

Quantifying Color Constancy: Evidence for Nonlinear Processing of Cone-Specific Contrast

MARCEL P. LUCASSEN,* JAN WALRAVEN*

Received 21 April 1992; in revised form 31 August 1992

Color constancy was studied by the method of comparing color samples under two different illuminants using a CRT color monitor. In addition to the classical approach in which one of the illuminants is a (standard) white, we performed experiments in which the range of differential illumination was extended by using pairs of lights that were *both* colored. The stimulus pattern consisted of an array of thirty-five color samples (including five neutral samples) on a white background. A trichromatic illuminant-object interaction was simulated analogous to that resulting from illumination by three monochromatic lights. The test samples, as seen under “test” and “match” illumination, were successively presented to the left and right eye (haploscopic matching). The data show systematic deviations from predictions on the basis of cone-specific normalization procedures like those incorporated in the Retinex algorithm and the von Kries transformation. The results can be described by a nonlinear response transformation that depends on two factors, receptor-specific sample/background contrast and the extent to which the illuminant stimulates the receptor system in question. The latter factor explains the deviations. These are mainly caused by the short-wave-sensitive system, as a consequence of the fact that this system can be more selectively stimulated than the other, spectrally less separated, cone systems.

Color vision Color constancy Trichromatic reflectance Cone-specific contrast Nonlinear response

INTRODUCTION

Object colors are perceived as more or less constant, despite considerable variations in the color of the ambient light. This is the well-known phenomenon of color constancy, a subject with a long history, but still an area of many unresolved issues. The central problem of color constancy is usually cast into terms of how the visual system is capable of decomposing the product of illuminance \times reflectance, that is, separating light from matter. Obviously, this is impossible when these two variables are spatially and temporally inseparable, as is the case for a homogeneous surface illuminated in a dark void. However, for a slightly more complex stimulus, it is already possible to develop models that, under certain constraints and assumptions, are capable of recovering surface reflectance (Buchsbaum, 1980; Dannemiller, 1989; D’Zmura, 1992; D’Zmura & Lennie, 1986; Gershon & Jepson, 1989; Lee, 1986; Maloney & Wandell, 1986; Yuille, 1987). These computational models typically try to estimate the illuminant based on the image spatial context (e.g. Buchsbaum, 1980), and for that purpose require sampling responses over large retinal areas. Such mechanisms run into problems when confronted with rapid local changes in illumination. This problem has been addressed by Rubin and Richards

(1982) who discuss an operator that responds to edges that are most likely due to reflectance changes only, and hence, provides a means for discriminating light from material changes.

A somewhat different approach to color constancy, not explicitly directed at estimating the illuminant is embodied in the well-known Retinex model(s) by Land and coworkers (e.g. Land, 1959, 1986a, b; McCann, 1971; Land & McCann, McKee & Taylor, 1976). The Retinex model incorporates an algorithm that calculates lightness values within each cone system. It has been used for describing the results of the color constancy experiments by McCann *et al.* (1976). The results from a study by Creutzfeldt, Lange-Malecki and Dreyer (1990) were similarly analyzed in terms of receptor-specific inputs that are scaled before contributing to the trichromatic color signal. Actually, the principle underlying the Retinex model is akin to a von Kries type recalibration (von Kries, 1905), as has been pointed out by various authors (e.g. Valberg & Lange-Malecki, 1990; Jameson & Hurvich, 1989). In the von Kries color transformation scheme the output of each receptor is recalibrated, so as to compensate for changes in the color signal elicited by a (perfect) white reflector. Insofar as this is a linear scaling relative to a (reference) white, such an adjustment implies responding to cone-specific lightness. It can be shown that this is not a solution for complete color constancy (Worthey, 1985; Worthey & Brill, 1986; Brill & West, 1986). On the other hand, the

*TNO Institute for Perception, P.O. Box 23, 3769 ZG, Soesterberg, The Netherlands.

visual system does not exhibit perfect color constancy, so this actually could be consistent with the performance of a von Kries operator.

There are many unresolved questions with regard to color constancy (cf. Jameson & Hurvich, 1989), which may at least partly be attributed to the fact that the results of most of the studies in this field are quantified in terms of CIE x, y chromaticity units, a rather indirect, and also incomplete measure (when luminance is not specified) of the physiological stimulus. Another major problem is methodology. Arend and Reeves (1986) have shown that there is quite a difference between matching on the basis of perception as opposed to recognition. That is, a particular sample may not exhibit color constancy (as judged by a color match), but may nevertheless be correctly identified on the basis of various cues or inferences about how an illuminant may change the color of the sample in question. In our experiments the observers were instructed to base their matches on perceived color rather than recognition. We deliberately chose this approach since we were interested in isolating a purely sensory response, not (yet) influenced by whatever cognitive cues the visual system might employ.

In a precursor to this study, it could be shown that, under the conditions of the particular color constancy paradigm employed (to be discussed in Methods), cone-specific contrast (rather than absolute light input) provides the relevant signal for color perception (Walraven, Benzschawel, Rogowitz & Lucassen, 1991). However, in that same study it was also shown that the short-wave-sensitive (S-) cones did not quite fit such a simple scheme, a discrepancy that can also be inferred from other studies as well (McCann & Houston, 1983; Troost, Wei & de Weert, 1992). One of the purposes of the present study was to further explore the discrepant behavior of the S-cones. We thereto extended the range of differential stimulation by comparing not only white vs colored, but also colored vs colored illuminant conditions. The data thus obtained not only enabled us to better probe the short-wave system, but also to further test the validity of the hypothesis that color constancy is at least partly mediated by cone-specific contrast processing.

METHODS

Illuminant-object interaction

The simulation of surface color on a CRT monitor requires generating screen luminances, Y_R , Y_G and Y_B , that produce the same visual effect (i.e. the same XYZ tristimulus values) as that of the light reflected from the surface in question. The color of that light is determined by two variables, illumination and reflectance. The interaction between these two variables can be computed if the emission and reflectance spectra involved have been defined, either for real lights and objects or synthetic stimuli. We chose for the latter type of stimuli, for reasons to be discussed at the end of this section.

The illuminant-object interaction that we simulated was the same as that used in an earlier study (Walraven

et al., 1991). The spectral interactions between light and matter are registered by few coefficients modulating the outputs of three light channels. This is illustrated in Fig. 1. In Fig. 1 three primary light channels are shown representing a red, green and blue luminance channel. The values Y_{R_w} , Y_{G_w} and Y_{B_w} denote the (reference) luminances required for producing white light. By introducing emission coefficients, a_R , a_G and a_B , the white light can be changed into colored light. Surface reflectance is described by the (fixed) reflection coefficients, b_R , b_G and b_B , that determine how much light a particular sample reflects within each luminance channel. So, for a given sample-illuminant combination, the reflected light, L_r , is given by

$$L_r = a_R b_R Y_{R_w} + a_G b_G Y_{G_w} + a_B b_B Y_{B_w}. \quad (1)$$

The six coefficients in equation (1) are constant, for a particular combination of illuminant and surface sample. However, the constants are confounded in the product $a_i \times b_i$ ($i = R, G, B$), just as light and matter are confounded in the wavelength product $E(\lambda) \times R(\lambda)$ in the wavelength domain.

The video implementation of the trichromatic principle described by equation (1) is achieved by letting Y_R , Y_G and Y_B correspond to the luminances of the red, green and blue light emitted by the RGB phosphors. That is, the visual stimulus is produced by the additive mixture of three monochrome images (red, green and blue), each of which varies in luminance only. The illuminant-object interaction that we created in this way can be best described as the video analog of the method used by McCann *et al.* (1976). In their study the test pattern, the well-known "Mondrian" pattern, was illuminated by three near-monochromatic lights. By varying the luminance ratio of these primary sources, a wide gamut of easily quantifiable colored illuminants could be created. The only difference with the stimuli used in the study of McCann *et al.* (1976) is that in their study each monochrome image was generated with near-monochromatic light (cf. Young, 1987).

In order to compute trichromatic emission and reflectance coefficients for the illuminants and the test

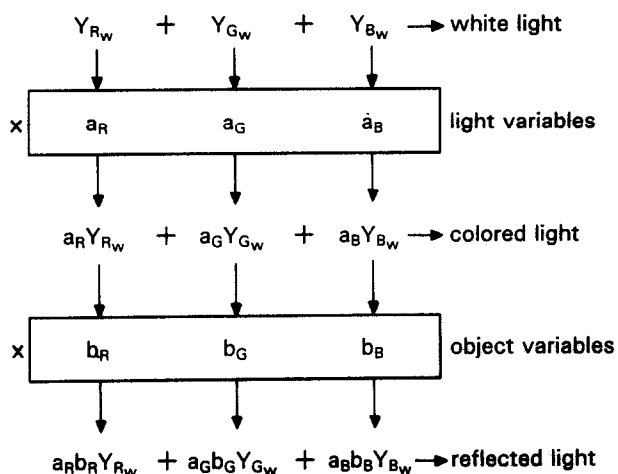


FIGURE 1. Diagram illustrating the principle of trichromatic light reflection.

samples, they were specified in CIE *XYZ* units (a normal procedure in CRT colorimetry). The procedure for computing the phosphor luminances Y_R , Y_G and Y_B for producing specified X , Y , Z tristimulus values (and vice versa) are detailed in Appendix B. For a choice of test and illuminant colors the computation proceeds as follows:

1. Transform X , Y , Z of the standard white illuminant into phosphor luminances. These are indicated as Y_{R_w} , Y_{G_w} and Y_{B_w} in Fig. 1.
2. Transform X , Y , Z of the (colored) illuminant into phosphor luminances $Y_{R,ill}$, $Y_{G,ill}$ and $Y_{B,ill}$. Calculate the emission coefficients a_i with

$$a_i = Y_{i,ill} / Y_{i_w} \quad i = R, G, B. \quad (2)$$

3. Transform X , Y , Z of a sample j under the standard white illuminant into phosphor luminances $Y_{R,j}$, $Y_{G,j}$ and $Y_{B,j}$. Calculate the reflection coefficients b_i with

$$b_i = Y_{i,j} / Y_{i_w} \quad i = R, G, B. \quad (3)$$

4. Compute the three phosphor luminances Y_R , Y_G , Y_B of the light reflected from the sample according to $Y_R = a_R b_R Y_{R_w}$, $Y_G = a_G b_G Y_{G_w}$ and $Y_B = a_B b_B Y_{B_w}$ [equation (1)].

In Appendix A a numerical example is given of the various steps involved in computing phosphor CRT luminance for a particular sample/illuminant combination. The matrix transformations from *XYZ* to $Y_R Y_G Y_B$, and vice versa, are detailed in Appendix B.

The main advantage of using a trichromatic reflectance paradigm is that one is no longer constrained by the limited choice of tabulated spectra (in particular illuminant spectra). So, if one wants to explore a wide color gamut—as in this experiment—the stimuli can be freely chosen in the *XYZ* domain. The only restriction, common to all CRT applications in color vision, is imposed by the boundaries of the color space covered by the color monitor.

Another, less obvious, advantage is that this method avoids the problem of illuminant metamerism (Worthey, 1985; Worthey & Brill, 1986). This is observed when changes in illumination cause different colors to be registered as identical in the cone pigments (or vice versa). Such departures from color constancy cannot be removed by the visual system, and may thus obscure how the visual system deals with the more interesting, i.e. soluble, problems of color constancy.

Finally, the trichromatic specification of reflectance and emission greatly simplifies the computation of the interaction between these two variables. This may not be of such importance for our relatively simple test pattern, but it does become a consideration for more complex or dynamic visual scenarios.

Equipment

A Hitachi 19 in. high resolution color monitor, driven by an 8-bit/gun video card of a Sun 3/260 computer was used for presenting the stimulus pattern. With the aid of

a SpectraScan PR-702AM spectroradiometer (which automatically converted spectral energy distributions into x , y , Y coordinates) and a Spectra Pritchard photometer (both from Photo Research), the standard procedure (e.g. Cowan, 1986) was used for the initial monitor calibration. Daily calibration checks were done with a simple recalibration algorithm, that was designed for use with the kind of stimuli to be discussed. The exact details of the algorithm are reported elsewhere (Lucassen & Walraven, 1990). In short, the purpose of the recalibration algorithm was to maintain accurate color reproduction, and to avoid the time consuming calibration measurements that are normally necessary whenever the contents of the displayed image are changed, or when a substantial amount of time (in the order of days) has elapsed since initial calibration. Before each trial a white reference stimulus was generated on the display, with phosphor luminance settings (according to initial monitor calibration) required for producing the desired white CIE x , y , Y values. The actual x , y , Y values that appeared on the screen were measured with the spectroradiometer. This single measurement provided an estimate of the extent the monitor was out of calibration. From the measured x , y , Y values, three scale factors were derived (one for each color gun) that were used to calculate the new set of (scaled) input–output relations that determine, with 8 bit/gun precision, the monitor's light output. The recalibration algorithm enabled color reproduction with an average error of about 0.005 CIE x , y units (Lucassen & Walraven, 1990).

A large viewing pyramidal box (blackened inside) with two viewing holes was placed in front of the monitor, so as to prevent the screen from illuminating the (dark) environment. At the 1 m viewing distance we used, the monitor's screen subtended a maximum visual angle of $20 \times 16^\circ$. With the aid of a mechanical shutter, under the observer's control, the left and right viewing hole could be alternately opened and closed in synchrony with the presentation of a test and a match pattern. In this way each pattern was only seen by either the left or the right eye. The two patterns were calculated to present a set of color samples illuminated successively by two different light sources. This method has been considered to be most satisfactory for studies on asymmetric matching (Eastman & Brecker, 1972).

Stimulus

Test samples. The stimulus pattern consisted of an array of 35 color samples displayed on a white surface. For the sake of simplicity, the latter was (arbitrarily) defined as an ideal white, reflecting 100% of the incident light. It could also be envisaged, of course, as reflecting, say, 50% in combination with an illuminant emitting twice as much light. The samples were presented as $1.3 \times 1.3^\circ$ squares, with 1.3° mutual separation, resulting in the arrangement shown in Fig. 2.

This is the same configuration as used in the study by Walraven *et al.* (1991), the precursor to the present one. Since the luminance of the background was always such

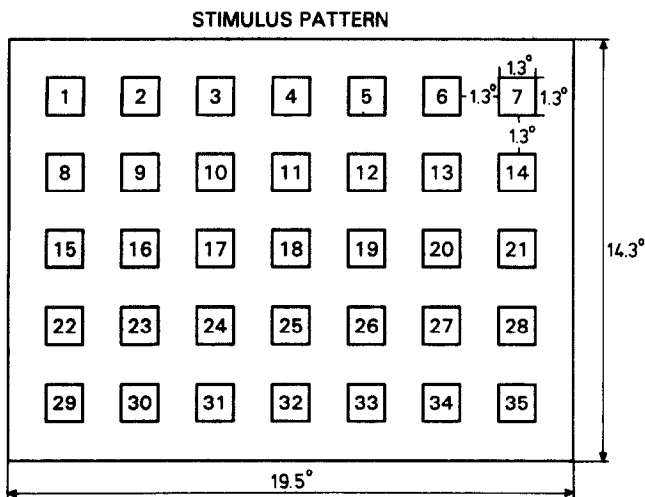


FIGURE 2. Geometry of the stimulus pattern used for both test and match stimulus. The pattern is a computer simulation of 35 color samples (1.3° squares), displayed on a piece of white paper (grid). The whole pattern can be illuminated by various (simulated) light sources. The sample numbers correspond to those in Table 1, where the corresponding color specifications are given.

that it simulated a 100% reflector, it had exactly the same luminance and color of the illuminant. The background (grid) thus conveys the illuminance and exact color of the illuminant. The x, y chromaticities of the 35 test samples are the x, y equivalents of a selection of Munsell samples under *white* light [Wyszecki and Stiles (1982) Table I (6.6.1)], and are plotted in Fig. 3.

The specifications of the color samples are listed in Table 1. There are 30 chromatic and five achromatic samples. The x, y chromaticities of the chromatic samples were selected from three loci of equal Munsell Chroma (6, 4 and 2) at Munsell Value 5. The luminance of the 30 chromatic samples was set to be consistent with 50% of the luminance of the white light. The remaining

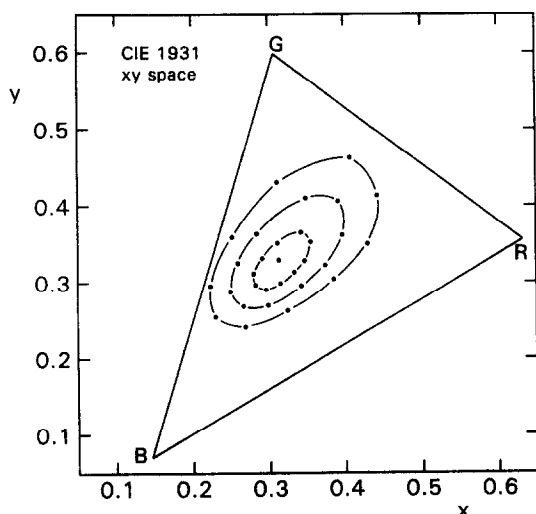


FIGURE 3. Chromaticities (CIE x, y coordinates) of the color samples (Table 1), under standard white illumination (RGB metamer of D_{65} , 12 cd/m^2). The triangle encloses the chromaticity space covered by our color monitor. The set consists of 30 samples, with chromaticities evenly distributed over three loci of equal Munsell Chroma (5/6, 5/4, 5/2), and five achromatic samples (represented as a single point in the center).

five achromatic samples covered a luminance range representing reflectances of 10, 25, 50, 75 and 90%. For example, under our standard white light (the RGB metamer of D_{65} , 12 cd/m^2), the luminance of the chromatic samples was 6 cd/m^2 , whereas the luminance of the achromatic samples was 1.2, 3, 6, 9 and 10.8 cd/m^2 , respectively. Strictly speaking the Munsell Color System does not allow for changing luminance separately from Munsell Value, since these are inextricably tied up with each other. Munsell Value 5, for instance, corresponds to a luminance factor of 19.77, i.e. about 20% reflection. However, the Munsell system enables selecting x, y equivalents to obtain a set of samples with chromaticities that are perceptually equi-distant (under white light) from the white point. The 50% reflectance was chosen in order to prevent the samples from appearing too dark. (In practical applications the Munsell chips are supposed to be displayed on a gray background of 20% reflectance.)

The sample numbers in Table 1 correspond to the array numbers in Fig. 2. Although this arrangement of sample colors might look completely random, there is a reason for the given distribution. Each sample in the array has its complementary color of equal saturation in the position that is mirrored through the center patch (No. 18). During an experiment, this center patch is the point of interest, that being the locus for an interocular match in the haploscopic protocol (to be discussed henceforth.) In this way, the local color average over a few neighboring patches of this center is more or less balanced, resembling the global average. To reduce the effect of (uncontrolled) adaptation through eye movements outside the matching area, the most saturated background samples were allocated to the most peripheral positions of the test pattern. The rationale behind these precautions was to study the more or less isolated effect of illuminant changes, that is, without contamination of (possible) interactions between samples. We might possibly also have used a single homogeneous background, as has been argued by Valberg and Lange-Malecki (1990), but we had planned to treat that as a separate condition in later experiments.

In order to be able to relate our experimental data to those obtained in the earlier study (Walraven *et al.*, 1991), we also used a stimulus pattern with only neutral samples (all 50% reflectance), except for the one in the center of the pattern. This pattern was used in order to keep one eye as neutrally adapted as possible. This was the eye in which the color of the central patch of the stimulus pattern was matched to the color of the corresponding sample, as seen by the other eye, under a different illuminant. The pattern with achromatic samples is called the "match pattern", whereas the pattern with colored samples is referred to as the "test pattern". This difference in test and match pattern was only used in the first series of experiments (see later Expt 1). In all other experiments the test and match pattern consisted of the same samples (same reflection properties), the illuminant being the only variable.

Illuminants. Following the classical approach in color constancy, our first series of experiments involved

TABLE 1. Specification of the 30 chromatic and five achromatic samples of the stimulus pattern shown in Fig. 2

Sample number in Fig. 2	CIE 1931 specification			Reflection coefficient in simulation			Munsell chip with same x, y
	x	y	Y (cd/m ²)	b_R	b_G	b_B	
21	0.3243	0.2630	6.0	0.85	0.34	0.76	10 P 5/6
12	0.3851	0.3039	6.0	1.03	0.31	0.48	10 RP 5/6
8	0.4299	0.3499	6.0	1.06	0.33	0.28	10 R 5/6
14	0.4428	0.4128	6.0	0.87	0.42	0.13	10 YR 5/6
9	0.4072	0.4621	6.0	0.57	0.53	0.08	10 Y 5/6
15	0.3108	0.4301	6.0	0.24	0.62	0.24	10 GY 5/6
24	0.2519	0.3587	6.0	0.10	0.64	0.49	10 G 5/6
28	0.2234	0.2952	6.0	0.06	0.62	0.77	10 BG 5/6
22	0.2299	0.2548	6.0	0.20	0.54	0.98	10 B 5/6
27	0.2686	0.2412	6.0	0.55	0.42	0.99	10 PB 5/6
4*	0.2986	0.2699	6.0	0.64	0.42	0.77	5 P 5/4
31*	0.3421	0.2954	6.0	0.81	0.38	0.58	5 RP 5/4
20*	0.3740	0.3220	6.0	0.88	0.38	0.44	5 R 5/4
29*	0.3968	0.3614	6.0	0.84	0.41	0.29	5 YR 5/4
34*	0.3915	0.4057	6.0	0.66	0.48	0.20	5 Y 5/4
32*	0.3482	0.4097	6.0	0.45	0.55	0.24	5 GY 5/4
5*	0.2841	0.3628	6.0	0.25	0.60	0.43	5 G 5/4
16*	0.2591	0.3246	6.0	0.21	0.59	0.59	5 BG 5/4
7*	0.2493	0.2879	6.0	0.24	0.56	0.76	5 B 5/4
2*	0.2662	0.2687	6.0	0.42	0.49	0.83	5 PB 5/4
25	0.3148	0.2986	6.0	0.62	0.44	0.61	10 P 5/2
33	0.3332	0.3131	6.0	0.68	0.43	0.53	10 RP 5/2
17	0.3465	0.3278	6.0	0.70	0.44	0.46	10 R 5/2
1	0.3546	0.3524	6.0	0.65	0.46	0.37	10 YR 5/2
6	0.3422	0.3648	6.0	0.55	0.50	0.35	10 Y 5/2
11	0.3110	0.3508	6.0	0.42	0.54	0.43	10 GY 5/2
3	0.2910	0.3310	6.0	0.37	0.54	0.52	10 G 5/2
19	0.2796	0.3111	6.0	0.36	0.53	0.61	10 BG 5/2
35	0.2821	0.2966	6.0	0.42	0.50	0.67	10 B 5/2
30	0.2959	0.2905	6.0	0.53	0.47	0.68	10 PB 5/2
23	0.3127	0.3290	1.2	0.10	0.10	0.10	—
10	0.3127	0.3290	3.0	0.25	0.25	0.25	—
18*	0.3127	0.3290	6.0	0.50	0.50	0.50	—
26	0.3127	0.3290	9.0	0.75	0.75	0.75	—
13	0.3127	0.3290	10.8	0.90	0.90	0.90	—

The samples are arranged in four blocks, the first three representing sets at decreasing levels of saturation (Munsell Chroma /6,4,2). All chromatic samples were presented as if reflecting 50% under white light (RGB metamer of D₆₅, 12 cd/m²). The fourth block contains the achromatic samples in the range 10–90% reflectance. The 11 samples of the test set are indicated by an asterisk.

comparing test samples under white and colored light, respectively. The (colored) illuminants produced equal grid luminances (12 cd/m²) and equal Munsell Chroma (Chroma/6). The x, y, Y specification and emission coefficients a_R, a_G, a_B for these illuminants, including that of the white reference illuminant (RGB metamer of D₆₅), are presented in Table 2. As shown in Fig. 4, the chromaticities of the set of (equi-luminant) light sources form complementary pairs located on lines passing through the RGB primaries and the white point. The effect of a change from white to colored illumination is to shift the chromaticities of the color samples in the direction of that of the color of the illuminant. An example of such an illuminant-induced color shift of the test-pattern is shown in Fig. 5. These are the chromaticity coordinates of the illuminant–reflectance products as described above (Appendix A). The luminance of the chromatic samples was 6 cd/m² in white light. Under colored illumination, however, these samples may reflect

more or less of the incident light, depending on the color of the illuminant. As a result, the sample luminances under colored illumination will no longer be the same. This interaction between light and sample color is correctly calculated by the reflectance model.

In the experiments, two differently illuminated sets of color samples (like the ones shown in Fig. 4) were successively shown to the observer. In one of the illuminant conditions, the match condition, the observer could vary the color of the central patch of the stimulus pattern (shown in Fig. 2), to match it to the color of the corresponding test sample seen under the other illuminant (test condition). The two illuminants will be referred to as the “match illuminant” and the “test illuminant”, respectively. In studies on color constancy the match or reference illuminant is usually white. This is also the paradigm we used for our first series of experiments.

In addition to the classical white vs colored illuminant comparison, we also compared a wide variety of colored

TABLE 2. CIE 1931 x, y chromaticities and luminance Y (at 100% grid reflectance) of the test illuminants (see also Fig. 4)

Test illuminant	CIE (1931) specification			Phosphor luminances			Emission coefficient in simulation		
	x	y	Y (cd/m ²)	Y_R	Y_G	Y_B	a_R	a_G	a_B
W(hite)	0.313	0.329	12.0	2.85	8.11	1.04	1.00	1.00	1.00
R(ed)	0.415	0.330	12.0	6.13	5.13	0.74	2.15	0.63	0.71
G(reen)	0.313	0.432	12.0	1.41	10.11	0.49	0.49	1.25	0.47
B(lue)	0.259	0.241	12.0	2.70	7.18	2.12	0.95	0.89	2.04
Y(ellow)	0.410	0.460	12.0	3.37	8.46	0.17	1.18	1.04	0.16
M(agenta)	0.310	0.256	12.0	4.47	5.80	1.72	1.57	0.72	1.65
C(yan)	0.227	0.308	12.0	0.30	10.22	1.47	0.11	1.26	1.41

Also shown are the phosphor luminances Y_R, Y_G, Y_B required for generating the x, y, Y specifications of these lights on our CRT (see Appendix B) and the emission coefficients a_R, a_G, a_B used in the reflected light simulation.

vs colored illuminant pairs. We were thus able to obtain a much larger range of differential chromatic stimulation than is possible with the classical white vs colored illuminant comparison. In total, with the inclusion of the six equi-luminant white vs colored pairs, we used 45 different illuminant pairs. Among this set were combinations of the colored illuminants listed in Table 2 [Y(ellow)/B(lue) for instance], and also illuminants with higher saturation. There is no need for presenting the specifications of all these lights. They were mainly selected for creating illuminants with predetermined cone input ratios. We were thus able to better distribute the illuminant pairs over the range of differential cone stimulation covered by the phosphors of our monitor. The assumptions and computations needed to convert phosphor luminance from the RGB domain to (relative) cone inputs in the LMS domain, are presented in Appendices B and C.

Procedure

After 5 min dark adaptation and a few more minutes for adapting to the average luminance of the test pattern

(about 10 cd/m²), the observer started the first presentation of the two illuminant conditions to be compared during a session. When viewing the test (left eye) and match pattern (right eye) the observer concentrated on the central patch. The latter, which could be controlled in the match pattern, was initially black. By pressing the space bar of the keyboard the observer could switch back and forth between test and match pattern. The color of the central patch was controlled by eight keys; four for increasing the luminance of the three RGB guns, either singly or in unison (brightness key), and another four keys for the opposite action. There were no restrictions on time, fixation and number of presentations.

In the first series of experiments all measurements were replicated. Since the reproducibility of the matches was quite satisfactory, we concluded that there was no need to do the same for all of the remaining conditions. Moreover, since the effect of an illuminant—which is actually the variable of interest—is registered by eleven color matches, these can be considered as measurements of the same variable. The two authors, both with normal color vision, served as subjects (ML and JW). However, 80% of the data was supplied by ML. The data of JW

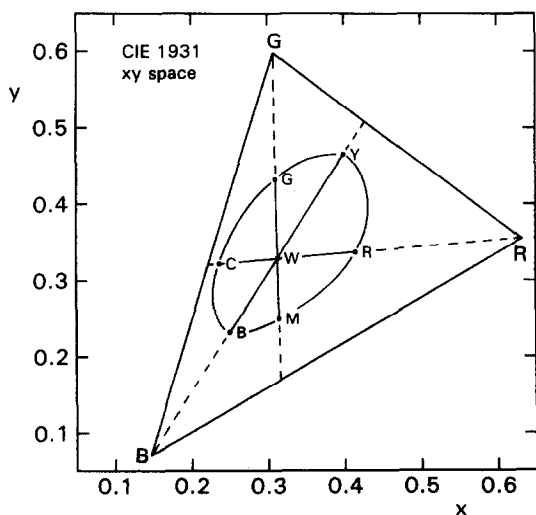


FIGURE 4. Chromaticities of the test illuminants, B(lue), C(yan), G(reen), Y(ellow), R(ed), M(agenta) and W(hite). The x, y coordinates of the colored lights are located on the Munsell 5/6 Chroma line and represent lights that are perceptually equi-distant from the white point. The corresponding chromaticities are listed in Table 2.

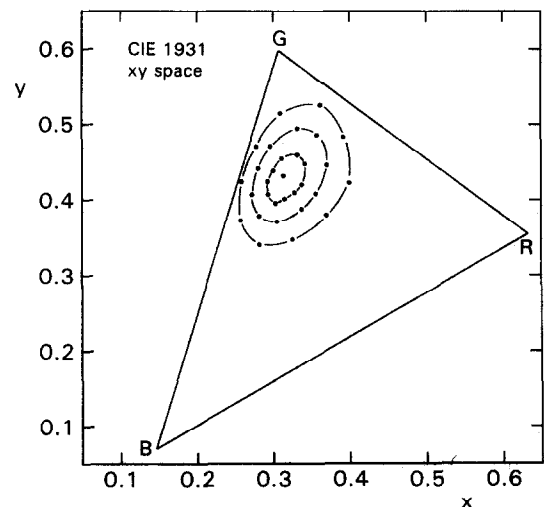


FIGURE 5. The same color samples as shown in Fig. 3, but now illuminated by green light. The samples are now centered around the chromaticity locus of the green illuminant, which coincides with that of the achromatic samples.

represent selected replications (or additions) evenly distributed over the whole stimulus range covered by ML.

Task

The observers adjusted the central patch in the match pattern so as to make it exactly match the hue, saturation and brightness of the corresponding sample in the test pattern. They were free to use as many test/match alternations as were necessary to obtain a satisfactory match, but were instructed to divide the presentations roughly equally between the left and right eye.

RESULTS

In the pilot stage of this study, we did an experiment in which we made (haplosopic) color matches without introducing a difference in illumination, the white versus white illuminant combination. We refer to this as the "trivial match" condition. It served to test the reliability of the method and also provided a check on possible interocular differences in chromatic sensitivity. No such differences were found for either observer. As for the precision of the haplosopic matching technique, the matches are less precise than what can be achieved in a (monocular) side by side comparison. We found an average error (for our particular set of eleven test colors) of $\bar{\Delta}xy = 0.008$, as computed with

$$\bar{\Delta}xy = \frac{1}{11} \sum_{i=1}^{11} (\Delta x_i^2 + \Delta y_i^2)^{1/2}. \quad (4)$$

This is about a factor of 6 less precise than the average lower limit defined by the MacAdam ellipses in this region of CIE color space (Wyszecki & Stiles, 1982). However, considering the size of the effects we measured, this precision is more than sufficient.

The actual color constancy results were obtained in three different experiments. We varied the color of the test illuminant only (Expt 1), the color of both the test and match illuminant (Expt 2), and, as an extension of Expt 2, the luminance of the match illuminant (Expt 3). The data will first be presented in terms of CIE chromaticity coordinates, so as to enable comparison with results from other studies and/or analyses by other investigators. In our further analysis, to be presented after the Results section, we shall first transform the CIE units to units that are more directly related to receptor stimulation.

Experiment 1: varying the color of the test illuminant

In this type of experiment, the six illuminants G(reen), Y(ellow), R(ed), M(agenta), B(lue) and C(yan), listed in Table 2, were used for illuminating the (colored) test pattern, whereas the white standard illuminant (W) was used to illuminate the neutral match pattern, i.e. the pattern consisting of only gray samples. We will denote the test/match illuminant combinations by writing G/W, Y/W, etc. The results of these six experimental conditions are shown in Fig. 6. Open squares represent the chromaticities of the color samples under the (colored) test illuminant, open circles those under the (white)

match illuminant. The solid circles indicate the chromaticities of the samples that the observer matched to the test samples. The chromaticities of the match samples, indicated by the hatched area, show a general tendency to resemble those of the colors under white, rather than under colored illumination. Thus, the physical color shift brought about by the change from white to colored illumination, is counteracted, in varying degrees, by the system's mechanism mediating color constancy. If perfect color constancy had been achieved, the match and test pattern would not have appeared as different to the observer, thus obviating the need for adjusting the color of the match samples. Perfect color constancy corresponds to exactly overlapping hatched and open areas in Fig. 6.

The compensatory color shift, along the line joining the chromaticities of the test illuminant and the white point, is mainly effective in procuring constancy of the neutral point, as is most strikingly shown in condition C/W in Fig. 6. In this case, the (cyan) illuminant is located near the boundary of the monitor's phosphor triangle (see Fig. 4). This implies that the contribution of the red phosphor primary is almost completely lost, a loss that the visual system fails to recover, despite its shift of the neutral point. This example illustrates the kind of problems encountered when dealing with certain artificial light sources. Due to incomplete coverage of the spectrum (e.g. sodium lamps) color information may simply get lost in the illuminant-reflectance product in those parts of the spectrum not covered by the artificial illuminant.

When trying to model the data shown in Fig. 6, it soon became clear that we needed a more extended stimulus range (in terms of interocular difference in illumination) if we were to adequately quantify the nonlinearities involved. That called for a more rigorous illuminant change than could be achieved by just comparing colored vs white light.

Experiment 2: varying the color of both the test and match illuminant

In this, and following experiments, the neutral samples of the match pattern were replaced by the same (colored) samples as used for the test pattern. Actually, there was no real need for doing so, since the matches were not systematically different. This is possibly due to the fact that the spatially averaged chromaticity of the stimulus pattern was not affected by the change. Still, we felt that the experiment was "cleaner" by manipulating just one variable, i.e. the color of the illuminant.

In order to sample an adequate range over which color constancy could be measured were selected ten different illuminant pairs. Representative results, as obtained for the illuminant pairs B/G, C/Y, and Y/B, are shown in Fig. 7. For each of these three pairs, we also plotted the data that resulted from interchanging the test and match illuminants. The various symbols have the same meaning as in Fig. 6, except that the open circles now represent chromaticities of match samples as seen under colored, instead of under white light.

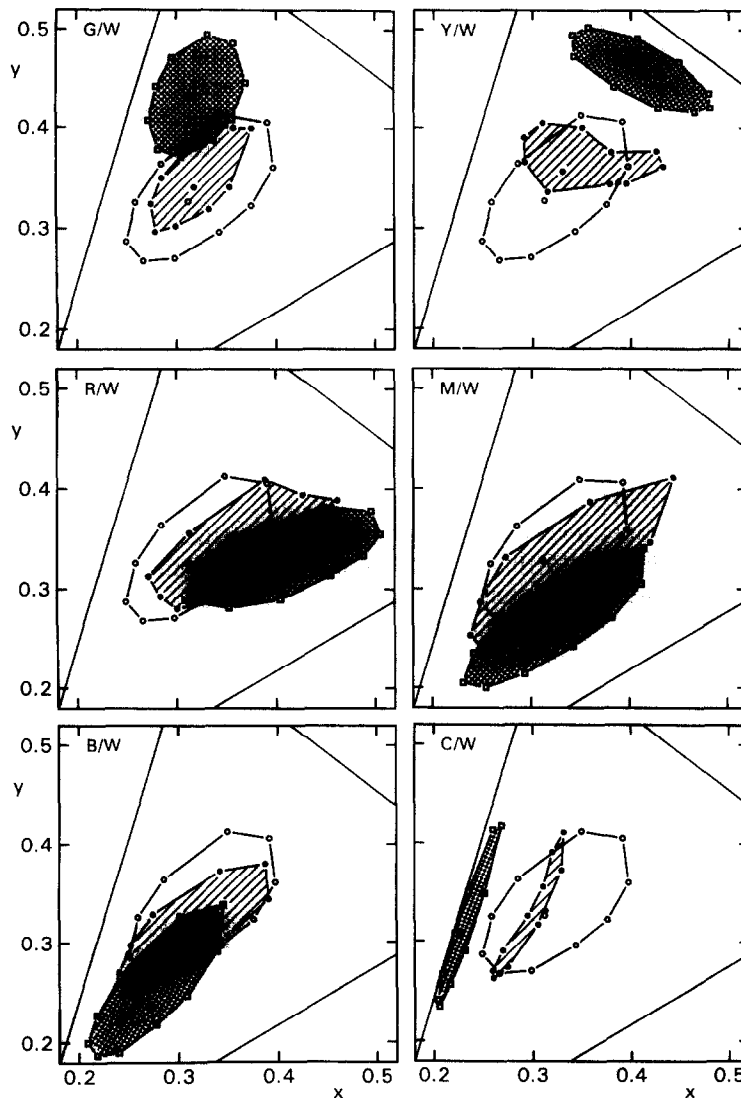


FIGURE 6. Data from Expt 1. The different plots relate to different combinations of test and match illumination. The six colored illuminants, listed in Table 2, were used for illuminating the test pattern, whereas the standard white illuminant (W) was used for illuminating the match pattern, thus producing the illuminant combinations G/W, Y/W, R/W, M/W, B/W and C/W. Each plot contains three sets of chromaticities. Open squares and circles indicate the chromaticities of the color samples under the test and match illuminants, respectively. Solid circles indicate the chromaticities of samples (under white light) that are matched by the observer to the corresponding samples seen under colored light. For clarity, area fill has been used to discriminate between test stimuli under colored light (dotted area) and their matches as made under white light (hatched area). Perfect color constancy would be indicated by superposition of hatched and open area.

Judged by the criterion that perfect color constancy required that matched chromaticities coincide with chromaticities of test samples under the match illuminant (superposed open and hatched areas), the results shown in Fig. 7 would seem to indicate that color constancy is much less effective in the conditions of Expt 2. However, the general pattern is actually not different from that observed before. What happens is that the matches now no longer represent the color of the samples as seen under white light, but those that are perceived under the colored light of the match condition. For example, in the B/G (test/match) condition, the achromatic test sample has the chromaticity of the blue illuminant, but will nevertheless be perceived as approximately white. Hence, it is matched by a sample that is rather greenish (see Fig. 7, condition B/G), simply because that is the sample that will appear as approximately white under

the green match light. Perfect color constancy would still be indicated, as in Fig. 6, by matches that coincide with the sample chromaticities under the match illuminant. The increase in mismatch in these two-color combinations is due to the combined effect of having incomplete color constancy under *both* the test and match illuminant.

Experiment 3: the effect of luminance

The illuminants used in the two experiments described above varied in chromaticity but were fixed in luminance to produce a grid luminance of 12 cd/m². In Expt 3 we repeated the B/Y, C/R and R/C experiments, but now with halved (6 cd/m²) and doubled (24 cd/m²) luminance of the match illuminant. The results showed—not surprisingly—that there was only a small effect of luminance on the chromaticity matches. This is illustrated in

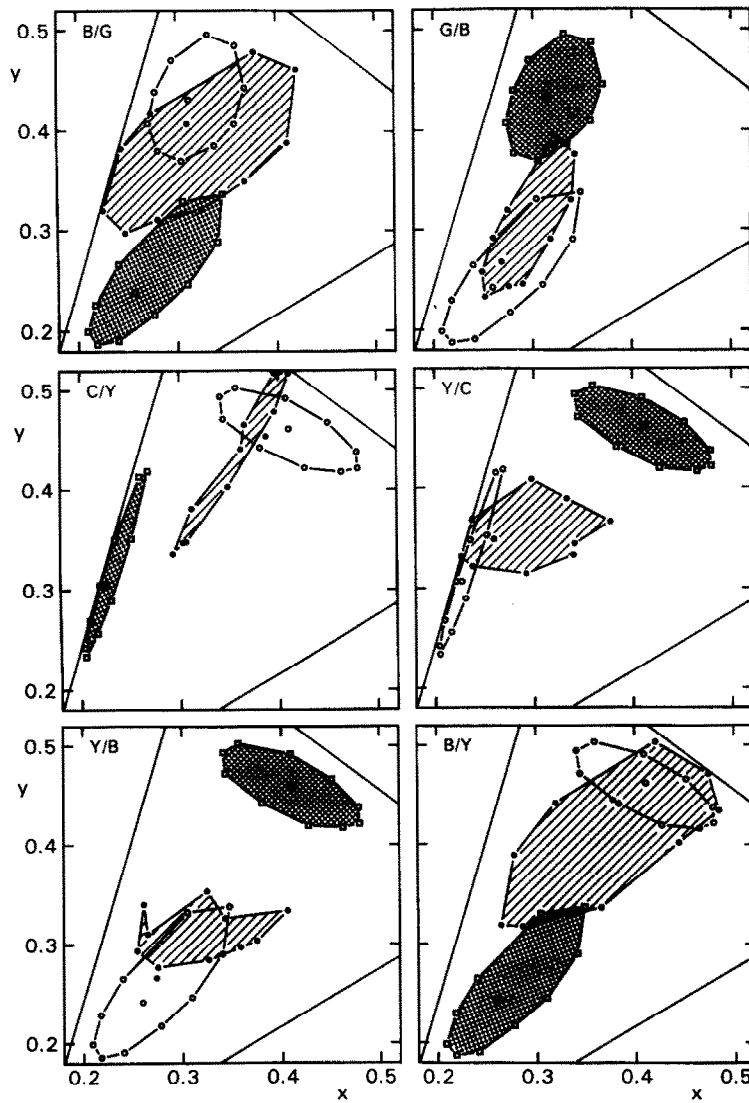


FIGURE 7. Data from Expt 2. The symbols and area fillers have the same meaning as in Fig. 6, except that the open circles now represent the chromaticities of matches made under *colored* instead of under white illumination. The left graphs show the results for the three illuminant combinations B/G, C/Y and Y/B, the graphs on the right relate to the conditions in which the role of "test" and "match" for these illuminant pairs was interchanged (G/B, Y/C, B/Y). Note that this results in quite different matching chromaticities. Perfect color constancy would be indicated again by coinciding hatched and open areas.

Fig. 8, which shows the results for the B/Y condition. The results obtained in the other conditions were similar. The data plotted in Fig. 8 only relate to the purely chromatic aspect of the color matches. They do not show

a possible effect of luminance. The luminance settings of the matches did indeed vary with the overall luminance level, but in such a way that luminance contrast (i.e. the ratio of sample luminance and grid luminance) was

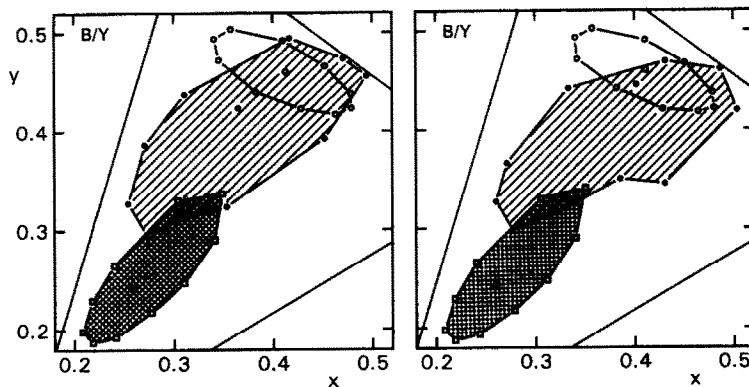


FIGURE 8. Example of data from Expt 3. Symbols and area fillers as in Figs 6 and 7. The experiment is identical to the B/Y condition in Expt 2 (see also Fig. 7), but now with the brightness of the match illuminant halved (left plot) or doubled (right plot). Note that these (overall) luminance changes hardly affect the chromaticities of the matching samples.

approximately maintained. However, there was a small but systematic deviation from the constant contrast prediction. This is illustrated in Fig. 9, which shows the pooled results from the B/Y, C/R and R/C conditions.

Figure 9 plots the luminance of the matching samples at three levels of illumination, producing grid luminances of 6, 12 and 24 cd/m^2 , that match test samples under constant illumination (grid luminance of 12 cd/m^2). The dashed lines indicate the luminance that would be required for an exact contrast match of the test samples. Note that only in the condition where test and match pattern have the same grid luminance (12 cd/m^2), the test and match contrasts are the same (open circles). When the illumination of the match pattern is increased, the samples have to be reduced in contrast, and vice versa. This means that an increase in illumination, while keeping contrast fixed, is nevertheless accompanied by a slight increase in brightness.

The results of the three experiments can be qualitatively summarized as follows.

- Color constancy is manifested in a shift of the neutral point in the direction of the color of the match illuminant. If the latter is white, the shift is in that direction, thus producing the typical color constancy effect.
- When using a colored match illuminant, it is the color of that light that now sets the neutral point, causing a corresponding color shift of the matches.
- The shift in neutral point is hardly affected by a change in luminance, but brightness can only be maintained as long as contrast is more or less maintained.

This summarizes the general trend of the data, but there is much left to be explained. It is still unclear what determines the varying degrees of color constancy observed from one condition to the other. What is needed is an analysis based on a more relevant stimulus representation. That is, a quantification of the stimulus in terms of units related to L-, M- and S-receptor inputs.

DATA ANALYSIS

In this section, we take a closer look at the data, but now expressed in a unit that may be assumed to provide a measure for the activation of the L-, M- and S-receptors. Boynton and Whitten (1972) already computed such a quantity (which they called "effective troland") by distributing troland values over receptor classes, in proportion to their relative sensitivities to the stimulus. The same approach was found to be useful for

*The bulk of our data was obtained with constant overall luminance (grid luminance 12 cd/m^2). In the presentations with doubled and halved grid luminance (Expt 3), direct observation of the pupil size, using a simple laboratory device, indicated a maximal variation in the order of 15–20%. Considering that the effect of overall illumination changes is already so small (as shown in Expt 3), it is unlikely that correction for pupil size would have any effect on the outcome of the data analysis.

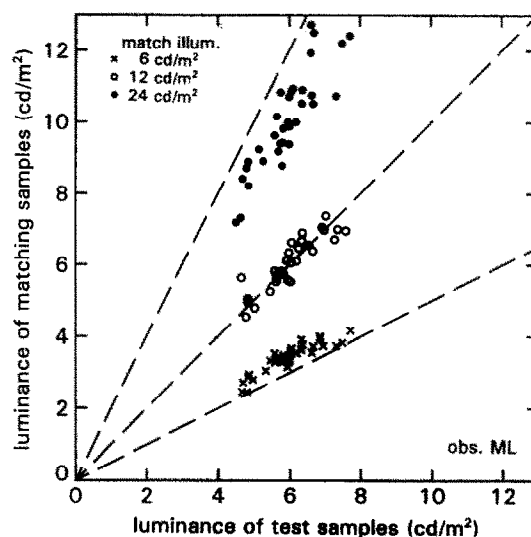


FIGURE 9. Comparison of the luminances of test and matching samples for illuminant combinations B/Y, C/R and R/C. The different symbols correspond to three levels of illumination. The test illuminant always produced a grid luminance of 12 cd/m^2 , whereas the match illuminant produced grid luminances of 6 cd/m^2 (crosses), 12 cd/m^2 (open circles) or 24 cd/m^2 (solid circles). The dashed lines show the predicted result if matches are based on the matching of contrast rather than luminance. The results show that the contrast of the matched sample is relatively low compared to that of the test sample when illuminated by more light than the test sample, and relatively high when illuminated by less light than the test sample.

analyzing gain mechanisms in chromatic adaptation (Walraven, 1981; Werner & Walraven, 1982). The troland value of a stimulus is, for a given pupil size, proportional to its luminance. Therefore, assuming pupil size to be of no consequence in our study* we shall use luminance as the basis for a quantity to be referred to as "receptor input" (cf. Appendix D). It will be denoted by the symbol Q , and has the dimension of cd/m^2 per receptor (L , M , S).

In order to compute receptor inputs, the x , y , Y specifications of the test and match stimuli have to be transformed into relative L , M , S units. This transformation, given in Appendix C, is based upon the Vos-Walraven cone spectral sensitivities (Vos & Walraven, 1971), as tabulated by Vos (1978), but normalized to yield equal quantum catches for the L-, M- and S-receptor at equal-energy white (Walraven & Werner, 1991). As an example, the resulting L , M , S values for the various illuminants (listed in Table 2) are given in Table 3. Note from Table 3 that, by changing from white to a colored illuminant, the change in L and M values is small compared to the change in S value, the latter ranging from about 1 to 8 cd/m^2 . This is due to the large overlap of the spectral sensitivities of the L- and M-cones, and the fact that the S-cone input can be quite independent of luminance (the luminance channel is virtually blind to the short-wave input, e.g. Eisner & MacLeod, 1980). Thus, if two colors are equi-luminant, their S values may differ considerably, whereas their L and M values will be highly correlated. This could mean that the S component is an important mediator for signaling changes in color, and hence, a critical factor for testing models on color constancy. As will be discussed

TABLE 3. Cone input values (in cd/m² per receptor) of the test illuminants from Table 2 as produced by transforming their x, y, Y values according to equations (C1)–(C3)

Test illuminant	Selective cone input (cd/m ²)		
	L	M	S
W(white)	3.92	4.09	4.16
R(red)	4.27	3.41	2.95
G(green)	3.19	4.33	2.27
B(blue)	3.82	4.20	7.85
Y(yellow)	4.04	3.91	1.09
M(magenta)	4.02	3.82	6.42
C(cyan)	3.63	4.63	5.76

below, this is exactly what we found when trying to analyze our data in the context of the Retinex model.

Comparison with Retinex/von Kries theory

From previous experience (Walraven *et al.*, 1991), we expected that receptor-specific contrast may be at the root of color “constancy” (as measured in our experimental paradigm). We therefore decided to test to what extent the Retinex theory, which may be reduced to a contrast model, could be used to describe our data. Actually, the Retinex algorithm does not compute the local contrast ratio between a sample and its surround, but rather the contrast of a unit element (j) relative to a spatially averaged mean input (per receptor class). Of the various Retinex algorithms developed in the course of time, we used the version in which each color sample j is represented by a point in a color three-space, with coordinates D_j^L, D_j^M, D_j^S , the so called *designators*. These can be calculated from

$$D_j^p = \log\left(\frac{Q_j^p}{\bar{G}^p}\right) \quad p = L, M, S \quad (5)$$

where the superscript p denotes receptor class, Q_j^p is a measure for the cone input of sample j and \bar{G}^p is the *geometric* mean of cone input values in the visual scene (cf. Brainard & Wandell, 1986). Dividing Q_j^p by the factor \bar{G}^p is the principle of the von Kries coefficient law, \bar{G}^p representing the coefficient, which acts to scale the receptor input. A so-called (complete) von Kries transformation—a practice used in illumination engineering (to assess the effect of chromatic adaptation)—implies coefficients that produce color constancy for a standard white surface. This is equivalent to computing \bar{G}^p for such a neutral standard.

Since \bar{G}^p and Q_j^p are both expressed in cd/m² per receptor, the ratio Q_j^p/\bar{G}^p is a measure for cone-specific reflectance. For the 100% reflecting standard white in our RGB world, \bar{G}^p corresponds to Q_w^p , and hence, reflectance is defined here as Q_j^p/Q_w^p . Note, that this luminance ratio also defines the contrast between a sample and the (white) grid surrounding it. So, in our experiments cone-specific contrast and reflectance are actually the same.

Assuming that, in our stimulus pattern (Fig. 2), a single patch (1.3° square) may be taken as representing one unit of area, the grid consists of $(15 \times 11) - 35 = 130$

units, whereas the color samples occupy 35 units (out of a total of 165). Thus, the value of \bar{G}^p (for a given cone class) can be computed according to

$$\bar{G}^p = \left((Q_w^p)^{130} \prod_{j=1}^{35} Q_j^p \right)^{1/165} \quad (6)$$

where Q_w^p is the cone input of a unit that belongs to the grid, and subscript w refers to “white reflector”. Since we modeled the grid as a 100% reflector (which acts like a mirror facing a homogeneous illuminant), Q_w^p actually represents the cone input value from the illuminant. For white light and a nominal sample reflectance of 50%, the average cone inputs become: $\bar{G}^L = 0.860Q_w^L$, $\bar{G}^M = 0.857Q_w^M$ and $\bar{G}^S = 0.849Q_w^S$. In general, not considering self-luminous surfaces, the spatially averaged (cone-specific) mean of a scene can always be expressed as a fraction α^p of the cone input values of the light source that illuminates the scene. (However, the fraction may be different for different illuminants.) The cone input value of each individual surface element (j), which we find in the numerator of the designator [equation (5)], may also be expressed as a fraction β_j^p (reflection coefficient) of the illuminant value Q_w^p . Therefore, the designator reduces to

$$D_j^p = \log\left(\frac{Q_j^p}{\bar{G}^p}\right) = \log\left(\frac{\beta_j^p Q_w^p}{\alpha^p Q_w^p}\right) = \log\left(\beta_j^p \alpha^p\right) \quad (7)$$

which is independent of the overall illumination. Note that for $\alpha^p = 1$, the case for a 100% reflectance white, the designator represents receptor-specific reflectance.

In Fig. 10, the experimental results from condition B/Y (see Fig. 7 also) are plotted together with the predicted chromaticities that follow from applying the Retinex model to the stimulus pattern (cf. Appendix E). It is clear from Fig. 10 that the Retinex model (crosses) does not predict the data (open circles) in this particular experimental condition. This is also the case, although usually less pronounced, for most of the other experimental conditions. Apparently, encoding color by taking the logarithm of a receptor-specific contrast ratio is not enough to account for the data. Walraven *et al.*

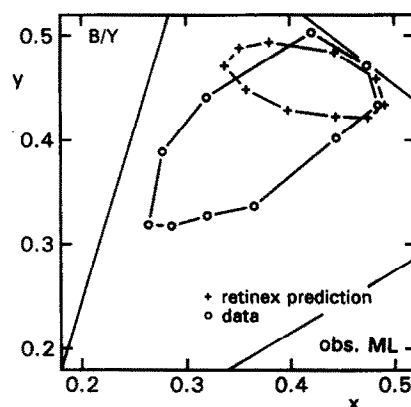


FIGURE 10. Example of experimental data (open circles) and their prediction on the basis of the Retinex model (crosses). The data represent the chromaticities of samples, under blue illumination that match the colors of the corresponding samples under yellow illumination (condition B/Y).

(1991), who tested a (local) contrast explanation of color constancy, concluded that at least one other factor has to be introduced to describe the data. The results of their study indicate that the additional factor should be traced to the short-wave system. However, due to the smaller (differential) illumination range employed in that study, the data were too limited for further investigation of the missing factor. The present data, covering a much larger stimulus range, are better suited for that purpose.

The second factor

The need for a second (S-cone) factor can be best illustrated by plotting the data from a large set of experimental conditions (36 illuminant combinations) in a *contrast ratio diagram* for the S-receptor. This is shown in Fig. 11 (left panel). Along the (logarithmic) axes are plotted receptor-specific contrast ratios C^t and C^m defined by

$$C^t = \frac{Q_j^t}{Q_w^t} \quad \text{and} \quad C^m = \frac{Q_j^m}{Q_w^m} \quad (8)$$

where Q_j is the S-cone input value of sample j , Q_w is the S-cone input value of the surrounding stimulus grid, and superscripts t and m refer to test and match pattern, respectively. Since the grid is physically identical to the illuminant, these local contrast ratios also represent sample reflectances (in cone space). Figure 11 illustrates that the observed (approximate) color constancy does not merely exhibit contrast constancy, at least for the S-cones. Perfect contrast constancy would result in data points located on the dashed line. By analyzing the data from each individual experimental condition (test/match illuminant pair) separately, we found that the S-cone behavior could be parameterized. This is

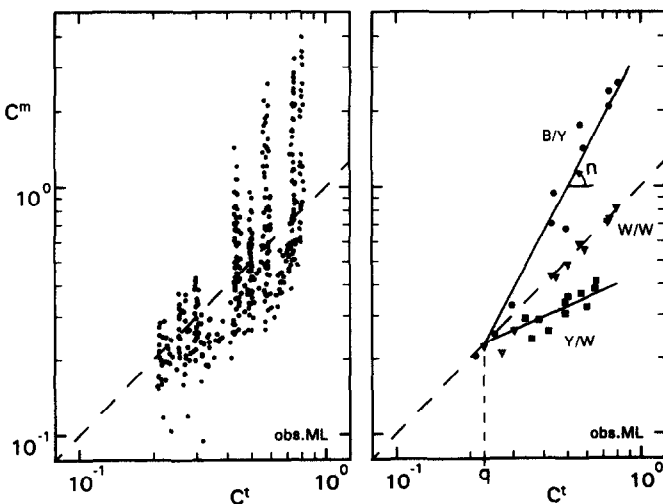


FIGURE 11. Comparison of receptor-specific contrast (S-cone system), for test and match samples (C^t and C^m). Contrast is defined as $C = Q_j/Q_w$. Left panel: pooled data from 36 experiments. Right panel: replotted data for experimental condition, B/Y (circles), W/W (triangles) and Y/W (squares). These data sets show that the variance observed in the pooled data is not due to noise, but can be attributed to an interaction between test/match contrast ratio (C^t/C^m) and illuminant ratio (Q_w^t/Q_w^m). See text for further explanation.

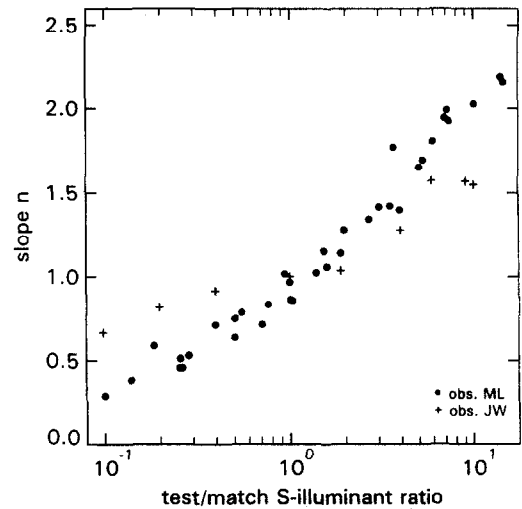


FIGURE 12. The coefficient n , i.e. the slope of the lines used for fitting the S-cone data shown in Fig. 11, plotted as a function of the interocular illuminant ratio Q_w^t/Q_w^m .

illustrated in the right panel of Fig. 11 for conditions B/Y, W/W and Y/W (circles, triangles and squares, respectively), where the data points are fitted with a straight line (slope n), intersecting the dashed identity line at coordinates $x = y = \log(q)$. Hence, for each condition the data can be described by

$$\log(C^m) - \log(q) = n(\log(C^t) - \log(q)) \quad (9)$$

which reduces to

$$\log(kC^m) = n \log(kC^t) \quad k = q^{-1}. \quad (10)$$

The value of q varied only slightly and not systematically between conditions. We obtained a mean value $\bar{q} = 0.23$, hence $\bar{k} = 4.35$.

The next step is to find an expression for the remaining unknown in equation (10), the coefficient n which represents the slope of the functions shown in Fig. 11 (right panel). We found that n correlated, to a fair approximation, with the interocular (S-cone) ratio Q_w^t/Q_w^m . Since Q_w is a measure for the cone input as produced by the grid (=illuminant), the Q_w^t/Q_w^m ratio relates to the differential stimulation of the S-cones, as produced by the test and match illuminant, respectively.

In order to describe our data with a fixed value for k , i.e. its mean value $\bar{k} = 4.35$, we computed for each condition the best fitting power function passing through the (average) point of intersection, as defined by k . We thus obtained for each condition a different value of n . These values, plotted as a function of Q_w^t/Q_w^m , are shown in Fig. 12. The two data sets, shown in Fig. 12, can both be approximated by the power function

$$n = (Q_w^t/Q_w^m)^r \quad (11)$$

with $0 < r < 1$. The best fits were obtained for $r = 0.33$ (observer ML) and $r = 0.20$ (observer JW). One could speculate that a difference in macular pigmentation (JW is 25 yr older than ML), and hence a difference in effective S-cone input, might at least be partly responsible for this difference in the value of r .

Returning to equation (10), we can now substitute $\bar{k} = 4.35$ and specify n according to equation (11). One thus obtains

$$\log(4.35C^m) = (Q_w^t/Q_w^m)^n \log(4.35C^t) \quad (12)$$

which can be rewritten as

$$(Q_w^m)^n \log(4.35C^m) = (Q_w^t)^n \log(4.35C^t). \quad (13)$$

The symmetry of equation (12) lends itself to a model in which the color signal depends on the two factors appearing at both sides of this equation. That is, a cone system response (R) that can be described (over the limited contrast range tested here) as

$$R^p \equiv (Q_w^p)^n \log(4.35C^p). \quad (14)$$

The factor C^p , the receptor-specific sample contrast (here also reflectance), is also incorporated in the Retinex model [see the factor β^p in equation (7)]. The

factor Q_w^p , the receptor input produced by the illuminant, is the second factor we were looking for. It is instructive to see what the improvement is when using equation (14) rather than equation (5) as basis for describing our data.

Data predictions

The expression given in equation (14) was derived on the basis of data relating to the S-cone input. Although equation (14) can also be applied to the L- and M-cone data, these do not provide a critical test for its validity. The reason for that is the large overlap of the L and M spectral sensitivities. This causes (Q_w^m) and (Q_w^l) to be fairly similar, even for the most extreme illuminant conditions in our experiment, and thus renders this variable to a factor that more or less cancels out in equation (12).

The data predictions are in terms of the quantity of Q_j^m , the cone input that is required for the match

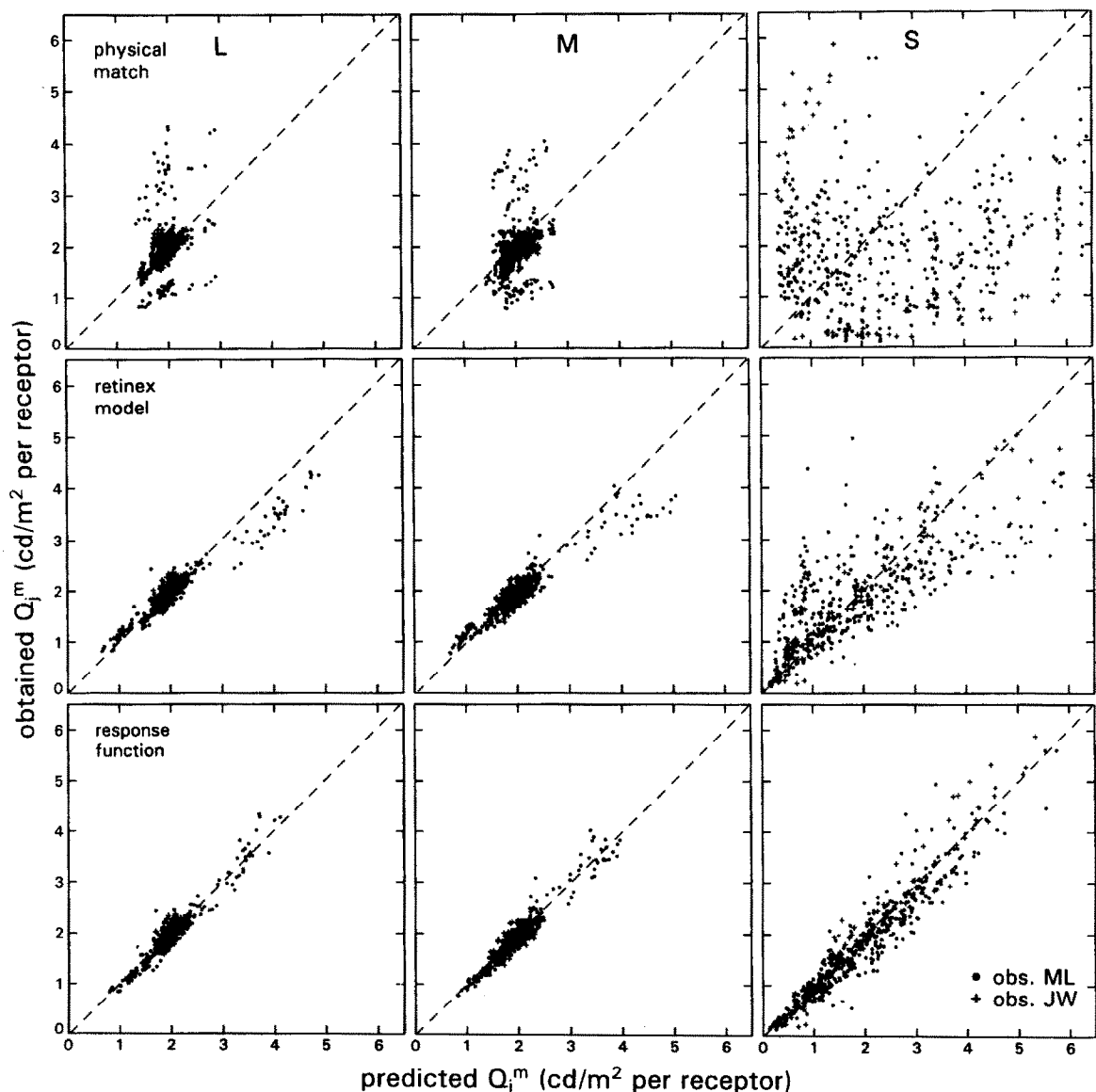


FIGURE 13. Predicted vs. obtained data (495 in all), in terms of Q_j^m , the quantity that provides a measure for the receptor inputs as required for the match sample (j), under the match illuminant (m). The panels are arranged in three rows, representing the predictions (for L-, M- and S-cones, respectively), as obtained with three different models: the physical match, the Retinex model and the response function described by equation (14).

sample j . The latter can be solved from equation (13) by defining test and match contrast ($C^{p,m}$ and $C^{p,t}$) according to their definition given in equation (8). One thus obtains

$$\log(Q_j^{p,m}) = \left(\frac{Q_w^{p,t}}{Q_w^{p,m}}\right)^r \log\left(4.35 \frac{Q_j^{p,t}}{Q_w^{p,t}}\right) + \log\left(\frac{Q_w^{p,m}}{4.35}\right) \quad (15)$$

with r taking the value 0.33 (for ML) or 0.20 (for JW). In Fig. 13, the comparison is shown between the predictions of the value of $Q_j^{p,m}$ on the basis of our description, as given in equation (15), and the predictions according to the Retinex theory. The latter are obtained by equating the designators of the match sample to those of the corresponding test sample, i.e.

$$D_j^{p,m} = D_j^{p,t} \quad p = L, M, S. \quad (16)$$

After expressing D_j^p according to equation (5) one can derive

$$Q_j^{p,m} = Q_j^{p,t} (\bar{G}^{p,m} / \bar{G}^{p,t}), \quad (17)$$

from which the predicted $Q_j^{p,m}$ can be obtained (for given $Q_j^{p,t}$) after computing $\bar{G}^{p,m}$ and $\bar{G}^{p,t}$ with equation (6).

Figure 13 is composed of nine graphs, each of which plots the predicted value of $Q_j^{p,m}$ on the abscissa and the actually obtained value of $Q_j^{p,m}$ on the ordinate. Each horizontal triplet of panels (top, middle and bottom row) relates to a different prediction, separately specified for each receptor waveband L, M and S (panels left to right). The plotted points represent the complete set of data (495 in all), as obtained in 45 experimental conditions, in which the 11 different test samples were presented (and matched) under widely varying illuminants. Note that perfect model predictions would be indicated by data points appearing on the dashed line. The graphs in the top panel of Fig. 13 relate to the naive physical model in which the predicted values of the match are those that reproduce the test sample ($Q_j^{p,m} = Q_j^{p,t}$). So, this "model" denies any stimulus transformation based on more than just the *local* receptor input. As expected, the predictions thus obtained are at variance with the data. It is of interest though, that the deviations are most pronounced for the S-cone data.

The graphs in the middle panel show the predicted values of $Q_j^{p,m}$ on the basis of Retinex theory, as computed with equation (17). This clearly results in quite an improvement compared to the simple model predictions displayed in the top panel. However, the Retinex algorithm still exhibits some serious shortcomings in describing the S-receptor's behavior, and

there are also small deviations from the highest and the lowest predicted values for the L- and M-receptor.

As shown in the previous section, the main difference* in color coding between the Retinex algorithm and the derived response function R^p , is the factor (Q_w^p) . The improvements upon the Retinex algorithm that can be achieved by including this factor, are shown in the bottom panel. For these graphs, we obtained correlation coefficients (ρ) of 0.960, 0.978 and 0.977 for L, M and S, respectively. This implies that equation (14) explains 92.2, 95.6 and 95.4% of the associated data variance (ρ^2). The improvement in data prediction when applying equation (15) rather than the Retinex prediction equation (17) can even be better appreciated in x, y chromaticity space (see Appendices for the computations in question). This is illustrated in Fig. 14, which reproduces Fig. 10, but now with the added prediction based on equation (15). The result shown in Fig. 14 relate to an experimental condition that causes a rather extreme illuminant change within the S-waveband. As can be seen in Fig. 13, failures in the Retinex predictions are mainly confined to the short-wave system. Actually, this can also be observed in the data of the quantitative retinex studies, notably those of the study of McCann and Houston (1983). This is a consequence of the fact that the second factor, the factor not incorporated in the Retinex model, mainly comes to the fore in the S-cone response. This is to be expected, as will be explained in the Discussion, when considering the relatively eccentric position of the S-cone's spectral window.

DISCUSSION

The results obtained in this study show that the human visual system does not achieve perfect color constancy. This is a common finding (e.g. Valberg & Lange-Malecki, 1990; Walraven *et al.*, 1991; Tiplitz Blackwell & Buchsbaum, 1988a), even in conditions in which performance is boosted by allowing the subject to identify rather than match the test sample (Arend & Reeves, 1986). The extent to which color constancy fails is usually expressed in terms of a difference in CIE

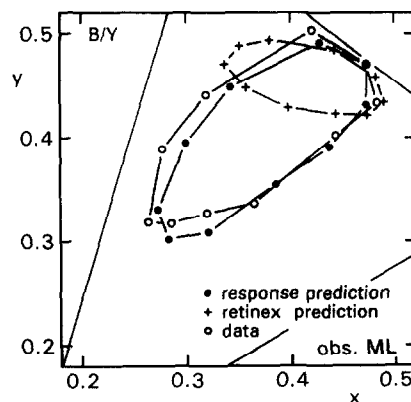


FIGURE 14. Experimental data as obtained for the B(lue)/Y(ellow) condition (open circles), compared to the predictions on the basis of the Retinex algorithm (crosses) and the response function (solid circles) described by equation (14).

*Another difference between the Retinex algorithm and the response function is that the former depends on the geometric mean. However, the geometric mean can be expressed as a fraction α of the cone input of the illuminant, as shown in equation (7): $\bar{G}^p = \alpha^p Q_w^p$. Since the contribution of the background in the computation of the geometric mean is large compared to that of the samples, the value α^p will not be very different for two different illuminants. This means that, when substituted in the Retinex prediction equation (17), α^p drops out of the equation. So, the prediction with the retinex algorithm reduces to a prediction on the basis of the contrast $C^p = Q_j^p / Q_w^p$.

chromaticity coordinates (e.g. Arend, Reeves, Schirillo & Goldstein, 1991). However, this is only the first step in quantifying color constancy. What still has to be explained is what *causes* the observed departures from color constancy.

As is shown by the x, y chromaticity plots of our data (Figs 6–8), color “constancy” varies quite a bit, and not very systematically, from one illuminant condition to the other. However, when analyzed in terms of receptor-specific contrast there seems to be a fairly simple mechanism underlying this apparent complexity: a non-linear response function [equation (14)], in which both (sample) contrast and illumination level are the input variables. The inclusion of the factor illumination in this function* implies that the color (contrast) signal is not independent of luminance. While, this may seem an undesirable property for a system that is supposed to strive for color constancy, as has been pointed out by Jameson and Hurvich (1989), the visual system is likely to be designed to convey information about invariant (reflectance) as well as varying (illumination) aspects of the visual stimulus. It is common experience, of course, that we can sense the level of illumination; we do so because of the absence of *complete* constancy. This is also confirmed by studies that show near perfect constancy of lightness—the achromatic manifestation of color constancy—but, at the same time, the lack of brightness constancy (e.g. Arend & Goldstein, 1987; Jacobsen & Gilchrist, 1988).

Although equation (14) shows the *single channel* color response to depend on illumination level (Q_w^p), an overall change in light level may nevertheless result in a *triple channel* color response, that signals a fairly constant chromaticity. This can be seen in Fig. 8, which shows that the observer’s matches (in terms of x, y chromaticity coordinates), are hardly affected by a change in illumination level. To understand this (common) finding, one only has to assume that the visual system assigns color on the basis of the *ratio* of the cone channel outputs. It can be easily shown that equation (14) predicts that, for a given illuminant/sample condition, the channel output ratio will not be affected by illumination level. This invariance principle will be jeopardized, of course, when, somewhere in the neural pathway, response saturation sets in. If, for example, the cones are driven to their upper limit (as can be achieved with flashed lights), a fixed (maximum) response ratio will result, presumably the ratio that generates white, whatever the color of the stimulus in question (Walraven & Werner, 1991).

The Retinex model predicts no effect of illumination level on perceived color. However, it does not account for the fact that we can nevertheless perceive changes in illumination level. Despite this weakness (which is common to most computational models), the Retinex algorithm has turned out to be of great value for the

development of computational approaches to color constancy (Hurlbert, 1986) or as a “sparring partner” for more sophisticated approaches. When pitted against one of the most recently developed constancy algorithms, Crule (Forsyth, 1990), it was found that the Retinex still outperformed the latter (Forsyth, 1990), provided the average surface color was not chromatically biased by adding large colored borders to the stimulus pattern. For example, a surround biased towards red, would yield, according to the Retinex algorithm, a color shift in the direction of green. Actually, this is the kind of “mistake” the visual system might make as well (chromatic induction). Nevertheless, the present results show that the Retinex can be significantly improved by introducing what we called the “second factor”, which causes the color signal to respond to illumination level.

We noted already that the effect of the second factor is best observed in the S-cone data (see Fig. 13). This can be attributed to the fact that this variable, insofar as it is due to *chromatic* stimulus changes, is subject to stronger variation in the short-wave system than in the other two systems. This can be inferred from the three upper panels in Fig. 13. For these panels the horizontal scale gives the cone-specific inputs of all the sample/illuminant combinations. Note, that the range over which the (spectrally manipulated) input to the L- and M-cones varies, is much smaller than that of the S-cones. This is only partly due to the restrictions imposed by the RGB phosphors; rather, it reflects the large overlap of the L- and M-spectral sensitivities. It is thus impossible, at a fixed level of illumination, to substantially modulate the input to the L- and M-systems. In contradistinction, the S-cones, which are spectrally much more isolated, may be driven into nearly complete darkness, when the illuminant mainly contains wavelengths beyond 540 nm. It is for that reason, that the effect of the second factor mainly shows up in the S-cone response.

The “first factor” of our color response function, receptor-specific contrast, is reminiscent of the lightness operator of the various Retinex models. However, it is important to recognize that the (local) contrast of a stimulus is not invariant since it covaries with the luminance of the adjacent surround. Lightness, on the other hand, may be considered as the perceptual attribute that correlates with reflectance, which is defined as the ratio of the reflected flux to the incident flux. Obviously, a visual system that responds to reflectance, a physical invariant, stands a better chance to register a stable (object) world than a system that responds to contrast. Actually, the present data do not exclude the possibility that reflectance rather than contrast is the relevant variable that has to be entered into equation (14). Our stimulus pattern simulated samples surrounded by a 100% reflecting grid, so the contrast of a sample relative to its surround, here also represents its reflectance (for the incident light in question).

Nevertheless, following Shapley (1986), we hesitate to reject the more simple assumption that the visual system responds to (local) contrast rather than reflectance, at

*By illumination level we mean absolute level of cone stimulation by the grid, which is a 100% reflector to the incident light. It is represented by the factor Q_w^p , the cone-specific input for the grid (background).

least under the conditions of our (and most other) laboratory experiments. Contrast has already been identified as the determinant of achromatic color, or lightness (Wallach, 1948; Shapley, 1986), in particular for coplanar surfaces (Gilchrist, 1977; Schirillo, Reeves & Arend, 1990). The importance of local contrast has actually been acknowledged in the most recent version of the Retinex algorithm (Land, 1986b), by the added feature of "small aperture" sampling of the surface pattern. The need for doing so has been demonstrated in studies showing, contrary to the prediction of the earlier Retinex version, that the perception of a sample in a Mondrian configuration is mainly determined by adjacent samples (Shapley, 1986; Creutzfeldt, Lange-Malecki & Wortmann, 1987; Valberg & Lange-Malecki, 1990; Tiplitz-Blackwell & Buchsbaum, 1988a, b). Many studies have also demonstrated the important role of local changes (edges) in the perception of color and brightness (e.g. Krauskopf, 1963; Arend, 1973; Walraven, 1973, 1977).

Support for the assumption that the visual system responds to contrast rather than reflectance (in a "Mondrian world") has been obtained by Walraven *et al.* (1991). In nearly the same experimental paradigm as in our study, a simulation was performed in which the grid was selectively illuminated with colored light whereas the samples were shown under white light. This configuration resembles the set-up for demonstrating the classical "colored-shadows" phenomenon, the samples being "shadowed" from the colored light source. Since only the grid was subjected to a change in color, the perceived color of the samples should have been truthfully signaled by a system that records reflectance. Alternatively, a system that responds to contrast would not be able to do so, but instead, signal a color change consistent with the altered contrast between sample and surround. This was confirmed by the results, which showed the expected effect of chromatic induction, i.e. a shift of the samples towards a color complementary to the grid. It could be shown that, apart from the effect of the "second factor", the color shift could be fully accounted for by assuming the color of a sample to be determined by its (receptor-specific) sample-surround contrast. This finding, which is corroborated by the results of Tiplitz-Blackwell and Buchsbaum (1988a, b), indicates that chromatic induction (simultaneous contrast) may be interpreted as the visual system's misdirected attempt at color constancy (Walraven, Benzschawel & Rogowitz, 1989). Misdirected, because the colored grid illumination triggered a (contrast) response that "compensated" for colored light that was incident on the grid, but not on the test samples.

The way in which contrast enters into our response function, that is, $\log(4.35C^p)$ (with $C^p = Q_p^p/Q_w^p$), raises questions as to the meaning of the logarithmic transformation and the coefficient 4.35. As for the latter, we are considering (and testing) the possibility that this factor is not constant but may depend on factors that were constant only in our experiments, i.e. a fixed spatial separation between the samples and a fixed nominal

contrast (50% under white light). It is also possible that this factor represents an additive noise term; consider in this respect that $\log(4.35C^p) = \log(C^p) + 0.64$. At this stage, it would be premature to go into any further speculation. One should also keep in mind that we do not know, as yet, what contrast range can be successfully described by the response function, simply because the present study was concerned with the effect of illumination, in which contrast was not the independent variable.

As for the logarithmic transformation in our response function, this would be functional for any visual system that responds to contrast, the key to maintaining invariance of perception in conditions of varying illumination (cf. Walraven, Enroth-Cugell, Hood, MacLeod & Schnapf, 1990). This does not mean that we have to assume a "hard-wired" logarithmic transducer function, as has sometimes been proposed (e.g. Cornsweet & Pinsker, 1965; Kelly, 1969). All that is needed, is a fast proportional gain control (e.g. Koenderink, van de Grind & Bouman, 1971; Ullman & Schechtman, 1982; Hayhoe, Benimoff & Hood, 1987). It can be shown that the output of such a gain control mimics the effect of a logarithmic transducer function (Koenderink *et al.*, 1971). A proportional gain control receiving input described by a Naka-Rushton (Naka & Rushton, 1966) type (receptor) signal, has been found to provide accurate quantitative accounts for various adaptation phenomena (Walraven, 1980; Walraven & Valetton, 1984).

It is quite possible that there are essential factors missing in the response function reported here. This follows from the way in which the function was derived, namely, by assuming that the symmetrical equation described by equation (13), identifies (at either side of the identity sign) a *complete* response function. Clearly, this is not justified, since any transformation or adding of terms that cancel when applied to both sides of equation (13), will go undetected. Keeping this restriction in mind, there can be little doubt that contrast and illumination level, the two variables that account for about 95% of the variance of our data, represent major determinants of the visual system's color response. It is of interest that these are also the two factors, that have been implemented (in essentially the same way as in our response function) in the color appearance model developed by Nayatani and co-workers (e.g. Nayatani, Takahama, Sobagaki & Hashimoto, 1990). That model was derived to account for results obtained by different authors, employing different experimental paradigms. Apparently, the stimulus variables that we found to be essential for describing the present data set are not specific for our particular visual test scenario, but can be identified in other studies as well.

Although the present data can be accounted for by assuming only receptor specific processing, this does not mean that there would be no opponent processing involved as well. It is quite possible that, because of the way we analyzed the data, opponent processes have been "back projected" on receptor processes. Another possibility is that our experimental conditions do not

sufficiently probe processes at the opponent level. Results of experiments that are now in progress suggest that we may have to introduce luminance normalized contrast signals in our analysis. This implies a stage of separate chromaticity coding, disconnected from luminance contrast.

REFERENCES

- Arend, L. E. (1973). Spatial differential integral operations in human vision: Implications of stabilized retinal image fading. *Psychological Review*, 80, 374–395.
- Arend, L. E. & Goldstein, R. (1987). Simultaneous constancy, lightness, and brightness. *Journal of the Optical Society of America A*, 4, 2281–2285.
- Arend, L. E. & Reeves, A. (1986). Simultaneous color constancy. *Journal of the Optical Society of America A*, 3, 1743–1751.
- Arend, L. E., Reeves, A., Schirillo, J. & Goldstein, R. (1991). Simultaneous color constancy: Papers with diverse Munsell values. *Journal of the Optical Society of America A*, 8, 661–672.
- Boynton, R. M. & Whitten, D. N. (1972). Selective chromatic adaptation in primate photoreceptors. *Vision Research*, 12, 855–874.
- Brainard, D. H. & Wandell, B. A. (1986). Analysis of the retinex theory of color vision. *Journal of the Optical Society of America A*, 3, 1651–1661.
- Brill, M. H. & West, G. (1986). Chromatic adaptation and color constancy: A possible dichotomy. *Color Research and Application*, 11, 196–204.
- Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310, 1–26.
- Cornsweet, T. N. & Pinsker, H. M. (1965). Luminance discrimination of brief flashes under various conditions of adaptation. *Journal of Physiology, London*, 176, 294–310.
- Cowan, W. B. (1986). CIE calibration of video monitors (available from National Research Council of Canada, Ottawa, Ontario, K1A 0R6, Canada).
- Creutzfeldt, O., Lange-Malecki, B. & Dreyer, E. (1990). Chromatic induction and brightness contrast: A relativistic color model. *Journal of the Optical Society of America A*, 7, 1644–1653.
- Creutzfeldt, O., Lange-Malecki, B. & Wortmann, K. (1987). Darkness induction, retinex algorithm and cooperative mechanism in vision. *Experimental Brain Research*, 63, 270–283.
- Dannemiller, J. L. (1989). Computational approaches to color constancy: Adaptive and ontogenetic considerations. *Psychological Review*, 96, 255–266.
- D'Zmura, M. (1992). Color constancy: Surface color from changing illumination. *Journal of the Optical Society of America A*, 9, 490–493.
- D'Zmura, M. & Lennie, P. (1986). Mechanisms of color constancy. *Journal of the Optical Society of America A*, 3, 1662–1672.
- Eastman, A. A. & Brecker, S. A. (1972). The subjective measurements of color shifts with and without chromatic adaptation. *Journal of the Illuminating Engineering Society*, 2, 239.
- Eisner, A. & MacLeod, D. I. A. (1980). Blue-sensitive cones do not contribute to luminance. *Journal of the Optical Society of America*, 70, 121–123.
- Estévez, O. (1979). On the fundamental data-base of normal and dichromatic color vision. Ph.D. dissertation, University of Amsterdam, The Netherlands.
- Forsyth, D. A. (1990). A novel algorithm for color constancy. *International Journal of Computational Vision*, 5, 5–36.
- Gershon, R. & Jepson, A. D. (1989). The computation of color constant descriptors in chromatic images. *Color Research and Application*, 14, 325–334.
- Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. *Science*, 195, 185–187.
- Hayhoe, M. M., Benimoff, N. I. & Hood, D. C. (1987). The time-course of multiplicative and subtractive adaptation process. *Vision Research*, 27, 1981–1996.
- Hurlbert, A. (1986). Formal connections between lightness algorithms. *Journal of the Optical Society of America A*, 3, 1684–1693.
- Jacobsen, A. & Gilchrist, A. (1988). The ratio principle holds over a million-to-one range of illumination. *Perception and Psychophysics*, 43, 1–6.
- Jameson, D. & Hurvich, L. M. (1989). Essay concerning color constancy. *Annual Review of Psychology*, 40, 1–22.
- Judd, D. B. (1951). Colorimetry and artificial daylight. Proceedings of the 12th Session CIE, Stockholm, Vol. 1 (TC7) Bureau Central de la CIE, Paris 11.
- Kelly, D. H. (1969). Flickering patterns and lateral inhibition. *Journal of the Optical Society of America*, 59, 1361–1365.
- Koenderink, J. J., van de Grind, W. A. & Bouman, M. A. (1971). Foveal information processing at photopic luminances. *Kybernetik*, 8, 128–144.
- Krauskopf, J. (1963). Effect of retinal image stabilization on the appearance of heterochromatic targets. *Journal of the Optical Society of America A*, 53, 741–744.
- von Kries, J. (1905). Die Gesichtsempfindungen. In Nagel, W. (Ed.), *Handbuch der Physiologie des Menschen* (Vol. 3, pp. 109–282). Braunschweig.
- Land, E. H. (1959). Color vision and the natural image, Part I. *Proceedings of the National Academy of Science U.S.A.*, 45, 115–129.
- Land, E. H. (1986a). Recent advances in retinex theory. *Vision Research*, 26, 7–21.
- Land, E. H. (1986b). An alternative technique for the computation of the designator in the retinex theory of color vision. *Proceedings of the National Academy of Science U.S.A.*, 83, 3078–3080.
- Land, E. H. & McCann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America A*, 61, 1–11.
- Lee, H.-C. (1986). Method for computing the scene-illuminant chromaticity from specular highlights. *Journal of the Optical Society of America A*, 3, 1694–1699.
- Lucassen, M. P. & Walraven, J. (1990). Evaluation of a simple method for color monitor recalibration. *Color Research and Application*, 15, 321–326.
- Maloney, L. T. & Wandell, B. A. (1986). Color constancy: A method for recovering surface spectral reflectance. *Journal of the Optical Society of America A*, 3, 29–33.
- McCann, J. J. & Houston, K. L. (1983). Calculating color sensation from arrays of physical stimuli. *IEEE Transactions on Systems, Man and Cybernetics*, 13, 1000–1007.
- McCann, J. J., McKee, S. P. & Taylor, T. H. (1976). Quantitative studies in retinex theory: A comparison between theoretical predictions and observer responses to the “color Mondrian” experiments. *Vision Research*, 16, 445–458.
- Naka, K. I. & Rushton, W. A. (1966). S-potentials from colour units in the retina of fish. *Journal of Physiology*, 185, 587–599.
- Nayatani, Y., Takahama, K., Sobagaki, H. & Hashimoto, K. (1990). Color appearance model and chromatic-adaptation transform. *Color Research and Application*, 15, 210–221.
- Rubin, J. M. & Richards, W. A. (1982). Color vision and image intensities: When are changes material? *Biological Cybernetics*, 45, 215–226.
- Schirillo, J., Reeves, A. & Arend, L. E. (1990). Perceived lightness, but not brightness, of achromatic surfaces depends on perceived depth information. *Perception and Psychophysics*, 48, 82–90.
- Shapley, R. (1986). The importance of contrast for the activity of single neurons, the VEP and perception. *Vision Research*, 26, 45–61.
- Sproson, W. N. (1983). *Colour science in television and display systems*. Bristol: Hilger.
- Tiplitz Blackwell, K. & Buchsbaum, G. (1988a). Quantitative studies of color constancy. *Journal of the Optical Society of America A*, 5, 1772–1780.
- Tiplitz Blackwell, K. & Buchsbaum, G. (1988b). The effect of spatial and chromatic parameters on chromatic induction. *Color Research and Application*, 13, 166–173.
- Troost, J. M., Wei, L. & de Weert, C. M. M. (1992). Binocular measurements of chromatic adaptation. *Vision Research*, 32, 1987–1997.
- Ullman, S. & Schechtman, G. (1982). Adaptation and grain normalization. *Proceedings of the Royal Society of London B*, 216, 299–313.
- Valberg, A. & Lange-Malecki, B. (1990). Colour constancy in Mondrian patterns: A partial cancellation of physical chromaticity shifts by simultaneous contrast. *Vision Research*, 30, 371–380.

- Vos, J. J. (1978). Colorimetric and photometric properties of a 2° fundamental observer. *Color Research and Applications* 3, 125–128.
- Vos, J. J. & Walraven, P. L. (1971). On the derivation of the foveal receptor primaries. *Vision Research*, 11, 799–818.
- Wallach, H. (1948). Brightness constancy and the nature of achromatic colors. *Journal of Experimental Psychology*, 38, 310–324.
- Walraven, J. (1973). Spatial characteristics of chromatic induction; the segregation of lateral effects from straylight artifacts. *Vision Research*, 11, 1739–1753.
- Walraven, J. (1977). Color signals from incremental and decremental light stimuli. *Vision Research*, 17, 71–76.
- Walraven, J. (1980). The derivation of nerve signals from contrast flash data; a re-analysis. *Biological Cybernetics*, 38, 23–29.
- Walraven, J. (1981). Perceived colour under conditions of chromatic adaptation: Evidence for gain control by π -mechanisms. *Vision Research*, 21, 611–620.
- Walraven, J. & Valetton, J. M. (1984). Visual adaptation and response saturation. In van Doorn, A. J., van de Grind, W. A. & Koenderink, J. J. (Eds), *Limits in perception*. Utrecht: VNU Science Press.
- Walraven, J. & Werner, J. S. (1991). The invariance of unique white; possible implications for normalizing cone action spectra. *Vision Research*, 31, 2185–2193.
- Walraven, J., Benzschawel, T. & Rogowitz, B. E. (1989). Color-constancy interpretation of chromatic induction. *Die Farbe*, 34, 269–273.
- Walraven, J., Benzschawel, T., Rogowitz, B. E. & Lucassen, M. P. (1991). Testing the contrast explanation of color constancy. In Valberg, A. & Lee, B. B. (Eds), *From pigments to perception* (pp. 369–378). New York: Plenum Press.
- Walraven, J., Enroth-Cugell, C., Hood, D. C., MacLeod, D. I. A. & Schnapf, J. L. (1990). The control of visual sensitivity: Receptor and postreceptor processes. In Spillmann, L. & Werner, J. S. (Eds), *Visual perception: The neurophysiological foundations* (pp. 53–101). San Diego, Calif.: Academic Press.
- Werner, J. S. & Walraven, J. (1982). Effect of chromatic adaptation on the chromatic locus: The role of contrast, luminance and background color. *Vision Research*, 22, 929–943.
- Worthey, J. A. (1985). Limitations of color constancy. *Journal of the Optical Society of America A*, 2, 1014–1026.
- Worthey, J. A. & Brill, M. H. (1986). Heuristic analysis of von Kries color constancy. *Journal of the Optical Society of America A*, 3, 1708–1712.
- Wyszecki, G. & Stiles, W. S. (1982). *Color science* (2nd edn). New York: Wiley.
- Young, R. A. (1987). Color vision and the Retinex theory. *Science*, 238, 1731–1732.
- Yuille, A. (1987). A method for computing spectral reflectance. *Biological Cybernetics*, 56, 195–201.

Acknowledgement—This study was supported by the Netherlands Organization for Scientific Research (NWO) through the Foundation for Biophysics.

APPENDIX A

Reflected Light Simulation: a Numerical Example

The following example delineates the steps involved in the colorimetric calculation of light reflection for the purpose of presentation on a CRT. All numerical values that appear throughout this example, except those for chromaticities x and y , were rounded at the second decimal for simplicity. We consider the case of the first sample in Table 1. That is, color 21 (a purple), displayed on a white background, and illuminated by white (RGB metamer of D_{65}) and green (G) light, respectively. The background is assumed to reflect 100%, and thus has the same color and luminance as the illuminant. The reflectance of this sample, like all chromatic samples, was set at 50% (under white light). The following specifications apply under white light (subscripts j and b refer to sample and background, respectively):

sample: $x_j = 0.3243$ $y_j = 0.2630$ $Y_j = 6$ cd/m²
background: $x_b = 0.313$ $y_b = 0.329$ $Y_b = 12$ cd/m²

and under green light:

sample: $x_j = ?$ $y_j = ?$ $Y_j = ?$ cd/m²
background: $x_b = 0.313$ $y_b = 0.432$ $Y_b = 12$ cd/m².

The requested sample values under green light are determined by the following steps:

1. Calculate the phosphor luminances (Y_R, Y_G, Y_B) required for producing the specified x, y, Y values of the sample and background under white light, and the x, y, Y values of the background under green light, respectively. The transformation from x, y, Y into Y_R, Y_G, Y_B (for our CRT) is presented in Appendix B. Denote the resulting luminances by $Y_{R,j}$ etc., with subscript j or b referring to sample or background, respectively. This should result in:

$$\begin{aligned} \text{white light: } Y_{R,j} &= 2.42, & Y_{G,j} &= 2.78, & Y_{B,j} &= 0.80 \\ & Y_{R,b} &= 2.85, & Y_{G,b} &= 8.11, & Y_{B,b} &= 1.04 \\ \text{green light: } Y_{R,b} &= 1.41, & Y_{G,b} &= 10.11, & Y_{B,b} &= 0.49. \end{aligned}$$

The latter luminance values are the ones shown in Table 2 (line 3).

2. Calculate the emission coefficients a_i ($i = R, G, B$), associated with the change in illuminant color, according to equation (2), i.e. $a_R = Y_{R,ill}(\text{green light})/Y_{R,w}$, etc. Since the background is a 100% neutral reflector, the subscript ill (for illuminant) and w (for white) can be replaced by b (for background). One thus obtains $a_R = 1.41/2.85 = 0.49$, $a_G = 10.11/8.11 = 1.25$ and $a_B = 0.49/1.04 = 0.47$ (these values also appear in Table 2, line 3).
3. Calculate the coefficients b_i ($i = R, G, B$) according to equation (3), i.e. $b_R = Y_{R,j}(\text{under white light})/Y_{R,w}$, etc. The phosphor luminances that create the standard white illuminant are $Y_{R,w} = 2.85$, $Y_{G,w} = 8.11$ and $Y_{B,w} = 1.04$. This yields $b_R = 2.42/2.85 = 0.85$, $b_G = 2.78/8.11 = 0.34$ and $b_B = 0.80/1.04 = 0.77$ (see also Table 1).
4. Calculate the luminance of the (simulated) reflected light according to equation (1):

$$\begin{aligned} Y_r &= Y_R + Y_G + Y_B = a_R b_R Y_{R,w} + a_G b_G Y_{G,w} + a_B b_B Y_{B,w} \\ &= (0.49 \times 0.85 \times 2.85) + (1.25 \times 0.34 \times 8.11) \\ &\quad + (0.47 \times 0.77 \times 1.04) \\ &= 1.19 + 3.45 + 0.38 = 5.02 \text{ cd/m}^2. \end{aligned}$$

With equation (B3), the three phosphor luminances $Y_R = 1.19$, $Y_G = 3.45$, $Y_B = 0.38$ of the reflected light are transformed into $X_r = 4.67$, $Y_r = 5.02$, $Z_r = 4.84$, hence $x_r = 0.3214$, $y_r = 0.3455$. So, the requested values of the sample under green light are: $x_j = 0.3214$, $y_j = 0.3455$, $Y_j = 5.02$ cd/m².

APPENDIX B

Transformation from CIE x, y, Y to Monitor Y_R, Y_G, Y_B Luminance and Vice Versa

Given the chromaticity coordinates (x, y) and luminance (Y) of the color to be displayed on a color monitor, the first step is to determine its tristimulus values X, Y, Z . The latter are given by $X = (x/y)Y$, $Y = Y$ and $Z = (z/y)Y$, with $z = 1 - x - y$. The phosphor luminances Y_R, Y_G, Y_B , required for producing these tristimulus values are related by (e.g. Sproson, 1983)

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} x_R/y_R & x_G/y_G & x_B/y_B \\ 1 & 1 & 1 \\ z_R/y_R & z_G/y_G & z_B/y_B \end{pmatrix} \begin{pmatrix} Y_R \\ Y_G \\ Y_B \end{pmatrix} \quad (\text{B1})$$

where x, y and z are the phosphors' chromaticity coordinates, and Y_R, Y_G, Y_B and X, Y, Z are expressed in cd/m². The phosphor chromaticities of our monitor, as measured with a SpectraScan PR-702AM (Photo Research) spectroradiometer, are $(x_R, y_R) = (0.6326, 0.3549)$, $(x_G, y_G) = (0.3065, 0.5984)$ and $(x_B, y_B) = (0.1459, 0.0701)$, hence equation (B1) becomes (matrix inversion)

$$\begin{pmatrix} Y_R \\ Y_G \\ Y_B \end{pmatrix} = \begin{pmatrix} 0.776 & -0.380 & -0.111 \\ -0.785 & 1.399 & 0.021 \\ 0.009 & -0.019 & 0.089 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (\text{B2})$$

The reverse transformation from Y_R, Y_G, Y_B to X, Y, Z , needed to transform the observer's R, G, B settings into x, y, Y units (for our particular set of phosphor chromaticities), is given by

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 1.782 & 0.512 & 2.055 \\ 1 & 1 & 1 \\ 0.035 & 0.159 & 11.184 \end{pmatrix} \begin{pmatrix} Y_R \\ Y_G \\ Y_B \end{pmatrix}. \quad (B3)$$

The values of x, y and Y are computed with $x = X/(X + Y + Z)$, $y = Y/(X + Y + Z)$ and $Y = Y$.

APPENDIX C

Transformation from CIE x, y, Y to Cone L, M, S Units and Vice Versa

The first step in the transformation from CIE to cone space incorporates Judd's (1951) modification of the x, y chromaticities (cf. Vos, 1978), to compensate for imperfections in the original CIE short wavelength region of the luminous efficiency function, $V(\lambda)$. This modification transforms x and y to slightly different chromaticities x' and y' . The transformation is quantified by Vos (1978) as follows

$$x' = \frac{1.0271x - 0.00008y - 0.00009}{0.03845x + 0.01496y + 1} \quad (C1)$$

$$y' = \frac{0.00376x + 1.0072y + 0.00764}{0.03845x + 0.01496y + 1}. \quad (C2)$$

The modified tristimulus values X', Y', Z' are then given by $X' = (x'/y')Y'$, $Y' = Y'$ and $Z' = (z'/y')Y'$.

In order to transform from tristimulus values to L, M, S receptor inputs, we used Vos-Walraven cone spectral sensitivity functions (Vos & Walraven, 1971), as tabulated by Vos (1978). Following Walraven and Werner (1991), we normalized the sensitivities of the receptor systems such that, the L-, M- and S-cones receive equal quantum catches at equal-energy white ($x = y = 0.33$). As a result the following matrix equation is obtained

$$\begin{pmatrix} L \\ M \\ S \end{pmatrix} = \begin{pmatrix} 0.0778 & 0.2722 & -0.0186 \\ -0.1562 & 0.4569 & 0.0297 \\ 0 & 0 & 0.3315 \end{pmatrix} \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix}. \quad (C3)$$

Starting with known L, M, S units, the modified X', Y', Z' tristimulus values are computed with the inverse of equation (C3):

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} 5.8746 & -3.5001 & 0.6424 \\ 1.9948 & 1 & 0.0221 \\ 0 & 0 & 3.0169 \end{pmatrix} \begin{pmatrix} L \\ M \\ S \end{pmatrix}. \quad (C4)$$

The modified chromaticities, as determined by $x' = X'/(X' + Y' + Z')$ and $y' = Y'/(X' + Y' + Z')$, can be transformed to x, y by using the reverse equations of equations (C1) and (C2):

$$x = \frac{1.00709x' + 0.00008y' + 0.00009}{-0.03867x' - 0.01537y' + 1.03450} \quad (C5)$$

$$y = \frac{-0.00347x' + 1.02710y' - 0.00785}{-0.03867x' - 0.01537y' + 1.03450}. \quad (C6)$$

In principle, one also requires a transformation for $Y' \rightarrow Y$. However, as long as stimuli are not located in the (far) blue corner of the CIE chromaticity diagram, one may safely assume $Y = Y'$.

APPENDIX D

A Unit for "Receptor Input" (cd/m^2 Per Receptor)

Estimating how the light entering the eye is (effectively) absorbed in the three classes of cones is still not possible without making a number of assumptions. At best, one can make an educated guess about how much absorbed quanta/second/cone correspond to a photopic or scotopic troland [see Boynton and Whitten (1972) for a discussion]. The troland unit, and hence, the unit of luminance (cd/m^2), may thus provide a measure for quanta incident on the retina, but the effect of the quanta can only be traced to their integrated action, that is, their contribution to the luminance "channel". The latter has an action spectrum, $V(\lambda)$, that can be described as the envelope of the separate L-, M- and S-cone action spectra (after appropriate weighing).

In order to obtain a unit that may provide a measure for the separate L-, M- and S-cone (luminous) inputs, we assume the stimulus energy (as registered in the quantities X, Y and Z) to be distributed over the cones according to the transformations given in equation (C3). The latter imply cone action spectra that are normalized—a still unresolved issue (Walraven & Werner, 1991)—so as to yield equal sensitivity at equal-energy white (for which $X = Y = Z$, and hence, $x = y = 0.33$). As a consequence, the contribution of the cone classes to (Judd-modified) luminance is given, as shown in equation (C4), by $Y' = 1.99L + M + 0.02S$. Even if the equal-energy normalization, which is not an uncommon one (e.g. Judd, 1951; Estévez, 1979), would turn out to be incorrect, this would hardly affect the data. These are analyzed in terms of receptor-specific contrast, a quantity that does not change if a different contribution of the cone classes to luminance would have to be assumed.

Given the above assumptions, the receptor input associated with a particular sample (Q_j), measured in terms of cd/m^2 per receptor (L, M or S), can be computed with quantities (C3). For example, a sample reflecting 12 cd/m^2 white light, with chromaticity coordinates $x = 0.313$ and $y = 0.329$, yields $X'Y'Z'$ quantities, expressed in cd/m^2 , of $X' = 11.3$, $Y' = 12.0$ and $Z' = 12.5$. Using these values as input to equation (C3) produces receptor inputs (Q_j) of $L = 3.92$, $M = 4.09$ and $S = 4.15 \text{ cd}/\text{m}^2$.

APPENDIX E

Data Predictions in Terms of CIE x, y Units

Data on chromatic adaptation or color constancy are usually plotted in the 1931 CIE xy chromaticity diagram. Although this may not provide the best metric from an analytical point of view, it is quite useful for purposes of color specification. We did so for both obtained data (Figs 6-8) and an example of predicted data (Fig. 14). The obtained data are available in terms of Y_R, Y_G, Y_B luminances of the match samples (as set by the observer), and thus can be readily transformed to CIE units by employing equation (B3).

As for the predicted data, either on the basis of the Retinex model [equation (17)] or the response we derived [equation (15)], one first has to compute the quantity $Q_j^{p,m}$, the receptor input for a given receptor class, associated with the match sample j . As discussed in Appendix D, $Q_j^{p,m}$ is measured in cd/m^2 per receptor. This implies that the predicted L-, M- and S-inputs can be transformed, using equation (C4), to X', Y' and Z' . The final step, the transformation from x', y' to x, y , is given by equations (C5) and (C6).