Quantifying Human Color Constancy

Jan Walraven and Marcel P. Lucassen
TNO Institute for Human Factors, Soesterberg, The Netherlands

Introduction

The visual system generates object color on the basis of spectral information mediated by reflected light. Although the color of this light, and hence, surface color, varies with the color and intensity of the ambient illumination, the visual system is capable of maintaining a fair degree of color fidelity, a phenomenon known as color constancy.

Color constancy was already discussed by Helmholtz,[1] but there still is surprisingly little known about the various underlying mechanisms. The current state of the art is rather confusing, since theoretical and experimental studies tend to follow different routes. Initiated by the work of Buchsbaum,[2] a new class of computational models has emerged,[3,4] but so far, no attempts have been made to confront these models with experimental data. As if in retribution, no attempts have been made in the recent experimental studies[5,6] to quantify the data within the context of these new models. As a matter of fact, modeling of color constancy data is hardly ever done. A notable exception is the study by McCann et al.[7] in which a sizeable data set was not only collected, but also theoretically accounted for, the latter within the context of the well-known retinex model.[8,9]

In the following we present some of the results of research that started as an "RGB analogue" of the McCann et al. study,[10] but that eventually turned into a much more extensive PhD project.[11] Here only two of the variables tested will be discussed: the chromaticity and the spectral composition of the illuminant. The data we obtained not only allowed the derivation of a simple (but accurate) quantitative model, but also provided the first experimental test of the basic principles underlying the aforementioned computational approaches to color constancy.

Methods

General Procedure

The experimental set-up, already detailed elsewhere,[11,12] can be summarized as follows. Using computer simulation of surface reflectance under various illuminants, a stimulus pattern was generated representing an array of color samples (according to Munsell specifications) displayed on a neutral (spectrally flat) background. This pattern, shown in Figure 1, could be displayed as being alternately illuminated by two different light sources, the "test" and "match" illuminant. A black box (1m length) with two viewing holes was mounted in front of the monitor. Mechanical shutters, located behind the viewing holes, and synchronized with the two display modes, locked left and right eye to the pattern under either test or match illumination.

The observer toggled between the two illuminant conditions (either at will or under computer control), and could thus evaluate to what extent the difference in illumination was attended by a difference in perceived surface color. This was quantified as follows. When the display image was shown under the match illuminant, the observer could control the color of the central patch (which was initially presented as black). The instruction was to match this sample to the sample appearing in the center of the test display. The latter was one out of a selection of 11 samples taken from the total of 35 colors shown in the display.

The matches provide colorimetric specifications of samples that are perceived as identical under different illuminants. These data are informative, therefore, as to how the visual system processes (surface) color under these conditions. Note that perfect color constancy would imply that the observer would not perceive a difference between the samples as seen under test and match illumination, respectively.

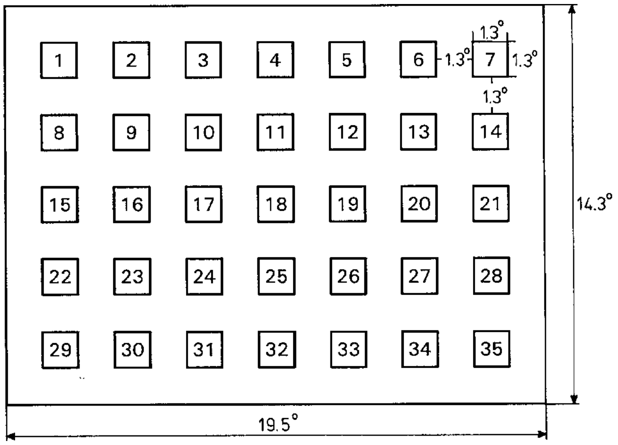

Figure 1. Test/match display. The stimulus pattern represents a computer simulation of surface color samples (squares) on a white or grey background. In the match mode the color of the central patch (here nr. 18) was adjusted by the observer. Sample numbers refer to x,y,Y specifications (cf. Lucassen & Walraven[12]).

Experimental Conditions

Two illuminant variables were tested, chromaticity and spectral composition.

Chromaticity (Experiment 1)

Using a trichromatic (RGB) light reflection paradigm,[11,12] 22 different test/match illuminant combinations were tested. Most of these combinations (16 pairs) were selected from equiluminant sources (12 cd/m^2) with the following colors: W(hite), x=0.313, y=0.329; R(ed), x=0.415, y=0.330; G(reen), x=0.313, y=0.432; B(lue), x=0.259, y=0.241; Y(ellow), x=0.410, y=0.460; M(agenta), x=0.310, y=0.256; C(yan), x=0.227, y=0.308. In three conditions, i.e., the test/match illuminant combinations B/Y, C/R, R/C, the experiments were repeated with either lower (6 cd/m^2) or higher (24 cd/m^2) luminance of the match illuminant. The light "reflected" from the test samples was always 50% of that of the background, which itself was modeled as reflecting 100% of the incident light.

Spectral Composition (Experiment 2)

Knowing the spectral reflectances of the 11 (Munsell) test samples, the color of the light reflected from these samples under respectively broad and narrow-band illumination (also with known spectra) was reproduced on the display.[11] Four different test illuminants were used: two phases of daylight, D_{40} (T_c=4000°K) and D_{250} (T_c=25000°K), and their respective two-wavelength metamers, M_{40} (λ_1=592 nm, λ_2=491.8 nm) and M_{250} (λ_1=560 nm, λ_2=433.7 nm). The match illuminant was the same for all four conditions, i.e. the CIE standard light D_{65} (T_c=6500°K). All illuminants were equated for luminance at 30.4 cd/m^2, resulting in sample luminances of 6 cd/m^2 and a background luminance of 13 cd/m^2, consistent with Munsell Values 5 and 7, respectively.

Results

Experiment 1

Representative examples of the results obtained in Experiment 1, i.e. the data from the test/match illuminant combinations G(reen)/W(hite), Y(ellow)/W(hite), B(lue)/G(reen), and G(reen)/B(lue), are shown in Figure 2.

The upper-left quadrant in Figure 2 illustrates that illuminant induced changes in sample color (compare open and dotted areas) are reasonably well offset by the effect of color constancy. This is indicated by the proximity of chromaticities of observer's matches (hatched area) and of samples under white light (open area). Note that perfect color constancy would imply that the colors of samples seen under the (white) match illumination would not have to be altered in order to match the same samples seen under (green) test illumination. That is, the hatched and open area would be superimposed. The other extreme, no color constancy, would be indicated by superposition of hatched and dotted area (physical match).

The upper-right quadrant in Figure 2, showing the Y/W test/match results, illustrates that the degree of color con-

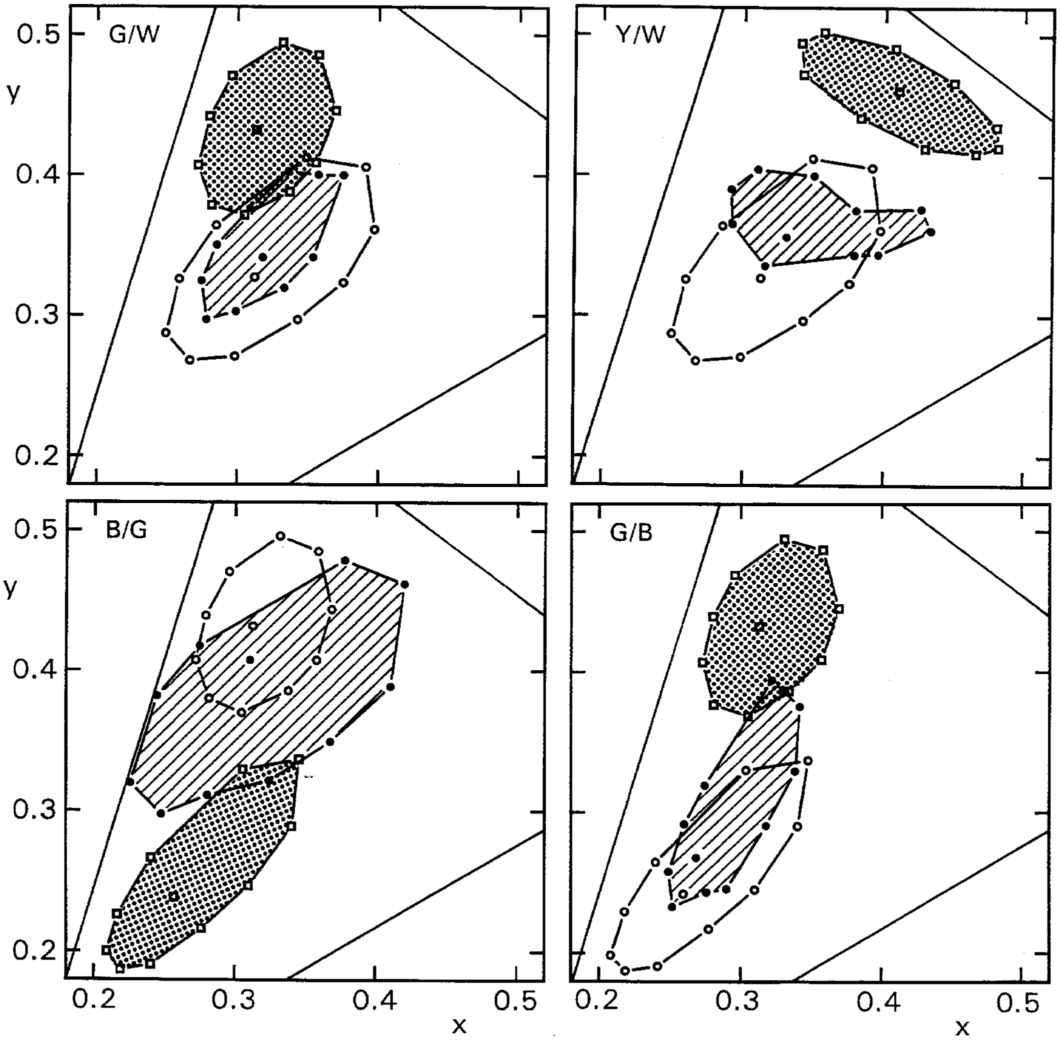

Figure 2. Data samples relating to four different combinations of test/match illumination (G/W, Y/W, B/G, G/B). The three clusters of x,y chromaticities relate to the same 11 test samples, i.e. test samples under test illumination (no area fill), the same samples under match illumination (dotted area fill), and the observer's matches (hatched area fill). Superposition of hatched and open areas would indicate perfect color constancy. Superposition of hatched and dotted areas would indicate no color constancy.

stancy can also be less than that observed in the G/W condition (less overlap of hatched and open area). Apparently, the degree of color constancy is not independent of the color of the illuminant.

The two lower quadrants in Figure 2, in which the results of the B/G and G/B test/match conditions are plotted, show that the degree of color constancy not only depends on the choice of the illuminants to be compared, but also on their allocation as test or match illuminant.

In order to analyze the seemingly complex results shown in Figure 2, we transformed the data from XYZ units to receptor units.[12] These units, cd/m^2 per L(ong)-, M(edium)-, or S(hort)-wave-sensitive cone, are based on Vos-Walraven[13] cone spectral sensitivities, normalized to yield equal L, M,S quantum catches at equal-energy white.[14]

The data analysis[12] yielded a fairly simple expression, transforming the cone-specific quantum catch (Q) into a response (R), according to

$$R = Q_w{}^n \log (k\, Q_j / Q_w)\,, \qquad (1)$$

where Q_j / Q_w represents the cone-specific contrast of a sample (j) relative to the white background (w). The exponent n is observer-dependent ($r \approx 0.3$); the coefficient k depends on sample/background luminance contrast,[11] and takes the value $k = 4.35$ for the 50% luminance contrast used in the experiments discussed here.

Using equation (1), we computed predictions of the observer's matches in terms of Q_j^p, the input per sample (j) for each cone class (p=L,M,S). A comparison of predicted versus obtained matches for all 22 illuminant test/match conditions is shown in Figure 3. Also shown is the same comparison, for a "model" that only computes a physical match (no color constancy), i.e. the cone inputs produced by the test samples.

From the data plots shown in Figure 3 it is clear that the observers make matches that are not just physical matches (top row), but matches that are the result from sensory processes that are apparently reasonably well described by equation (1). We obtained predicted *versus* obtained correlation coefficients (for L, M and S cone inputs) of 0.960, 0.978, and 0.977, which implies an associated explained data variance of 92.2%, 95.6%, and 95.4%, respectively.

Experiment 2

As discussed in the Methods section, this experiment investigates the effect of spectral distribution (narrow-band *versus* broad-band illumination) on color constancy. The motivation for doing so is to evaluate whether current computational models[3,4] — for which spectral distribution is the variable of prime interest — are better suited for predicting this kind of data than equation (1), which is concerned only with cone-specific contrast.

The computational model we tested is a generalized model, based on principles common to all models of its kind. We refer to it as the Judd-Cohen model, because it reconstructs the illuminant spectrum E(λ) and the reflectance function R(λ) on the basis of the principle components analyses by Judd et al.[15] and Cohen,[16] respectively. Given foreknowledge regarding the first three principal components (spectral basis functions) and an estimate of the color coordinates (*not* the spectrum) of the illuminant, the

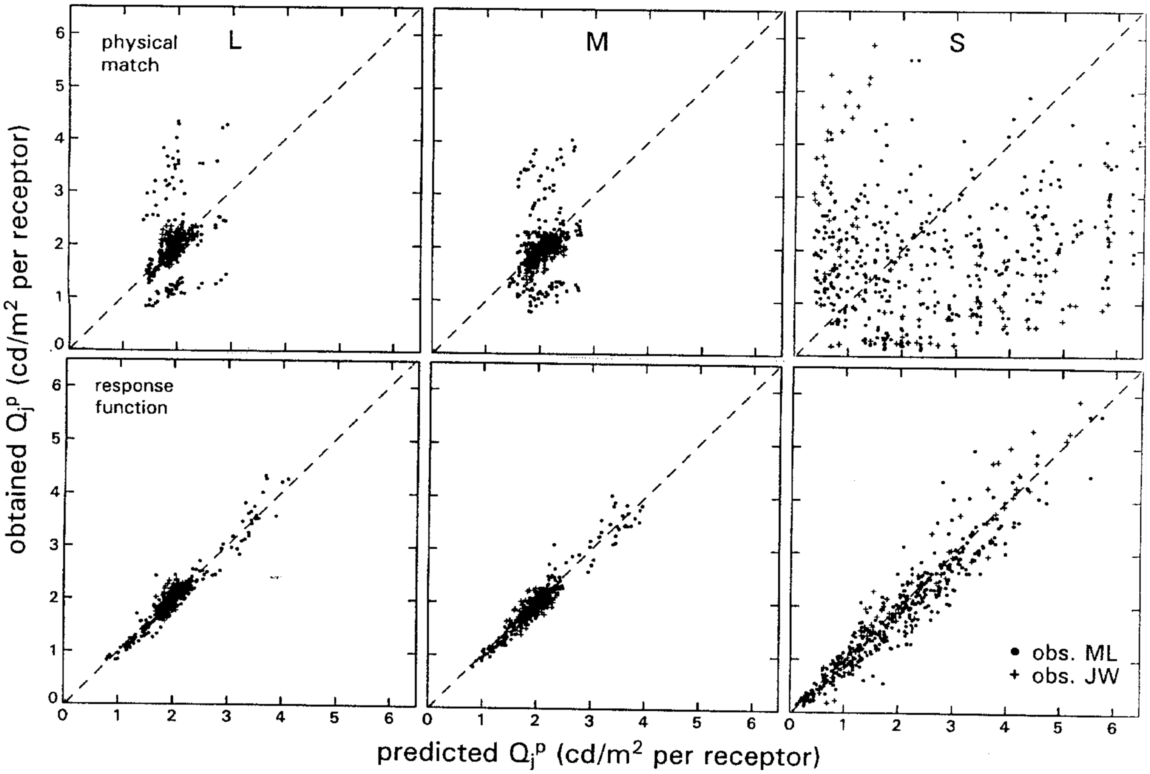

Figure 3. Predicted versus obtained data (matches) for all illuminant combinations tested. The matches are expressed as Q_j^p, the input (cd/m^2) per sample (j) and cone class (p=L,M,S). The upper row shows predictions on the basis of a physical match (reproducing the test sample), whereas the bottom row shows the predictions according to a sensory match, as computed with equation (1).

latter can be eliminated, and (invariant) surface reflectance can be recovered. (The product R(λ) • E(λ) can be decomposed.) Following Buchsbaum,[2] we used the average color of the visual scene as an estimate of the color of the illuminant, an approach commonly known as the "grey-world assumption". This turned out to be quite valid for our stimulus.

Representative examples of the results and their predictions are shown in Figure 4. The data obtained under narrow-band illumination (M_{250}) show the expected break-down of color constancy. The chromaticity space covered by the test samples is collapsed onto a single line, and so are the matches.

Comparison of obtained and predicted results shows that the performance of the response function (equation (1) is somewhat better than that of the more sophisticated Judd-Cohen computational model.

We computed a prediction error in terms of distances (d) in the CIE u',v' chromaticity diagram, a perceptually more uniform representation than the x,y diagram. For the prediction based on the response function, as obtained for the 4 different test/match conditions D_{40}/D_{65}, D_{250}/D_{65}, M_{40}/D_{65}, M_{250}/D_{65}, the average prediction error $d_{u',v'}$ took values of 0.0074, 0.0049, 0.0073, and 0.0072. The corresponding errors for the Judd-Cohen model were 0.0119, 0.0227, 0.0210, and 0.0263, values that are 1.6 to 4.6 times higher than those obtained for the response function.

Discussion

The results of this study, as well as those of other labora-tory studies,[5,6] indicate that the visual system does not achieve perfect color constancy. This is a common, but often ignored finding. It may even be denied in the Judd-Cohen type of computational approaches, these typically being preoccupied with obtaining a very high accuracy in the recovery of surface reflectance.[17] This may well be the reason for our finding that such a type of model predicts the data less well than the simple response func-tion described by equation (1). On the other hand, models of the Judd-Cohen type may be quite useful in machine vision applications. In this field one may also consider other options for machine color constancy.[18]

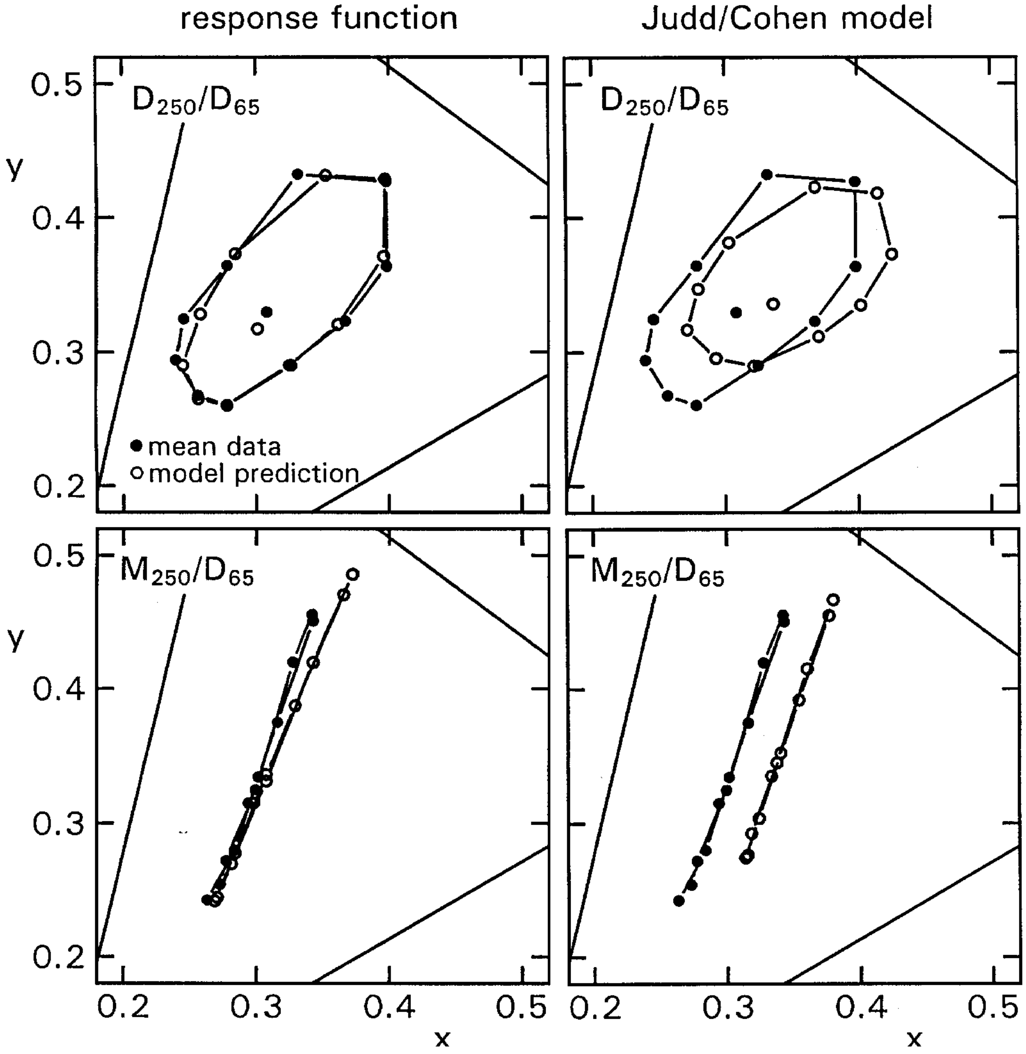

Figure 4. Averaged observers' matches (filled circles) and their predictions (open circles), as computed with equation (1) or the Judd-Cohen computational model. Upper and lower panels show the results for a broad-band test light (D_{250}) and its two-wavelength metamer (M_{250}), respectively. The matches are made under D_{65} broad-band illumination.

The response function we derived is no more than a first attempt at quantification of data obtained under the usual, simplified laboratory conditions. Still, it probably correctly identifies relative cone input—the term Q_j/Q_w in equation (1)—as an important mediator of color constancy.[10] We thus confirm a basic principle underlying visual sensitivity control,[19] as is also embodied in the normalizing procedure underlying the retinex model.[9] An interesting question is whether the mechanisms involved operate on globally sampled information,[20,21] or mainly on a more or less locally determined contrast signal.[22,23]

Color constancy may not be perfect, but it keeps the (perceived) variance of surface colors within acceptable limits. It is worthwhile, therefore, to consider it in the context of the fidelity of (electronic) color reproduction. The quality of color reproduction is usually quantified in terms of a mismatch in the mapping of original and reproduced colors, in a suitable three-dimensional color space. If the mismatch resembles the effect of a change in illumination (cf. Figure 2)—or can be made to do so—then one may expect color constancy to remove at least part of the perceived color differences between original and reproduced image.

References

1. H. L. F. von Helmholtz, *Handbuch der physiologischen Optik* (3rd ed., 1st ed. 1866), W. Nagel, A. Gullstrand, and J. von Kries, Eds, Voss, Hamburg/Leipzig, 1911.
2. G. Buchsbaum, A spatial processor model for object colour perception, *Journal of the Franklin Institute,* **310**, 1-26 (1980).
3. L. T. Maloney and B. A. Wandell, Color constancy: a method for recovering surface spectral reflectance, *Journal of the Optical Society of America A,* **3**, 29-33(1986).
4. M. D'Zmura and P. Lennie, Mechanisms of color constancy, *Journal of the Optical Society of America A,* **3**, 1662-1672(1986).
5. L. E. Arend and A. Reeves, Simultaneous color constancy, *J. of the Optical Society of America A,* **3**, 1743-1751 (1986).
6. A. Valberg and B. Lange-Malecki, Colour constancy in Mondrian patterns: a partial cancellation of physical chromaticity shifts by simultaneous contrast, *Vision Research,* **30**, 371-380 (1990).
7. J. J. McCann, S. P. McKee, and T. H. Taylor, Quantitative studies in retinex theory: a comparison between theoretical predictions and observer responses to the "color Mondrian" experiments, *Vision Research,* **16**, 445-458(1976).
8. E.H. Land, The retinex, *American Scientist,* **52**, 247-264 (1964).
9. E. H. Land, Recent advances in retinex theory, *Vision Research,* **26**, 7-21(1986).
10. J. Walraven, T. Benzschawel, B. E. Rogowitz, and M. P. Lucassen, Testing the contrast explanation of color constancy, in *From Pigments to Perception*, A. Valberg and B. B. Lee, Eds, Plenum Press, New York, 1991, pp. 369-378.
11. M. P. Lucassen, *Quantitative studies of color constancy*, Thesis, University of Utrecht, 1993, 181 p.
12. M. P. Lucassen and J. Walraven, Quantifying color constancy: evidence for nonlinear processing of cone-specific contrast, *Vision Research,* **33**, 739-757 (1992).
13. J. J. Vos and P. L. Walraven, On the derivation of the foveal receptor primaries, *Vision Research,* **11**, 799-818 (1971).
14. J. Walraven and J. S. Werner, The invariance of unique white; possible implications for normalizing cone action spectra, *Vision Research,* **31**, 2185-2193(1991).
15. D. B. Judd, D. L. MacAdam, and G. Wyszecki, Spectral distribution of typical daylight as a function of correlated color temperature, *Journal of the Optical Society of America,* **54**, 1031-1040(1964).
16. J. Cohen, Dependency of the spectral reflectance curves of the Munsell color chips, *Psychonomic Science,* **1**, 369-370 (1964).
17. J. L. Dannemiller, Spectral reflectance of natural objects: how many basis functions are necessary? *Journal of the Optical Society of America A,* **9**, 507-515 (1992).
18. C. van Trigt, Smoothest reflectance functions. I. Definition and main results, *Journal of the Optical Society of America A,* **7**, 1891-1904(1990).
19. J. Walraven, C. Enroth-Cugell, D.C. Hood, D. I. A. MacLeod, and J. L. Schnapf, The control of visual sensitivity: receptoral and postreceptoral processes, in *Visual Perception: the Neurophysiological Foundations*, L. Spillmann and J. S. Werner, Eds, Academic Press, San Diego, 1990, pp. 53-101.
20. J. J. McCann, Local/global mechanisms for color constancy, *Die Farbe*, **34**, 275-283 (1987).
21. J. J. McCann and R. L. Savoy, Measurements of lightness: dependence on the position of a white in the field of view, *Human vision, visual processing, and digital display II, SPIE* Vol. **1453**, 402-411 (1991).
22. J. Walraven, Spatial characteristics of chromatic induction; the segregation of lateral effects for straylight artifacts, *Vision Research*, **11**, 1739-1753 (1973).
23. R. Shapley, The importance of contrast for the activity of single neurons, the VEP and perception, *Vision Research*, **26**, 45-61 (1986).

published previously in the IS&T 1993 Color Imaging Conference Proceedings, page 60