order as the color difference between the two global illuminants. Moreover, performance was no worse when both the sky and the sphere were eliminated from the field of view. The mean levels of color constancy of 0.71 and 0.70 found here were compatible with those reported for a much larger population of rural and urban scenes (Foster et al., 2006*a*).

The simplest conclusion is that judging surface color in natural scenes does not depend on explicit knowledge of the color of the illumination (cf. Yang & Maloney, 2001). This conclusion is consistent with another experiment (Amano et al., 2005) in which observers made asymmetric color matches between pairs of simultaneously presented Mondrian-like patterns of colored papers under different daylights. The patterns had either 49 surfaces or a minimal 2 surfaces, too few for an accurate estimate of the illuminant to be formed. Yet color-constancy indices were almost identical (0.73 and 0.72 for 1 deg paper squares and closely similar to the mean values reported here).

As has been argued elsewhere (Foster, 2003), a possible explanation for the insensitivity of surface-color judgments to information about the illuminant on a scene is observers' use of "relational color constancy" (Foster & Nascimento, 1994). This refers to the constancy of perceived color relations between surfaces under different illuminants, as distinct from color constancy, which refers to the constancy of perceived colors of surfaces. Thus, when discriminating between illuminant and material changes in scenes, observers simply compare how the color of the test surface relates to the color of one or more other surfaces in the scene or, indeed, to the scene as a whole, first under the one illuminant and then under the other.

There is a ready physiological substrate for these comparisons: the ratios of cone-photoreceptor excitations generated in response to light reflected from pairs of surfaces or groups of surfaces. Such ratios, which can also be calculated across postreceptoral combinations and spatial averages of cone signals, have the remarkable property of being almost exactly invariant under changes in illuminant both with natural surfaces (Nascimento et al., 2002) and with colored papers (Foster & Nascimento, 1994).

More generally, when surface-color judgments were measured with a much larger group of scenes (Foster et al., 2006*a*), the variation in color-constancy indices across the scenes, from 0.69 to 0.97, was best explained by spatial ratios of cone excitations being calculated globally for each scene, that is, over all possible pairs of points, rather than just between the test surface and one or more other surfaces in the scene.

In everyday viewing, we are able to make rapid and relatively accurate judgments about the colors of things, an ability that seems not to require attentional effort (Foster et al., 2001*b*). Yet the illuminant is often not visible, and our percepts are much the same as when the illuminant is visible, an intuition that seems confirmed by the present experiments.

Acknowledgments

This research was supported by the EPSRC (grant nos. GR/R39412/01 and EP/B000257/1).

References

Amano, K. & Foster, D.H. (2004). Colour constancy under simultaneous changes in surface position and illuminant. *Proceedings of the Royal Society of London Series B* **271**, 2319–2326.

Amano, K., Foster, D.H. & Nascimento, S.M. (2005). Minimalist surface-colour matching. *Perception* **34**, 1009–1013.

Arend, L. & Reeves, A. (1986). Simultaneous color constancy. *Journal of the Optical Society of America A. Optics, Image Science, and Vision* **3**, 1743–1751.

Arend, L.E., Jr., Reeves, A., Schirillo, J. & Goldstein, R. (1991). Simultaneous color constancy: Papers with diverse Munsell values. *Journal of the Optical Society of America A. Optics, Image Science, and Vision* **8**, 661–672.

Bramwell, D.I. & Hurlbert, A.C. (1996). Measurements of colour constancy by using a forced-choice matching technique. *Perception* **25**, 229–241.

Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute* **310**, 1–26.

Cleveland, W.S. & Devlin, S.J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association* **83**, 596–610.

Craven, B.J. & Foster, D.H. (1992). An operational approach to colour constancy. *Vision Research* **32**, 1359–1366.

de Almeida, V.M.N., Fiadeiro, P.T. & Nascimento, S.M.C. (2004). Color constancy by asymmetric color matching with real objects in three-dimensional scenes. *Visual Neuroscience* **21**, 341–345.

D'Zmura, M. & Iverson, G. (1993*a*). Color constancy. I. Basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces. *Journal of the Optical Society of America A. Optics, Image Science, and Vision* **10**, 2148–2165.

D'Zmura, M. & Iverson, G. (1993*b*). Color constancy. II. Results for two-stage linear recovery of spectral descriptions for lights and surfaces. *Journal of the Optical Society of America A. Optics, Image Science, and Vision* **10**, 2166–2180.

D'Zmura, M. & Iverson, G. (1994). Color constancy. III. General linear recovery of spectral descriptions for lights and surfaces. *Journal of the Optical Society of America A. Optics, Image Science, and Vision* **11**, 2389–2400.

Efron, B. & Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. New York: Chapman and Hall.

Finlayson, G.D., Hordley, S.D. & Hubel, P.M. (2001). Color by correlation: A simple, unifying framework for color constancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23**, 1209–1221.

Foster, D.H. (2003). Does colour constancy exist? *Trends in Cognitive Sciences* **7**, 439–443.

Foster, D.H., Amano, K. & Nascimento, S.M. (2006*a*). Color constancy in natural scenes explained by global image statistics. *Visual Neuroscience*, **23**, 341–349.

Foster, D.H., Amano, K. & Nascimento, S.M.C. (2001*a*). Colour constancy from temporal cues: Better matches with less variability under fast illuminant changes. *Vision Research* **41**, 285–293.

Foster, D.H., Amano, K., Nascimento, S.M.C. & Foster, M.J. (2006*b*). Frequency of metamerism in natural scenes. *Journal of the Optical Society of America A. Optics, Image Science, and Vision* (in press).

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

Foster, D.H., Nascimento, S.M.C. & Amano, K. (2004). Information limits on neural identification of colored surfaces in natural scenes. *Visual Neuroscience* **21**, 331–336.

Foster, D.H., Nascimento, S.M.C., Amano, K., Arend, L., Linnell, K.J., Nieves, J.L., Plet, S. & Foster, J.S. (2001*b*). Parallel detection of violations of color constancy. *Proceedings of the National Academy of Sciences of the U.S.A.* **98**, 8151–8156.

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

Judd, D.B., MacAdam, D.L. & Wyszecki, G. (1964). Spectral distribution of typical daylight as a function of correlated color temperature. *Journal of the Optical Society of America* **54**, 1031–1040.

Kraft, J.M. & Brainard, D.H. (1999). Mechanisms of color constancy under nearly natural viewing. *Proceedings of the National Academy of Sciences of the U.S.A.* **96**, 307–312.

Land, E.H. & McCann, J.J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America* **61**, 1–11.

Linnell, K.J. & Foster, D.H. (1997). Space-average scene colour used to extract illuminant information. In *John Dalton's Colour Vision Legacy*, eds. Dickinson, C., Murray, I. & Carden, D., pp. 501–509. London: Taylor and Francis.

Linnell, K.J. & Foster, D.H. (2002). Scene articulation: Dependence of illuminant estimates on number of surfaces. *Perception* **31**, 151–159.

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

Nascimento, S.M.C., Ferreira, F.P. & Foster, D.H. (2002). Statistics of spatial cone-excitation ratios in natural scenes. *Journal of the Optical Society of America A* **19**, 1484–1490.

Smithson, H.E. (2005). Sensory, computational and cognitive components of human colour constancy. *Philosophical Transactions of the Royal Society B* **360**, 1329–1346.

von Helmholtz, H. (1867). *Handbuch der Physiologischen Optik, Vol. 2*. Second edition. Leipzig: Leopold Voss. Translated as *Helmholtz's Treatise on Physiological Optics*, ed. Southall, J.P.C., pp. 286–287. Third edition. Washington, DC: Optical Society of America, 1924. Reprint, New York: Dover Publications, 1962.

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