

Quantitative Studies of Color Constancy

The cover reproduces the image of grey squares illuminated by yellow light, as can be verified by placing the hole in the book marker over one of the squares. That the squares are nevertheless perceived as grey, is due to the effect of color constancy.

ISBN 90-9005902-4

Quantitative Studies
of
Color Constancy

Kwantitatief Onderzoek naar Kleurconstantie

(met een samenvatting in het Nederlands)

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR AAN DE RIJKSUNI-
VERSITEIT TE UTRECHT OP GEZAG VAN DE RECTOR MAGNIFICUS,
PROF. DR. J.A. VAN GINKEL, INGEVOLGE HET BESLUIT VAN
HET COLLEGE VAN DEKANEN IN HET OPENBAAR TE VERDEDIGEN
OP DINSDAG 20 APRIL 1993 DES NAMIDDAGS TE 16.15 UUR

DOOR

Marcel Petrus Lucassen

GEBOREN OP 15 JULI 1964, TE OOSTERBEEK

Promotor : Prof. Dr. D. van Norren
Faculteit der Geneeskunde
R.U. Utrecht

Co-promotor : Dr. J. Walraven
Instituut voor Zintuigfysiologie TNO

Dit proefschrift werd mede mogelijk gemaakt met financiële steun van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), het Instituut voor Zintuigfysiologie TNO en Sun Microsystems Nederland B.V.

Voorwoord

Het onderzoek dat in dit proefschrift is beschreven, werd verricht op het Instituut voor Zintuigfysiologie (IZF-TNO) te Soesterberg. Als (NWO) Onderzoeker in Opleiding genoot ik daar een grote vrijheid, waardoor ik naast mijn eigen fundamentele werk tevens in aanraking kon komen met toegepast wetenschappelijk onderzoek. De interactie met onderzoekers van uiteenlopend pluimage, van fysici tot psychologen, heb ik als zeer stimulerend ervaren, en beschouw ik als een *must* voor iedere onderzoeker. Ik heb op het IZF van uitstekende faciliteiten gebruik kunnen maken, waarvoor ik de directie zeer erkentelijk ben.

Zonder goede begeleiding zou mijn pad van fundamenteel onderzoek ongetwijfeld meerdere malen tot het moeras hebben geleid. Ik mag gerust zeggen dat jij, Jan Walraven, als co-promotor me langs goed begaanbare paadjes hebt gevoerd, op kleurrijke wijze. De urenlange discussies over data en kleur, het gefilosofeer (over het visuele systeem en nog veel meer) en gespeculeer (wie de onwillende reviewers zouden zijn deze keer) zullen me altijd bijblijven. Ik heb me verbaasd over het gedegen overzicht dat je bezit van de kleur-literatuur, waar je mij van hebt laten profiteren. Verbaasd heb ik me ook over de wijze waarop je het weer iedere keer voor elkaar kreeg om mijn manuscripten (in prachtige laser-print) met doodgewoon potlood en gum om te zetten in een nòg gestroomlijndere versie (in goed Engels), en over je pogingen om daarbij te ontsnappen aan de wet van de afnemende meeropbrengst. Ik ben je bovenal dankbaar voor het feit dat je me de vrijheid en tijd gaf om mijn eigen, soms vruchteloze, ideeën op de computer (“zand waar wat stroom doorheen gaat”) los te laten.

Mijn dank voor goede raad en vrijheid gaat ook uit naar mijn promotor, Dick van Norren, die dan toch nog foutjes uit de manuscripten wist te vissen, ondanks het passeren van het Walraven-filter. Ik verwacht, na het verschijnen van dit boekje, nog wel eens deel te nemen aan je eenmans-college “how to deal with editors”. Naast zijn functie als hoogleraar zie ik hem als een goed voorbeeld van onderzoeks-manager. Hij was ook degene die tijdens een zeildag met de Visuologen op de Loosdrechtse plassen mij bekend maakte met het begrip “dessert-wijn”.

Met technisch getinte - maar ook heel gewone - vragen vond ik altijd een

luisterend oor bij Johan Alferdinck, Aart Everts en Jan Varkevisser. Ook hen, en de overige Visuologen wil ik bedanken voor hun tijd en bereidheid tot (technische) ondersteuning. Koos "Quickplot" Wolff en Walter van Dijk bedank ik voor hun onmisbare hulp bij de vervaardiging van de illustraties.

Jan Theeuwes, van een geheel andere discipline, leerde me in de catacomben van het IZF wat "visual pop-out" nu eigenlijk is, en hoe je dat kan meten. Misschien hadden we toch in de Methods van ons paper moeten vermelden dat het er, t.g.v. een overstroming, naar bloemkool rook. Gelukkig was hij niet te beroerd om, na een lange avond meten, die geur weg te werken met een ander geestverruimend middel in de Utrechtse binnenstad. Sindsdien is Jan maar al te goed bekend met het begrip "one for the road".

De laatste maanden van het onderzoek, waarin de promovendus zich onherroepelijk in de schrijf-mode moet begeven, hadden bij mij een nogal turbulent karakter, zeker wat betreft de modellering van de laatste onderzoeksresultaten. Maar ook de hardware problemen bleven niet uit. Graag zou ik nog eens van iemand vernemen wat het optimale aantal back-ups van een proefschrift is, en hoe je die over het land moet verdelen om risicospreiding te garanderen. Het onophoudelijke geklingel van de klokken in de Utrechtse Domtoren, op twee steenworpen afstand van mijn bureau en bed, dreef mij naar Delft, waar ik een rustige werkkamer vond bij mijn vriendin Titia. Zij, en haar kat Brutus - een opmerkelijk levendig beest, zelfs na zijn operatie - bezorgden mij de nodige hoeveelheid relativerend vermogen en ondersteuning om dit boekje op tijd af te krijgen. Onder de goedkeurende blik van Brutus, bovenop de monitor gelegen, werden ontelbare toetsaanslagen geproduceerd. Eénmaal werd het Brutus wat te veel. Hij kreeg het voor elkaar om Control-X te bedienen, de nachtmerrie van iedere gebruiker van de Tempus-tekstverwerker, zonder eerst "gesaved" te hebben.

Tot slot bedank ik mijn familie, een fijne thuishaven, voor de geestelijke steun bij de totstandkoming van dit proefschrift.

Contents

Voorwoord	i
Contents	iii
1 Introduction	1
1.1 Color constancy in historical perspective	2
1.2 The color constancy problem	4
1.3 The mathematical approach: recovering spectral reflectance	6
1.4 The physiological approach	8
1.4.1 Chromatic adaptation	8
1.4.2 Lightness models	10
1.5 Scope of this thesis	11
1.6 Preview	12
2 Evaluation of a Simple Method for Color Monitor Recalibration	15
2.1 Introduction	17
2.2 Method	18
2.2.1 Colorimetry	18
2.2.2 Measuring the input-output relation	18
2.2.3 The recalibration algorithm	20
2.3 Evaluation	21
2.3.1 Constant-configuration case	21
2.3.2 Effect of background luminance and color	25
2.4 Discussion	26
3 Quantifying Color Constancy: Evidence for Nonlinear Processing of Cone-Specific Contrast	29
3.1 Introduction	31

3.2	Methods	33
3.2.1	Illuminant-object interaction	33
3.2.2	Equipment	35
3.2.3	Stimulus	36
3.2.4	Procedure	43
3.2.5	Task	44
3.3	Results	44
3.3.1	Experiment 1: Varying the color of the test illuminant	45
3.3.2	Experiment 2: Varying the color of both the test and match illuminant	47
3.3.3	Experiment 3: The effect of luminance	49
3.4	Data analysis	52
3.4.1	Comparison with Retinex/von Kries theory	53
3.4.2	The second factor	55
3.4.3	Data predictions	58
3.5	Discussion	62
3.6	Appendices	67
3.6.1	A: Reflected light simulation: a numerical example	67
3.6.2	B: Transformation from CIE x,y,Y to monitor Y_R, Y_G, Y_B luminance and vice versa	68
3.6.3	C: Transformation from CIE x,y,Y to cone L,M,S units and vice versa	69
3.6.4	D: A unit for “receptor input” (cd/m^2 per receptor)	70
3.6.5	E: Data predictions in terms of CIE x,y units	71
4	Color Constancy under Natural and Artificial Illumination	73
4.1	Introduction	75
4.2	Methods	77
4.2.1	Surface reflectances	77
4.2.2	Illuminants	77
4.2.3	Stimulus presentation	80
4.2.4	Procedure	81
4.2.5	Task	81
4.3	Results	81
4.4	Data Predictions	85
4.4.1	Response function	85

4.4.2	Computational model	87
4.5	Discussion	91
5	Separate Processing of Chromatic and Achromatic Contrast in Color Constancy	95
5.1	Introduction	97
5.2	Methods	99
5.2.1	General	99
5.2.2	Experiment 1: Varying luminance contrast	101
5.2.3	Experiment 2: Varying illumination level	104
5.3	Theoretical preamble	105
5.3.1	Presentation of models	105
5.3.2	Mathematical equivalence of the JH and LW model	106
5.3.3	The modified LW model	108
5.4	Data predictions	108
5.5	Results	109
5.5.1	Experiment 1: Varying luminance contrast	109
5.5.2	Experiment 2: Varying illumination	113
5.6	Modifying equation (5.1): evidence for luminance-normalized cone signals	117
5.7	Discussion	123
5.7.1	Luminance-free cone signals	123
5.7.2	Luminance and luminance contrast	124
5.7.3	Comparison with models for achromatic vision	125
5.7.4	Other quantitative accounts of color constancy	125
5.7.5	Model implications	126
6	Color Constancy under Conditions Varying in Spatial Configuration	129
6.1	Introduction	131
6.2	Method	133
6.2.1	Stimuli	133
6.2.2	Experimental conditions	133
6.2.3	Illuminants	134
6.2.4	Simulating surface color on CRT	137
6.2.5	Stimulus presentation	137
6.2.6	Task	138

6.3	Model predictions	139
6.3.1	Response function	139
6.3.2	Retinex algorithm (Land, 1986b)	140
6.3.3	Data format	143
6.4	Results	144
6.4.1	Precision of matches and experimental variables	144
6.4.2	Experiment 1	146
6.4.3	Experiment 2	148
6.5	Discussion	149
	Summary	153
	Samenvatting	157
	Epilogue	161
	References	167
	Bibliography	179
	Curriculum Vitae	181

Chapter 1

Introduction

1.1 Color constancy in historical perspective

Among the many impressive achievements of the human visual system, there is one that recently has become the focus of renewed scientific interest. It is usually referred to as color constancy, the ability of the visual system to perceive object colors as fairly constant, despite considerable changes in the spectral composition of the illuminant.

Although awareness of the phenomenon of color constancy may have existed ever since visual scientists realized that (surface) color is mediated by light, Helmholtz (1866) was probably the first who discussed the problem of how the visual system should get rid of the confounding effect of the illuminant. Important contributions followed shortly thereafter, notably by Hering (1878), von Kries (1905) and Katz (1911).

In the twentieth century the research on color constancy gradually advanced (e.g. Krauss, 1926; Helson, 1938, 1943; Judd, 1940), but progress was relatively slow. Until, in 1959, an outsider provoked the scientific establishment into renewed thinking. Edwin Land (1959a,b) gave his first “two-color projection” demonstrations. Using Karp’s (1959) words to describe the experiment: “The technique is to superimpose on a screen the images of two black-and-white slides contained in two projectors having individual colored light sources. The slides for the two projectors are always prepared by photographing a scene through red and green filters so that the transparent areas of each slide correspond to light reflected from the scene in the spectral intervals red-to-yellow, and yellow-to-cyan.” When, for instance, red and white light are used for the projection, the image would be expected to generate red, white and shades of pink, the additive mixtures of red and white light. However, a whole gamut of colors could be perceived, resembling that of the colors present in the original scene. Not only did this experiment show that colored images can be fairly well reproduced with only two primaries, it also showed that the perceived color of an object depends on more than just the spectral distribution of the light reflected from it.

The latter aspect was more thoroughly studied by Land and coworkers in the, by now, classical experiment with the “Color Mondrian”, a collage of rectangular sheets of colored paper, resembling the paintings of the artist Piet Mondriaan. Two identical Mondrians were placed side by side, each one illuminated with its own set of three projector illuminators equipped with band-pass filters (more or less mimicking the spectral sensitivities of the three cone types) and independent

luminance controls, so that the long-wave (“red”), middle-wave (“green”) and short-wave (“blue”) illumination could be mixed in any desired ratio. When the illuminators were so adjusted that the same triplet of radiant energies were reflected to the eye from, for instance, a white paper in the left Mondrian and a red paper in the right Mondrian, the two papers in the Mondrians kept their original color, despite the different illuminations. This procedure could be repeated for any paper in the second Mondrian, and with the same result. So, physically identical stimuli can nevertheless provide many different color sensations. These results indicate that surface colors retain their color identity under a great variety of lighting conditions. Land (1977) concluded: “This constancy is not a minor second-order effect but is so fundamental as to call for a new description of how we see color”. This was a too strong statement, for which Land was heavily criticized (e.g. Judd, 1960; Walls, 1960), but it certainly meant progress. The new approach that Land advocated was embodied in the Retinex (retina-and-cortex) theory of color vision (e.g. Land, 1964, 1974; Land & McCann, 1971). In short, it states that color is the end-product of the independent processing of three black-and-white images from the retina, each image “seen” by one cone type, independent of the absolute flux of radiant energy, but correlated with the (cone-specific) reflectance of objects. In the course of time, several versions of the Retinex (lightness) algorithm have been proposed, never leaving the basic assumption of independent cone signal processing, but with modifications on the spatial sampling of the retinal image.

The Retinex theory has recently been criticized (Brainard & Wandell, 1986), for the reason that it would not achieve adequate color constancy. On the other hand, it is well-known that the visual system does not exhibit perfect color constancy (e.g. Arend & Reeves, 1986; McCann, McKee & Taylor, 1976; Reeves, Arend & Schirillo, 1989; Tiplitz-Blackwell & Buchsbaum, 1988a; Troost & de Weert, 1991b; Valberg & Lange-Malecki, 1990; Walraven *et al.*, 1991), as is also evidenced by the effect of simultaneous contrast. The latter effect can be demonstrated by comparing the color of a surface sample on different backgrounds. What happens is that the color of the sample moves in the direction complementary to that of the background. This indicates, as argued by Shapley (1986), that the visual system does *not* calculate reflectance, but rather, responds to (local) contrast. Although we know now that the Retinex model is not the answer to color constancy, it opened up new ways of thinking about color and provided computational vision theorists with a prototype for new lightness algorithms (see

Hurlbert, 1986).

The most recent advances in color constancy research are mainly of a theoretical nature. In a 1986 issue of the *Journal of the Optical Society*, a Feature Section was devoted to computational approaches to color vision. Eight of the twelve papers dealt with color constancy, and the primary aim was to find algorithms and/or computational approaches that would provide the means for deriving constant surface reflectance properties of objects for different and initially unknown illuminants. Boosted by the progress in automation, color constancy now also has become of interest for the design of machine vision systems. So, there are now two groups of scientists active in this field: those who share the wish to understand the fundamentals of human color vision, and those who would like to implement color constancy into machine vision systems (e.g. robots) so as to mimic, or even outperform, the human visual system. This duality is reflected in the literature, where experimental and theoretical studies have surprisingly little in common. The experimental studies typically lack a quantitative theoretical framework, whereas the theoretical (i.e. computational) studies appear to be hesitant at confronting models with real data. This does not only apply to the field of color constancy, of course. But as pointed out by Troost and de Weert (1991a), it seems that computational studies of color constancy tend to show a complete disregard for the constraints imposed by the visual system.

1.2 The color constancy problem

The main issue in color constancy is how the visual system may recover the color of an object, considering that the most relevant physical information that enters the eye, the spectral distribution of the light emanated from that object, is the univariant wavelength product of illumination, $E(\lambda)$, and reflectance, $R(\lambda)$. Since “light” and “matter” both contribute to the product $E(\lambda) \times R(\lambda)$, the image that is projected onto the retina is one of a dualistic nature. In order to arrive at color constancy, the visual system has to disentangle these two elements, that is, separate the contributions from $E(\lambda)$ and $R(\lambda)$ that are so thoroughly confounded at the input of the visual system. That input, the light absorbed in the photopigments, is obtained by integrating $E(\lambda) \times R(\lambda) \times P(\lambda)$ over the visual spectrum (about 400-700 nm), where $P(\lambda)$ denotes the spectral sensitivity function for the photoreceptor in question. When applied to the human visual system, $P(\lambda)$ can be either $L(\lambda)$, $M(\lambda)$, or $S(\lambda)$, the spectral sensitivities of the

long-, middle-, and short-wavelength-sensitive cones, respectively. It are these cones that perform the (rough) spectral analysis of the incoming light, and hence, are the starting point for all color constancy models.

Unfortunately, the terms *color* and *constancy* may cause some misunderstandings. Although *constancy* would seem to suggest that object colors are perceived as truly constant under varying illumination, it has long been known that color constancy is not perfect. Actually, as Buchsbaum (1980) already pointed out, *inconstancy* would be the more appropriate designation when referring to the perception of light reflected from objects under changing illumination.

More importantly perhaps, *color* can also be interpreted in more than one way, which, when placed in the context of color constancy, is also reflected in the aforementioned division in the literature. According to Thompson, Palacios and Varela (1992), the different explanations favor different philosophical positions: psychophysics and neurophysiology are more compatible with subjectivism (the color is in the brain), whereas computational vision is more compatible with objectivism (the color is in the object). Both interpretations are correct, of course. The object interpretation relates to the stimulus, which can be measured and expressed in physical units. The subject interpretation relates to the way in which the visual system transforms the stimulus into a response. Since spectral reflectance is the invariant physical property of an object under changing illumination, this may be considered as the most relevant stimulus variable in the context of color constancy. However, it still takes a biological system to generate the associated visual percept. The latter, the response, is an arbitrary property of the system, which explains why a color normal and a person with (partial) color blindness may not perceive the same color when looking at the same spectral reflectance. However, as long as changes in reflectance consistently correlate with changes in the visual response, it is irrelevant what different individuals see. More important is that the possible variations in response are matched to the possible variations in the stimulus. This is not true for color defective vision, but neither for normal (trichromatic) vision. Therefore, everybody is seriously "color blind", although some more than others. Interestingly, a color defective person should have relatively good color constancy, for in a (subjective) world where color differences are very limited, color changes due to illuminant changes are limited as well.

Consistent with the duality of the definition of color, two approaches for attacking the problem of color constancy may be considered. Their common goal is

the search for a color descriptor that is invariant to the illuminant, but they may differ in the interpretation of the word color. One approach could be indicated as “following the cone signal processing upstream to the visual cortex”. The other approach could be indicated as “backward calculation to the physical stimulus”. The first approach is directed at tracking down cone signal transformations that occur in the subsequent stages of color vision, the second is directed at recovering the spectral reflectance function of the object (the mathematical solution), which, being invariant, would yield perfect color constancy. In the following, the latter approach - a fairly recent one - is discussed first.

1.3 The mathematical approach: recovering spectral reflectance

The mathematical solution to color constancy involves the recovery of the spectral reflectance functions, $R(\lambda)$, of the objects in the visual scene. Thereto, the illuminant component, $E(\lambda)$, has to be removed from the product $E(\lambda) \times R(\lambda)$ that is registered in the photopigments. Two problems arise here. First, the spectral power distribution of the illuminant may be a completely unknown variable to the visual system. Second, the detailed spectral distribution of $E(\lambda) \times R(\lambda)$ is not available to the visual system. It is reduced to a caricature, due to the fact that in each cone pigment the different wavelengths lose their identity in the process of (integrated) light absorption, the “principle of univariance” as Rushton (1972) called it. Even in the condition where $E(\lambda)$ is known, there is insufficient information to unambiguously recover $R(\lambda)$. However, the problem may be simplified by employing a number of assumptions.

To solve the first problem, the unknown illuminant, Buchsbaum (1980) proposed a strategy to estimate the color of the illuminant. Under the assumptions that the illuminant is common to the entire visual field, and that the field in question has always the same spatial reflectance-average (e.g. the grey-world assumption in which the average reflectance is neutral), the illuminant chromaticity can be estimated from the average chromaticity in the scene. Another way to estimate the chromaticity of the illuminant is to make use of specular highlights (D’Zmura & Lennie, 1986; Lee, 1986), which convey the color of the ambient light. When the chromaticity of the illuminant is estimated, it still has to be transformed into a spectral power distribution. This can be achieved by taking

advantage of the fact, that principal components analysis has shown that only three spectral basis functions are necessary to account for most of the variance of daylight (Judd, MacAdam & Wyszecki, 1964; Dixon, 1978). Hence, within this restricted set of possible illuminants (the phases of daylight) each illuminant chromaticity is determined by a unique triplet of basis coefficients that reconstruct the spectral distribution function for that chromaticity.

To solve the second problem, the decomposition of $E(\lambda) \times R(\lambda)$, an additional assumption about the variety of spectral reflectance functions has to be employed. Cohen (1964) showed that the spectral reflectances of a set of 150 Munsell chips can be described by a linear combination of only three basis functions, covering about 99% of the variance. However, Parkkinen, Hallikainen and Jaaskelainen (1989), using a much larger test set, claim that as many as eight basis functions may be needed. More recently, Dannemiller (1992) concluded that three basis functions are necessary, and probably sufficient, for representing the spectral reflectance functions of natural objects. The number of basis functions will always be a more or less arbitrary decision, since it can become very large, depending on the desired accuracy with which one wants to reconstruct the test set.

A model that approximates reflectance and illuminant spectra by a linear combination of a few basis spectral functions is commonly referred to as a “linear model”, or as a “Sällström-Buchsbaum” model (Brill & West, 1986; Troost & de Weert, 1991a). A visual system that has foreknowledge with respect to these basis functions (those of the illuminant and those of the surface reflectances) may then reduce the problem of color constancy to solving a set of mathematical equations for the basis coefficients that recover an object’s spectral reflectance (e.g. Brill, 1978; Buchsbaum, 1980; Dannemiller, 1989; D’Zmura & Lennie, 1986; Forsyth, 1990; Maloney & Wandell, 1986; Sällström, 1973). In a way, all assumptions concerning the spectral constraints of both the illuminant and the surface reflectances reduce the number of possible combinations of light source and surface reflectances that constitute the same visual input, i.e. the same triplet of cone absorptions. Without these assumptions, there is no unique solution to the problem of color constancy, a property also known as metamerism (Wyszecki & Stiles, 1982). The various algorithms that have been proposed for recovering surface reflectances all rely on the availability of *a priori* information in addition to the quantum catches by the three cone types. They mainly differ in the set of constraints that have to be met. For example, Maloney and Wandell (1986) require the number of spectral sensors to be greater than the number of basis

reflectance functions.

For machine vision systems, in which the parameters for color vision (the photopigments) can be adjusted at will, it may be desirable to implement the mathematical solution to color constancy in order to recognize objects on the basis of recovered spectral reflectance, possibly under a very restricted set of illuminant conditions. It is questionable, however, whether the proposed methods can be considered as serious candidates for modeling human color constancy. Given the coarse sampling of spectral information with three broad cone sensitivity functions, the recovery of a spectral reflectance function is a strategy that intuitively would seem unnecessary complicated (and time consuming) to be profitable for the human visual system. Moreover, human color constancy is known to be imperfect, whereas the methods discussed above enable perfect color constancy within their restricted world of illuminants and surface reflectances. It is therefore not surprising that the applicability of such computational approaches to the human visual system has recently become subject of dispute (Troost & de Weert, 1991a; Dannemiller, 1991).

1.4 The physiological approach

Exactly how the human visual system samples the retinal image and transforms photoreceptor outputs into a robust color code is still a controversial and unresolved issue (e.g. Jameson & Hurvich, 1989; Hurlbert, 1991). There is no physiological evidence, yet, for neural structures that could perform the computations needed for recovering spectral reflectance functions (discussed in the previous section). However, there are less complicated alternatives, one of which is usually referred to as chromatic adaptation.

1.4.1 Chromatic adaptation

One of the earliest and still most used adaptation models, initiated by the ideas of Helmholtz (1866), is the so-called “coefficient rule” of von Kries (1905). It states that the sensitivities of the three cone systems are regulated by cone-specific coefficients (gain-factors), which are inversely proportional to the level of adaptation. Put in the context of color constancy, consider a visual scene illuminated by, for example, either white or blue light. Suppose that under blue light the blue (S) cones are stimulated twice as much as under white light, whereas the red (L)

and green (M) cone activations remain fixed. When changing from white to blue illumination, the von Kries coefficient law predicts that the colors in the scene are changed according to a transformation scheme in which the L, M, and S-cone sensitivities are divided by a factor (coefficient) 1, 1, and 2, respectively. So, for a (non-selective) white object, which has the same color as the illuminant, the cone responses remain unaltered under the illuminant change, implying that von Kries adaptation permits perfect color constancy for that particular white. However, for all other colored objects in the scene, perfect color constancy is not automatically guaranteed. The problem is that the spectral interaction between the blue illuminant and a particular, non-white, surface reflectance in question may cause the ratio (blue/white) of energies reflected in the S-cone waveband to be different from 2. It is for this reason that the role of von Kries adaptation has been debated (Worthey, 1985; Worthey & Brill, 1986; Brill & West, 1986).

Helson (1938, 1943; Helson & Jeffers, 1940) proposed an adaptation model in which the visual system is adapted to a medium grey level. Samples with reflectances above that of the adaptation level take on the hue of the illuminant, whereas samples with reflectances below that of the adaptation level take on the hue complementary to the hue of the illuminant. This principle of how sample colors are related to the color of the illuminant is commonly referred to as the "Helson-Judd" effect.

Another chromatic adaptation model is that of Judd (1940). Instead of scaling the fundamental tristimulus values of object colors, as with von Kries adaptation, the Judd adaptation translates the chromaticities of object colors so as to compensate for the shift of the white point (the achromatic locus) under changes of the illuminant. The difference with the coefficient rule is, that the compensatory effect involves a subtractive instead of a multiplicative operation. As pointed out by Buchsbaum (1980), this model consists of a set of rather arbitrary empirical functions for evaluating hue, saturation and lightness. Again, white plays an important role. In cases where the (new) illuminant is unknown, the best estimate for it is provided by a white reflector since that one conveys the exact color of the illuminant. As with the von Kries adaptation model, perfect color constancy is limited.

The lack of perfect color constancy, as displayed by the above adaptation models, is actually the common result of experimental studies on color constancy (e.g. Arend & Reeves, 1986; McCann, McKee & Taylor, 1976; Reeves, Arend & Schirillo, 1989; Tiplitz-Blackwell & Buchsbaum, 1988a; Troost & de Weert,

1991b; Valberg & Lange-Malecki, 1990; Walraven *et al.*, 1991; this thesis). There may actually be a good reason for allowing departures of perfect color constancy. A visual system with perfect color constancy would not be able to discriminate, so to speak, between day and night since it is capable of complete discounting of the illuminant. One could imagine a visual system that performs less well in this respect, thereby exchanging the loss of constancy for gaining other useful information. For instance, a system that is tuned to discount 90% of the illuminant component may use the information of the remaining 10% for keeping track of the absolute level of illumination. Consider in this respect our capability to perceive different phases of daylight illumination (blue sky versus sundown).

The von Kries and Judd adaptation models are examples of linear chromatic adaptation models. Also, nonlinear (compressive) models exist (e.g. Hunt, 1991; Nayatani *et al.*, 1990), which are often called color appearance models. These models are typically designed to provide data fits to a constrained set of stimuli, and may do so very well. Despite attempts to link the models to receptor physiology (Hunt, 1991; Nayatani, 1990), these models have only found acceptance in the field of applied color science.

1.4.2 Lightness models

Color constancy implies that the visual system is capable of decomposing the product $E(\lambda) \times R(\lambda)$, that is, separate light from matter. This could be interpreted as one of the benefits of color vision. However, the problem of distinguishing between light and matter - e.g. noting the difference between a white, dimly lit, and a grey, brightly lit paper - is not confined to the world of color. The same problem has to be solved in an achromatic (black-and-white) world; for example, when seeing in the dark (scotopic vision) or watching black-and-white television. In those conditions, we have no problem in perceiving invariant what is physically invariant (reflectance) and perceiving as variant what may be physically variant (illumination). In this context, we speak of lightness and brightness (Gilchrist, Delman & Jacobsen, 1983; Jacobsen & Gilchrist, 1988). Lightness is the perceptual correlate of reflectance, whereas brightness refers to the perception of light. The results of various experimental studies (Arend & Goldstein, 1987; Gilchrist, 1988; Jacobsen & Gilchrist, 1988) indicate that the visual system exhibits lightness constancy (i.e. recovers achromatic reflectance), but not brightness constancy.

The problem of color constancy lends itself to be treated as the three-dimensional extension to the one-dimensional problem of lightness and brightness, whereby each cone type deals with the constancy problem in its own black-and-white world. The best known (trichromatic) lightness algorithm is the Retinex algorithm of Land (1983; Land & McCann, 1971). When applied to the experiments with the Mondrian stimulus, it could be shown that the computation of cone-specific lightnesses correlates with the color perception of the human visual system. More specifically, the algorithm computes lightness values (so called “designators”) within each of the three cone wavebands. For each point in an image, the three designators are computed as the intensity at that point relative to the maximum intensity in the entire visual field, or, in another Retinex version, relative to the spatially averaged intensity of the field. The perceived color of that point is then registered by three designator values as a point in a three-dimensional lightness space. The designators are quite resistant against changes in (natural) illumination, and hence, have the property of being illuminant-invariant.

Hurlbert (1986) showed that several other lightness algorithms, all having the Retinex algorithm as their precursor, are formally connected by one and the same mathematical formula. It is of interest to note, however, that the main principle underlying the Retinex computations is equivalent to that of von Kries’ (1905) chromatic coefficient rule (e.g. Jameson & Hurvich, 1989; Valberg & Lange-Malecki, 1990; Walraven *et al.*, 1991). In the von Kries adaptation scheme, the cone-specific intensity at each point is scaled by the intensity (in that same waveband) of a standard white, whereas in the Retinex algorithm it is scaled by the maximum intensity or the spatially averaged intensity of the visual field. The maximum intensity in a scene is often provided by a white, and, under the grey world assumption, the average intensity is that of a grey, only differing from white by a multiplication factor. This equivalence is further discussed in the work presented in this thesis.

1.5 Scope of this thesis

Experimental studies of human color constancy have traditionally been carried out in paradigms employing a flat, two-dimensional stimulus, under homogeneous illumination. Whereas the earlier experiments were performed with real lights and real objects (e.g. Helson, 1938; Land, 1964; McCann *et al.*, 1976), many re-

searchers have turned to using simulations of illuminant-object interactions on a computer controlled color monitor. This clearly has the advantage of flexibility in the design of stimulus patterns, the selection of surfaces and illuminants to be investigated (either realistic or synthetic), and also allows a good deal of automated data processing. However, the use of a color monitor requires repeated colorimetric calibration, and puts certain constraints on the spectral domain and luminance level that can be studied. Nevertheless, we did not hesitate to use a color CRT (Cathode-Ray Tube) in the experiments presented in this thesis. It provided us with the means to systematically study the effects of various experimental parameters that would otherwise not have been as easily accessible in the classical paradigm. In addition, it reduced the calibration problem to a one-measurement procedure.

In real life, when confronted with more complex stimuli than those that we simulated, the mechanism underlying color constancy may use a variety of cues for correctly recovering object color. The strategy may include 3D object features (shape, depth) and illuminant cues (highlights, shading), but also higher cognitive processes (foreknowledge, memory) may be involved. From an analytical point of view it seems a logical first step to start investigating the visual system's behavior with simple 2D-stimuli (such as the Mondrian), stripped from as many cues as possible. This offers the possibility to separately quantify the most primitive mechanism that the visual system employs for achieving color constancy. Then, if that hard-ware contribution is understood, the effect of gradually adding cues to the visual stimulus can be studied in a very systematic manner, possibly leading to a model of human color constancy in which, eventually, cognitive processes are incorporated as well.

This thesis is limited to quantitative studies that are directed at the measuring and modeling of the simple, sensory aspects of human color constancy. The experiments do not sample the full potential of color constancy, but may serve as a baseline for further studies. We refer to the recent work of Troost and de Weert (Troost, 1992) for studies that, following up on the work of Arend and Reeves (1986), attempt to come to grips with the cognitive aspects of color constancy.

1.6 Preview

In this section we present a brief overview of the studies comprising this thesis. With the exception of Chapter 2, which deals with the technical issue of monitor

calibration, each study focuses on a separate variable. That is, the color of the illuminant, the spectral distribution of the illuminant (using pairs of metameric lights), the contrast between sample and background, the over-all illumination level, and the effect of spatial configuration.

In all these experiments the method for quantifying color constancy was basically the same. Two patterns, a “test” and a “match” pattern, representing simulations of an array of Munsell chips under two different illuminants (a “test” and a “match” illuminant), were successively presented to the observer’s left or right eye, so that each pattern was locked to a different eye. In this way, each eye viewed the same pattern, but with colors rendered under its “own” illuminant. The observers were instructed to focus their attention on a central patch (the “test”), but were free to make eye movements. In the match pattern, the observer had to adjust the color and brightness of the central patch so that it matched the color and brightness of the corresponding patch in the test pattern. The comparison of test and match samples thus enables the modeling of how perceived color depends on the variables discussed above.

Chapter 2 addresses the problem of colorimetric calibration of a computer controlled color CRT used for generating the stimuli. As full calibration of such devices is time-consuming, and offers only temporal validity, a simple recalibration procedure was developed that only requires one measurement of a reference white, just prior to an experimental session. The main contribution of this study is not just the method itself, but rather the experimental evaluation of its validity.

In Chapter 3, the effect of illuminant chromaticity on the perceived color of simulated surface samples is studied. The computer simulation employs a trichromatic illuminant-object interaction, which represents a simplification of the interaction in the real world. In the latter, the light reflected from an object is characterized by the product of the spectral distribution of the light source and surface reflectance, $E(\lambda) \times R(\lambda)$. Hence, it requires the complete wavelength functions to be known. In contrast, the trichromatic simulation only requires three illuminant-specific emission coefficients, and three sample-specific reflectance coefficients, which allows a considerable freedom of choice for illuminants and objects to be investigated. From the data analysis, a cone-specific response function could be derived, that successfully describes the experimental data in terms of cone-specific contrast and absolute level of cone stimulation.

Chapter 4 deals with color constancy under natural and artificial illumination. The experimental method was the same as in Chapter 3, except that the

illuminant-object interaction was now simulated in the wavelength domain. So, the stimuli now represented “real” Munsell samples under either natural illumination (phases of daylight) or under (metameric) artificial illumination (light composed of only two wavelengths). In the latter condition, color constancy may be expected to break down completely, due to the impoverished spectral characteristics of the illuminant. Data predictions are shown, computed on the basis of two models. One of the models was the response function derived in Chapter 3, the other was a computational model, based on an algorithm for generating the smoothest (i.e. natural) reflectance functions (van Trigt, 1990). It could be shown that the simple response function was a better predictor of the data than the much more complex computational model.

In Chapter 5, two experiments are discussed that expand on the stimulus range employed in the experiments of Chapter 3. In one experiment, illumination is fixed, but the reflectance of the samples and the background are varied such that the luminance contrast between samples and background - not varied in earlier experiments - covers a two log-unit range. In a second experiment, the color and intensity of the illuminants is manipulated to specifically stimulate the S-cone system, or all three cone types simultaneously. We present a mathematical equivalence - only valid over a limited stimulus range - between our response function (derived in Chapter 3) and the model of Jameson and Hurvich (1964) for brightness contrast. The two experiments of Chapter 5 go beyond the range over which the equivalence between the two models holds, and therefore allow a critical test of the two models. A modified version of our response function is presented, that performs better than the Jameson and Hurvich (1964) model, and also accounts for the data of the experiments discussed in Chapters 3 and 4.

Chapter 6, finally, focuses on the spatial parameter in color constancy. It describes a set of experiments with a variable spacing between the samples in the stimulus pattern. Also, the effect of changing the local surround of the test samples is studied. Again, the experimental data are compared with two model predictions, one of which is the improved response function derived in Chapter 5. The other model is the most recent version of the Retinex algorithm (Land, 1986b), which includes spatial sampling over the entire visual scene. The latter model had not yet been confronted with experimental data, possibly because this requires considerable computational effort. Although this model certainly seems worth pursuing, it was nevertheless outperformed by the much less sophisticated response function.

Chapter 2

Evaluation of a Simple Method for Color Monitor Recalibration

Lucassen, M. P. & Walraven, J. (1990). *Color Research and Application*, 15, 321-326.

Part of the work described in this Chapter was presented at the Tenth International Display Research Conference (Eurodisplay '90; September 25-27, 1990, Amsterdam, The Netherlands).

Abstract

An algorithm for recalibrating a color monitor's RGB input-output relations is presented that requires only a single measurement of a properly chosen reference stimulus. For the application under concern, i.e. reproduction of 35 different colored patches that were used as stimuli for psychophysical experiments on color constancy, the reference stimulus was a white (D_{65}) presented at a luminance corresponding to the mean of the test stimuli.

Three sets of data were obtained for evaluating the algorithm's error reduction power for a given stimulus configuration. These relate to different ways in which the monitor can get out of calibration. That is, slow, but cumulative changes over time, fast changes due to gun interaction (resulting from changed stimulus conditions), and error introduced by a different setting of the monitor's brightness control. Additional experiments were performed to evaluate the effect of background intensity and color.

The algorithm was found to be quite effective in dealing with the instantaneous changes (gun interaction, brightness control), and also for keeping track of the slow changes that may finally necessitate a full recalibration of the monitor.

2.1 Introduction

Computer controlled CRTs are used for a wide range of applications, from displaying text to complex animated graphics. We use our color monitor as a stimulus generator for psychophysical studies on color constancy. Typical for this purpose is the need for a well-defined input-output calibration, i.e. the relation between the CRT's digital input (digital to analog converter value, DAC value) and the screen's light output (luminance) for each of the three R,G,B guns.

When a computer controlled color monitor has been calibrated for a certain stimulus configuration, there is no guarantee that after a period of time, or after a change of configuration, the calibration is still valid. Depending on the application, display hardware and photometric equipment, many adjustments may be needed to reach the desired accuracy for color reproduction.

Recently, several authors reported their findings from monitor calibration efforts (Post & Calhoun, 1987, 1989; Brainard, 1989; Cowan 1983, 1986). Post and Calhoun (1987, 1989) compared seven models for generating colors with specific CIE chromaticity coordinates and luminances on CRTs. They conclude that a piecewise linear interpolation method is most accurate, and found that 16 calibration points per gun are sufficient to reconstruct the input-output relation. However, their work does not solve the common problems of gun interaction and temporal instability. Brainard (1989) focussed on finding a minimal set of assumptions that limit the number of measurement points for monitor calibration, including assumptions of spatial interaction.

A full monitor calibration can be very time consuming, so it is worthwhile to find out when recalibration really becomes necessary. For most applications, a "measure and adjust" algorithm as proposed by Post and Calhoun (1987, 1989) may be used, but again, this involves a lot of measurements.

In this communication we report on the results obtained with a recalibration algorithm that reduces measurements to a minimum. We found that, for a given stimulus condition, a single measurement, i.e. the measurement of the average stimulus chromaticity (usually white) at an intermediate luminance level, may already result in an acceptable recalibration. Recalibration here means shifting the R,G,B input-output relations along the log luminance axis. The chromaticity coordinates of the monitor's phosphors are assumed to remain constant (as was also confirmed by measurement). In the following we shall present data that show both the need for continuous calibration and the efficacy of the method proposed.

2.2 Method

2.2.1 Colorimetry

In principle, that is, assuming additive color mixing to apply, one only needs the input-output relations (luminance vs DAC value) and the three phosphor chromaticity coordinates to calculate the DAC values for the red, green and blue gun, required for producing specified XYZ (CIE 1931) tristimulus values. The colorimetric equation for deriving the monitor's luminance output ($Y_R + Y_G + Y_B$) is given by

$$\begin{pmatrix} Y_R \\ Y_G \\ Y_B \end{pmatrix} = \begin{pmatrix} x_R/y_R & x_G/y_G & x_B/y_B \\ 1 & 1 & 1 \\ z_R/y_R & z_G/y_G & z_B/y_B \end{pmatrix}^{-1} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (2.1)$$

where x, y and z are the 1931 CIE chromaticity coordinates with subscripts R, G, B referring to the appropriate phosphor. The assumption of phosphor constancy implies that the matrix in eq. (2.1) has fixed elements. Note the inversion sign on the matrix in eq. (2.1). The DAC values for the three guns are calculated by

$$\begin{pmatrix} DAC_R \\ DAC_G \\ DAC_B \end{pmatrix} = \text{INTERPOLATION} \begin{pmatrix} Y_R \\ Y_G \\ Y_B \end{pmatrix} \quad (2.2)$$

where the INTERPOLATION operation stands for interpolating the input-output curve on a logarithmic scale. A smaller interpolation error results this way, because the logarithmic input-output curves show less curvature than the linear curves. Applying eq. (2.2) after eq. (2.1) will be referred to as "generating" colors, whereas applying the inverse of eq. (2.1) after the inverse of eq. (2.2) will be referred to as "analyzing" colors. Thus, "generating" involves transforming XYZ to RGB space, whereas "analyzing" implies the opposite transformation.

2.2.2 Measuring the input-output relation

Before a recalibration algorithm can be used, the original set of R,G,B input-output relations must be known. The monitor we used was a high resolution Hitachi 19 inch color monitor (1152×900 pixels, 24 bit/pixel), controlled by a Sun 3/260 computer. Measurements of the CRT's light output were performed with

a SpectraScan PR-702AM (Photo Research) spectroradiometer and a Spectra Pritchard (Photo Research) photometer. The photometer was used for measuring at low luminance levels.

Following the practice recommended by Cowan (1986) and Brainard (1989), the pattern we used for measuring the calibration curves, (spatially) resembled the pattern that was used in the psychophysical experiments. Here, the calibration pattern displayed 35 square patches (70×70 pixels), arranged in a 5×7 array, on a black background. The patches' centers were separated by a (square) grid distance of 140 pixels. The luminance of the central patch, located at the screen's center, was measured with all 35 patches displayed in the same color. The DAC values were chosen so as to produce roughly equal luminance intervals on a logarithmic scale. Each R,G,B curve was measured while the other two guns were disconnected, to exclude their residual contributions (McManus & Hoffman, 1985; Walraven, 1988).

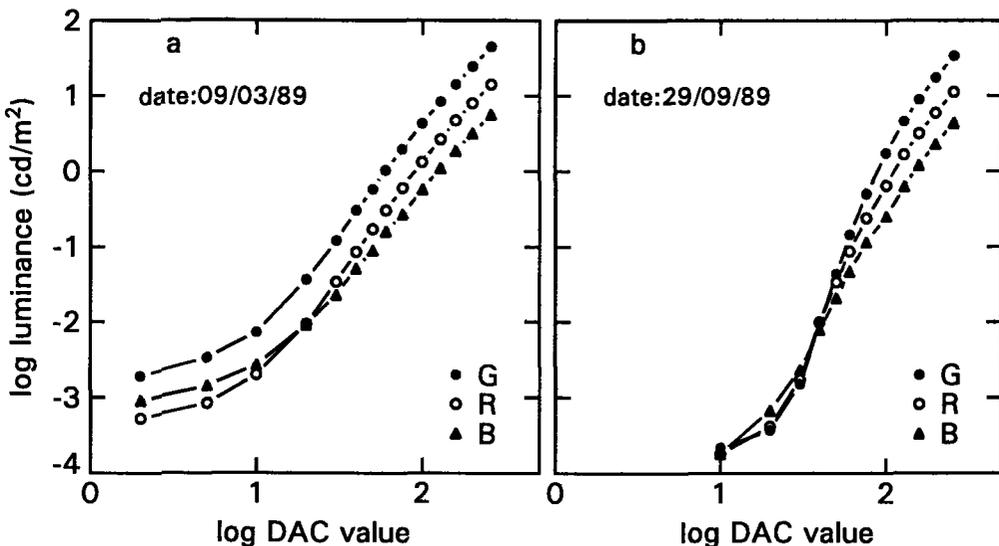


Figure 2.1: a: Luminance vs DAC-value characteristics measured at installation date. b: The same curves measured after about six months of display use.

Figure 2.1a shows the input-output relations, measured at the central patch, whereas Fig. 2.1b shows the same measurements six months later. An-

ticipating the results, to be discussed in the next section, it is clear that the monitor's calibration curves changed quite a bit over time (especially at the lower DAC values). This might be due to aging of the phosphors, although we found, confirming Brainard (Brainard, 1989), that their chromaticity coordinates had hardly changed. We initially measured, at the highest DAC values (255), the following set of (x,y) values for R, G and B: (0.6312,0.3550), (0.3076,0.5957), (0.1473,0.0697), whereas six months later we obtained: (0.6326,0.3549), (0.3065,0.5984), (0.1459,0.0701). For sure, the two sets of curves in Fig. 2.1a and Fig. 2.1b are not related by a single scale factor and thus show the monitor's state to be complex over time.

Note that, on a log-log scale, the input-output relations show an almost linear relationship for the greater part of the DAC values that are used. This is the more or less expected result, considering the exponential relationship between gun voltage and beam current.

Apart from long term variations in screen luminance, also short term effects, like those following a stimulus change (gun interaction), may alter the input-output relations. These are the more day-to-day calibration problems that ask for a simple solution.

2.2.3 The recalibration algorithm

When colors are generated on a CRT screen, in a configuration that is quite different from the one used for calibrating the display, the screen voltage may not remain constant and thus affect the R,G,B beam currents. Other effects may have to be considered as well, but whatever the mechanisms involved, the net result is a change in the input-output relation. In other words, loading the DAC values calculated from (2.1) and (2.2) may not produce the desired luminances Y_R, Y_G and Y_B . The basic idea behind the recalibration algorithm is to compensate for such effects, in as far as they can be treated as gain changes in the DAC-to-luminance conversion. The adjustment consists of a vertical shift (offset) of the three input-output curves on the logarithmic scale, consistent with a scaling of the luminance (Y_R, Y_G, Y_B). The adjustments are made on the basis of a single reference, i.e. an achromatic stimulus (D_{65}) of medium luminance, presented in the center of the screen.

The recalibration procedure thus requires three steps:

1. Generate the white reference stimulus (x_0, y_0, Y_0) using eq. (2.1) and eq. (2.2), and determine the required phosphor luminances, Y_{R_0} , Y_{G_0} and Y_{B_0} .
2. Measure the reference stimulus (x, y, Y) which will probably deviate from its nominal values (x_0, y_0, Y_0) , and calculate the required phosphor luminances, Y_R , Y_G and Y_B .
3. Calculate the correction factors C_R, C_G and C_B , (using $C_R = Y_{R_0}/Y_R$ etc.) and correct the luminances Y_R, Y_G, Y_B of the original input-output curves accordingly. That is, the original input-output relations have their outputs Y_R, Y_G and Y_B divided by the factors C_R, C_G and C_B , respectively.

2.3 Evaluation

2.3.1 Constant-configuration case

The recalibration algorithm was evaluated in the course of psychophysical studies on color vision. Its main purpose was to correct for the gun interaction that occurred when changing from a dark background (as used for calibration) to the light backgrounds used for the stimulus pattern. In addition, the calibration provided information over the gradual change in the light output of the monitor. In order to test the precision of the recalibration, 20 colors were selected out of the 35 that build up the test grid. These 20 colors, located on two different loci of equal Munsell Chroma (see Fig. 2.2), were presented successively in the central patch at target luminance 10 cd/m^2 on a white (D_{65}) background of 12 cd/m^2 . From the remaining 15 colors, 10 colors were located on a third locus of equal Munsell Chroma (10 cd/m^2), whereas the other five were neutrals in the luminance range $1\text{-}11 \text{ cd/m}^2$.

The chromaticities (x, y) and luminances (Y) of the test colors were measured with the spectroradiometer, and then compared with their nominal values (x_0, y_0, Y_0) . The chromatic error, Δxy , and percent luminance error, $\%|\Delta Y|$, were calculated with

$$\Delta xy = \left((x_0 - x)^2 + (y_0 - y)^2 \right)^{1/2} \quad (2.3)$$

$$\%|\Delta Y| = 100 \frac{|Y_0 - Y|}{Y_0} . \quad (2.4)$$

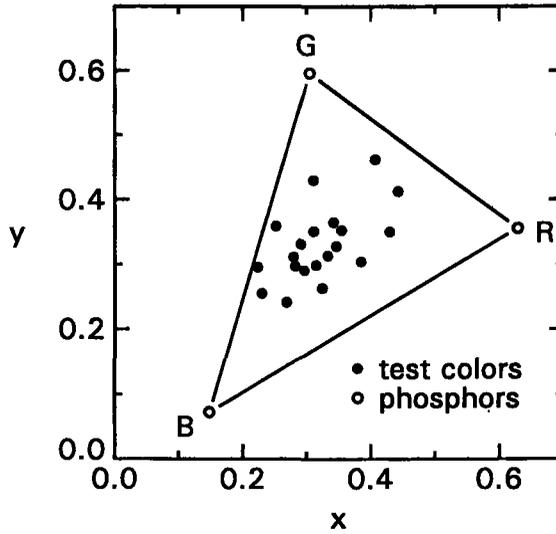


Figure 2.2: Chromaticities (x, y) of the 20 test colors used for evaluating the recalibration algorithm. These colors, located on two loci of equal Munsell Chroma (Value 5/), were presented at luminance 10 cd/m^2 in the central patch of the stimulus pattern.

Transformation into the uniform CIE 1976 $L^*u^*v^*$ color space enables expression of these errors in terms of a color difference ΔE^* :

$$\Delta E^* = \left((\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2 \right)^{1/2}. \quad (2.5)$$

The errors were calculated for the set of test colors, when generated respectively with the original set of calibration functions or the set that resulted after applying the recalibration algorithm. The stimulus pattern for the reference measurement, as demanded by the algorithm, displayed the 35 patches at averaged chromaticity and luminance (D_{65} , 10 cd/m^2) on a white background. This background (D_{65} , 12 cd/m^2) is also the averaged background of the psychophysical experiments.

The recalibration algorithm was applied to three different sets of data. The first set (Set 1) relates to the situation where the same input-output curves are still used after a year's monitor use. It turned out, as shown already in Fig. 2.1, that over this period of time the gradual changes in the monitor had culminated in quite a drastic change of its input-output characteristics. The second set (Set

2) relates to the standard usage of the algorithm, that is, with up-to-date calibration curves, but not necessary applicable to the experimental condition in question (i.e. light background, rather than the dark background used during calibration). In the third set (Set 3) the data were generated in a condition where the brightness control of the monitor was deliberately changed. This is the kind of error that may be introduced when the monitor has different users.

The results obtained in the three test conditions are shown in Table 2.1. What is shown is a comparison of the average error and standard deviation of the 20 test colors when using either the original or recalibrated (scaled) R,G,B input-output curves.

data set	scale factor			$\% \Delta Y $		Δxy		ΔE^*	
	R	G	B	mean	sd	mean	sd	mean	sd
1	1.00	1.00	1.00	54.18	1.65	0.0265	0.0123	13.19	0.89
	1.95	2.45	2.16	31.37	5.13	0.0163	0.0094	9.18	2.90
2	1.00	1.00	1.00	15.04	0.84	0.0028	0.0019	3.09	0.29
	1.18	1.17	1.19	0.63	0.55	0.0031	0.0016	1.02	0.50
3	1.00	1.00	1.00	35.66	0.99	0.0106	0.0057	7.67	0.42
	1.53	1.54	1.52	5.15	1.21	0.0088	0.0046	3.58	1.84

Table 2.1: Comparison of mean error and standard deviation of 20 test colors, either without RGB recalibration (scale factors 1) or with RGB recalibration (scale factor variable). The different data sets relate to different conditions as described in the text.

The results of Table 2.1 are plotted in Fig. 2.3. Note that the error reduction for data sets 2 and 3 is mainly in the luminance direction and that, exactly for that reason, the effect of recalibration is quite effective, reducing the error by 15% and 30% respectively. The small change in chromatic error is reflected in roughly equal scale factors for R, G and B (see Table 2.1). The error reduction for the data of Set 1 is large in both the luminance and chromatic direction and the remaining errors cannot be neglected. Note (in Table 2.1) that the scale factors are different now. This is the expected result in view of the change in shape of the input-output curves over a six month period. Whether such errors are allowed depends on the application. Often, chromatic errors are compared with the size of a MacAdams ellipse, which provides an estimate of the minimum error due to

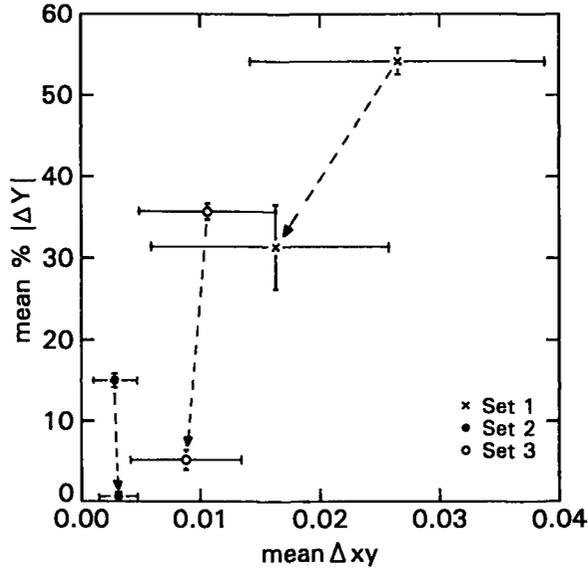


Figure 2.3: Means and standard deviations for chromatic (Δxy) and luminance (ΔY) errors, measured with and without the recalibration algorithm (data from Table 2.1). The dashed arrows indicate the error reduction due to the recalibration algorithm.

the limitations of the visual system. On the basis of tabulated MacAdam ellipses (Wyszecki & Stiles, 1982), we obtained a rough estimate of the average minimum error in the chromaticity space covered by the color monitor. Considering only the error in the direction of the major axis of the relevant ellipses (i.e. those located within the monitor's RGB space), we arrived at an average (Δxy) of 0.005. This means that, for data sets 2 and 3, the accuracy of color reproduction (obtained with interpolation of the input-output curve and the recalibration algorithm) can be in the order of a just perceptible chromaticity difference. The same conclusion can be drawn from the analysis of ΔE^* in Table 2.1, since a just noticeable difference can be estimated to be of the order of 2 to 3 CIELUV units (Sproson, 1983). On the other hand, data set 1 shows that a full monitor recalibration is necessary.

2.3.2 Effect of background luminance and color

So far we have considered only one change of stimulus configuration, that is, changing the background from dark to D_{65} at 12 cd/m^2 . We also performed some additional experiments (at a later time) to evaluate the effects of changing the luminance or color of the background (grid). If such effects could be described by a simple relationship between scale factors and background parameters, this might possibly obviate the need for a reference stimulus for each new experimental condition. First, a full calibration, as described earlier, was performed because of our monitor's continuing decline. We then set out to measure the 20 test colors, presented as the central patch of the 35 patches, using different grid luminances. For each color, the R,G,B scale factors (C_R , C_G and C_B) were calculated that would be required to exactly reproduce the nominal x_0 , y_0 , Y_0 values. The effect of the D_{65} background luminance on the obtained scale factors is shown in Fig. 2.4.

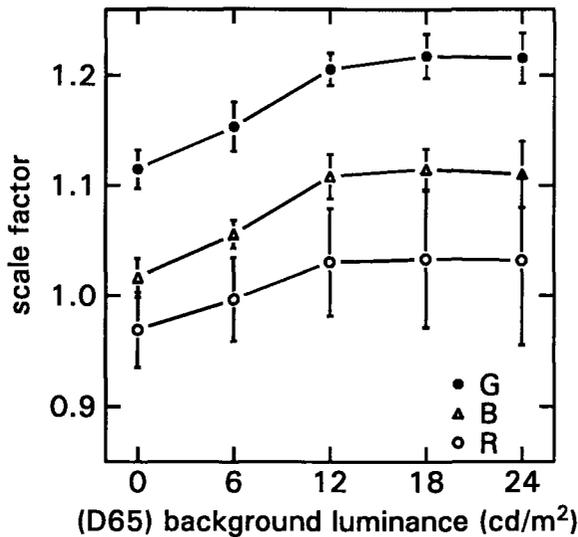


Figure 2.4: Calculated R,G,B scale factors (mean values \pm standard deviation) as a function of the D_{65} background (grid) luminance. The same procedure and the same 20 test colors (see Fig. 2.2) were used.

At the time these measurements were made the monitor had still changed somewhat more, resulting in somewhat different scale factors from those shown

in Table 2.1. Figure 2.4 shows that the scale factors increase with increasing grid luminance, but only up to a value of about 12 cd/m². The over-all pattern is regular enough to suggest a procedure for a more general recalibration algorithm, that could utilize (average) screen luminance for deriving the associated R,G,B scale factors.

To test whether the color of the background might also be an important variable, we compared the scale factors obtained for white light (D_{65} at 12 cd/m²) to those obtained for equi-luminant red ($x=0.4150$; $y=0.3300$) and green ($x=0.3127$; $y=0.4320$) backgrounds, respectively. These backgrounds introduce a change in the R,G,B luminance distributions in the direction of either the red or green gun, and might thus reveal a possible gun-specific effect of background on scale factor. One should observe, then, that the red scale factor (C_R) is more affected by the red than the green background, and vice versa.

The results we obtained with the two colored backgrounds were all in the direction of a reduction of the R,G,B scale factors relative to those obtained in the condition with a white background. That is, for the red background : 2.9% for C_R , 2.5% for C_G and 1.8% for C_B . For the green background the reductions are : 1.9% for C_R , 3.3% for C_G and 2.7% for C_B . These results indicate a (small) gun-specific effect (the largest reductions of C_R and C_G are found for the correspondingly colored backgrounds), but since the effect of color is small anyway (here 2 to 3%), a gun-specific parameterization of the CRT image does not seem very profitable. The main effect (see Fig. 2.4) is due to the luminance step, irrespective of the color of the background.

2.4 Discussion

The simple recalibration algorithm we proposed turned out to be well suited for the purpose it was developed for, that is, compensating for non-additivity of the (separately measured) color guns. In general, this method is only suitable for correcting errors, that can be described in terms of vertical translation of the log RGB vs DAC value functions. It is of interest though, that our results show that this is the kind of error that is likely to be encountered on a CRT display.

The results from our experiments in which we varied the luminance and color of the background are too limited to allow firm conclusions. Still, they indicate that the R,G,B scale factors vary with the over-all rather than the gun-specific display luminance. This might indicate an interaction between screen-voltage and

beam current, that is fairly insensitive to the particular ratio of the constituent R,G,B beam currents. This should make it easier to adapt the recalibration algorithm for use in conditions with varying stimulus conditions.

If, in the course of time, the algorithm shows error reduction to be less complete, this is a warning signal. Values from 1.1 to 1.2 are normally found, but when the scale factors become too large a full monitor recalibration is needed. This is illustrated by the data of Set 1, which relate to the condition where the shape of the input-output curves had changed with time. So, regularly checking the scale factors is also effective to discover slow drifts in the monitor's output.

The fact that the scale factors are greater than 1, means that the measured output is less than would be expected from the calculations. Several factors (e.g. phosphor aging, gun interaction) may contribute to this loss in effective output, but these are nevertheless handled by the simple scaling procedure of the recalibration algorithm. This is particularly helpful when different stimulus configurations, requiring different correction factors, have to be displayed. The fact that a single measurement (of the reference white) was found to be sufficient for the recalibration procedure, does not necessarily apply to all stimulus conditions. However, if it does, as can be tested in the way we have shown, much time and effort can be saved in maintaining accurate stimulus control in complex stimulus scenarios. Moreover, measuring just a single white point on the screen, could be done with a simple (but reliable) chromaticity meter, which is much less expensive and cumbersome than using the spectroradiometer that would be needed for measuring colored stimuli.

Chapter 3

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

Lucassen, M. P. & Walraven, J. (1992a). *Vision Research* (in press).

Part of the work described in this Chapter was presented at the Thirteenth European Conference on Visual Perception (September 4-7, 1990, Paris, France).

Abstract

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 35 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.

3.1 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 and 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, 1959a, 1986a,b; Land & McCann, 1971; 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. In as far 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 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 the Methods section), cone-specific contrast (rather than absolute light input) provides the relevant signal for color perception (Walraven *et al.*, 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 versus colored, but also colored versus 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 subserved by cone-specific contrast processing.

3.2 Methods

3.2.1 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. 3.1.

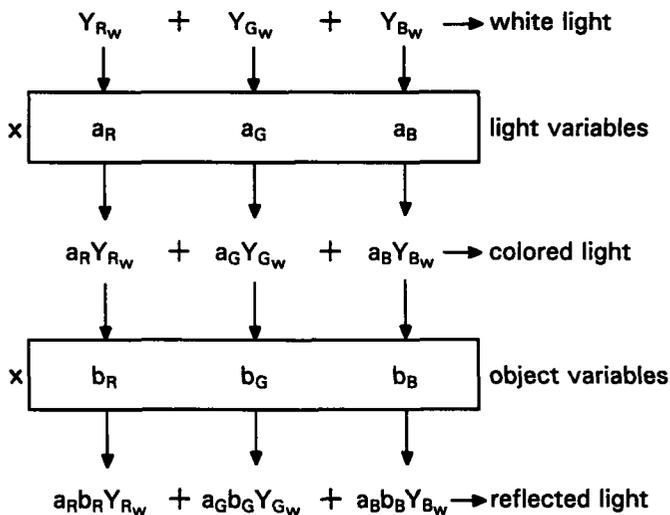


Figure 3.1: Diagram illustrating the principle of trichromatic reflection.

In Fig. 3.1 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

$$L_r = a_R b_R Y_{R_w} + a_G b_G Y_{G_w} + a_B b_B Y_{B_w}. \quad (3.1)$$

The six coefficients in eq. (3.1) 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 eq. (3.1) 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 analogue 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 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 XYZ 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. 3.1.
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

$$a_i = Y_{i,ill} / Y_{i_w} \quad i = R, G, B. \quad (3.2)$$

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

$$b_i = Y_{i,j}/Y_{i,w} \quad i = R, G, B. \quad (3.3)$$

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 $Y_B = a_B b_B Y_{B,w}$ (eq. (3.1)).

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 for 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 computer simulations, 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.

3.2.2 Equipment

A Hitachi 19 inch 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 to what 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 meter 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 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).

3.2.3 Stimulus

Test samples

The stimulus pattern consisted of an array of 35 color samples displayed on a white surface. For 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. 3.2. This is the same

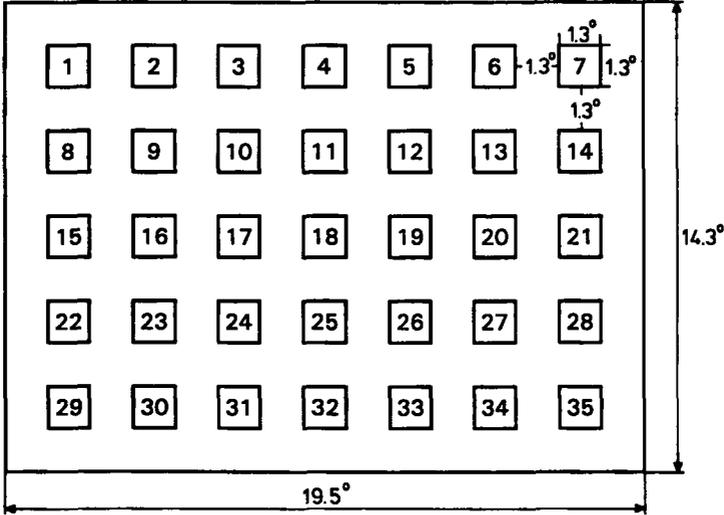


Figure 3.2: Geometry of the stimulus pattern used for both test and match stimulus. The pattern is a computer simulation of 35 color samples (1.3° squares), displayed on a piece of white paper (grid). The sample numbers correspond to those in Table 3.1, where the corresponding color specifications are given under standard white light. The 11 test samples (marked by an asterisk in Table 3.1) were always presented at the location of sample 18.

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 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 & Stiles, 1982, Table I(6.6.1)), and are plotted in Fig. 3.3. The specifications of the color samples are listed in Table 3.1. 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 first three blocks in Table 3.1). The luminance of the 30 chromatic samples was set

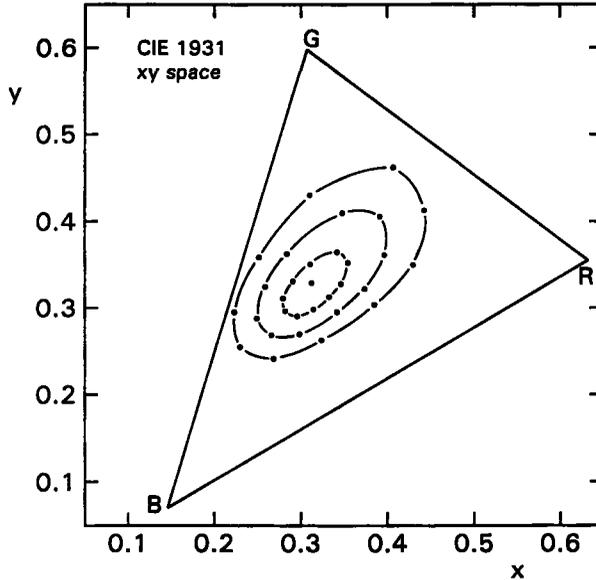


Figure 3.3: Chromaticities (CIE x, y coordinates) of the color samples (Table 3.1), under standard white illumination (RGB metamer of D_{65} , 12 cd/m^2). The triangle encloses the chromaticity space covered by our color monitor. The set consists of 30 samples, with chromaticities evenly distributed 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).

to be consistent with 50% of the luminance of the white light. The remaining five achromatic samples (fourth block in Table 3.1) 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

sample number in Fig. 3.2	CIE 1931 specification			reflection coeff. 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	-

Table 3.1: Specification (under standard white light) of the 30 chromatic and five achromatic samples of the stimulus pattern shown in Fig. 3.2. The 11 samples of the test set are indicated by an asterisk.

displayed on a grey background of 20% reflectance.) The sample numbers in Table 3.1 correspond to the array numbers in Fig. 3.2. 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 Experiment 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 comparing test samples under white and colored light, respectively. The (colored) illuminants produced equal grid luminances (12 cd/m^2) and equal Munsell Chroma (Chroma /6). The x, y, Y specification and emission coefficients a_R, a_G, a_B that appear in the reflection simulation of these illuminants, including that of the white reference illuminant (D_{65}), are presented in Table 3.2.

test illuminant	CIE 1931 specification			phosphor luminances			emission coeff. 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
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

Table 3.2: CIE 1931 x, y chromaticities and luminance Y (at 100% grid reflectance) of the test illuminants (see also Fig. 3.4). 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.

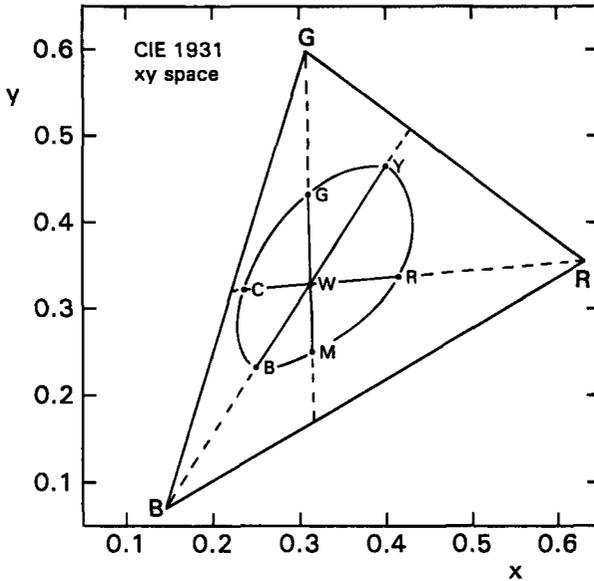


Figure 3.4: Chromaticities of the test illuminants, B(lue), C(yan), G(reen), Y(ellow), R(ed), M(agenta) and W(hite). The x, y coordinates of the colored lights are located on the Munsell 5/6 Chroma line and thus represent lights that are perceptually equi-distant from the white point. The corresponding chromaticities are listed in Table 3.2.

As is shown in Fig. 3.4, 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. 3.5. These are the chromaticity coordinates of the illuminant-reflectance products as described above (Appendix A). The luminance of the chromatic samples was 6 cd/m^2 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.

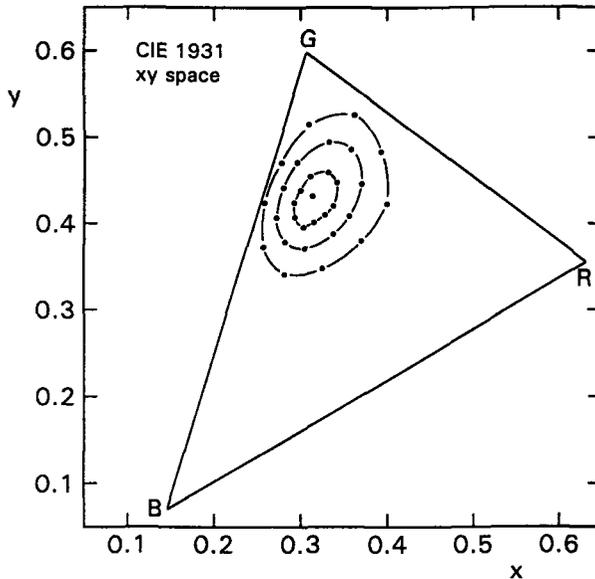


Figure 3.5: The same color samples as shown in Fig. 3.3, but now illuminated by green light. The samples are now centered around the chromaticity locus of the green illuminant, which coincides with that of the achromatic samples.

In the experiments, two differently illuminated sets of color samples (like the ones shown in Fig. 3.5) 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. 3.2), to match it to

the color of the corresponding test sample seen under the other illuminant (test condition). The two illuminants in question 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 versus colored illuminant comparison, we also compared a wide variety of colored versus colored illuminant pairs. We were thus able to obtain a much larger range of differential chromatic stimulation than is possible with the classical white versus colored illuminant comparison. In total, with the inclusion of the six equi-luminant white versus colored pairs, we used 45 different illuminant pairs. Among this set were combinations of the colored illuminants listed in Table 3.2 (Y(ellow)/B(lue) for instance), and also illuminants with a 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 our 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.

3.2.4 Procedure

After five minutes of dark adaptation and a few more minutes to adapt to the average luminance of the test pattern (about 10 cd/m^2), 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 11 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 represent

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

3.2.5 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. The only constraint was to divide the presentations roughly equally between the left and right eye.

3.3 Results

In the pilot stage of this study, we did an experiment in which we made (haploscopic) 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 haploscopic matching technique, the matches are less precise than what can be achieved in a (monocular) side by side comparison. We found an average deviation (for our particular set of 11 test colors) of $\bar{\Delta}xy = 0.008$, as computed with

$$\bar{\Delta}xy = \frac{1}{11} \sum_{i=1}^{11} (\Delta x_i^2 + \Delta y_i^2)^{1/2}. \quad (3.4)$$

This is about a factor six 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 (Experiment 1), the color of both the test and match illuminant (Experiment 2), and, as an extension of Experiment 2, the luminance of the match illuminant (Experiment 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.

3.3.1 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 3.2, 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 grey 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. 3.6. Open squares represent the chromaticities of the color samples under the (colored) test illuminant, open circles those under the (white) match illuminant. The filled 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. That is, 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. 3.6.

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. 3.6. In this case, the (cyan) illuminant is located near the boundary of the monitor's phosphor triangle (see Fig. 3.4). 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 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. 3.6, 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 versus white light.

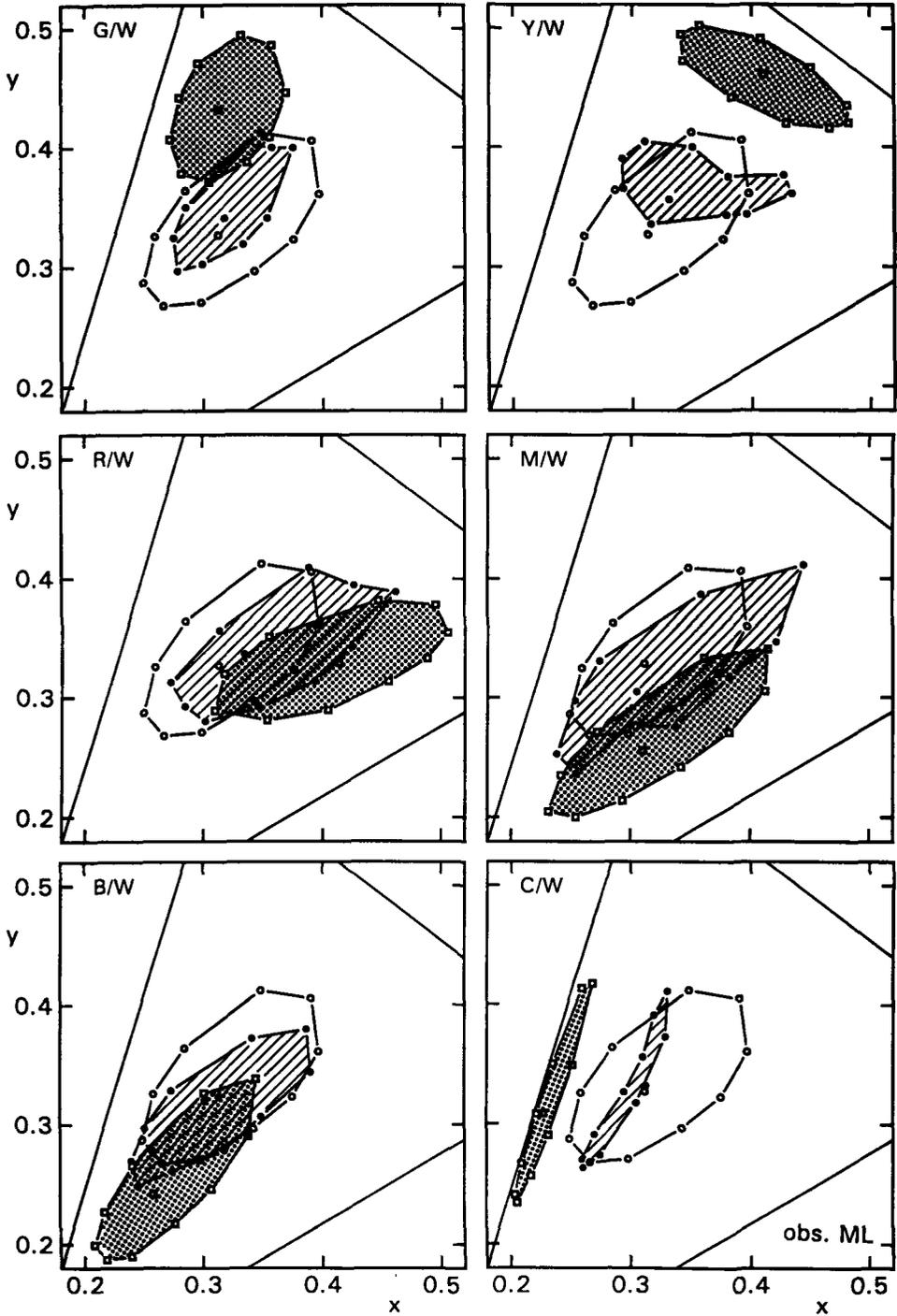


Figure 3.6: (previous page) Data from Experiment 1. The different plots relate to different combinations of test and match illumination. The six colored illuminants, listed in Table 3.2, were used for illuminating the test pattern, whereas the standard white illuminant (W) was used for illuminating the 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. Filled 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.

3.3.2 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, it turned out that there was no real need for doing so, since it turned out that 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 we selected 10 different illuminant pairs. Representative results, as obtained for the illuminant pairs B/G, C/Y, and Y/B, are shown in Fig. 3.7. 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. 3.6, except that the open circles now represent the chromaticities of match samples as seen under colored, instead of under white light.

Judged by the criterion that perfect color constancy requires that matched chromaticities coincide with chromaticities of test samples under the match illuminant (superposed open and hatched areas), the results shown in Fig. 3.7 would seem to indicate that color constancy is much less effective in the conditions of

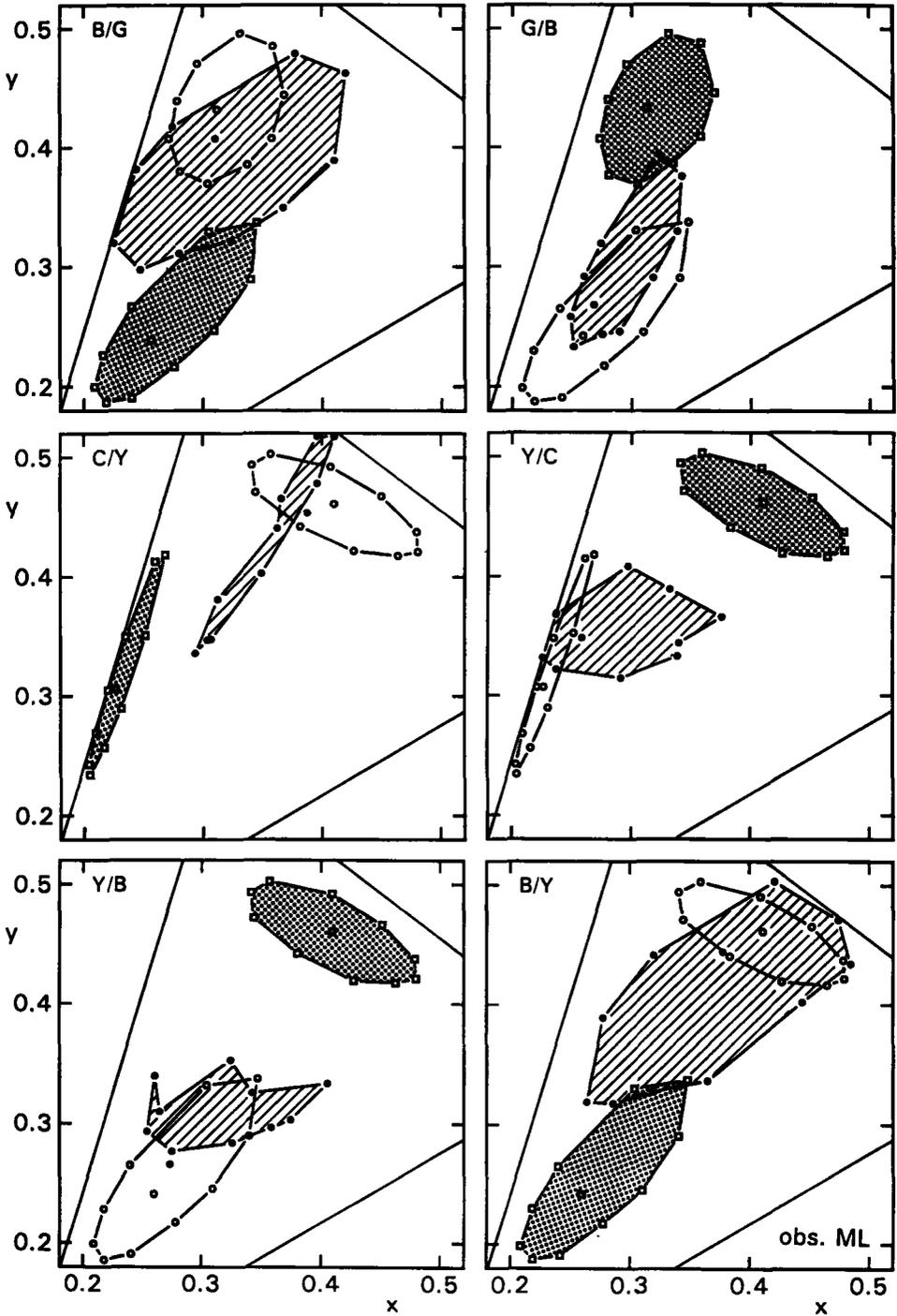


Figure 3.7: (previous page) Data from Experiment 2. The symbols and area fillers have the same meaning as in Fig. 3.6, except that the open circles now represent the chromaticities of matches made under *colored* instead of under white illumination. The left graphs show the results for the three illuminant 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.

Experiment 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. 3.7, 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. 3.6, 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.

3.3.3 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/m^2 . In Experiment 3 we repeated the B/Y, C/R and R/C experiments, but now with halved (6 cd/m^2) and doubled (24 cd/m^2) 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 Fig. 3.8, which shows the results for the B/Y condition. The results obtained in the other conditions were similar.

The data plotted in Fig. 3.8 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

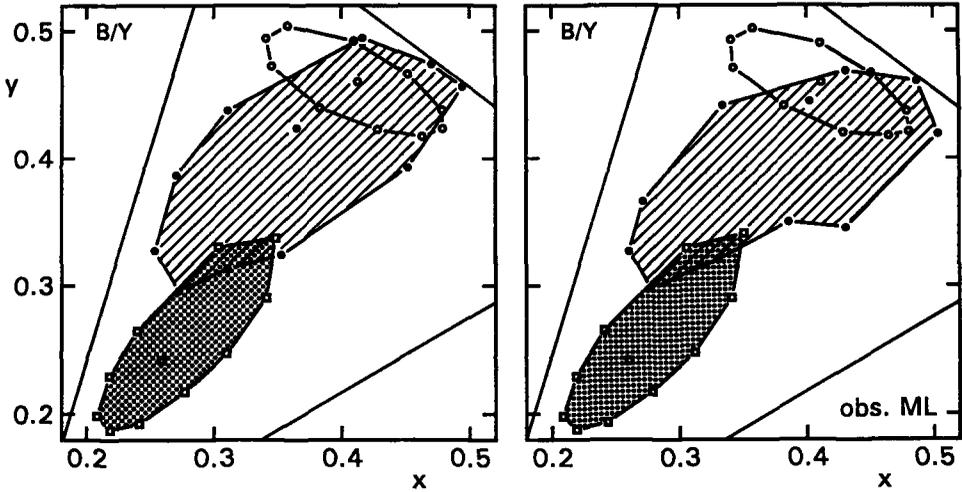


Figure 3.8: Example of data from Experiment 3. Symbols and area fillers as in Figs. 3.6 and 3.7. The experiment is identical to the B/Y condition in Experiment 2 (see also Fig. 3.7), but now with the brightness of the match illuminant halved (left plot) or doubled (right plot). Note that these (overall) luminance changes hardly affect the chromaticities of the matching samples.

grid luminance) was approximately maintained. However, there was a small but systematic deviation from the constant contrast prediction. This is illustrated in Fig. 3.9, which shows the pooled results from the B/Y, C/R and R/C conditions.

Figure 3.9 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 attended by a slight increase in brightness.

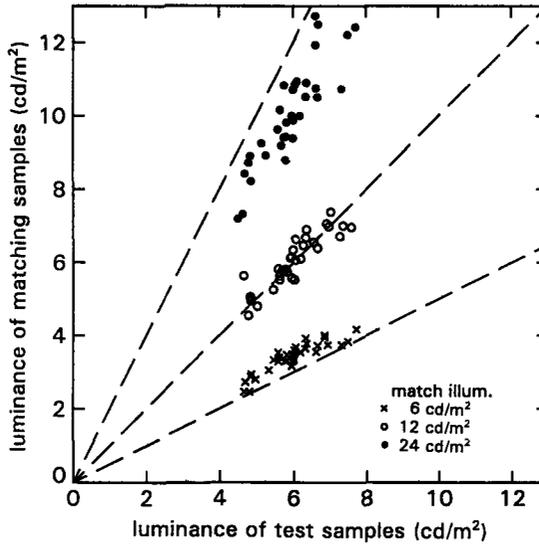


Figure 3.9: Comparison of the luminances of test and matching samples for illuminant combinations B/Y, C/R and R/C. The different symbols relate to three levels of illumination. The test illuminant always produced a grid luminance of 12 cd/m^2 , whereas the match illuminant produced grid luminances of 6 cd/m^2 (crosses), 12 cd/m^2 (open circles) or 24 cd/m^2 (filled 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.

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.

3.4 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 for the stimulus in question. The same approach was found to be useful for 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 study¹ 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 3.2) are given in Table 3.3. Note from Table 3.3 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.

¹The 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 (Experiment 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 Experiment 3), it is unlikely that correction for pupil size would have any effect on the outcome of the data analysis.

test illuminant	selective cone input (cd/m ²)		
	L	M	S
W(hite)	3.92	4.09	4.16
R(ed)	4.27	3.41	2.95
G(reen)	3.19	4.33	2.27
B(lue)	3.82	4.20	7.85
Y(ellow)	4.04	3.91	1.09
M(agenta)	4.02	3.82	6.42
C(yan)	3.63	4.63	5.76

Table 3.3: Cone input values (in cd/m² per receptor) of the test illuminants listed in Table 3.2 as produced by transforming their x, y, Y values according to eqns. (3.21)-(3.23).

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

3.4.1 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 reduce 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

$$D_j^p = \log \left(\frac{Q_j^p}{\bar{G}^p} \right) \quad p = L, M, S \quad (3.5)$$

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/m^2 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. 3.2), 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} (for a given cone class) can be computed according to

$$\bar{G}^p = \left((Q_w^p)^{130} \prod_{j=1}^{35} Q_j^p \right)^{1/165} \quad (3.6)$$

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 modelled 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: $\bar{G}^L = 0.860 Q_w^L$, $\bar{G}^M = 0.857 Q_w^M$ and $\bar{G}^S = 0.849 Q_w^S$. 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 illuminating 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 eq. (3.5), may also be expressed as a fraction β_j^p (reflection coefficient) of the illuminant value Q_w^p . Therefore, the designator reduces to

$$D_j^p = \log \left(\frac{Q_j^p}{\bar{G}^p} \right) = \log \left(\frac{\beta_j^p Q_w^p}{\alpha^p Q_w^p} \right) = \log \left(\frac{\beta_j^p}{\alpha^p} \right) \quad (3.7)$$

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

In Fig. 3.10, the experimental data from condition B/Y (see Fig. 3.7 also) is 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. 3.10 that the Retinex model (crosses) does not predict the data (open circles) in this

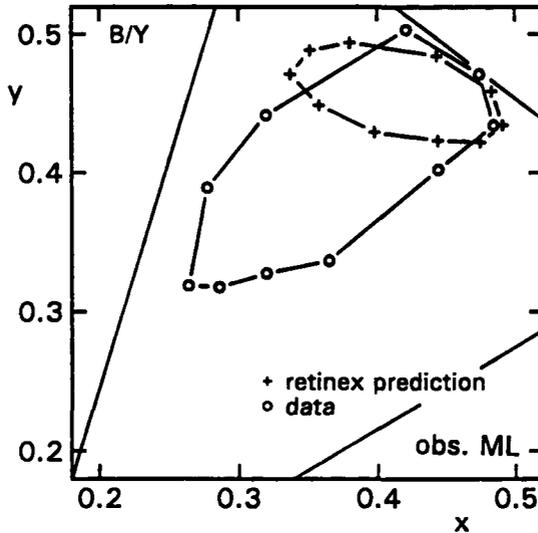


Figure 3.10: Example of experimental data (open circles) and their prediction on the basis of the Retinex model (crosses). The data represent the chromaticities of samples, under blue illumination that match the colors of the corresponding samples under yellow illumination (condition B/Y).

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.* (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.

3.4.2 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. 3.11 (left panel). Along the (logarithmic) axes are plotted receptor-specific contrast ratios C^t and

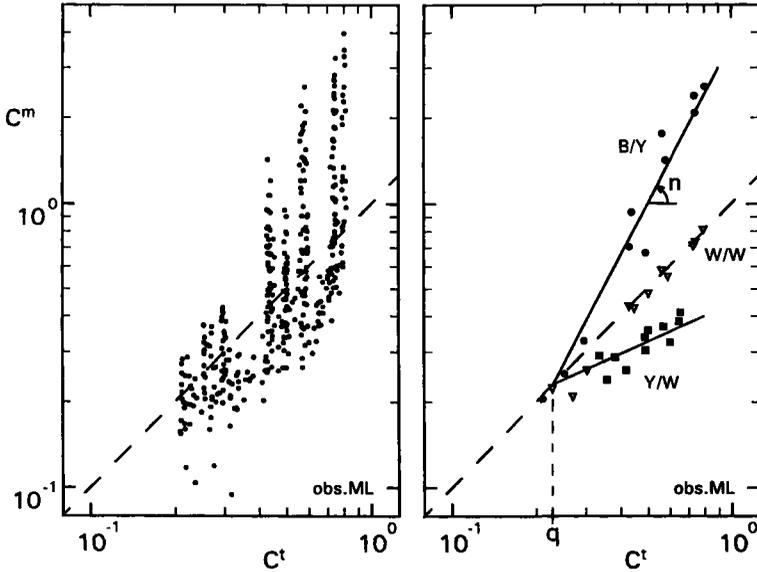


Figure 3.11: Comparison of receptor-specific contrast (S-cone system), for test and match samples (C^t and C^m). Contrast is defined as $C = Q_j/Q_w$. Left panel: Pooled data from 36 experiments. Right panel: Replotted data for experimental condition: B/Y (circles), W/W (triangles) and Y/W (squares). 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.

C^m defined by

$$C^t = \frac{Q_j^t}{Q_w^t} \quad \text{and} \quad C^m = \frac{Q_j^m}{Q_w^m} \quad (3.8)$$

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 3.11 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 illustrated in the right panel of Fig. 3.11 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

$$\log(C^m) - \log(q) = n \left(\log(C^t) - \log(q) \right) \quad (3.9)$$

which reduces to

$$\log(k C^m) = n \log(k C^t) \quad k = q^{-1}. \quad (3.10)$$

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 eq. (3.10), that is, the coefficient n , which represents the slope of the functions shown in Fig. 3.11 (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. 3.12. The two data sets, shown in Fig. 3.12, can both be approximated by the power function

$$n = \left(Q_w^t / Q_w^m \right)^r \quad (3.11)$$

with $0 < r < 1$. The best fits were obtained for $r = 0.33$ (obs. ML) and $r = 0.20$ (obs. JW). One could speculate that a difference in macular pigmentation (JW is 25 years 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 eq. (3.10), we can now substitute $\bar{k}=4.35$ and specify n according to eq. (3.11). One thus obtains

$$\log(4.35 C^m) = \left(Q_w^t / Q_w^m \right)^r \log(4.35 C^t) \quad (3.12)$$

which can be rewritten as

$$\left(Q_w^m \right)^r \log(4.35 C^m) = \left(Q_w^t \right)^r \log(4.35 C^t) . \quad (3.13)$$

The symmetry of eq. (3.12) 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

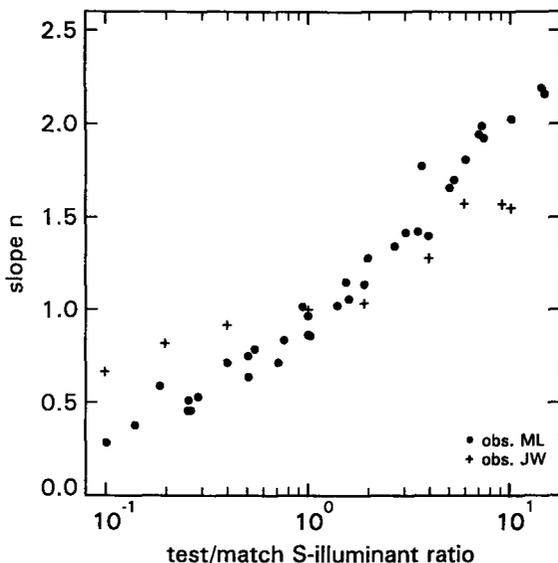


Figure 3.12: The coefficient n , i.e. the slope of the lines used for fitting the S-cone data shown in Fig. 3.11, plotted as a function of the interocular illuminant ratio Q_w^t/Q_w^m .

cone system response (R) that can be described (over the limited contrast range tested here) as

$$R^p \equiv (Q_w^p)^r \log(4.35 C^p). \quad (3.14)$$

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 eq. (3.7)). 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 eq. (3.14) rather than eq. (3.5) as basis for describing our data.

3.4.3 Data predictions

The expression given in eq. (3.14) was derived on the basis of data relating to the S-cone input. Although eq. (3.14) 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 eq. (3.12).

The data predictions are in terms of the quantity of $Q_j^{p,m}$, the cone input that is required for the match sample j . The latter can be solved from eq. (3.13) by defining test and match contrast ($C^{p,m}$ and $C^{p,t}$) according to their definition given in eq. (3.8). One thus obtains

$$\log(Q_j^{p,m}) = \left(\frac{Q_w^{p,t}}{Q_w^{p,m}} \right)^r \log \left(4.35 \frac{Q_j^{p,t}}{Q_w^{p,t}} \right) + \log \left(\frac{Q_w^{p,m}}{4.35} \right) \quad (3.15)$$

with r taking the value 0.33 (for ML) or 0.20 (for JW). In Fig. 3.13, the comparison is shown between the predictions of the value of $Q_j^{p,m}$ on the basis of our description, as given in eq. (3.15), 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, that is,

$$D_j^{p,m} = D_j^{p,t} \quad p = L, M, S. \quad (3.16)$$

After expressing D_j^p according to eq. (3.5) one can derive

$$Q_j^{p,m} = Q_j^{p,t} (\bar{G}^{p,m} / \bar{G}^{p,t}) \quad (3.17)$$

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 eq. (3.6).

Figure 3.13 is composed of nine graphs, each of which plots the predicted value of Q_j^m on the abscissa and the actually obtained value of Q_j^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. Perfect model predictions would be indicated by data points appearing on the dashed line. The graphs in the top panel of Fig. 3.13 relate to the naive physical model in which the predicted values of the match are those that reproduce the test sample ($Q_j^m = Q_j^t$). 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.

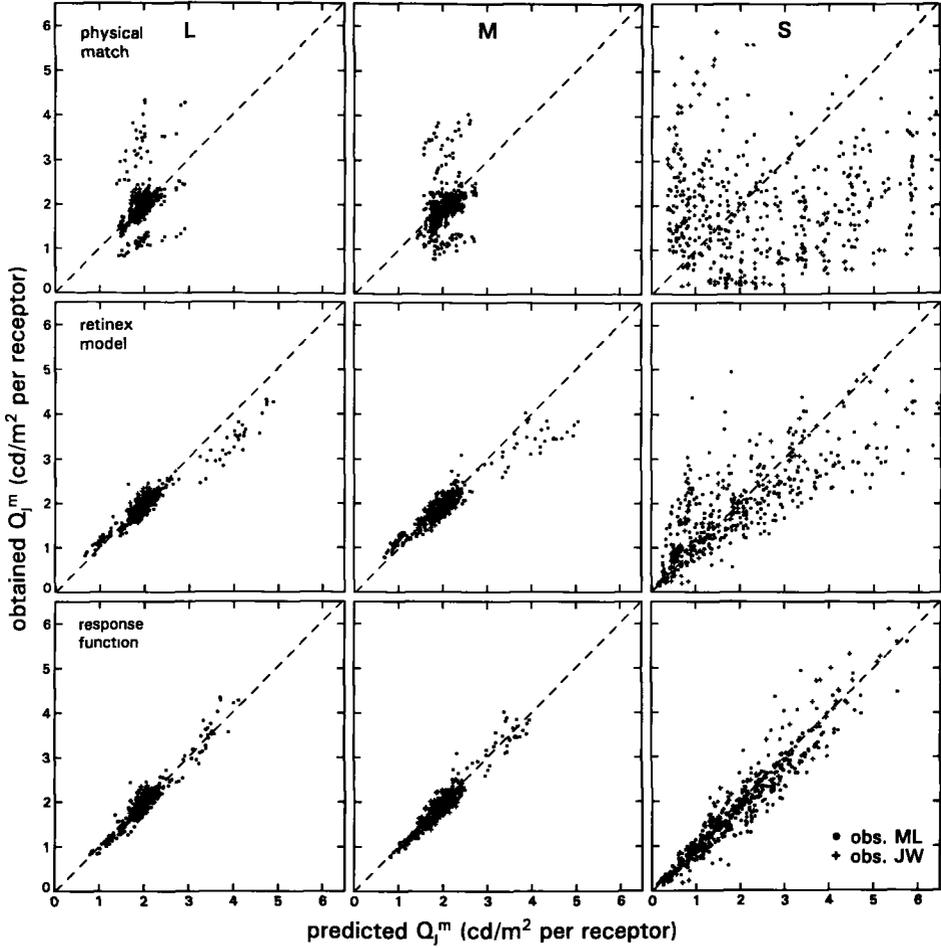


Figure 3.13: Predicted versus obtained data (495 in all), in terms of Q_j^m , the quantity that provides a measure for the receptor inputs as required for the match sample (j), under the match illuminant (m). The panels are arranged in three rows, representing the predictions (for L, M and S-cones, respectively), as obtained with three different models: the physical match, the Retinex model and the response function described by eq. (3.14).

The graphs in the middle panel show the predicted values of $Q_j^{p,m}$ on the basis of Retinex theory, as computed with eq. (3.17). 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 lowest predicted values for the L and M-receptor.

As shown in the previous section, the main difference² in color coding between the Retinex algorithm and the derived response function R^p , is the factor $(Q_w^p)^r$. 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 eq. (3.14) explains 92.2%, 95.6% and 95.4% of the associated data variance (ρ^2). The improvement in data prediction when applying eq. (3.15) rather than the Retinex prediction eq. (3.17) can even be better appreciated in x, y chromaticity space (see Appendices for the computations in question). This is illustrated in Fig. 3.14, which reproduces Fig. 3.10, but now with the added prediction based on eq. (3.15). The result shown in Fig. 3.14 relate to an experimental condition that causes a rather extreme illuminant change within the S-waveband. As can be seen in Fig. 3.13, 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.

²Another 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 eq. (3.7) : $\bar{C}^p = \alpha^p Q_w^p$. Since the contribution of the background in the computation of the geometric mean is large compared to the 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 eq. (3.17), α^p drops out of the equation. So, the prediction with the Retinex algorithm reduces to a prediction on the basis of the contrast $C^p = Q_j^p / Q_w^p$.

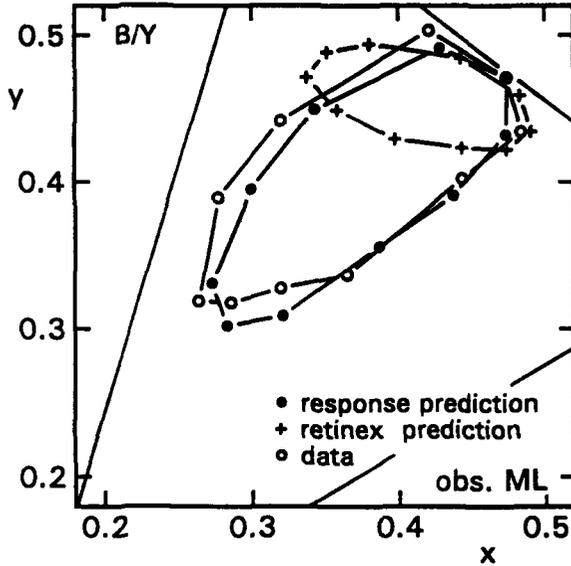


Figure 3.14: Experimental data as obtained for the B(lue)/Y(ellow) condition (open circles), compared to the predictions on the basis of the Retinex algorithm (crosses) and the response function (filled circles) described by eq. (3.14).

3.5 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 chromaticity coordinates (e.g. Arend *et al.*, 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. 3.6-3.8), 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 nonlinear response function (eq. (3.14)), in which both (sample) contrast and illumination level are the input variables. The inclusion of the

factor illumination in this function³ 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 eq. (3.14) 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. 3.8, 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 eq. (3.14) 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

³By 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).

the Retinex can be significantly improved upon by the introduction of what we called the “second factor”, the factor that 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. 3.13). 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. 3.13. For these panels the horizontal scale plots 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 plunged 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 eq. (3.14). 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 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 signalled by a system that records reflectance. Alternatively, a system that responds to contrast would not be able to do so, but instead, signal a change in color, 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, $\log(4.35 C^p)$ (with $C^p = Q_j^p/Q_w^p$), 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 $\log(4.35 C^p) = \log(C^p) + 0.64$. 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 *et al.*, 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 still essential factors missing in the response function reported here. This follows from the way in which the function was derived, *viz.*, by assuming that the symmetrical equation described by eq. (3.13), 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 eq. (3.13), 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 *et al.*, 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 already indicate 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.

3.6 Appendices

3.6.1 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 3.1. 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: $x_b=0.3130$ $y_b=0.3290$ $Y_b=12$ cd/m²,

and under green light:

sample : $x_j=?$ $y_j=?$ $Y_j=?$ cd/m²
 background: $x_b=0.3130$ $y_b=0.4320$ $Y_b=12$ 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 $Y_{j,R}$, etc, with subscript j or b referring to sample or background, respectively. This should result in:
 white light: $Y_{R,j}=2.42, Y_{G,j}=2.78, Y_{B,j}=0.80$

$$Y_{R,b}=2.85, Y_{G,b}=8.11, Y_{B,b}=1.04$$

green light: $Y_{R,b}=1.41, Y_{G,b}=10.11, Y_{B,b}=0.49$.

The latter luminance values are the ones shown in Table 3.2 (line 3).

2. Calculate the emission coefficients a_i ($i = R, G, B$), associated with the change in illuminant color, according to eq. (3.2), i.e. $a_R = Y_{R,ill}$ (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 3.2, line 3).

3. Calculate the coefficients b_i ($i = R, G, B$) according to eq.(3.3), i.e. $b_R = Y_{R,j}$ (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 3.1).
4. Calculate the luminance of the (simulated) reflected light according to eq. (3.1):
- $$Y_r = Y_R + Y_G + Y_B = a_R b_R Y_{R_w} + a_G b_G Y_{G_w} + a_B b_B Y_{B_w} =$$
- $$(0.49 \times 0.85 \times 2.85) + (1.25 \times 0.34 \times 8.11) + (0.47 \times 0.77 \times 1.04) =$$
- $$1.19 + 3.45 + 0.38 = 5.02 \text{ cd/m}^2.$$
- With eq. (3.20), 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:
- $$x_j=0.3214, \quad y_j=0.3455, \quad Y_j=5.02 \text{ cd/m}^2.$$

3.6.2 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)

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} x_R/y_R & x_G/y_G & x_B/y_B \\ 1 & 1 & 1 \\ z_R/y_R & z_G/y_G & z_B/y_B \end{pmatrix} \begin{pmatrix} Y_R \\ Y_G \\ Y_B \end{pmatrix} \quad (3.18)$$

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/m^2 . 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 eq. (3.18) becomes (matrix inversion)

$$\begin{pmatrix} Y_R \\ Y_G \\ Y_B \end{pmatrix} = \begin{pmatrix} 0.776 & -0.380 & -0.111 \\ -0.785 & 1.399 & 0.021 \\ 0.009 & -0.019 & 0.089 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}. \quad (3.19)$$

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

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 1.782 & 0.512 & 2.055 \\ 1 & 1 & 1 \\ 0.035 & 0.159 & 11.184 \end{pmatrix} \begin{pmatrix} Y_R \\ Y_G \\ Y_B \end{pmatrix}. \quad (3.20)$$

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

3.6.3 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

$$x' = \frac{1.0271 x - 0.00008 y - 0.00009}{0.03845 x + 0.01496 y + 1} \quad (3.21)$$

$$y' = \frac{0.00376 x + 1.0072 y + 0.00764}{0.03845 x + 0.01496 y + 1}. \quad (3.22)$$

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

$$\begin{pmatrix} L \\ M \\ S \end{pmatrix} = \begin{pmatrix} 0.0778 & 0.2722 & -0.0186 \\ -0.1562 & 0.4569 & 0.0297 \\ 0 & 0 & 0.3315 \end{pmatrix} \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix}. \quad (3.23)$$

Starting with known L,M,S units, the modified X', Y', Z' tristimulus values are computed with the inverse of eq. (3.23):

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} 5.8746 & -3.5001 & 0.6424 \\ 1.9948 & 1 & 0.0221 \\ 0 & 0 & 3.0169 \end{pmatrix} \begin{pmatrix} L \\ M \\ S \end{pmatrix}. \quad (3.24)$$

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 eq. (3.21) and eq. (3.22):

$$x = \frac{1.00709 x' + 0.00008 y' + 0.00009}{-0.03867 x' - 0.01537 y' + 1.03450} \quad (3.25)$$

$$y = \frac{-0.00347 x' + 1.02710 y' - 0.00785}{-0.03867 x' - 0.01537 y' + 1.03450} \quad (3.26)$$

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

3.6.4 D: A unit for “receptor input” (cd/m² 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 & Whitten (1972) for a discussion). The troland unit, and hence, the unit of luminance (cd/m²), 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 eq. (3.23). 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 eq. (3.24), by $Y' = 1.99 L + M + 0.02 S$. 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 eq. (3.23). 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 eq. (3.23) produces receptor inputs (Q_j) of $L=3.92$, $M=4.09$ and $S=4.15 \text{ cd/m}^2$.

3.6.5 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. 3.6-3.8) and an example of predicted data (Fig. 3.14). 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 eq. (3.20).

As for the predicted data, either on the basis of the Retinex model eq. (3.17) or the response function we derived eq. (3.15), 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 B, $Q_j^{p,m}$ is measured in cd/m^2 per receptor. This implies that the predicted L, M and S inputs can be transformed, using eq. (3.23), to X', Y' and Z' . The final step, the transformation from x', y' to x, y , is given by eq. (3.25) and eq. (3.26).

Chapter 4

Color Constancy under Natural and Artificial Illumination

Lucassen, M. P. & Walraven, J. (1992b). Submitted for publication in *Vision Research*.

Abstract

We studied color constancy under natural and extremely artificial illumination. Four test illuminants were used: two broad-band phases of daylight (correlated color temperatures 4000 and 25000 °K) and two spectrally impoverished metamers of these lights, each consisting of only two wavelengths. A computer controlled color monitor was used for reproducing the chromaticities and luminance of an array of Munsell color samples rendered under these illuminants. An asymmetric haploscopic matching paradigm was used in which the same stimulus pattern, either illuminated by one of the test illuminants, or by a standard broad-band daylight (D_{65}), was alternately presented to the left and right eye. Subjects adjusted the RGB settings of the samples seen under D_{65} (match condition), to match the appearance of the color samples seen under the test illuminant. The results show the (expected) failure of color constancy under two-wavelengths illumination, and approximate color constancy under natural illumination. Quantitative predictions of the results were made on the basis of both a computational approach to color constancy and a model that assumes the color response to be determined by cone-specific contrast and absolute level of stimulation (Lucassen & Walraven, 1992a). The latter model was found to provide the most accurate predictions under all illuminant conditions.

4.1 Introduction

In this paper we report experiments in which we compared the visual system's response to computer simulations of Munsell chips that are illuminated by either broad-band light, or light composed of only two wavelengths. This was done in the context of color constancy, the ability to perceive object colors as fairly stable, independent of the color of the illuminant. In recent studies of color constancy, it is customary to employ a more or less "natural" illuminant-object interaction, usually Munsell chips illuminated by incandescent light or different phases of daylight (e.g. Arend & Reeves, 1986; Arend *et al.*, 1991; Foster, Craven & Sale, 1992; Ho, Funt & Drew, 1990; Tiplitz-Blackwell & Buchsbaum, 1988a). The reason why we chose to also measure color constancy under extremely impoverished spectral conditions is twofold. First, we wanted to test the general applicability of an undoubtedly too simple (but accurate) model, derived in a preceding study (Lucassen & Walraven, 1992a). That model was based on data from a rather synthetic world, characterized by a trichromatic illuminant-object interaction commonly used in computer graphics (cf. Borges, 1991). The present study now provides "real world" data, both for natural and artificial lighting conditions.

Our second reason for doing these experiments is the lack of experimental research concerning the physiological relevance of recent computational approaches to color constancy (e.g. Brill & West, 1986; Buchsbaum, 1980; D'Zmura & Lennie, 1986; Forsyth, 1990; Maloney, 1986, 1992; Maloney & Wandell, 1986; van Trigt, 1990). These models typically aim at recovering the spectral information that is lost in the process of light absorption in the photopigments. This implies decomposing the light reflected from a surface, into its two constituent spectral distributions, i.e. the spectral power distribution of the illuminant and the reflectance function of the surface in question. The underlying principle relies on the spectral constraints that have been found to apply to our own natural environment. That is, it can be shown, by principal component analysis, that the phases of daylight can be adequately described by only three basis functions (Judd, MacAdam & Wyszecki, 1964). A similar simplicity underlies surface reflectance (Cohen, 1964), for which three basis functions may also account for most of the variance (Danne-miller, 1992; Maloney, 1986). Given the two sets of basis functions and knowledge (or an estimate) of the color of the illuminant, the latter can be eliminated (e.g. Buchsbaum, 1980), and hence, surface reflectance extracted. This is, in a nut-

shell, the rationale underlying computational approaches to color constancy. For a more detailed discussion, see the comprehensive introductions by D'Zmura and Lennie (1986) or Thompson, Palacios and Varela (1992).

For flat, homogeneously illuminated surfaces, and within the spectral constraints of naturally occurring surface reflectance functions and illuminant spectral power distributions, computational models can be quite successful in recovering surface reflectance. When these preconditions are not met, the models may be expected to fail. This would apply to our test condition in which the test samples are illuminated with a two-wavelengths light. However, these failures should be precisely predictable, for a given choice of model and illuminant-surface interaction (Maloney, 1992). Therefore, confronting such predictions with the findings from psychophysical experiments would seem a logical step in the validation of computational models. To our knowledge, this study is the first in that direction.

Although the primary goal of this study is to show the general applicability of our earlier data analysis (that is, without having to consider spectral constraints) we shall also present predictions that are representative for a computational approach to color constancy. The latter is based on one of the most recently developed algorithms for transforming a trichromatic input (XYZ) into a natural reflectance function (van Trigt, 1990). We found this model to predict near perfect color constancy under broad-band illumination, and, as expected, lack of constancy under the two-wavelengths illumination. These predictions were in qualitative agreement with the experimental findings, but they were not as accurate as those obtained with the much simpler model described in our earlier study (Lucassen & Walraven, 1992a). That model, which takes the form of a cone-specific response function, operates without any considerations regarding the spectral make-up of the visual scene.

As in most studies on color constancy we only address the purely sensory aspect of color vision. The subjects are asked to match the brightness, hue and saturation of samples seen under different illuminants. This task can be performed with good reproducibility (Lucassen & Walraven, 1992a) and requires no long training sessions. Other methods might have been used as well (see the Discussion), but since we wanted to extrapolate the results of our previous study, we decided to stick to the same method. As for the secondary purpose of this study, the comparison of different model predictions, there is no *a priori* reason why the choice of method for obtaining the data would favor one or the other model.

4.2 Methods

4.2.1 Surface reflectances

The spectral reflectance, $R(\lambda)$, of 36 samples from the Munsell Book of Color (glossy finish) were measured in the range $390 \leq \lambda \leq 730$ nm at 2-nm wavelength intervals with a SpectraScan PR-702AM spectroradiometer (Photo Research). The reflectances were measured relative to a BaSO₄ white, in the 0/45° measuring geometry. The CIE x, y chromaticities and luminance factor β (relative to white) of these samples under various illuminations, $E(\lambda)$, were computed by first calculating the X, Y, Z tristimulus values, using the numerical procedure:

$$X = \sum_{\lambda=390}^{730} E(\lambda)R(\lambda)\bar{x}(\lambda)\Delta\lambda \quad (4.1)$$

$$Y = \sum_{\lambda=390}^{730} E(\lambda)R(\lambda)\bar{y}(\lambda)\Delta\lambda \quad (4.2)$$

$$Z = \sum_{\lambda=390}^{730} E(\lambda)R(\lambda)\bar{z}(\lambda)\Delta\lambda \quad (4.3)$$

where $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ represent the CIE 1931 color matching functions and $\Delta\lambda=2$ nm. The colorimetric specifications of the 36 Munsell samples under illuminant C (the standard illuminant for viewing the Munsell Book) are listed in Table 4.1. We used 30 chromatic and six achromatic samples, presented as a 5×7 matrix, the same stimulus pattern as used in our earlier studies on color constancy and chromatic induction (Lucassen & Walraven, 1992a; Walraven *et al.*, 1991). The chromatic samples were selected from three loci of equal Munsell Chroma ($/6$, $/4$ and $/2$) at Munsell Value $5/$, the neutrals ranged from Value $2.5/$ through $7.0/$. The samples were presented on a neutral background ($N 7.0/$), resulting in a relative reflectance (sample to background) of 46 %. The numbers of the samples in Fig. 4.1 correspond to those in Table 4.1. Eleven samples (10 chromatic and one neutral), indicated by an asterisk in Table 4.1, were used as test stimuli.

4.2.2 Illuminants

Two classes of illuminants were simulated: three (natural) broad-band daylights and two (artificial) two-wavelength compositions. One of the broad-band illumi-

sample number in Fig. 4.1	simulated Munsell chip	x, y, β equivalents under illuminant C		
		x	y	β
1	10 YR 5/2	0.3546	0.3524	0.1977
2 *	5 PB 5/4	0.2662	0.2687	0.1977
3	10 G 5/2	0.2910	0.3310	0.1977
4 *	5 P 5/4	0.2986	0.2699	0.1977
5 *	5 G 5/4	0.2841	0.3628	0.1977
6	10 Y 5/2	0.3422	0.3648	0.1977
7 *	5 B 5/4	0.2493	0.2879	0.1977
8	10 R 5/6	0.4299	0.3499	0.1977
9	10 Y 5/6	0.4072	0.4621	0.1977
10	N 3.5/	0.3103	0.3163	0.0900
11	10 GY 5/2	0.3110	0.3508	0.1977
12	10 RP 5/6	0.3851	0.3039	0.1977
13	N 6.5/	0.3103	0.3163	0.3620
14	10 YR 5/6	0.4428	0.4128	0.1977
15	10 GY 5/6	0.3108	0.4301	0.1977
16 *	5 BG 5/4	0.2591	0.3246	0.1977
17	10 R 5/2	0.3465	0.3278	0.1977
18 *	N 5/	0.3103	0.3163	0.1977
19	10 BG 5/2	0.2796	0.3111	0.1977
20 *	5 R 5/4	0.3740	0.3220	0.1977
21	10 P 5/6	0.3243	0.2630	0.1977
22	10 B 5/6	0.2299	0.2548	0.1977
23	N 2.5/	0.3103	0.3163	0.0461
24	10 G 5/6	0.2519	0.3587	0.1977
25	10 P 5/2	0.3148	0.2986	0.1977
26	N 6.0/	0.3103	0.3163	0.3005
27	10 PB 5/6	0.2686	0.2412	0.1977
28	10 BG 5/6	0.2234	0.2952	0.1977
29 *	5 YR 5/4	0.3968	0.3614	0.1977
30	10 PB 5/2	0.2959	0.2905	0.1977
31 *	5 RP 5/4	0.3421	0.2954	0.1977
32 *	5 GY 5/4	0.3482	0.4097	0.1977
33	10 RP 5/2	0.3332	0.3131	0.1977
34 *	5 Y 5/4	0.3915	0.4057	0.1977
35	10 B 5/2	0.2821	0.2966	0.1977
36	N 7.0/	0.3103	0.3163	0.4306

Table 4.1: Munsell rennotations and CIE x, y, β equivalents (under white light) of the 30 chromatic and six achromatic samples of the stimulus shown in Fig. 4.1. β represents luminance reflectance relative to white. The 11 samples of the test set are indicated by an asterisk. These were always presented at the location of sample 18.

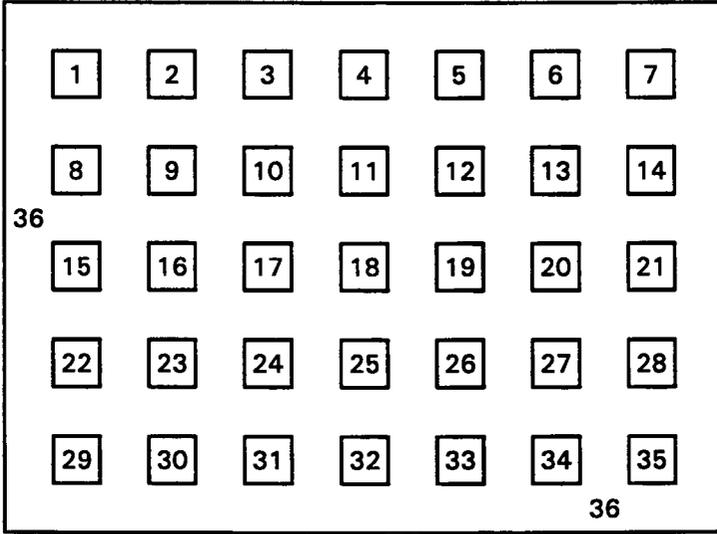


Figure 4.1: Stimulus geometry. The 1.3° squares are separated by a 1.3° square grid. Background (grid) dimensions: $19.5 \times 14.3^\circ$. See Table 4.1 for colorimetric specifications of the numbered samples.

nants (D_{65}) was used for illuminating the match (reference) pattern, the other four served as test illuminants for the test pattern.

The relative spectral radiant power distributions of the daylight illuminants were generated by the CIE method - derived from the principal components analysis of Judd *et al.* (1964) - as described in Wyszecki and Stiles (1982). This method takes as input the correlated color temperature (T_c) of a daylight illuminant D , where T_c may range from 4000 to 25000 °K. The output is a spectrum $E(\lambda)$, with λ in steps of 10 nm. In order to obtain the same spectral resolution as in the reflectance measurements (2 nm) we interpolated $E(\lambda)$ at 2-nm intervals.

In our simulation, the standard (white) illuminant D_{65} ($T_c=6500$ °K, $x=0.3127$, $y=0.3290$) was used for illuminating the match (reference) pattern. The two other daylight illuminants, D_{40} ($T_c=4000$ °K, $x=0.3823$, $y=0.3838$) and D_{250} ($T_c=25000$ °K, $x=0.2499$, $y=0.2548$), were used as test illuminants (two of the four). Strictly speaking, the CIE method for generating the spectral power distribution of daylight illuminants requires the illuminant's x -coordinate to satisfy $0.25 < x < 0.38$. The x -chromaticities of D_{40} and D_{250} (0.3823 and 0.2499) violate these boundary conditions, but the differences are so small that we may safely assume that this does not affect the reality aspect of our simulation.

The other two test illuminants, designated by M_1 and M_2 , were each composed of two wavelengths, λ_1 and λ_2 . For M_1 , $\lambda_1=592$ nm and $\lambda_2=491.8$ nm, and for M_2 , $\lambda_1=560$ nm and $\lambda_2=433.7$ nm¹. The relative intensities (power ratio $I_{\lambda_2}/I_{\lambda_1}$) of these wavelengths were 1.566 for M_1 and 1.254 for M_2 , so as to yield the same x, y chromaticities of illuminant M_1 and M_2 as for D_{40} and D_{250} . Thus, M_1 was metameric with D_{40} , and M_2 was metameric with D_{250} . The overall intensity of the (homogeneous) illuminants was equal to 30.4 cd/m², resulting in a luminance of the chromatic samples (Value 5/) of 6 cd/m² under D_{65} , consistent with our earlier studies.

The x, y, Y values of the Munsell samples under the broad-band illuminants, D_{40} and D_{250} , are those that can be found when viewing the Munsell Book in outdoor illumination (ignoring atmospheric effects etc.), that is, they are realistic (natural) values. The x, y chromaticities of the Munsell samples rendered under the two-wavelengths lights, M_1 and M_2 , fall on the lines that connect the corresponding wavelengths in CIE x, y chromaticity space. Although rather unnatural, such stimuli are physically realizable in the laboratory by using laser light, interference filters or monochromators.

4.2.3 Stimulus presentation

The x, y, Y equivalents of the samples under the various illuminants were displayed on a calibrated high resolution color monitor (Sony, 1152×900 pixels) that was controlled by a Sun 3/260 computer (24 bit/color). For the human eye, the video RGB metamers are physically indistinguishable (as far as color is concerned) from their paper counterparts. The calibration procedure for the monitor, and the colorimetric equations required for displaying specified x, y, Y values on a color monitor, have been published elsewhere (Lucassen & Walraven, 1990).

In each experimental condition, two displays were used: a test pattern, i.e. the samples as arranged in Fig. 4.1 under one of the test illuminants D_{40} , D_{250} , M_1 or M_2 , and a match pattern of identical geometry, illuminated by D_{65} . A pyramidal box (1 m length) with two viewing holes was placed in front of the monitor. A mechanical shutter system, located just behind the two viewing holes, alternately occluded the left and right viewing hole. In this way, each eye was locked to one or the other of the two successive illuminant conditions (test or match) to be compared. The colors of the test and match pattern were changed during the switching time of the shutters, which only took a fraction of a second. The

¹In order to compute X, Y, Z tristimulus values according to eqs. (4.1-4.3) we interpolated the color matching functions and the reflectance spectra at 0.1 nm steps and used $\Delta\lambda=0.1$ nm for these two illuminants.

presentation time of each pattern was set at five seconds. This was long enough for the stimulus to “settle” (at these relatively low light levels) and short enough for not disrupting the comparison of test and match sample.

4.2.4 Procedure

After about five minutes of dark adaptation and a few more minutes for adapting to the average luminance and color of the test pattern, the observer started the first presentation of the two alternating illuminant conditions. When viewing the test (left eye) and match pattern (right eye) the observer concentrated on the central patch. The color of the matching sample, which was initially black, was under mouse control. The movements of the mouse were interpreted by the computer as movements through CIE x, y color space. Two of the three mouse buttons were pressed for increasing or decreasing the luminance of the patch at constant x, y chromaticities. The third mouse button was pressed for indicating that a satisfactory match had been obtained, after which the next test patch was presented (in total 11 samples, in pseudo-random order).

Even for unexperienced subjects, this matching procedure was easy to comprehend and required only a few training sessions for obtaining reliable results. The first author (ML), and two naive observers (AV and EG), all with normal color vision, served as subjects.

4.2.5 Task

The observers adjusted the central patch in the match pattern to make it match the perceived hue, saturation and brightness of the corresponding sample in the test pattern. They were free to make eye movements and to use as many test/match alternations as were necessary to obtain a satisfactory match. None of the observers reported that it was impossible to find a match.

4.3 Results

Each subject made 11 color matches in each of the four test illuminant conditions which we shall abbreviate as D_{40}/D_{65} , D_{250}/D_{65} , M_1/D_{65} and M_2/D_{65} , thus indicating the test/match illuminant combination. In Figs. 4.2-4.4, the matches for the separate observers, are plotted in the CIE x, y diagram. The top panels relate to the two conditions with the broad-band daylight illuminants, the bottom panels to the conditions with the two-wavelengths illuminants. The straight lines in the plots represent the boundaries of the triangular color space covered by the

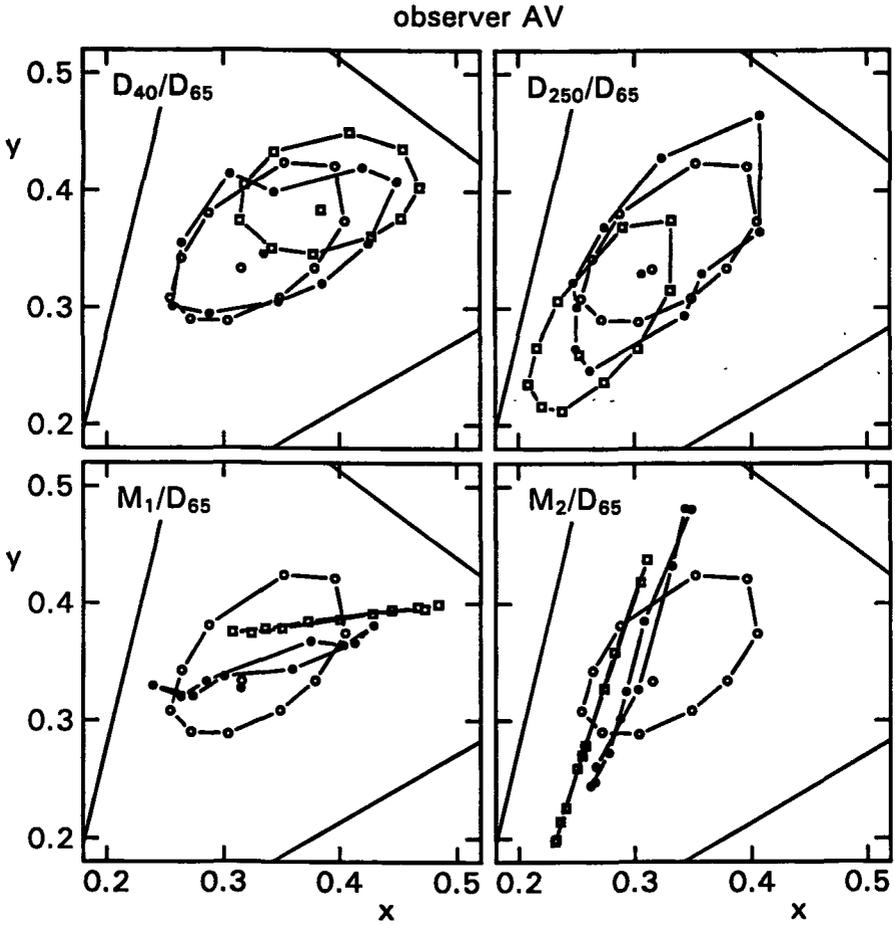


Figure 4.2: Experimental results with the daylight illuminants (top panels) and the two-wavelengths illuminants (bottom panels), for observer AV. Open circles: chromaticities of the test samples under D_{65} . Open squares: chromaticities of the test samples under the test illuminant in question. Closed circles: chromaticities of the observer's matches (under D_{65}) to the test samples under the test illuminant.

phosphors of our CRT. Open squares represent the chromaticities of the 11 test samples under the test illuminant, open circles those under the (D_{65}) match illuminant. When comparing top and bottom panels, note the difference in the

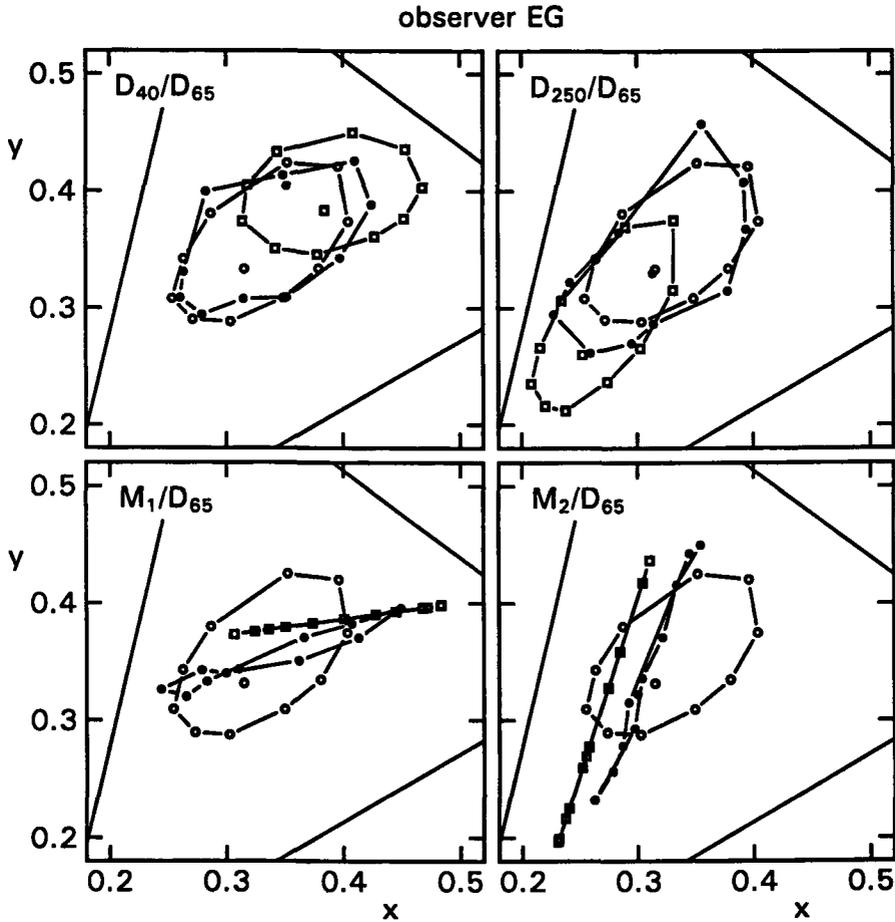


Figure 4.3: Same as Fig. 4.2, but now for observer EG.

chromaticities of the colors rendered under the test illuminant. The closed circles indicate the chromaticities of the observer's matches to the test samples.

The effect of changing the illuminant from broad-band (upper panel) to two-wavelengths illuminants (lower panel), is to collapse the chromaticity locus of the test colors onto a straight line (see the open squares). This is the line connecting the chromaticities of the two wavelengths of the M_1 or M_2 light source. Under monochromatic light, all chromaticities would project onto a single point, the condition that one may encounter under sodium (street) lighting.

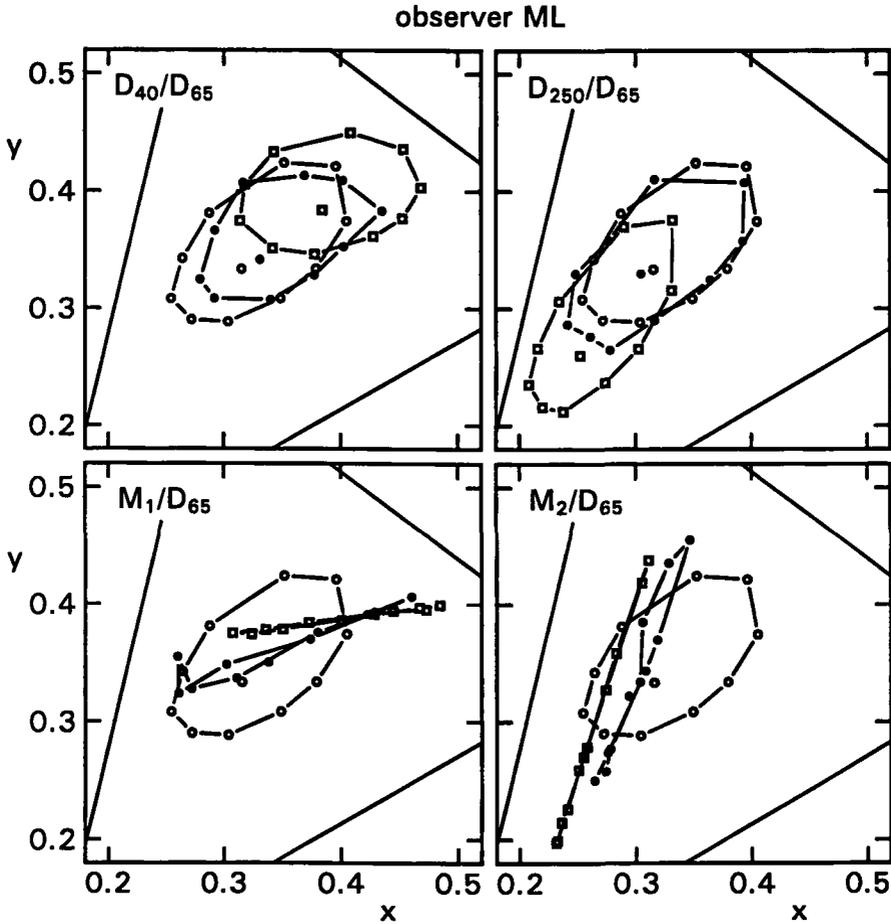


Figure 4.4: Same as Fig. 4.2, but now for observer ML.

In Figs. 4.2-4.4 perfect color constancy would have been indicated by coinciding closed and open circles, but this is never the case. As expected, the deviations from perfect constancy are smaller for the daylight illuminants than for the two-wavelengths illuminants. There is a general tendency for the neutral test sample to be shifted back in the direction of the chromaticities of the neutral sample under white light (the closed circle in the middle). Such a shift is in accordance with an incomplete von Kries color transformation scheme (von Kries, 1905). The chromatic test samples are shifted back in the same direction, but

for conditions M_1/D_{65} and M_2/D_{65} the loss of the original chromaticity spacing cannot be undone. For all observers, the color matches fall on a single line (within experimental spread). That line is translated (and rotated for condition M_1/D_{65}), away from the line that connects the physical chromaticities under the test illuminant. In the following we show that both the moderate and gross violations from perfect color constancy are adequately predicted by the simple contrast-based model, derived in our previous study.

4.4 Data Predictions

In this section we present predictions of the experimental data on the basis of the model that we derived in our previous study (Lucassen & Walraven, 1992a). In addition we shall present the predictions that can be obtained by using the quite different approach of recovering surface reflectance. The latter we shall refer to as the “computational model” whereas our own model will be referred to as “response function”.

4.4.1 Response function

The results from our earlier study, which were obtained with the same stimulus configuration (but situated in an “RGB world”), could be described by the response function

$$R^p = (Q_w^p)^r \log(4.35 \frac{Q_j^p}{Q_w^p}) \quad p = L, M, S, \quad (4.4)$$

where Q^p represents the quantum catch per cone class, as denoted by the superscript p . Additional subscripts j and w , indicate the input from test sample and (white) background, respectively. The spectral reflectance of the latter is flat, so the background conveys the chromaticity of the illuminant. The exponent r is observer dependent ($r \approx 0.3$). The response function presented in eq. (6.1) has to be applied to both the test and match eye. The prediction of the match, in terms of cone inputs (Q_j^p), is obtained by equating the test and match eyes’ responses, $R^{p,t} = R^{p,m}$, where superscripts t and m denote test and match, respectively. These superscripts have to be applied to each element in eq. (6.1). The cone inputs required for the matching sample, $Q_j^{p,m}$, can be computed by substitution of eq. (6.1) into $R^{p,t} = R^{p,m}$. One can thus derive

$$\log(Q_j^{p,m}) = \left(\frac{Q_w^{p,t}}{Q_w^{p,m}} \right)^r \log \left(4.35 \frac{Q_j^{p,t}}{Q_w^{p,t}} \right) + \log \left(\frac{Q_w^{p,m}}{4.35} \right). \quad (4.5)$$

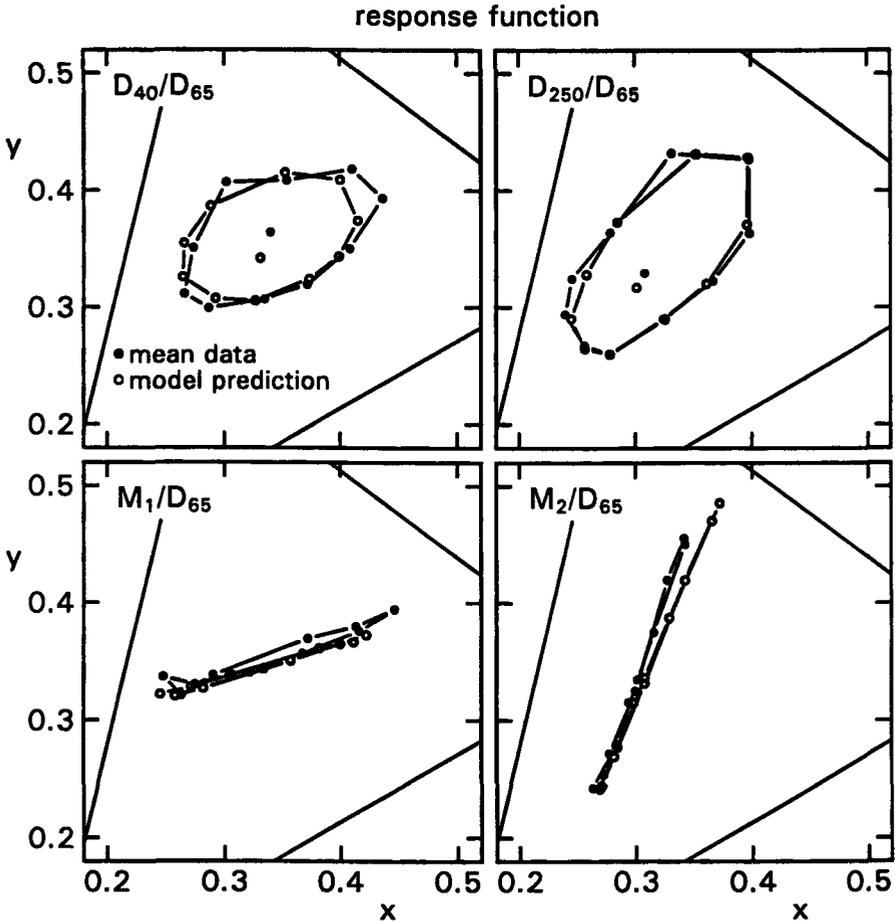


Figure 4.5: Mean observer matches (closed circles) and predictions (open circles) based on the response function derived by Lucassen & Walraven (1992a). Predictions computed with eq. (4.5) for $r=0.33$.

The predictions that are obtained by applying eq. (4.5) to the data, have been cast into terms of CIE x, y chromaticity coordinates (see Lucassen & Walraven (1992a) for details). In Fig. 4.5 these predicted chromaticities (open circles) are shown together with the experimentally obtained values averaged over the three observers (closed circles). It is clear from Fig. 4.5 that eq. (4.5) provides a good description of the results. We computed an average chromatic difference, $\bar{\Delta}_{xy}$,

between the predicted and experimentally obtained chromaticities according to

$$\bar{\Delta}_{xy} = \frac{1}{11} \sum_{i=1}^{11} \left((x_{pred,i} - x_{exp,i})^2 + (y_{pred,i} - y_{exp,i})^2 \right)^{1/2}. \quad (4.6)$$

The values of $\bar{\Delta}_{xy}$ for conditions D_{40}/D_{65} , D_{250}/D_{65} , M_1/D_{65} and M_2/D_{65} so computed were 0.0114, 0.0131, 0.0161 and 0.0204, respectively. Such small values, associated with about 95% explained variance of the data, were also obtained in our earlier study, for a much wider range of illuminants (covering the complete hue circle).

Actually, we are not concerned here with the predictive power of eq. (4.5). The important result is that the data from the daylight illuminants are about equally well predicted as the data from the two-wavelengths illuminants. Note that the only information required for these predictions is that provided by the cone inputs from sample and surround. There are no spectral constraints on either the reflectance or illuminant spectra involved.

4.4.2 Computational model

Computational models of color constancy typically aim at recovering the spectral reflectance function of a surface, on the basis of trichromatic information only. The latter should be representative for natural, that is, a smoothly varying function. The computational model we used is based on the algorithm derived by van Trigt (1990) for generating mathematically smoothest reflectance functions. This model requires the illuminant's spectral power distribution, $E(\lambda)$, to be known, and further needs the X, Y, Z tristimulus values of the sample under that specific illuminant. It generates a reflectance function, $R(\lambda)$ (which can be written as a sum of three elementary functions), that is spectrally the smoothest according to the criterion that the square of the reflectance function's derivative, integrated over the visual range, is minimal. $R(\lambda)$ is a solution to the following equations:

$$0 \leq R(\lambda) \leq 1 \quad 390 \leq \lambda \leq 730 \quad (4.7)$$

$$X = \int E(\lambda)R(\lambda)\bar{x}(\lambda)d\lambda \quad (4.8)$$

$$Y = \int E(\lambda)R(\lambda)\bar{y}(\lambda)d\lambda \quad (4.9)$$

$$Z = \int E(\lambda)R(\lambda)\bar{z}(\lambda)d\lambda \quad (4.10)$$

$$\int \left(\frac{dR}{d\lambda} \right)^2 d\lambda = \text{minimal} \quad (4.11)$$

where $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ are the CIE 1931 color matching functions. Here we shall not discuss the various steps involved in the computation of the reflectance functions; they are fully detailed in van Trigt (1990).

When applied to the Munsell chips that we used under the two daylight illuminants D_{40} and D_{250} , the spectral reflectances that were generated by van Trigt's algorithm globally resemble those that we measured with the spectroradiometer. As an example, the smoothest reflectance functions under D_{40} and D_{250} for one of the test samples (No. 2 in Table 4.1) are shown in Fig. 4.6, together with the reflectance function that we measured. The results for the other test samples are

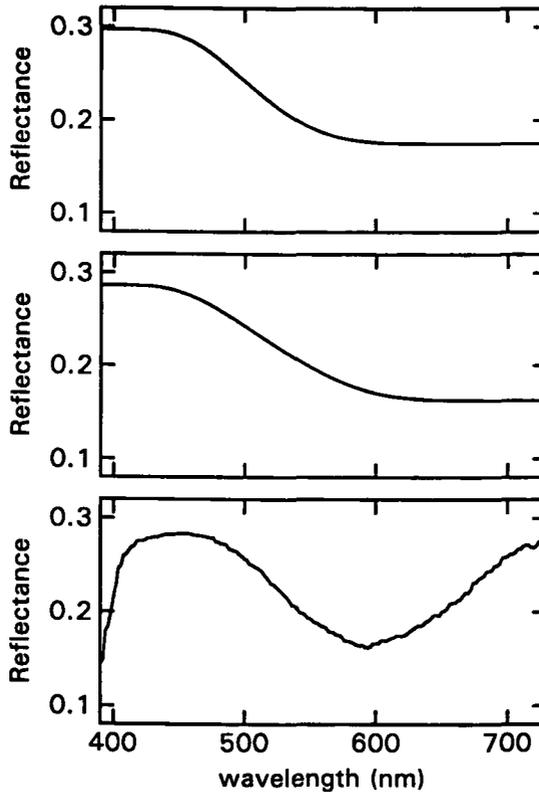


Figure 4.6: Spectral reflectance of the 5 PB 5/4 Munsell sample (No. 2 in Table 4.1) as measured with the spectroradiometer (bottom) and metameric smoothest reflectance function generated by the van Trigt (1990) algorithm under D_{40} (top) and D_{250} (middle) illumination.

similar. In the range from 450 to 600 nm the resemblance is best, which is also the most important spectral part for the human visual system (the cone sensitivities at the spectral boundaries of Fig. 4.6 are effectively zero). Actually, it is not important whether the reflectance function that is generated by the algorithm closely resembles the real reflectance function of the object under consideration. It should only have the property that it remains a metamer of that sample, over the range of illuminants for which it is supposed to exhibit color constancy (like the phases of daylight, for example). The smooth reflectance functions generated by the van Trigt algorithm perform very well in that respect, as follows from the extensive study of Troost and de Weert (1992). They tested this for a set of 2734 Munsell samples under nine different phases of daylight, and found that the color rendering performance of the van Trigt model is comparable to that of a linear model that employs Cohen's (1964) spectral basis functions, the fundamental basis for many computational approaches to color constancy (see Introduction for references). We chose van Trigt's (1990) model because it puts minimal constraints on the input (tristimulus values of objects as defined under any illuminant), and also because it is easy to implement in software.

Impressed by the fact that the recovered reflectance functions under D_{40} and D_{250} were quite color constant (see upper panels in Fig. 4.6), we were interested in finding out how well this model would predict observer responses under our two-wavelengths illuminants M_1 and M_2 . Since the illuminant's spectral power distribution has to be known, the question arises how this information becomes available. That is, if this model is implemented in a machine vision system equipped with color sensors that mimic the spectral sensitivities of the human cones, how would it be capable of recovering the illuminant spectrum from only the tristimulus values (or, the linearly related cone responses) of the stimulus? This problem, estimating the illuminant, is the central issue in computational models of color constancy. We shall not go into the various solutions proposed for this problem. It may be sufficient to mention that it can be shown that under the *grey world assumption* (Buchsbaum, 1980), the illuminant can be estimated quite successfully from the average chromaticities in the visual scene. Since our stimulus pattern should closely obey the grey world assumption (the samples are regularly spaced over the spectrum, and the grey background occupies about 80% of the stimulus area), we may assume that the color of the illuminant (its tristimulus values), and hence, the associated spectrum, can be made available to the model. That is what we did, but the model could not "know" of course, that, in the case of the M_1 and M_2 artificial illuminants, the XYZ values were not associated with real (broad-band) daylight spectra. So, the computational model may be expected to predict, then, the surface reflectances of samples that yield the same XYZ values under broad-band light, as do the original samples

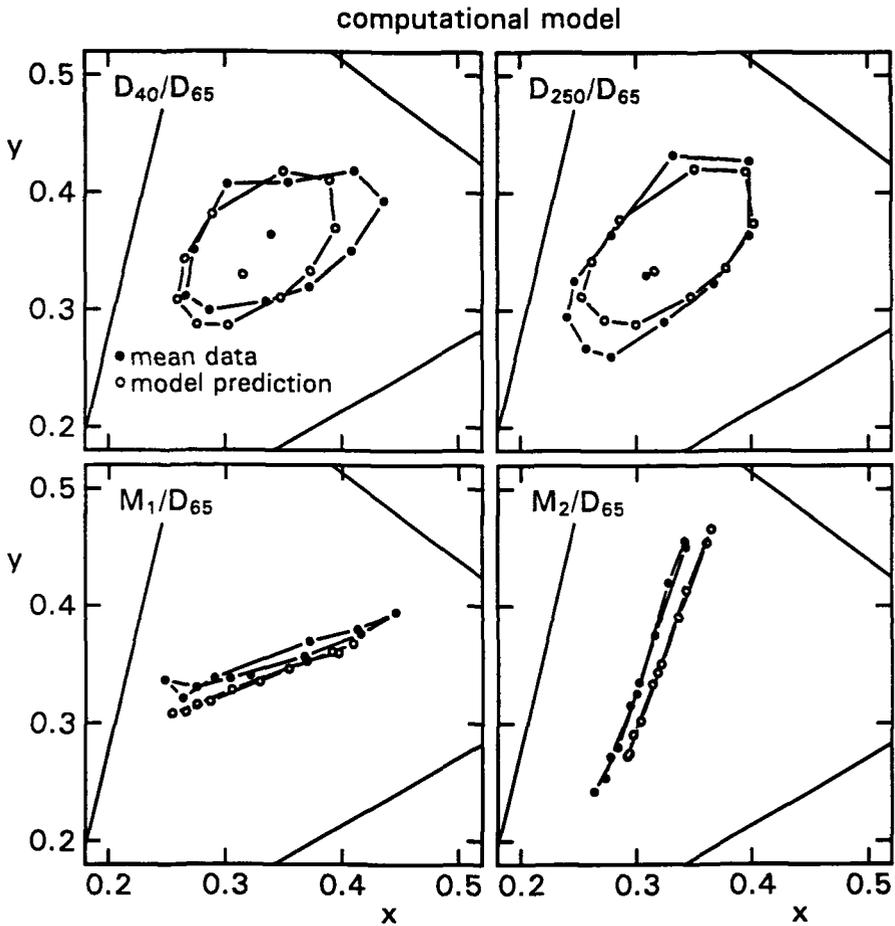


Figure 4.7: Mean observer matches (closed circles) and predictions (open circles) of a computational model based on the van Trigt (1990) algorithm for generating smoothest reflectance functions.

under two-wavelengths illumination.

Our computational predictions are again in terms of the x, y chromaticities of the matching sample under the D_{65} reference illuminant. They are computed by

the X, Y, Z tristimulus values defined by

$$X_{pred} = \sum_{\lambda=390}^{730} D_{65}(\lambda)R(\lambda)\bar{x}(\lambda)\Delta\lambda \quad (4.12)$$

$$Y_{pred} = \sum_{\lambda=390}^{730} D_{65}(\lambda)R(\lambda)\bar{y}(\lambda)\Delta\lambda \quad (4.13)$$

$$Z_{pred} = \sum_{\lambda=390}^{730} D_{65}(\lambda)R(\lambda)\bar{z}(\lambda)\Delta\lambda \quad (4.14)$$

where $R(\lambda)$ represents the smoothest reflectance function for the test illuminant and sample in question. The results are shown in Fig. 4.7. Figure 4.7 shows that the predictions of the computational model are less accurate than those obtained with our response function (Fig. 4.5), in particular for the natural conditions D_{40}/D_{65} and D_{250}/D_{65} . It also shows that under the two-wavelengths illuminants (lower panels in Fig. 4.7) the model is not capable of recovering the original spectra that define the Munsell Chroma /4 circle under D_{65} . In that respect, it is interesting to see that it makes the same kind of “mistake” as human observers. The values for the chromatic difference (the prediction error), as defined in eq. (4.6), are 0.0233, 0.0239, 0.0245 and 0.0286 (in the same sequence as presented before), that is, about twice as large as those obtained when using our response function. This does not mean that the computational model performs less in terms of color constancy. It actually performs better than the visual system, but that is exactly the reason why it is a less representative model for the way in which the visual system operates.

4.5 Discussion

Previous studies on color constancy typically employed illuminant conditions that were chosen to demonstrate the efficacy of the effect. The present study deviates in this respect by also including spectrally impoverished illuminants, the two-wavelengths metamers of D_{40} and D_{250} (M_1 and M_2). By doing so we were able to measure the deterioration of color constancy, that is specifically due to the lack of spectral “capacity” of the illuminant. The particular way in which color constancy breaks down under these conditions, is informative as to how spectral information is processed by the visual system. This information can be used for developing or testing models of color constancy, which is what we did. We thereby focussed on the spectral variable, for the reason that this has become an important consideration in current computational models of color constancy.

An important finding of our study is that despite the entirely different spectral power distributions of the illuminants, the associated differences in degree of color constancy can be adequately accounted for by a relatively simple model. No spectral considerations other than necessary for computing the visual input, enter into this model. This result would seem to argue against the much more complex computational approach, that aims at complete recovering of surface reflectance.

Our results also show, for the first time, the performance of a computational model under both natural and unnatural illuminant conditions. This model takes the tristimulus values of reflected light as input and then generates the smoothest (i.e. natural) reflectance function that, in conjunction with the illuminant spectrum, reproduces that input (van Trigt, 1990). We found that even under extreme phases of daylight (D_{40} and D_{250}), the model is capable of near perfect color constancy. For the asymmetric matching paradigm of our experiment, this amounts to match settings (under D_{65}) that are identical to the symmetric match (i.e. the condition with both eyes viewing the samples under D_{65}). This can be checked by comparing the computational matches (Fig. 4.7, open circles) with the chromaticities of the test samples under D_{65} (Figs. 4.2-4.4, open circles).

The good color constancy performance of the computational model we tested is actually the reason for its relatively less accurate predictions of what human observers see. Other computational models may face the same problem, since these too should exhibit near perfect color constancy (e.g. Brainard & Wandell, 1991). The reason for that is the fact that naturally occurring surface reflectances and illuminant spectra can be accurately described by linear combinations of a small number of basis-functions (Cohen, 1964; Judd *et al.*, 1964; Parkkinen, Hallikainen & Jaaskelainen, 1989).

It is conceivable that the visual system would be capable of better color constancy - thus confirming the predictions of computational models - when measured under more appropriate conditions and/or with better methods. As for our "normal" (broad-band) stimulus conditions, these are indeed somewhat synthetic, in the sense that the visual scene lacks a third dimension (no shadows and shading), and that the appearance of the samples is consistent with perfectly diffusing surfaces under a spatially uniform illumination. However, these are exactly some of the most important constraints - see Forsyth (1990) for a complete list - that have to be met when applying the present generation of computational models.

As for methodology, there are indeed different ways for measuring color constancy. One could argue that testing for invariance of the purely sensory aspect of color perception (hue, saturation, lightness) is not necessarily the best approach. A possible alternative is to test for the correct recognition (rather than perception) of surface samples, thereby ignoring possible deviations from sensory

invariance. This method was introduced by Arend and Reeves (1986), who asked their subjects to adjust the color of a match sample as if “cut from the same paper” as the test sample. The subject is thereby instructed to take into consideration that the samples are shown under different illuminants and thus may not necessarily appear as having the same color. Subjects are apparently able to follow that instruction, and do so on the basis of just cognitive skills (rather than using contextual cues), as follows from the fact that the same result can be obtained with a simple disk-annulus stimulus configuration (Arend *et al.*, 1991). In spite of relaxing the definition of color constancy (sensory invariance is no longer required) the latter studies did not achieve more than moderate color constancy, i.e. 60% in terms of a chromatic index analogous to the Brunswik ratio (Arend *et al.*, 1991). Other studies also show that incomplete color constancy is the rule rather than the exception (Reeves, Arend & Schirillo, 1989; Tiplitz-Blackwell & Buchsbaum, 1988a; Valberg & Lange-Malecki, 1990). So far, therefore, the perfect color constancy that computational models are capable of, does not seem to be paralleled by the human visual system.

The computational model was quite successful at predicting color constancy (or rather its failure) under artificial illumination. This is probably true for other models as well, since most of that prediction simply follows from physical constraints. It is not surprising that color constancy fails in conditions where color is rendered by only two wavelengths. Consider what would happen under monochromatic illumination. It does not take a sophisticated model to predict, then, that color discrimination will be completely impossible.

Taken altogether, the results of our study do not support the idea that the computational approach to color constancy is implemented in the visual system. This would square with the lack of neuro-physiological evidence for structures performing the estimates of surface reflectance required by computational models (D’Zmura & Lennie, 1986; Troost & de Weert, 1991a). We feel that our results are still too limited for rejecting the physiological relevance of computational models, but they do seem to warn against simply accepting such models without experimental evidence and/or considering alternative approaches.

Our own model of color constancy is basically a trichromatic extension of contrast or lightness models (e.g. Hurlbert, 1986). It thus resembles the Retinex model (Land, 1986; McCann, McKee & Taylor, 1976), but with some important modifications as discussed in Lucassen and Walraven (1992a). The simple model we used for describing the data may be only valid for our particular experiment. Still, its essential feature, responding to contrast - the key to visual constancy in a world where luminance varies over more than 10 decades - is consistent with the results from other psychophysical studies on invariant (a)chromatic vision (e.g. Arend & Goldstein, 1990; Shapley, 1986; Wallach, 1948; Walraven *et al.*,

1991). A system that responds to contrast can be implemented by a resetting mechanism or automatic gain control (Koenderink, van de Grind & Bouman, 1971; Rushton, 1965; Walraven & Valetton, 1984). Such a mechanism has also the effect of removing a steady-state signal, for which there is also psychophysical evidence (Tiplitz-Blackwell & Buchsbaum, 1988a; Walraven, 1976; Whittle & Challands, 1969).

As discussed elsewhere (Shapley *et al.*, 1990; Walraven *et al.*, 1990), the importance of contrast can also be demonstrated at the physiological level (e.g. Enroth-Cugell & Robson, 1966; Reid & Shapley, 1988; Shapley & Enroth-Cugell, 1984). As for our model's assumed cone-specificity of the contrast response, one may expect this to be reflected in receptive fields driven by single cone classes. Physiological evidence for this notion is available, but does not always exclude other interpretations. However, recently Reid and Shapley (1992) have provided unambiguous evidence for cone-specific inputs in both centers and surrounds of parvocellular neurons.

In conclusion, we have studied color constancy in both favorable and adverse illuminant conditions for demonstrating the effect. The data show that the associated variation in the degree of color constancy can be better accounted for by a mechanism that responds to cone-specific contrast, than by a computational model that recovers illuminant and surface reflectance spectra.

Chapter 5

Separate Processing of Chromatic and Achromatic Contrast in Color Constancy

Lucassen, M. P. & Walraven, J. (1992c). Submitted for publication in *Vision Research*.

Part of the work described in this Chapter was presented at the annual meeting of the Association for Research in Vision and Ophthalmology (May 3-8, 1992, Sarasota, Florida, USA).

Abstract

In a preceding study (Lucassen & Walraven, 1992a) we measured color constancy in experimental conditions in which (simulated) illuminants and surface colors were varied in the chromatic domain only. That is, both illumination level and sample reflectance were fixed. In the present study we focus on the achromatic dimension, both with respect to luminance contrast (Experiment 1), and overall illumination (Experiment 2). Sample/background contrast was varied over a two log-unit range, covering both luminance decrements and increments. Illumination level was varied either for the short-wave-sensitive (S) cones only, or for all three cone types simultaneously. Data predictions on the basis of a cone-specific response function, derived in our preceding study, indicate that this model has difficulty in accommodating the results obtained with varying luminance contrast. However, a modified version of the response function, incorporating separate processing of color and luminance contrast, correctly predicts the data from both the present and the previous study. We also show that, over a limited stimulus range, our earlier response function is mathematically equivalent to Jameson and Hurvich's (1964) model of brightness contrast. The latter model, cast into a trichromatic format, performs equally well or better than our original response function, but is less accurate than our modified model.

5.1 Introduction

Reflected light conveys information about changes in both surface reflectance and ambient illumination. How the visual system keeps track of these two variables is a problem typically treated in the context of color constancy. However, the problem is not confined to the chromatic domain. It has also to be solved in monochromatic vision or by trichromats that are forced to operate in the monochromatic mode. These conditions can be encountered in daily life, like, for example, under illumination by sodium light, seeing in the dark (scotopic vision), or when watching black-and-white TV or movie pictures. Despite the lack of spectral information under these circumstances, the visual system may nevertheless still be capable of recovering (achromatic) reflectance, as is also evidenced by studies on lightness constancy (e.g. Arend & Goldstein, 1987; Gilchrist, 1988; Jacobsen & Gilchrist, 1988).

The close functional connection between color constancy and lightness constancy invites a theoretical approach in which color constancy is treated as the trichromatic extension of lightness constancy (cf. Hurlbert, 1986). That is, the processing of lightness within three cone-specific channels, as is also the basic principle underlying the well-known Retinex model (e.g. Land, 1964, 1986). There is evidence showing that (local) contrast, rather than lightness, is the stimulus variable of interest (e.g. Fairchild & Lennie, 1992; Shapley, 1986; Walraven *et al.*, 1991), but that does not affect the rationale of the trichromatic lightness approach.

We have shown (Lucassen & Walraven, 1992a) that color constancy involves more than just the processing of (trichromatic) contrast. Contrary to the predictions of contrast or lightness models, we found an effect of the *absolute* level of (cone-specific) illumination. This effect, which has also been noticed in the study of McCann, McKee & Taylor (1976), could be quantified by assuming a cone-specific "response function"

$$R = (Q_w)^r \log \left(k \frac{Q_j}{Q_w} \right) \quad (5.1)$$

where Q_j/Q_w represents the cone-specific contrast of samples (j) relative to a white background (w). The term $(Q_w)^r$, which varies in proportion with the illuminant, makes R dependent on absolute light level. With $r=0.33$ and $k=4.35$, eq. (5.1) provides an accurate description (95% explained variance) of data obtained over a wide range of colored illuminants (Lucassen & Walraven, 1992a). It

could also be shown that eq. (5.1) correctly predicts the break-down of color constancy in conditions where the illuminant is spectrally impoverished (Lucassen & Walraven, 1992b).

The data from which eq. (5.1) was derived came from experiments in which color was the main stimulus variable of interest. That is, the test illuminants and test samples varied in chromaticity, but not in luminance (except for variations due to the spectral interaction between samples and illuminant). Variations in cone contrast (Q_j/Q_w) could thus only be produced by spectral modulation of illuminants and sample reflectances.

In the present study we address the question whether eq. (5.1) may also accommodate the results from experiments in which cone contrast is modulated by luminance rather than color. Thereto we presented the same set of test samples (Q_j) on backgrounds (Q_w) that varied in reflectance. We also did the reverse experiment, by varying sample reflectance while keeping background reflectance fixed. Having already extensively studied the effect of illuminant color, we now only compared the samples under a fixed pair of illuminants, blue and yellow. This combination, which strongly modulates the short-wave-sensitive (S) cones, had been found to be most critical for testing the validity of eq. (5.1).

In addition to studying the effect of contrast (Experiment 1) we also measured the effect of overall illumination level (Experiment 2). We did this for a single sample contrast (50% under white light), and used the same blue and yellow lights as in Experiment 1. In Experiment 2, however, these lights were not necessarily equi-luminant, but could be varied, so as to produce sizable inter-ocular differences in (cone-specific) background luminance Q_w .

The experimental results we obtained turned out to be difficult to account for on the basis of eq. (5.1). We therefore started searching for alternative ways of quantifying the data, thereby going back to some of the earlier studies on brightness contrast. We found, that a trichromatic implementation of Jameson and Hurvich's (1964) model for brightness contrast provided a reasonable good description of our data. This was quite unexpected, considering the quite different nature of the stimulus transformations involved. We shall show, however, that, over a limited stimulus range, these functions are mathematically equivalent. Beyond that range, the equivalence no longer holds, which is the reason why the Jameson and Hurvich model compares favorably to eq. (5.1) when it comes to predicting the results for the wide range of luminance contrasts, employed in our present study.

Although the Jameson and Hurvich model (later to be referred to as JH model) performs better than eq. (5.1), it still requires a rather complex adjustment of its adaptation constant k (to be discussed). We therefore decided to give eq. (5.1) a second chance, and searched for possible modifications. As a result we arrived at a normalized form of eq. (5.1), that resulted in a simpler and more accurate description of the data. The normalization takes place in the luminance domain, that is, dividing the (cone) input variables, Q_j and Q_w , of the response function eq. (5.1) by the associated luminance values (Y_j and Y_w). This implies deriving a luminance-free, i.e. purely chromatic, cone-specific contrast signal. Such a step would seem quite natural in view of the growing evidence for separate chromatic and achromatic postreceptoral mechanisms (e.g. Creutzfeldt, Lee & Elepfandt, 1979; Kaplan & Shapley, 1986; de Monasterio & Gouras, 1975; Wiesel & Hubel, 1966). The role of the luminance variable in color constancy (and color vision in general) has always been a difficult issue. We therefore decided not too lightly dismiss other approaches than our own, and consequently present the data together with the predictions derived from both the modified version of eq. (5.1) and those computed with the trichromatic version of the Jameson and Hurvich (1964) model.

5.2 Methods

Since the methods were essentially the same as in a preceding study (Lucassen & Walraven, 1992a), we here present only a brief summary of the general aspects that apply to both Experiment 1 and 2. Specific details regarding illuminants and stimulus configurations are discussed separately for Experiments 1 and 2.

5.2.1 General

Equipment

The stimuli were generated on a daily calibrated Sony high-resolution color monitor (8 bit luminance resolution per gun), controlled by a Sun 3/260 computer. A simple recalibration algorithm was used that enabled color reproduction within an average error of 0.005 CIE x, y units (Lucassen & Walraven, 1990). Manipulation of color was under mouse control. Movements of the mouse were interpreted by the computer as movements in CIE x, y space. Two of the three mouse buttons were used for increasing or decreasing luminance; the third mouse button

was pressed for indicating that a satisfactory match was obtained (of match to test sample), after which a new test sample was presented. A large box-shaped hood, fitted to the front of the display, restricted the field of view to the stimulus pattern, which, at the observation distance of 1 meter, subtended a visual angle of $14.3 \times 19.5^\circ$. The two viewing holes in the box could be alternately opened and closed by mechanical shutters, that operated in synchrony with the presentation of the “match” and “test” pattern. In this way, the two patterns were successively seen, each locked to a different eye.

Simulation of object-illuminant interaction

We used the same trichromatic reflection paradigm as in preceding studies (Lucassen & Walraven, 1992a; Walraven *et al.*, 1991). It describes reflected light, L_r , in terms of phosphor primary luminances (Y_R , Y_G , Y_B) that are independently modulated by surface-specific reflection coefficients (a_R , a_G , a_B) and illuminant-specific emission coefficients (b_R , b_G , b_B). So, for any sample-illuminant combination, the light reflected from the sample is given by

$$L_r = a_R b_R Y_{R_w} + a_G b_G Y_{G_w} + a_B b_B Y_{B_w} \quad (5.2)$$

where Y_{R_w} , Y_{G_w} and Y_{B_w} are the phosphor luminances required for producing white light. This is the case for a white surface ($a_R = a_G = a_B$) under white light ($b_R = b_G = b_B$). Note that the computation of reflected light on the basis of known reflectance and emission spectra, also implies the application of eq. (5.2), but then with the three primaries replaced by the wavelengths (and associated coefficients) constituting the reflectance and emission spectra. The details and specific advantages of the trichromatic reflection scheme are extensively discussed in the preceding paper (Lucassen & Walraven, 1992a). The question whether color constancy can be studied with the kind of synthetic stimuli we used, has been specifically addressed in a study employing both natural and artificial object-illuminant interactions (Lucassen & Walraven, 1992b). In that study we show that the visual response as described by eq. (5.1), is dependent only on the quantum catches of the cones, irrespective of the spectral characteristics of the illuminants.

Samples

The chromaticities of the simulated surface samples were the same as used before (Lucassen & Walraven, 1992a). That is, five achromatic samples (10, 25, 50, 75

and 90% reflectance) and 30 chromatic samples. The latter were selected Munsell samples, with chromaticities regularly spaced around the white point (see Fig. 3 in Lucassen & Walraven, 1992a). All chromatic samples were presented at the same reflectance (under white light). From the 35 samples, 11 were selected as test samples, 10 chromatic (all from the Munsell 4/5 series) and one neutral (of the same reflectance as the others). These test samples are indicated by an asterisk in Table 5.1, where the CIE x, y chromaticities of the light “reflected” by the 35 samples are shown under blue and yellow illumination. The latter two are the illuminants used in Experiment 1 and 2. Also shown in Table 5.1 are the luminances (Y) of the samples and the background of the stimulus patterns (to be discussed). These are the luminances found in the condition where the illuminants produce a luminance of 12 cd/m² (at 100% reflectance), and where sample and background have reflectances (under white light) of 50 and 100%, respectively. The simulated sample and background reflectances are varied in Experiment 1, hence their luminances vary accordingly.

Procedure

Three observers, the two authors and a naive subject, participated in the experiments. Their task was to match the color and brightness of a patch in the “match” pattern to the corresponding patch in the “test” pattern. As described above (Equipment), the two patterns were seen successively by the left and right eye (haploscopic matching). From condition to condition, the function of the eyes (test or match) was interchanged, in order to prevent possible long-term effects of chromatic adaptation. The test and match pattern were switched every five seconds, to keep both eyes equally exposed to their respective illuminant conditions. The observer was allowed to switch from one pattern to the other until he was completely satisfied with the match. The precision of the match, which is actually a short-term memory task, was quite satisfactory. Average deviations in terms of CIE chromaticity coordinates ($\bar{\Delta}xy$) are in the order of 1% (Lucassen & Walraven, 1992a).

5.2.2 Experiment 1: Varying luminance contrast

In Experiment 1, a simple center-surround stimulus was used, as shown in Fig. 5.1. It has been demonstrated in other studies (Valberg & Lange-Malecki, 1990; Arend & Reeves, 1986) that the results obtained with the center-surround stimuli are not qualitatively different from those obtained with complex stimulus configurations. The test and match pattern were illuminated by blue ($x=0.259, y=0.241$)

sample number	blue light			yellow light		
	x	y	Y (cd/m^2)	x	y	Y (cd/m^2)
1	0.2981	0.2715	5.86	0.4421	0.4503	6.18
2*	0.2189	0.1859	6.39	0.3832	0.4461	5.66
3	0.2426	0.2390	6.02	0.3824	0.4758	5.93
4*	0.2410	0.1905	6.33	0.4285	0.4242	5.80
5*	0.2411	0.2666	5.91	0.3583	0.5006	5.97
6	0.2897	0.2813	5.83	0.4232	0.4647	6.16
7*	0.2091	0.1980	6.29	0.3452	0.4759	5.66
8	0.3695	0.2869	5.77	0.5090	0.4124	6.43
9	0.3875	0.4275	5.51	0.4361	0.4805	6.42
10	0.2590	0.2410	3.00	0.4100	0.4600	3.00
11	0.2609	0.2607	5.92	0.3965	0.4752	6.03
12	0.3141	0.2311	6.01	0.5000	0.4026	6.23
13	0.2590	0.2410	10.8	0.4100	0.4600	10.8
14	0.4091	0.3697	5.58	0.4851	0.4414	6.49
15	0.2743	0.3452	5.68	0.3613	0.5183	6.14
16*	0.2186	0.2283	6.09	0.3419	0.4951	5.80
17	0.2862	0.2461	5.96	0.4478	0.4390	6.12
18*	0.2590	0.2410	6.00	0.4100	0.4600	6.00
19	0.2320	0.2203	6.13	0.3779	0.4700	5.84
20*	0.3085	0.2459	5.95	0.4783	0.4206	6.21
21	0.2585	0.1883	6.33	0.4655	0.4023	5.90
22	0.1949	0.1724	6.54	0.3293	0.4639	5.44
23	0.2590	0.2410	1.20	0.4100	0.4600	1.20
24	0.2164	0.2565	5.96	0.3165	0.5227	5.85
25	0.2563	0.2152	6.14	0.4302	0.4370	5.95
26	0.2590	0.2410	9.00	0.4100	0.4600	9.00
27	0.2183	0.1668	6.58	0.4054	0.4184	5.56
28	0.1928	0.2003	6.29	0.2968	0.5048	5.58
29*	0.3397	0.2911	5.78	0.4758	0.4342	6.33
30	0.2414	0.2060	6.21	0.4117	0.4430	5.85
31*	0.2769	0.2166	6.12	0.4632	0.4185	6.05
32*	0.3053	0.3326	5.69	0.4077	0.4860	6.22
33	0.2727	0.2305	6.04	0.4424	0.4365	6.05
34*	0.3480	0.3416	5.65	0.4486	0.4612	6.34
35	0.2323	0.2090	6.20	0.3895	0.4573	5.81
background	0.2590	0.2410	12.0	0.4100	0.4600	12.0

Table 5.1: CIE x, y chromaticities and luminance Y of the 35 samples and the background used, either under blue or yellow illumination. In this particular condition, the background (shown in the last entry in the table) conveys the exact color and luminance of the illuminant (100% reflectance). The 11 test samples used for obtaining the color matches are indicated by an asterisk in the first column.

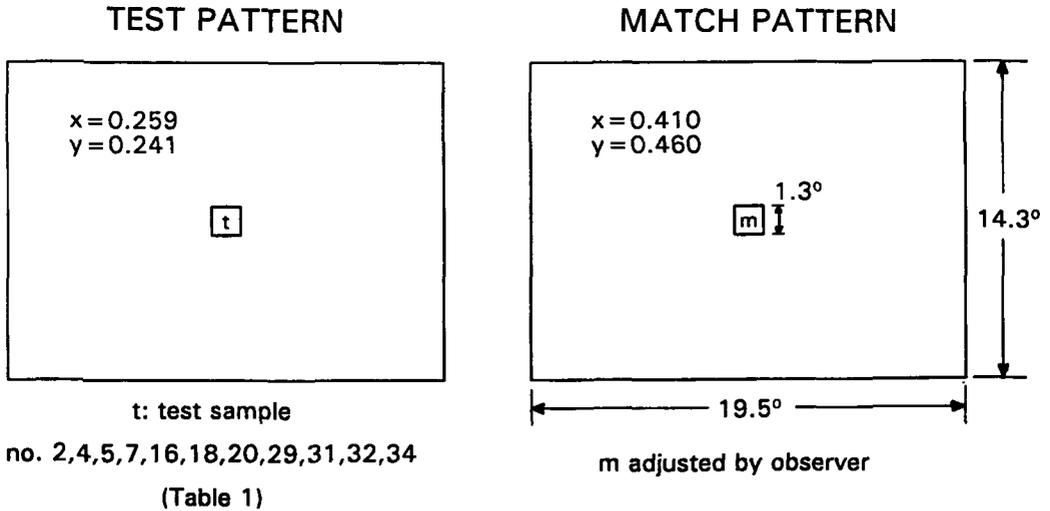


Figure 5.1: Stimulus configuration (center-surround) of Experiment 1.

and yellow ($x=0.410$, $y=0.460$) light, respectively. The samples under blue light (“test” illuminant) had to be matched by samples under yellow light (“match” illuminant). In this way, the short-wave system (S-cones) - which provides the most critical test for the validity of eq. (5.1) - receives a differential stimulation of about a factor seven, whereas the middle-wave-sensitive (M) and long-wave-sensitive (L) system remain at about the same level of activation. The x, y chromaticities of the test samples under blue light, presented in the center of the test pattern, and those of the backgrounds (surrounds), are given in Table 5.1.

As shown in Fig. 5.2, the luminance contrast between sample and (white) background (Y_j/Y_w) was varied by either manipulating the luminance of the test samples (upper panel) or the luminance of the background (lower panel). The range of contrasts in Fig. 5.2 may be related to reflectance in various ways depending on what reflectance is to be assigned to a particular luminance. In our previous study a (fixed) background luminance of 12 cd/m^2 was taken as representing 100% reflectance. Doing the same for the present stimuli would imply that luminances in excess of 12 cd/m^2 represent fluorescent stimuli, reflecting more than 100%. If, on the other hand, 100% reflectance is assigned to, say, 40 cd/m^2 , then all stimuli can be considered as surface reflectances in the range from 1.5 to 90%. However, this reflectance interpretation is just as arbitrary as the former one, so we prefer to describe the stimulus in terms of contrast (Y_j/Y_w),

rather than reflectance.

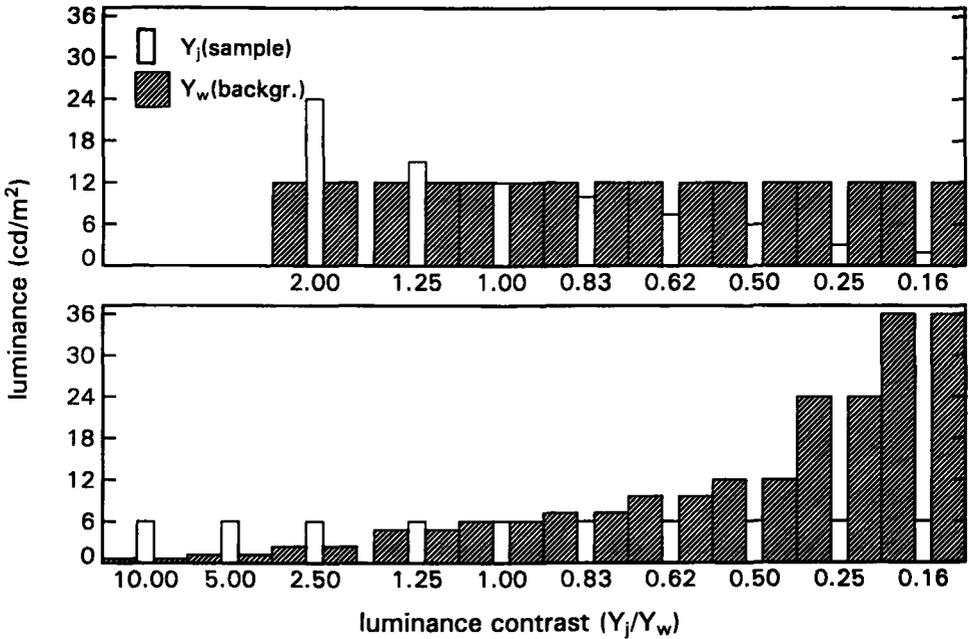


Figure 5.2: Schematic overview of the two sets of luminance profiles (sample surrounded by the background) used in Experiment 1. A range of corresponding luminance contrasts (Y_j/Y_w) was obtained by either varying sample reflectance around a fixed background reflectance (upper panel) or vice versa (lower panel). The luminance values shown here apply to samples and background under white light. The actual values under blue or yellow light are listed in Table 5.1.

5.2.3 Experiment 2: Varying illumination level

In Experiment 2, the stimulus pattern was a rectangular array consisting of 35 square samples on a homogeneous background (Fig. 5.3). This stimulus was also used in our preceding studies (Lucassen & Walraven, 1992a,b; Walraven *et al.*, 1991). The numbers (1-35) of the samples shown in Fig. 5.3 correspond to the numbers in the first column of Table 5.1, where the colors are specified under

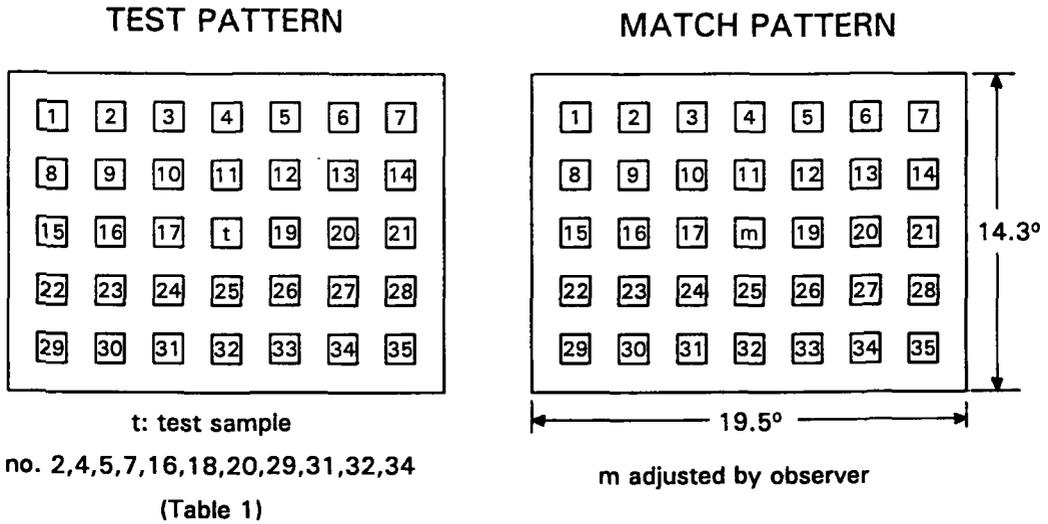


Figure 5.3: Stimulus configuration (array) of Experiment 2.

yellow and blue illumination. In contrast to Experiment 1, the reflectances of the samples and the background are now fixed at 50 and 100%, respectively, the standard configuration of our previous study. Four illuminant conditions are used in Experiment 2, each with a different combination of test and match illuminant. The x, y chromaticities and luminance levels of these four illuminant combinations are listed in Table 5.2, which is presented in the Results section. The different (cone-specific) illumination levels in test and match eye, were implemented by either a color difference, a luminance difference, or both.

5.3 Theoretical preamble

5.3.1 Presentation of models

In this section we discuss three different models, each one representing a different attempt at quantifying the data. These are, in “evolutionary” order: the original Lucassen and Walraven (1992a) model, the Jameson and Hurvich (1964) model, and the modified Lucassen and Walraven model. For brevity, the models will be respectively referred to as JH or LW models, whereby the latter may represent the original as well as the modified version.

As already mentioned in the Introduction, the LW and JH model are mathematically equivalent over a limited stimulus range. This finding is of theoretical significance considering the radically different ways in which the test (Q_j) and background (Q_w) are treated by the JH and LW model. The JH performs a subtractive operation, whereas the LW model is multiplicative in nature (taking ratios). As will be discussed henceforth, the reason for the mathematical equivalence can be traced to the fact that taking the logarithm of the ratio of two signals (as applies to the LW model), may not be that different from taking the difference of these signals after being transformed by a cube root transformation (as applies to the JH model).

5.3.2 Mathematical equivalence of the JH and LW model

Both the LW and JH model describe a response function in which test and surround stimulus (Q_j and Q_w) provide the input of a single channel. In case of the JH model this is the "brightness channel", but we use it here for describing the output of cone-specific pathways. Using our nomenclature, the JH response function takes the form

$$R_{JH} = \frac{c(Q_j^n - k Q_w^n)}{(1 - k^2)} \quad (5.3)$$

where k is a constant dependent on the particular stimulus configuration, and the exponent n typically takes the value $n = 0.33$. As for the proportionality factor c , Jameson and Hurvich (1964) could describe their results by simply assuming $c = 1$.

The LW response function is described by eq. (5.1), with $r=0.33$ and $k = 4.35$, that is

$$R_{LW} = (Q_w)^{0.33} \log \left(4.35 \frac{Q_j}{Q_w} \right) \quad (5.4)$$

The mathematical equivalence of eqs. (5.3) and (5.4) follows from the fact that, as is shown in Fig. 5.4, one may write

$$\log(x) \approx x^{0.33} - 1 \quad (0.5 \leq x \leq 6). \quad (5.5)$$

Upon substitution of $x = 4.35 Q_j/Q_w$ in eq. (5.5) this equation reads

$$\log \left(4.35 \frac{Q_j}{Q_w} \right) \approx \left(4.35 \frac{Q_j}{Q_w} \right)^{0.33} - 1 \quad (5.6)$$

which can be rewritten as

$$\log \left(4.35 \frac{Q_j}{Q_w} \right) \approx 1.624 \frac{(Q_j^{0.33} - 0.617 Q_w^{0.33})}{Q_w^{0.33}} \quad (5.7)$$

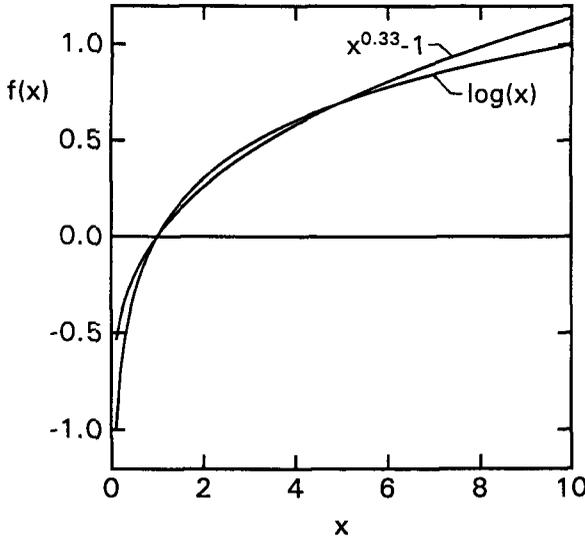


Figure 5.4: Comparison of the functions $\log(x)$ and $x^{0.33} - 1$. Over a certain range of x , the two functions closely resemble each other.

Multiplying both sides of eq. (5.7) with $(Q_w)^{0.33}$ yields

$$(Q_w)^{0.33} \log(4.35 \frac{Q_j}{Q_w}) \approx 1.624 (Q_j^{0.33} - 0.617 Q_w^{0.33}) . \quad (5.8)$$

The left hand side of eq. (5.8) is the LW response function, whereas the right hand side corresponds to the JH function, with $n = 0.33$, $k = 0.617$ and $c = 1.006$. These values closely correspond to the values cited by Jameson and Hurvich (1964), who found $n = 0.33$, $k = 0.6$ and $c = 1$ in order to describe the data of the classical Hess and Pretori (1894) experiments on brightness contrast.

The question that is raised by the above equivalence is whether the visual system encodes *differences* between center and surround (Jameson & Hurvich) or contrast *ratios* between center and surround (Lucassen & Walraven). The answer to this question cannot be given as long as the approximation in eq. (5.5) holds. As can be seen in Fig. 5.4, this is the case for the range $0.5 \leq x \leq 6$. With $4.35 Q_j/Q_w$ substituted for x , this range is $0.23 \leq Q_j/Q_w \leq 1.38$, a range that comprises both increments and decrements. Within each cone class, the value of Q_j/Q_w is determined by the chromatic difference as well as the luminance difference between sample and background. However, the chromatic difference is more or less limited by the spectral compositions of the center and surround, whereas the luminance difference can be manipulated almost unrestrictedly. Therefore, a

critical test for discriminating between the LW and JH model would require that the luminance contrast between sample and background is varied well outside the range $0.23 \leq Q_j/Q_w \leq 1.38$. This is what we did in our Experiment 1.

5.3.3 The modified LW model

As will be shown in the Results section, the LW response function, eq. (5.1), is not well suited to describe the present data. We therefore tried to modify eq. (5.1), and thereby arrived at a model in which the cone-specific inputs Q_j and Q_w are normalized for luminance. The modified response function can be written as

$$\tilde{R}_{LW} = (\tilde{Q}_w)^{0.33} \log \left(2.6 \frac{\tilde{Q}_j}{\tilde{Q}_w} \right) \quad (5.9)$$

where \tilde{Q}_w and \tilde{Q}_j are the luminance-normalized cone inputs

$$\tilde{Q}_w = \frac{Q_w}{Y_w}; \quad \tilde{Q}_j = \frac{Q_j}{Y_j}. \quad (5.10)$$

The luminance normalization makes \tilde{R}_{LW} unresponsive to variations in the cone input that result from changes in luminance. Any change in Y will cause a proportional change in Q , but since Y enters the denominator in eq. (5.10), there will be no change in \tilde{Q} . Equation (5.9) can also be used to describe the data of our earlier study. The reason why we arrived at eq. (5.1) instead, will be addressed in the Discussion section.

5.4 Data predictions

The data predictions can be derived by assuming that the subjects receive equal responses (R) from test and match samples, so that

$$R^m = R^t. \quad (5.11)$$

The superscripts m and t denote “match” and “test”, respectively. For the original LW model, as described by eq. (5.1), substitution of the latter into eq. (5.11) yields

$$(Q_w^m)^r \log \left(k \frac{Q_j^m}{Q_w^m} \right) = (Q_w^t)^r \log \left(k \frac{Q_j^t}{Q_w^t} \right). \quad (5.12)$$

From this equation the cone input of the matching sample, Q_j^m (for either L, M or S) can be solved, that is

$$\text{LW} : Q_j^m = \frac{Q_w^m}{k} \left[k \frac{Q_j^t}{Q_w^t} \right]^{(Q_w^t/Q_w^m)^r} \quad (5.13)$$

with $k = 4.35$ and $r=0.33$. For the modified version of the LW model the prediction is the same, but now in terms of \tilde{Q} , so that

$$\tilde{LW} : \tilde{Q}_j^m = \frac{\tilde{Q}_w^m}{k} \left[k \frac{\tilde{Q}_j^t}{\tilde{Q}_w^t} \right]^{(\tilde{Q}_w^t/\tilde{Q}_w^m)^r} \quad (5.14)$$

with $k = 2.6$ and $r=0.33$. Note that \tilde{Q}_j^m is a luminance-normalized cone input, in accordance with eq. (5.10). To compute the absolute cone input Q_j^m , one has to multiply \tilde{Q}_j^m by Y_j^m . However, Y_j^m is an unknown variable (to be set by the subject). This problem will be addressed when discussing the results. It will be shown that the data indicate a relationship between Y_j^m and the other (known) luminance variables (Y_w^m, Y_j^t, Y_w^t) from which Y_j^m can be solved.

For the JH model, we follow the same procedure, substitution of eq. (5.3) into eq. (5.11), and thus obtain

$$JH : Q_j^m = \left[(Q_j^t)^n - k \left((Q_w^t)^n - (Q_w^m)^n \right) \right]^{1/n} \quad (5.15)$$

with $k=0.617$ and $n = 0.33$. Since $Q_j^m \geq 0$, the term inside the brackets on the right hand side of eq. (5.15) should also be ≥ 0 .

All predictions will be expressed in terms of Q_j^m , the cone input (per receptor system) required for matching the test stimulus (Q_j^t). The quantity Q has the dimension of cd/m^2 per receptor system. It can be derived from the CIE x, y, Y specifications of the stimuli, for a given set of cone spectral sensitivities (e.g. Vos & Walraven, 1971; Walraven & Werner, 1991). A discussion of the Q unit, and the procedures for deriving Q from the CIE x, y, Y specifications, is presented in Appendices C and D of Lucassen and Walraven (1992a).

5.5 Results

5.5.1 Experiment 1: Varying luminance contrast

In Experiment 1 the test/background contrast (Q_j/Q_w) was varied by changing the luminance of either the test field (Y_j) or the background (Y_w). Figure 5.5 shows a comparison of predictions and results, as obtained with variable test and fixed background (see top panel in Fig. 5.2). The left three panels relate to the original LW model predictions (for L, M and S-cones, respectively), whereas the three right panels show the same for the JH model.

When comparing the predictions of the LW and JH model in Fig. 5.5, it is clear that these only differ for the S-cones. The reason for that is that the blue

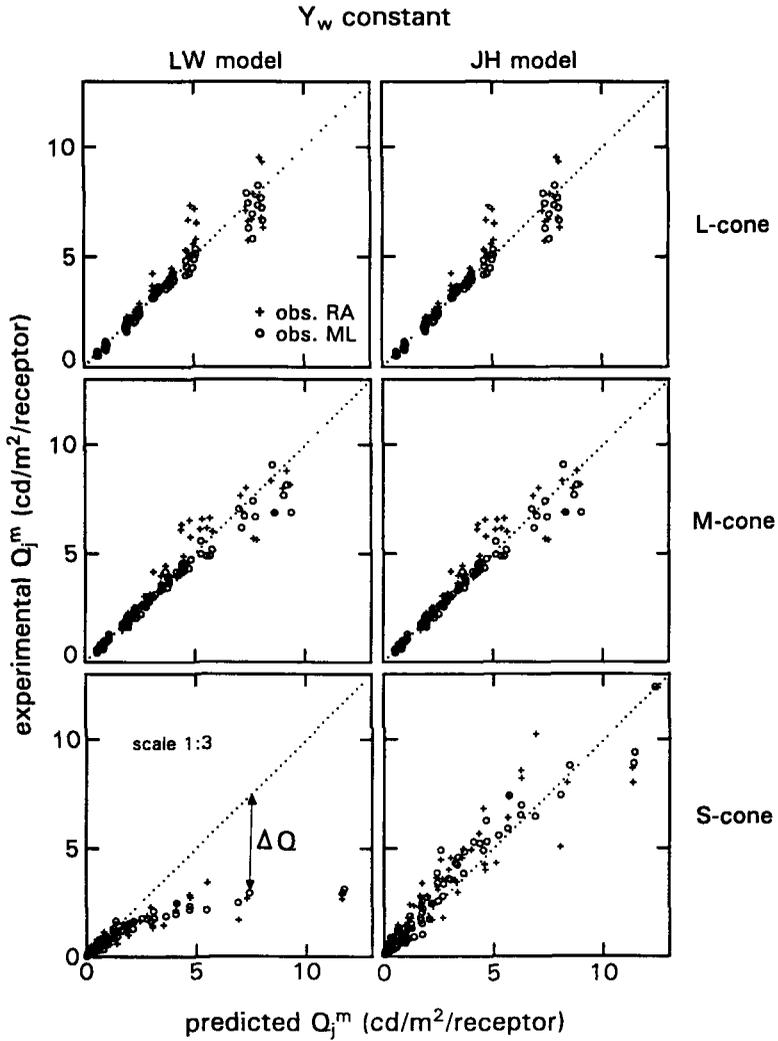


Figure 5.5: Results of Experiment 1, for conditions with fixed background luminance. Comparisons of predicted and experimentally obtained cone inputs of the color matches, shown separately for the L, M and S-cones (top to bottom). Graphs on the left relate to the predictions with the LW model, eq. (5.13); graphs on the right relate to those with the JH model, eq. (5.15). Note the different scale for the lower left graph.

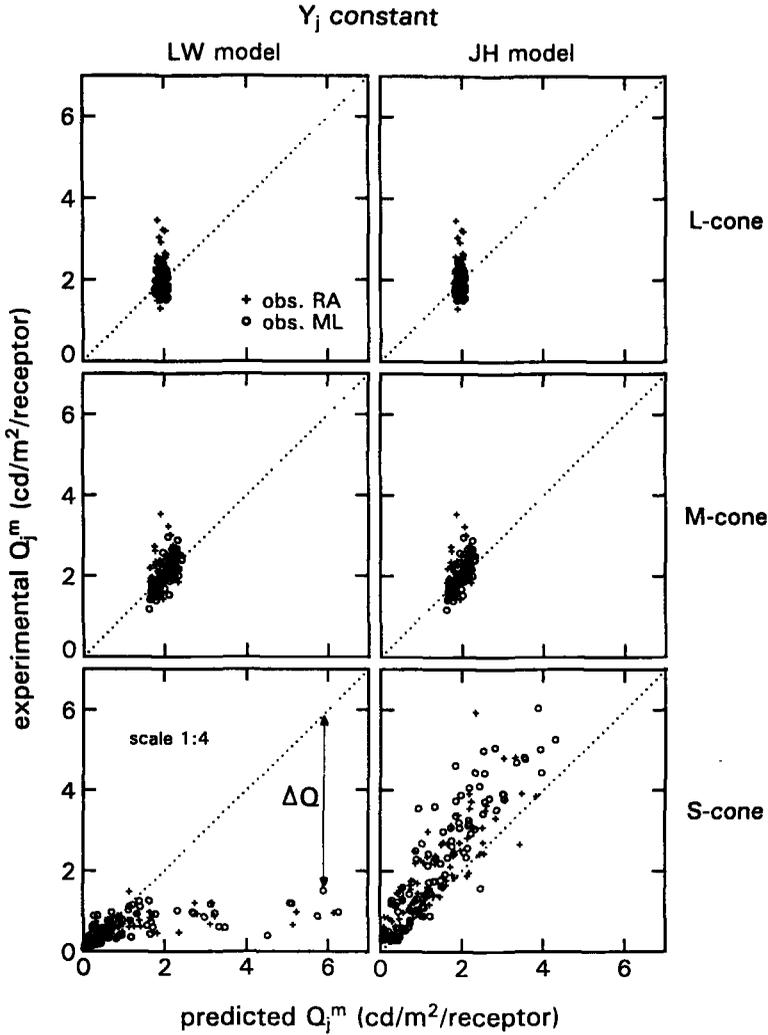


Figure 5.6: Same as Fig. 5.5, but now for the conditions with fixed sample luminance.

(test) and yellow (match) illuminant produce a differential effect in the S-cones only.

The L and M-cones hardly register a difference in illumination when changing from the yellow to the (equi-luminant) blue background. Consequently, the LW and JH model will concur in their predictions, that is, interocular identity matches

(but not without some scatter) for both the L and M-cone systems. Therefore, different predictions will thus only show up in the S-cone data, as is also apparent from the data obtained with fixed sample luminance and varying background. The latter are shown in Fig. 5.6. The data plotted in Fig. 5.6 show the same pattern as those presented in Fig. 5.5, and for the same reason. Here again the difference in model predictions are only found for the S-cones, the latter being strongly modulated when switching between blue and yellow light.

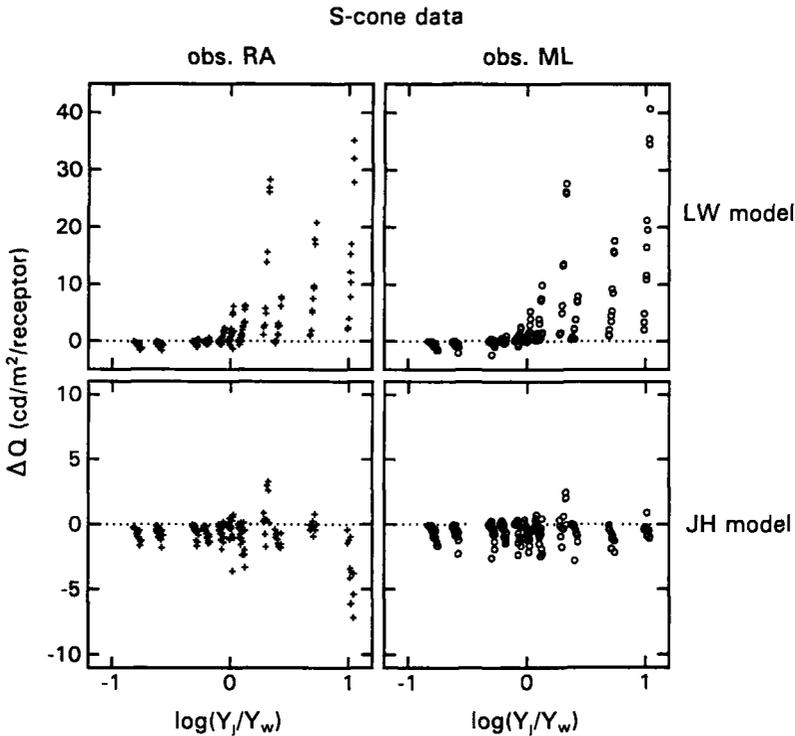


Figure 5.7: Mispredictions of the two models (ΔQ) as a function of sample to background luminance contrast (Y_j/Y_w). Pooled S-cone data from Figs. 5.5 and 5.6. The dotted lines indicate perfect model predictions.

When comparing the performance of the LW and JH model, it is clear that the latter can better account for the (S-cone) data. The problem with the LW model is that it does not correctly predict the results from samples that are much brighter than the background ($Y_j > Y_w$). This is shown in Fig. 5.7, where we plotted ΔQ_j^m , the error in predicting Q_j^m , as a function of Y_j/Y_w , the sample/background

contrast (in the test pattern). These are the pooled data from Figs. 5.5 and 5.6, so increasing values of Y_j/Y_w may either be due to an increase in Y_j or a decrease in Y_w . Figure 5.7 shows that for decrements ($\log(Y_j/Y_w) < 0$) the LW model performs slightly better than the JH model. However, for increments the LW model shows increasing mispredictions with increasing contrast, whereas the JH model does not show that tendency. Clearly, the LW model, as described by eq. (5.1), has to be modified in order to be applicable over the whole two log-unit contrast range investigated here.

5.5.2 Experiment 2: Varying illumination

In the previous experiment we studied only the effect of relative luminance variation (contrast). The absolute level of (simulated) illumination was not varied when switching between the blue test and yellow match illuminant condition. In Experiment 2 we focus on varying the illumination, thereby testing the four conditions mentioned in the Methods section. The specifications of the four combinations of test (Q_w^t) and match (Q_w^m) illuminant are shown in Table 5.2. These are given in terms of the luminance of the background (Y , cd/m^2) and the associated L, M and S-cone inputs (Q_j , cd/m^2 per receptor). In all conditions the sample/background contrast was $Y_j/Y_w=0.50$, corresponding with 100% reflectance of the background and 50% reflectance of the test sample (under white light). All experiments were performed with the 5×7 array configuration.

The conditions shown in Table 5.2 were so chosen in order to introduce pre-determined changes in cone activation (test/match ratio), either by just a color change (Condition 2), a luminance change (Condition 3), or both (Condition 4). For the sake of completeness, there was also a “no-change” condition (Condition 1), which tests the absence of any effects when both eyes see the same stimulus. This condition also allows an estimate of the precision of making (haplosopic) identity matches. Note that in Conditions 3 and 4, the same factor of increase in cone input (about a factor 7) is achieved, but by different means (a change in luminance and color respectively).

Since the results of the three observers participating in Experiment 2 were not systematically different, the averaged data were used for the analysis. For Conditions 1 and 2, these data and their model predictions are compared in Fig. 5.8. As can be seen from the close correspondence between obtained and predicted data, both the LW and JH model have no difficulty predicting the results of these two conditions. Recalling that in Experiment 2 the samples were always presented as decrements ($Y_j/Y_w = 0.50$), this was the expected outcome. We already found in Experiment 1 that mispredictions, as were obtained before with the LW model, only occur for incremental stimuli ($Y_j > Y_w$).

Condition	test illuminant	match illuminant	cone-specific ratios (test/match)
1	$x=0.2590$ $y=0.2410$ $Y=12.0 \text{ cd/m}^2$	$x=0.2590$ $y=0.2410$ $Y=12.0 \text{ cd/m}^2$	
	$L=3.815$ $M=4.204$ $S=7.850$	$L=3.815$ $M=4.204$ $S=7.850$	1.000 1.000 1.000
2	$x=0.2590$ $y=0.2410$ $Y=12.0 \text{ cd/m}^2$	$x=0.4100$ $y=0.4600$ $Y=12.0 \text{ cd/m}^2$	
	$L=3.815$ $M=4.204$ $S=7.850$	$L=4.038$ $M=3.911$ $S=1.087$	0.945 1.075 7.221
3	$x=0.2590$ $y=0.2410$ $Y=21.08 \text{ cd/m}^2$	$x=0.2590$ $y=0.2410$ $Y=2.92 \text{ cd/m}^2$	
	$L=6.710$ $M=7.385$ $S=13.79$	$L=0.928$ $M=1.023$ $S=1.910$	7.221 7.219 7.220
4	$x=0.2590$ $y=0.2410$ $Y=21.08 \text{ cd/m}^2$	$x=0.4100$ $y=0.4600$ $Y=2.92 \text{ cd/m}^2$	
	$L=6.710$ $M=7.385$ $S=13.79$	$L=0.983$ $M=0.952$ $S=0.264$	6.817 7.757 52.24

Table 5.2: Specification of the four illuminant conditions used in Experiment 2.

In Conditions 1 and 2 the change from test to match pattern was not accompanied by a change in the level of illumination ($Y_w^t = Y_w^m = 12 \text{ cd/m}^2$). In Conditions 3 and 4, however, the levels of test ($Y_w^t = 21.08 \text{ cd/m}^2$) and match ($Y_w^m = 2.92 \text{ cd/m}^2$) illumination differed by about a factor 7. As can be seen in Fig. 5.9, which plots the results for Conditions 3 and 4, this poses a problem for both the LW and JH model. The predictions from both models fall below the (dotted) identity diagonal in Fig. 5.9, which implies that the models overestimated the cone inputs required for matching the test samples. The failure of the two models might be possibly attributed to different pupil sizes (the match eye received less illumination). However, this possibility is ruled out, since control experiments with 4-mm artificial pupils for both eyes, yielded the same results.

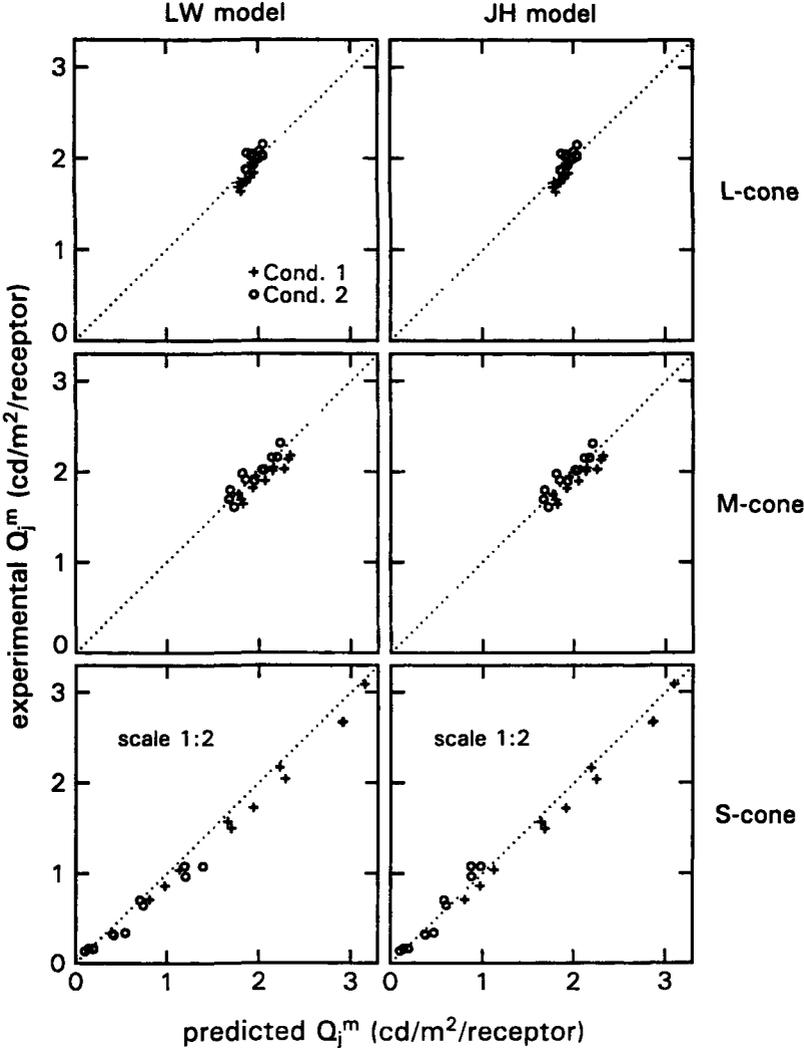


Figure 5.8: Results of Conditions 1 and 2 of Experiment 2 (mean data of three observers). Comparisons of predicted and experimentally obtained cone inputs of the color matches, shown separately for the L, M and S-cones (top to bottom). Graphs on the left relate to the predictions with the LW model, eq. (5.13); graphs on the right relate to those with the JH model, eq. (5.15).

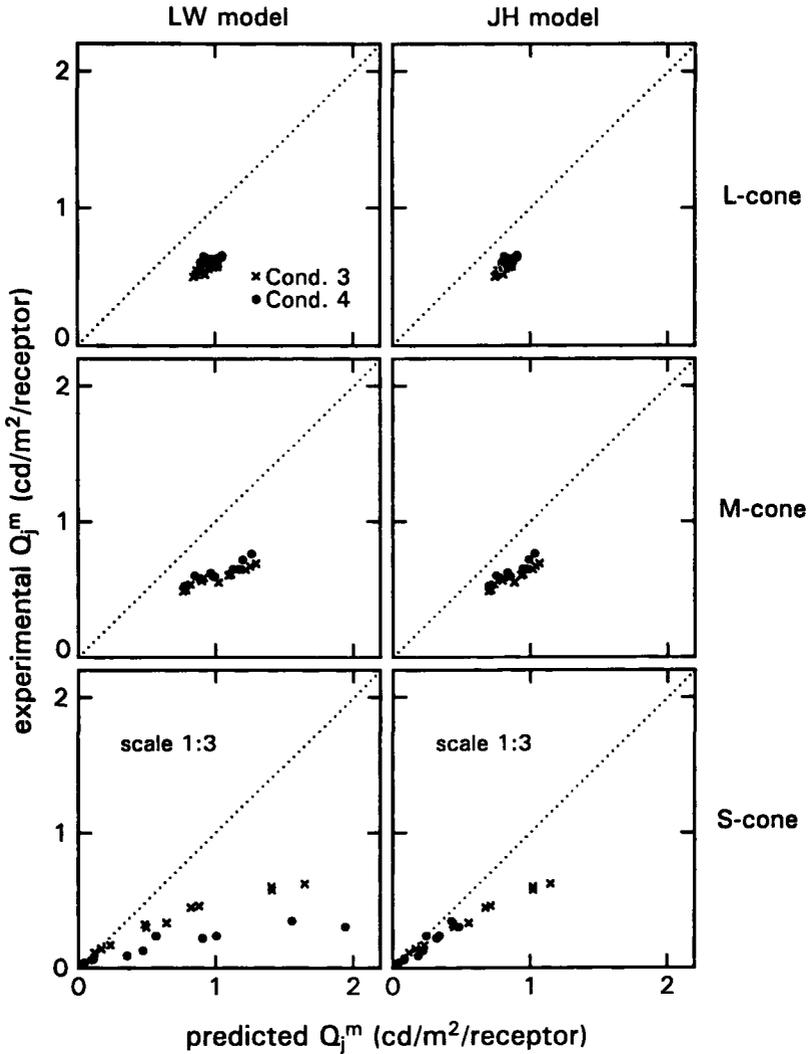


Figure 5.9: Same as Fig. 5.8, but now for Conditions 3 and 4.

(The 4-mm pupil is smaller than the natural pupil at the illumination levels in question.) The variables tested in Experiments 1 and 2, luminance contrast (Y_j/Y_w) and absolute illumination level (Q_w), clearly pose a problem for our model. This problem was not encountered in the study from which we derived the model (Lucassen & Walraven, 1992a), because these were exactly the variables

that were kept constant in that study. Consequently, the effect of these variables was reduced to a constant, a constant probably hidden somewhere within the constants of eq. (5.1). Taking that as a starting point, we shall now take a closer look at the constants of eq. (5.1).

5.6 Modifying equation (5.1): evidence for luminance-normalized cone signals

The results of Experiment 1 suggest that luminance contrast (Y_j/Y_w) is a variable that somehow must enter eq. (5.1) in order to be able to better predict the results that we found for increments ($Y_j > Y_w$). In eq. (5.1) the constant k acts as a multiplier on the cone-specific contrast (Q_j/Q_w), so this would make it a likely vehicle for implementing the effect of luminance contrast. That is, k should no longer be constant, but depend on luminance contrast, according to the (unknown) function

$$k = f\left(\frac{Y_j}{Y_w}\right). \quad (5.16)$$

In order to derive the function relating k to Y_j/Y_w , we need to know what value of k would yield correct predictions on the basis of eq. (5.1). Thereto we have to return to eq. (5.13), which is used for computing the predicted Q_j^m . Using eq. (5.13) it can be easily shown that k must satisfy

$$k_{LW} = \left(\frac{Q_j^t}{Q_w^t}\right) \left[(Q_w^m/Q_w^t)^{0.33} - 1\right]^{-1} \left(\frac{Q_j^m}{Q_w^m}\right) \left[(Q_w^t/Q_w^m)^{0.33} - 1\right]^{-1} \quad (5.17)$$

where we labeled k with subscript LW for denoting the LW model. The same procedure, but now applied to eq. (5.15), can be followed for the constant k_{JH} of the JH model

$$k_{JH} = \frac{(Q_j^t)^{0.33} - (Q_j^m)^{0.33}}{(Q_w^t)^{0.33} - (Q_w^m)^{0.33}}. \quad (5.18)$$

Both k_{LW} and k_{JH} depend on Q_j^m , an experimentally obtained value. If the model predictions would be perfect, then the obtained values of Q_j^m would show k to be constant. However, in reality we may expect k to depend on luminance contrast. Note that for $Q_w^t = Q_w^m$, that is, no illuminant change (for the receptor system in question), the JH model and the LW model predict k_{JH} and k_{LW} to reach infinity, respectively. As mentioned before, the condition $Q_w^t = Q_w^m$ more or less applies to the L and M-cones in our experiment, so only the S-cone data are of interest for getting information on the function $k = f(Y_j/Y_w)$.

The S-cone data obtained in Experiment 1 - the experiment in which Y_j/Y_w was varied - were used for computing both k_{LW} and k_{JH} . In Fig. 5.10, the values of k_{LW} and k_{JH} , averaged over the 11 samples per condition, are shown as a function of the inverse luminance contrast (Y_w/Y_j), also averaged per condition. The left panel shows the result for k_{LW} computed with eq. (5.17), the right panel the result for k_{JH} computed on the basis of eq. (5.18). The two different data symbols used in Fig. 5.10 relate to the conditions where either the background luminance (Y_w) was fixed (upper panel in Fig. 5.2) or the sample luminance (Y_j) was fixed (lower panel in Fig. 5.2).

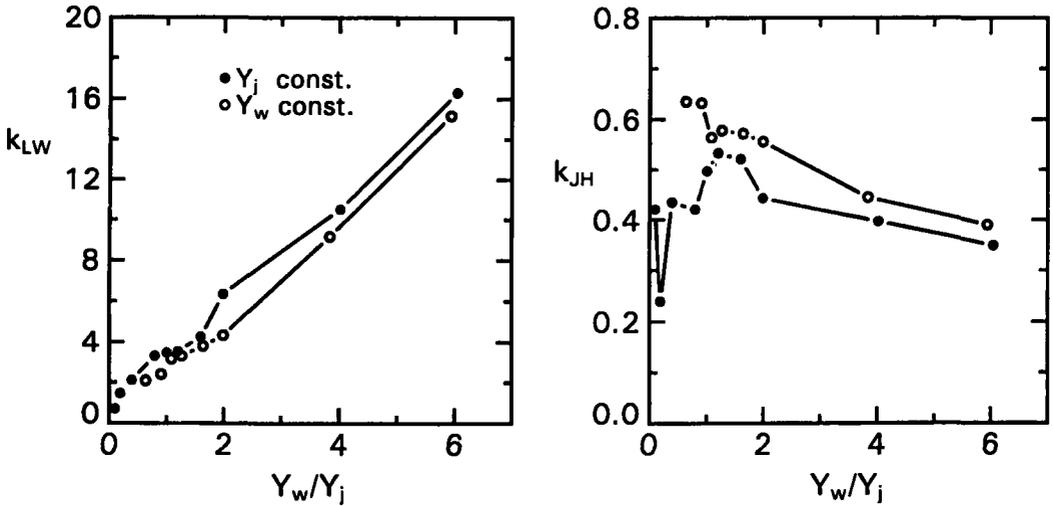


Figure 5.10: Value of k_{LW} (left panel) and k_{JH} (right panel), computed with eq. (5.17) and eq. (5.18), respectively, as a function of the background to sample luminance contrast (Y_w/Y_j). Data points represent averages over 11 samples per condition (S-cone data from Experiment 1). Open and closed circles relate to conditions with fixed background and fixed sample luminance, respectively.

The results for the Jameson and Hurvich model show a rather complex relationship between k_{JH} and Y_w/Y_j , which does not lend itself to a simple mechanistic interpretation. The behavior of k_{LW} is better behaved in this respect. It linearly increases with Y_w/Y_j , for both the conditions where Y_w =constant and

$Y_j = \text{constant}$. We fitted these data points with

$$k_{LW} = 2.6 \frac{Y_w}{Y_j} . \quad (5.19)$$

The significance of this result becomes immediately clear when substituting eq. (5.19) into eq. (5.1). The response function now becomes

$$R_{LW} = (Q_w)^r \log \left(2.6 \frac{Y_w Q_j}{Y_j Q_w} \right) = (Q_w)^r \log \left(2.6 \frac{Q_j/Y_j}{Q_w/Y_w} \right) \quad (5.20)$$

and thus shows the cone contrast Q_j/Q_w to be normalized for luminance. This implies that R_{LW} only responds to contrast changes that are due to a change in chromaticity. Any change in luminance contrast (Y_j/Y_w) is negated because of the normalizing factor Y_w/Y_j .

When we tried to predict the data on the basis of eq. (5.20), we still found this to be unsatisfactory for the data obtained in Conditions 3 and 4. These are the conditions in which we employed different illumination levels (Q_w) for the test and match eye. It turned out that also Q_w had to be normalized in order to fit the data. So, the cone inputs to eq. (5.1) all have to be converted from absolute (Q) to luminance-normalized inputs

$$\tilde{Q} = \frac{Q}{Y} . \quad (5.21)$$

Consequently, eq. (5.1) can now be rewritten as

$$\tilde{R} = (\tilde{Q}_w)^r \log \left(k \frac{\tilde{Q}_j}{\tilde{Q}_w} \right) \quad (5.22)$$

with $r=0.33$ and $k=2.6$.

Since eq. (5.22) applies to luminance-normalized contrast (\tilde{Q}_j/\tilde{Q}_w), its predictions with respect to the predicted input of the match sample (j), are in the same terms. That is, in analogy to eq. (5.13),

$$\tilde{Q}_j^m = \frac{\tilde{Q}_w^m}{k} \left[k \frac{\tilde{Q}_j^t}{\tilde{Q}_w^t} \right]^{(\tilde{Q}_w^t/\tilde{Q}_w^m)^r} . \quad (5.23)$$

On the basis of eq. (5.23) we can predict \tilde{Q}_j^m for each cone system, but not Q_j^m . Thereto we have to know Y_j^m , so as to "denormalize" Q_j^m according to

$$Q_j^m = Y_j^m \tilde{Q}_j^m \quad (5.24)$$

in which \tilde{Q}_j^m can be predicted with eq. (5.23). However, for Y_j^m , the luminance variable, we need a separate predictor.

Intuitively, one might expect that matching the achromatic aspect of test and match sample would involve the matching of luminance contrast ($Y_j^m/Y_w^m = Y_j^t/Y_w^t$). However, in our earlier study (Lucassen & Walraven, 1992a), we already noted that in the (few) experiments in which we used different illumination levels for the test and match illuminant, there were deviations from strict proportionality between test and match contrast. This is also shown in Fig. 5.11a, in which we plotted the data from conditions with unequal background luminances ($Y_w^t \neq Y_w^m$), both from our earlier study and the present one (Conditions 3 and 4).

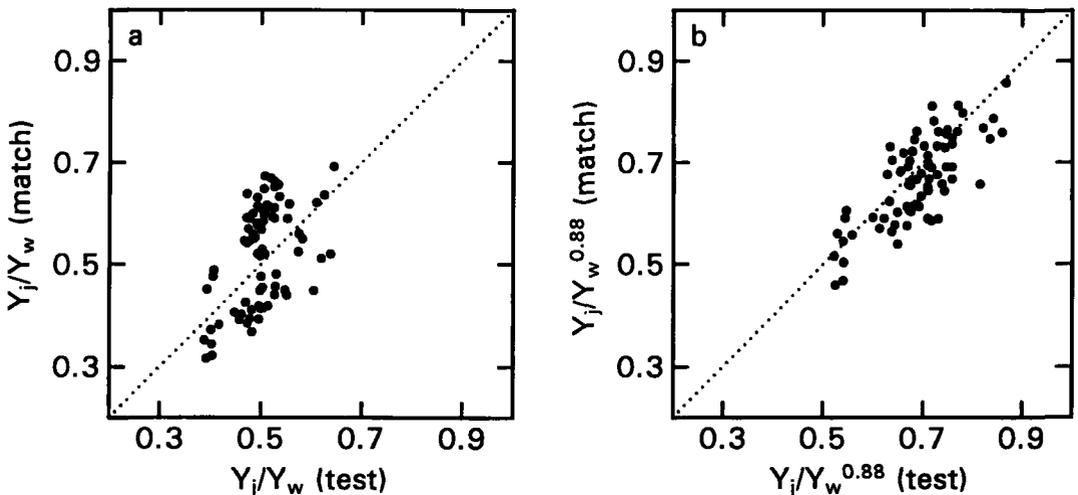


Figure 5.11: Comparison of luminance contrast in the test and match pattern. Luminance contrast is defined as either Y_j/Y_w (a) or $Y_j/(Y_w)^{0.88}$ (b). Shown are pooled data from a previous study (Lucassen & Walraven, 1992a) and Conditions 3 and 4 of Experiment 2. All data relate to conditions with asymmetric test and match illumination.

We found that the relatively low correlation between test and match contrast, shown in Fig. 5.11a, could be improved by redefining luminance contrast (C_Y) according to

$$C_Y = \frac{Y_j}{(Y_w)^{0.88}} \cdot \quad (5.25)$$

The improved correlation is shown in Fig. 5.11b. The exponent for compressing Y_w was found by minimizing data variance as a function of the power of Y_w . The implication of this result is that the response to a given luminance contrast increases with illumination level, an effect to be addressed in the Discussion.

From eq. (5.25) it follows that

$$Y_j^m = Y_j^t \left(\frac{Y_w^m}{Y_w^t} \right)^{0.88} \quad (5.26)$$

which, upon substitution in eq. (5.24), allows Q_j^m to be calculated, once \tilde{Q}_j has been computed with eq. (5.23).

Predictions of Q_j^m , as based on the modified response function, are shown in Fig. 5.12. The left panels in Fig. 5.12 relate to the pooled data of Experiment 1, the panels on the right to the pooled data of Experiment 2. The predictions shown here are for the data averaged over the observers. Although the predictions show some scattering around the identity line, the overall result is quite satisfactory, as is confirmed by the high correlation coefficients between prediction and experiment that we obtained. The correlation coefficients for the graphs on the left in Fig. 5.12 (Experiment 1) are 0.979, 0.981 and 0.929, for the L-, M- and S-cone data, respectively. For the graphs on the right (Experiment 2) they are 0.995, 0.994 and 0.991, respectively.

The data of our previous study (Lucassen & Walraven, 1992a) are also described by the modified response function. The correlation coefficients for the comparisons of the experimentally obtained cone inputs for the match and those obtained by prediction on the basis of the modified response function, are 0.972, 0.970 and 0.961, respectively, for that particular set of data.

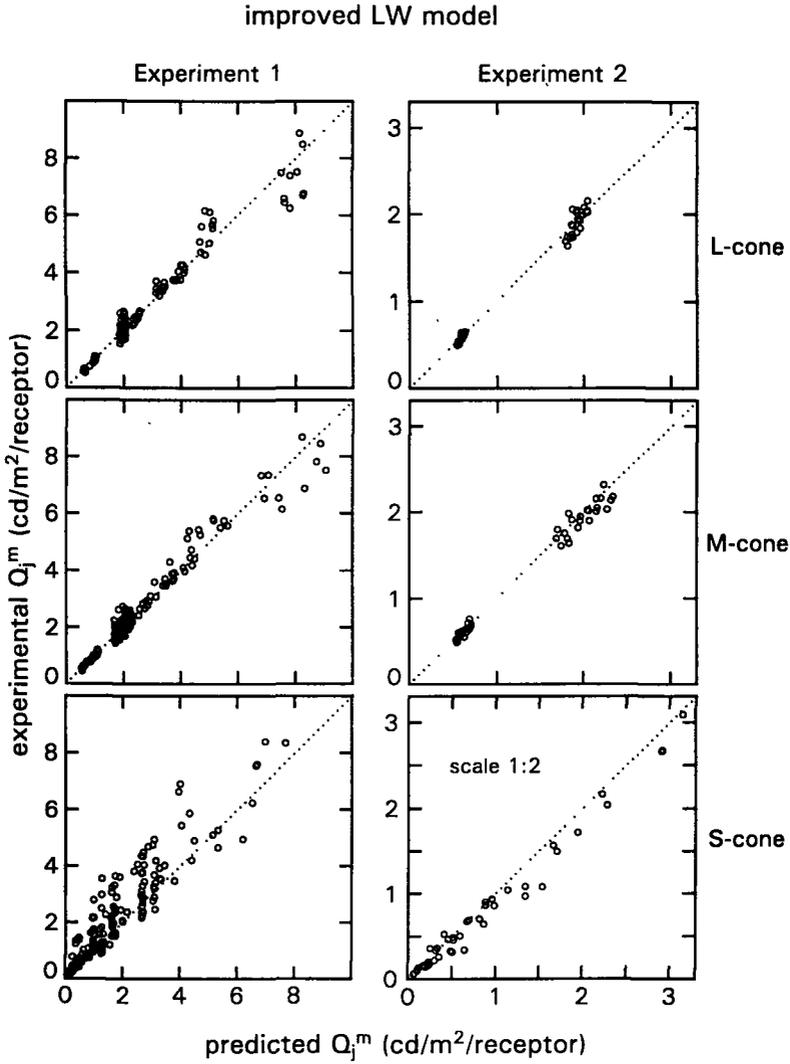


Figure 5.12: Predictions with the improved LW model. The graphs on the left relate to the (mean) data of Experiment 1, those on the right to Experiment 2.

5.7 Discussion

5.7.1 Luminance-free cone signals

We have shown that color constancy is not independent of the achromatic stimulus variable, contrast and illumination level. It is for that reason that we had to modify eq. (5.1), the response function that so accurately described the results of our earlier study (Lucassen & Walraven, 1992a). That it did so is not surprising, since the experimental variable of interest in that study was the color of the illuminant; contrast was fixed at $Y_j/Y_w=0.50$, and so was illumination level (except for small variations in a few pilot experiments).

The modification of eq. (5.1) involves replacing the cone-specific sample and background inputs (Q_j and Q_w) by their luminance-normalized values ($\tilde{Q}_j = Q_j/Y_j$ and $\tilde{Q}_w = Q_w/Y_w$). This normalization removes an ambiguity of the cone-specific contrast input (Q_j/Q_w). That is, a change in cone contrast may result from a change in luminance contrast, chromaticity contrast, or both. For example, if the L-cones signal a dark sample on a light background, this could represent a grey sample on a white background, but also, a green sample on a red background. The luminance normalization makes the contrast signal insensitive to luminance changes, so what is left, only reflects changes in chromaticity.

The effect of luminance normalization may not be very conspicuous when the spectral sensitivities of the luminance channel and a receptor channel are largely overlapping. In the extreme case, complete overlap (a monochromat), Q and Y will be equal and vary in tandem, resulting in a constant (chromatic) output of unity. As a matter of fact, both the L and M-cones have spectral sensitivities that are largely overlapping, both with respect to each other and with respect to V_λ . Consequently, the chromatic contrast signalled by the L and M-cones varies over a relatively small range. This explains why in Experiment 2, where luminance contrast was fixed, the variation in Q_j^m (which here reflects only the variation in the color of the test samples), is so much smaller for the L and M-cones than for the S-cones (see Figs. 5.8 and 5.9). The latter, therefore, are the most critical for testing models of color constancy. It is for that reason that our experimental conditions were designed for maximally exploring the chromatic signal range of the S-cones.

Luminance normalization of cone contrast implies that the activity generated within a cone system is put into relationship to that generated by other cone systems. The latter activity, the luminance signal (Y), represents the activities of the three cone systems according to $Y = 2L + M + 0.02S$ (Vos & Walraven, 1971; Vos, 1978). A more straightforward way for normalizing the L, M and S signals, would be to use the unweighed L, M and S sum, in the way this is

done in the CIE XYZ domain, where chromaticity coordinates x, y and z are defined according to $x = X/(X + Y + Z)$, etc. In the same way, one might also define “receptor coordinates” l, m and s according to $l = L/(L + M + S)$, etc. (see also Fairchild, 1991). The associated l, m diagram (analogous to the CIE x, y diagram) has been constructed by Walraven and Benzschawel (Benzschawel, 1992; Benzschawel, Walraven & Rogowitz, 1987), and this turned out to be a useful tool for analytical purposes. When analyzing the data on the basis of $(L + M + S)$ normalization, the predictions were not much different from those using luminance $(2L + M + 0.02S)$ normalization. We nevertheless chose for the latter approach, because it might be the more relevant choice from the physiological point of view. We think in this respect of the well-established notion of a “luminance-channel”, with its spectral sensitivity (V_λ) matching the combination $2L + M + 0.02S$ rather than $L + M + S$. Why the visual system chose for unequal weighing of the three cone systems is an interesting question, but one which is outside the scope of this study.

5.7.2 Luminance and luminance contrast

A visual system that encodes color in terms of luminance-free cone signals, pays the penalty of losing information about luminance contrast and absolute luminance level. The possession of a luminance channel allows the separate processing of these variables, of course, but it is not that self-evident how the visual system extracts a contrast signal, without losing information about absolute light level. A contrast signal, in our nomenclature Y_j/Y_w , remains invariant under varying illumination, because the light reflected from sample (j) and background (w), increases in the same proportion. So, a mechanism for extracting contrast implies eliminating absolute light level.

A way out of this problem is to process the approximate rather than the exact luminance ratio. This may explain why we found that the (interocular) contrast matches showed a better correlation when expressed as $Y_j/Y_w^{0.88}$ instead of Y_j/Y_w (see Fig. 5.11). This means that (apparent) contrast is not independent of illumination, but gradually increases with illumination, in proportion with $Y_w^{0.12}$. For the moment we consider this to be more of an empirical than a mechanistic description of how the visual system resolves the dilemma of encoding both relative and absolute luminance information. Probably, a more detailed model is necessary, possibly involving considerations regarding “noise” or “dark light”.

5.7.3 Comparison with models for achromatic vision

As discussed in the Introduction, color constancy may be analyzed in terms of a trichromatic extension of achromatic signal processing. A point in case is our finding that the brightness contrast model of Jameson and Hurvich (1964), applied to cone-specific contrast, can be used for describing at least part of our data (Figs. 5.5 and 5.8). It is of interest that this model, which treats contrast in terms of a *difference* signal, may yield similar results as our model, in which contrast is defined as a *ratio*. We have shown that this can be understood because of the equivalence (over a limited range) of $f(x) = x^{0.33}$ and $f(x) = \log(x) + 1$. So, the difference $Q_j^{0.33} - Q_w^{0.33}$ may yield approximately the same result as $\log(Q_j) - \log(Q_w) = \log(Q_j/Q_w)$.

A system that operates like a logarithmic analyzer will exhibit lightness constancy, since it will transmit a ratio (a reflectance) as a luminance invariant signal. As pointed out by Koenderink, van de Grind and Bouman (1971) it also has the advantage that it can easily remove (by subtractive filtering) the multiplicative noise introduced into the retinal image by the overlay of blood vessels and neural tissue. A plausible physiological mechanism for implementing the logarithmic transformation is a fast gain control that aims at a fixed output (Koenderink *et al.*, 1971). Such a mechanism, essentially Rushton's (1965) "automatic gain control", can account for a great variety of psychophysical data (Koenderink *et al.*, 1971; Walraven, 1980; Walraven & Valetton, 1984). It is for that reason that we prefer to analyze our data in terms that are compatible with a logarithmic rather than a cube root transformation.

In addition to the model of Jameson and Hurvich (1964) we tried various other models from the achromatic domain for describing our results. These include models by Burkhardt *et al.* (1984), Georgeson (1984), Kingdom and Moulden (1991), Semmelroth (1970), Stevens and Stevens (1963) and Whittle (1986). However, none of these gave better data predictions than our own equations, even when optimizing free parameters that are incorporated in some of those models.

5.7.4 Other quantitative accounts of color constancy

There are surprisingly few studies on color constancy in which experimental data are confronted with model predictions. Best known is the study of McCann *et al.*, (1976), which tested the validity of the (early) Retinex model (Land, 1964). Although the Retinex model(s) may be considered as the first analytical (rather than empirical) attempts at quantifying color constancy, we have shown that it performed less well than our own model (Lucassen & Walraven, 1992a). The main

shortcoming of the Retinex model, even in its most recent version (Land, 1986b), is that it is too relativistic. It makes no allowance for the absolute illumination level.

Other, more recent, studies on color constancy in which a model was presented to account for experimental data are rare. A notable exception is a study by Brainard and Wandell (1992). They showed that their data can be described by assuming cone-specific sensitivity adjustments, consistent with the von Kries (1905) coefficient law. Moreover, the coefficients (gain factors) were found to be proportional with the change in illumination. In that sense the predictions would not be different from the Retinex model which is essentially a von Kries model, as has been pointed out already by various authors (e.g. Jameson & Hurvich, 1989; Valberg & Lange-Malecki, 1990; Walraven *et al.*, 1991; Worthey, 1985).

A von Kries type of adaptation that varies in proportion with illumination changes, would result in perfectly discounting illuminant level. This is contrary to our daily experience, and is also contradicted by laboratory experiments on brightness or lightness constancy (Jameson & Hurvich, 1961; Jacobsen & Gilchrist, 1988). Our own results, showing that the matching of luminance contrast (Y_j/Y_w) between test and match image is not independent of illumination level, also argues against gain adjustment in proportion with stimulus intensity. In our opinion the proportionality found by Brainard and Wandell (1992) is probably due to small range linearity. When analyzing the illuminants they used we noted that these produced cone inputs that varied by a factor 2 to 3 at most.

5.7.5 Model implications

The present data, and those of our earlier studies (covering a wide gamut of illuminants), can be quantitatively accounted for. The correlation coefficients of predicted and experimentally obtained cone inputs for our match samples (Q_j^m) are in excess of $\rho = 0.9$ for all three cone classes. We do not think, however, that this already allows strong conclusions about the mechanisms subserving color constancy (as manifested in our experiments). All we can say is that the data are consistent with separate processing of luminance contrast (Y_j/Y_w) and chromatic contrast (\tilde{Q}_j/\tilde{Q}_w). The latter information may be conceived of as being captured by subtraction of the logarithmically transformed inputs \tilde{Q}_j and \tilde{Q}_w (i.e. $\log(\tilde{Q}_j/\tilde{Q}_w) = \log(\tilde{Q}_j) - \log(\tilde{Q}_w)$). Furthermore, within this context, the color signal \tilde{R} , as defined by eq. (5.22), is not only determined by this difference signal, but also by the overall chromatic response, as represented by the cube power of the input from the background ($\tilde{Q}_w^{0.33}$).

According to our data, luminance contrast is not signalled on a completely relative basis. As indicated by eq. (5.25), the contrast signal may result

from incomplete subtraction of the log transformed background signal (Y_w), i.e. $\log(Y_j/Y_w^{0.88}) = \log(Y_j) - 0.88 \log(Y_w)$. In terms of the “Weber-machine” of Koenderink *et al.* (1971) this would imply that the gain control does not aim at complete removal of the background signal.

An important consideration in all analyses of color processing in the visual system, is the transformation from cone outputs into chromatic opponent channels. This stage of signal processing, which may serve the decorrelation and optimal processing of the cone signals (Buchsbaum & Gottschalk, 1983), has been recognized in most models of color vision (e.g. Boynton, 1979; Guth, 1991; Jameson & Hurvich, 1955; Walraven, 1962). Our results can be cast into an opponent-color system, in the same way as this has been done by Worthey (1985) in his analysis of the data of McCann *et al.* (1976). However, since the data predictions do not necessitate this step, this could mean that processes underlying color constancy (as isolated in our experiments) operate prior to the transformation of cone signals into opponent signals. A similar conclusion, based on the same argument, was reached by Brainard and Wandell (1992), who explicitly tested whether a more complex transformation than just a von Kries transformation, was necessary for predicting the data.

Chapter 6

Color Constancy under Conditions Varying in Spatial Configuration

Lucassen, M. P. & Walraven, J. (1992d). Submitted for publication in *Vision Research*.

Abstract

We studied color constancy in relation to variations in the spatial configuration of the stimulus pattern. The latter was generated on a CRT color monitor, and simulated CIE x, y, Y values of Munsell chips (Value 5/), as rendered under two extreme phases of daylight, D_{40} and D_{250} (correlated color temperatures 4000 and 25000 °K). The stimulus pattern consisted of a rectangular 5×7 array of 1.3° square Munsell samples, displayed on a neutral background that was either lighter (N 7.0/), or darker (N 2.5/) than the color samples. The mutual separation, d , between the samples was either 0, 0.037, 0.37 or 1.67° . The stimulus pattern was presented under two alternating illuminant conditions (test and match), that were presented successively to the left and right eye. Two observers matched the color of 11 test samples, presented under D_{40} illumination, to the color of corresponding samples as seen under the D_{250} illuminant. The reflectance of the background adjacent to the central test sample, called the "local surround", was either identical to (Experiment 1) or different (Experiment 2) from that of the background. The observer responses, analyzed in terms of cone-specific inputs to the visual system, were compared with predictions on the basis of a slightly modified version of Land's (1986b) most recent Retinex algorithm, and with the response function derived by Lucassen & Walraven (1992c). The mismatch between model predictions and experiment was expressed as a mean color difference, $\bar{\Delta}E_{uv}^*$. It was found that both models perform reasonably well for decremental stimuli, but somewhat less so for increments. The predictions obtained with the Retinex algorithm were about a factor two less accurate than those derived with the response function. Since the latter takes local contrast across borders as input, the results suggest that color constancy is predominantly determined by processing of local spatial information.

6.1 Introduction

Color constancy is the ability of a visual system to perceive object colors as largely independent of the spectral composition of the illuminant. One of the most challenging, and perhaps least understood, issues in color constancy is how spatially distributed (chromatic) information on the retina is processed by the visual system and transformed into a robust color code. Land's Retinex theory of color vision (Land, 1959, 1986a,b) initiated the first algorithm for the spatial sampling of the retinal light distribution (Land & McCann, 1971). The primary function of the algorithm was to normalize the input of each retinal "pixel" to that of the maximum in the image, and thus arrive at a correlate of reflectance. When separately applied to the three cone systems, the resulting cone-specific reflectances provide the basis for a more or less illuminant-invariant color code. It can be easily shown, by demonstrating the effect of simultaneous contrast (e.g. Shapley, 1986; Todorovic, 1987; Walraven *et al.*, 1991), that the visual system does not compute reflectance, but rather something related to local contrast. This problem has been acknowledged in the most recent version of the Retinex algorithm (Land 1986b), in which more weight is given to the flux from fields adjacent to the image element in question. Since our own modeling of color constancy also relies on local contrast (Lucassen & Walraven, 1992a,b,c), one of the aims of this study is to test to what extent the Retinex and our model yield comparable predictions.

In previous color constancy studies we investigated the individual effects of several experimental parameters (Lucassen & Walraven, 1992 a,b,c; Walraven *et al.*, 1991), but without paying attention to the spatial parameter. Briefly summarizing the results of those studies, it was found that color constancy was never perfect, but the deviations from perfect color constancy could be adequately accounted for by a receptor-specific response function. This nonlinear function, to be presented as eq. (6.1), depends for its input on two variables: chromatic sample to background contrast (where achromatic contrast is filtered out by luminance-normalization), and the relative selectivity of cone stimulation. The latter variable mainly comes to the fore in the short-wave-sensitive (S) cone system, due to its spectrally more isolated spectral sensitivity. Implicit in our model is the assumption that the visual system's behavior could be described in terms of responding to the contrast between a sample and its surround without further considering the activity generated elsewhere in the stimulus pattern. Although

this assumption may seem rather naive, it enabled us to accurately describe observer responses, over a wide range of colored illuminants (Lucassen & Walraven, 1992a), irrespective of their spectral composition (Lucassen & Walraven, 1992b). Also, the effect of contrast and light level, the achromatic variables, usually ignored in models of color constancy, could be accounted for by this simple model (Lucassen & Walraven, 1992c).

In the present study we address the question whether our model may also be used for predicting the effect of illuminant changes for stimulus configurations that are different from those that we used for deriving the model. We used the same experimental method as in our previous studies, that is, simulation (on a color CRT) of an array of Munsell samples under two different illuminants, and successive haploscopic matching of perceived color (sensory matches). In one experiment the main variable of interest is the mutual separation, d , between the samples. For $d=0$, the samples are adjacent to each other (the Mondrian arrangement), whereas for $d > 0$, each sample is surrounded by the (homogeneous) neutral background, configured as a rectangular grid. A limited version of this experiment, using a fixed value for $d > 0$, has been reported by McCann (1987, 1989).

In a second experiment, we manipulated the (simulated) reflectance of the local background adjacent to the test sample (shown in the center of the background). The reflectance could be either higher or lower than that of the background. This experiment can be more or less regarded as the color analogue of a study by Creutzfeldt, Lange-Malecki and Wortmann (1987), who used an achromatic stimulus pattern. That study, like others (Tiplitz-Blackwell & Buchsbaum, 1988b; Walraven, 1973), showed the importance of local rather than global processes. However, other studies warn against ignoring the effect of mechanisms operating over the whole retinal image (McCann, 1987, 1989, 1992). In particular the role of a white or brightest spot in the visual scene may be of importance (McCann, 1992; McCann & Savoy, 1991), as is also most strikingly demonstrated by the well-known Gelb effect (Gelb, 1929).

In recent computational approaches to color constancy (see D'Zmura & Lennie (1986) and Thompson, Palacios & Varela (1992) for reviews), it is typically assumed that spectral information is sampled over large areas, so as to be able to estimate the color of the illuminant (e.g. Buchsbaum, 1980; D'Zmura, 1992). So, in this context too, it is important to test to what extent global (as opposed to local) sampling of chromatic information plays a role in color constancy.

6.2 Method

The experimental paradigm is largely the same as used in one of our earlier studies (Lucassen & Walraven, 1992b). We only introduced two new variables: the mutual separation of the color samples in the stimulus array, and the reflectance of the background area adjacent to the (central) test sample.

6.2.1 Stimuli

The spectral reflectance, $R(\lambda)$, of 36 samples from the Munsell Book of Color (glossy finish) were measured in the range $390 \leq \lambda \leq 730$ nm at 2-nm wavelength intervals with a SpectraScan PR-702AM spectroradiometer (Photo Research). The reflectances were measured relative to a BaSO_4 white, in the $0/45^\circ$ measuring geometry. We selected 30 chromatic samples from three loci of equal Munsell Chroma ($/6$, $/4$ and $/2$) at Value $5/$, and six achromatic samples ranging from Value $2.5/$ through $7.0/$ (see Table 6.1 for a colorimetric specification of the samples under the two illuminants we used.) The samples were arranged in the spatial configuration as shown in Fig. 6.1. It consisted of a rectangular 7×5 array of 1.3° square patches displayed on a homogeneous background ($19.5 \times 14.3^\circ$). The separation between the patches, d , was either 0, 0.037° , 0.37° or 1.67° . For spacing $d = 1.67^\circ$, the outer patches of the stimulus pattern were just inside the boundaries of the background. For $d = 1.3^\circ$, which corresponds to the spacing used in our earlier experiments (but not used here), the stimulus configuration is the same as that shown in Fig. 6.1. The area separating the central patch (No. 18) from its neighboring samples will be referred to as “local surround” (see the hatched area in Fig. 6.1).

6.2.2 Experimental conditions

Two experiments were performed. In Experiment 1 the (neutral) background was presented at two different reflectances, corresponding to that of Munsell chips N $7.0/$ and N $2.5/$. The local surround was presented at the same reflectance, making it indistinguishable from the background. The separation (d) between the samples was varied, in the aforementioned steps from 0 to 1.67° . The sole purpose of this experiment was to study the effect of sample separation, either in the increment mode (low reflectance background), or decrement mode (high reflectance background).

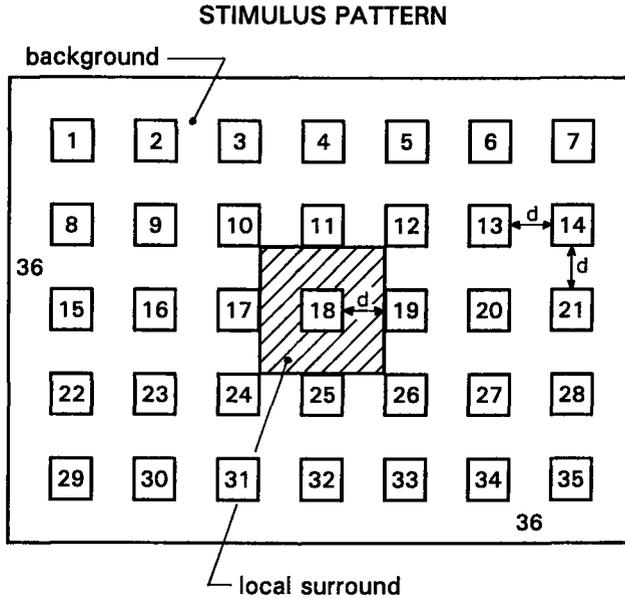


Figure 6.1: Stimulus geometry. The colorimetric specifications of the numbered Munsell samples are presented in Table 6.1. The (square) spacing between the samples, d , was either 0, 0.037, 0.37 or 1.67°. The background area of thickness d ("local surround") around the central patch (No. 18) was either identical to the background (Experiment 1), or different from the background (Experiment 2). Background dimensions: 19.5 × 14.3°.

In Experiment 2 the background had also reflectances of N 7.0/ and N 2.5/, but now the local surround was contrasting with the background, either darker (N 2.5/, combined with background N 7.0/) or brighter (N 7.0/, combined with background N 2.5/). The same values for d were used, except that $d = 0$ (no local surround) was omitted.

Summarizing, in total 14 conditions were used; eight conditions in Experiment 1 (four spacings × two backgrounds), and six conditions in Experiment 2 (three spacings × two background/local surround combinations).

6.2.3 Illuminants

Two daylight illuminants, D_{40} and D_{250} (correlated color temperatures of 4000 and 25000 °K, respectively), were used in our surface color simulation. The relative

spectral radiant power distributions of the daylight illuminants were generated by the CIE method - derived from the principal components analysis of Judd, MacAdam & Wyszecki (1964) - as described in Wyszecki & Stiles (1982). This method takes as input the correlated color temperature (T_c) of a daylight illuminant D, where T_c may range from 4000 to 25000 °K. The output is a spectrum $E(\lambda)$, with λ in steps of 10 nm. In order to obtain the same spectral resolution as in the reflectance measurements (2 nm) we interpolated $E(\lambda)$ at 2-nm intervals. The D_{250} and D_{40} illuminants were used for respectively illuminating the “test” and “match” pattern (the geometry shown in Fig. 6.1). Note, that whereas in most color constancy paradigms the illuminant that is used for the match pattern is a reference white (e.g. Arend & Reeves, 1986; Arend *et al.*, 1991; McCann, McKee & Taylor, 1976; Troost & de Weert, 1991b), we used the (yellow) D_{40} illuminant for the match pattern so as to create a maximum chromatic difference between the test and match pattern (within the range $4000 \leq T_c \leq 25000$ °K).

As discussed before (Lucassen & Walraven, 1992a,c), we chose for a blue/yellow illuminant combination, because this allows for the most critical testing of models of color constancy. Illuminants located along the red/green axis in color space produce relatively little chromatic polarization, due to the large spectral overlap of long- (L) and middle-wave-sensitive (M) cones. Variation of light along the blue-yellow dimension, which strongly modulates the S-cones, is more effective in this respect. This is illustrated in Fig. 6.2.

In Fig. 6.2 we plotted the L, M and S-cone inputs (cf. Lucassen & Walraven, 1992a, Appendix C) for the range of blue (25000 °K) to yellow (4000 °K) color temperatures that comprise a sizable range of daylight chromaticities. It shows the importance of the S-cones for registering the color of the phases of daylight.

The CIE x, y chromaticities and luminance Y of the stimuli under the “test” (D_{250}) and “match” (D_{40}) illuminants are presented in Table 6.1. We computed their X, Y, Z tristimulus values according to the normal procedure (Wyszecki & Stiles, 1982). The numbers that label the samples in Table 6.1 correspond to those in Fig. 6.1. Eleven samples (10 chromatic and one neutral), indicated by an asterisk in Table 6.1, were used as test stimuli.

sample number in Fig. 6.1	simulated Munsell chip	x, y, Y equivalents under illuminant D_{250}			x, y, Y equivalents under illuminant D_{40}		
		x	y	Y	x	y	Y
1	10 YR 5/2	0.2886	0.2972	5.72	0.4254	0.4027	6.02
2 *	5 PB 5/4	0.2193	0.2158	6.25	0.3414	0.3536	5.93
3	10 G 5/2	0.2395	0.2746	5.93	0.3587	0.4028	5.73
4 *	5 P 5/4	0.2385	0.2140	6.00	0.3788	0.3479	6.00
5 *	5 G 5/4	0.2348	0.3081	6.13	0.3433	0.4304	5.79
6	10 Y 5/2	0.2802	0.3131	5.68	0.4109	0.4158	5.83
7 *	5 B 5/4	0.2082	0.2353	6.21	0.3127	0.3740	5.65
8	10 R 5/6	0.3634	0.3050	5.16	0.5031	0.3860	6.10
9	10 Y 5/6	0.3582	0.4590	5.55	0.4578	0.4731	5.90
10	N 3.5/	0.2515	0.2561	2.67	0.3845	0.3839	2.68
11	10 GY 5/2	0.2552	0.2974	5.85	0.3783	0.4148	5.78
12	10 RP 5/6	0.3091	0.2485	5.44	0.4705	0.3566	6.17
13	N 6.5/	0.2509	0.2569	11.04	0.3834	0.3846	11.05
14	10 YR 5/6	0.3849	0.3945	5.55	0.4961	0.4294	6.30
15	10 GY 5/6	0.2642	0.3933	5.95	0.3687	0.4775	5.74
16 *	5 BG 5/4	0.2161	0.2668	6.08	0.3174	0.4040	5.59
17	10 R 5/2	0.2819	0.2720	5.73	0.4233	0.3878	6.06
18 *	N 5/	0.2514	0.2576	6.01	0.3839	0.3851	6.02
19	10 BG 5/2	0.2295	0.2506	6.08	0.3475	0.3864	5.82
20 *	5 R 5/4	0.3032	0.2674	5.58	0.4541	0.3756	6.16
21	10 P 5/6	0.2565	0.2074	5.73	0.4136	0.3344	6.02
22	10 B 5/6	0.1918	0.2022	6.50	0.2858	0.3433	5.69
23	N 2.5/	0.2513	0.2571	1.39	0.3837	0.3850	1.39
24	10 G 5/6	0.2108	0.3049	6.33	0.3017	0.4356	5.69
25	10 P 5/2	0.2534	0.2395	5.75	0.3937	0.3676	5.85
26	N 6.0/	0.2507	0.2567	9.31	0.3832	0.3845	9.32
27	10 PB 5/6	0.2176	0.1856	6.31	0.3473	0.3238	6.07
28	10 BG 5/6	0.1891	0.2390	6.54	0.2706	0.3823	5.64
29 *	5 YR 5/4	0.3311	0.3172	5.43	0.4678	0.4028	6.05
30	10 PB 5/2	0.2377	0.2298	6.02	0.3712	0.3635	5.94
31 *	5 RP 5/4	0.2746	0.2389	5.60	0.4277	0.3590	5.96
32 *	5 GY 5/4	0.2915	0.3714	5.74	0.4094	0.4506	5.82
33	10 RP 5/2	0.2700	0.2553	5.84	0.4133	0.3767	6.09
34 *	5 Y 5/4	0.3317	0.3752	5.59	0.4533	0.4367	5.99
35	10 B 5/2	0.2297	0.2375	6.12	0.3530	0.3740	5.90
36	N 7.0/	0.2507	0.2572	13.22	0.3830	0.3848	13.23
36	N 2.5/	0.2513	0.2571	1.39	0.3837	0.3850	1.39

Table 6.1: Munsell renotations and CIE x, y, Y equivalents of the 30 chromatic and six achromatic samples of the stimulus shown in Fig. 6.1, rendered under test and match illumination. The 11 samples of the test set are indicated by an asterisk.

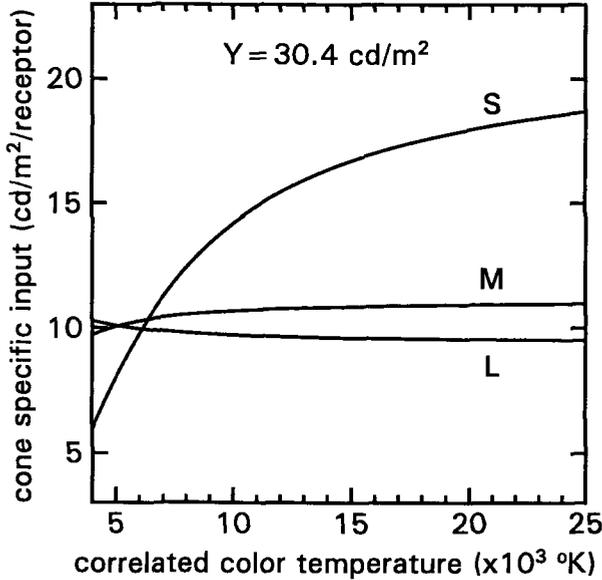


Figure 6.2: Cone stimulation, in units of $\text{cd/m}^2/\text{receptor}$, as a function of the correlated color temperature (T_c) of phases of daylight. Note that, over the range shown here, only the blue (S) cone system is differentially stimulated by the daylight phases.

6.2.4 Simulating surface color on CRT

The appearance of the samples in our simulation is consistent with perfectly diffusing surfaces under a spatially uniform illumination. The x, y, Y equivalents of the samples rendered under the daylight illuminants were displayed on a calibrated high resolution color monitor (Sony, 1152×900 pixels), controlled by a Sun 3/260 computer (24 bit/color). The calibration procedure for the monitor, and the colorimetric equations required for displaying specified x, y, Y values on a color monitor, have been published elsewhere (Lucassen & Walraven, 1990).

6.2.5 Stimulus presentation

Two displays were used: a test pattern (the samples as arranged in Fig. 6.1) under the D_{250} illuminant, and a match pattern of identical geometry illuminated by D_{40} . A viewing box (1 m length), blackened inside, was placed in front of the monitor. A mechanical shutter system was located just behind the two viewing holes. When the left viewing hole was open, the right one was occluded by

the shutter, and vice versa. The interchangement of the two viewing conditions was under control of the computer. The colors belonging to the test and match pattern were changed during the switching time of the shutters, which only took a fraction of a second. In this way, each eye was locked to one of the two successive illuminant conditions (test or match) to be compared. The role of the two eyes, as test or match eye, was alternated from trial to trial, to prevent possible long term effects of chromatic adaptation. Also, the sequence of the different experimental conditions of Experiments 1 and 2 were randomized for the two observers.

After about five minutes of dark adaptation and a few more minutes for adapting to the average luminance and color of the test pattern, the observer started the first presentation of the two illuminant conditions to be compared during a session. When viewing the test (left or right eye) and match pattern (right or left eye) the observer concentrated on the central patch. The color of the latter, which could be controlled during the presentation of the match pattern, was initially black. Every five seconds, the computer activated the shutter system for switching between test and match condition. The color of the matching sample was adjusted by mouse movements, which were translated by the computer as movements through CIE x, y color space. Two of the three mouse buttons were pressed for increasing or decreasing the luminance of the patch at constant x, y chromaticities. The third mouse button was pressed for indicating that a satisfactory match had been obtained, after which the next test patch was presented (in total 11 samples, in pseudo-random order). Even for unexperienced subjects, this matching procedure is easy to comprehend and requires only a few training sessions to obtain reliable results.

One of the authors (ML), and a (naive) observer who had participated before in one of our earlier color constancy experiments (AV), served as subjects. They both had normal color vision as confirmed by the standard color tests.

6.2.6 Task

The observer adjusted the central patch in the match pattern to match the perceived hue, saturation and brightness of the corresponding sample in the test pattern as well as possible. He was free to make eye movements and to use as many test/match alternations as were necessary to obtain a satisfactory match. Neither of the observers found it difficult to always find a satisfactory match.

6.3 Model predictions

Algorithms for dealing with the spatial variable in color constancy are typically rooted in the Retinex model and/or related lightness models (Land, 1983, 1986b; Land & McCann, 1971; McCann & Houston, 1983; McCann et al., 1976). The inputs to these algorithms is derived from each (unit) retinal location within the separate classes of photoreceptors. The algorithm samples the retinal receptor array in search of discontinuities that are then ranked along a common lightness scale. See for a more in depth discussion, reports by Hurlbert (1986) and Brainard and Wandell (1986).

In our preceding studies on color constancy we were able to quantify the data without having to introduce any parameters for describing the spatial configuration of our stimulus. This was not that surprising since spatial configuration was not a variable in our experiments. This no longer applies to the data obtained in the present study, so we can now test to what extent the model needs to be expanded to deal with the spatial variable. More specifically, we shall compare the predictions of our model, which only responds to local contrast, with the prediction of a model that computes color on the basis of the light distribution over the whole area of the stimulus pattern. The latter model is Land's latest version of the Retinex (Land, 1986b). In the following we shall discuss the procedures for deriving the predictions on the basis of our own "response function" and those obtained with Land's Retinex model.

6.3.1 Response function

As discussed in Lucassen & Walraven (1992c), the results of our earlier experiments could be described by a cone-specific response function (\tilde{R}) of the form

$$\tilde{R} = (\tilde{Q}_w)^{0.33} \log(2.6 \frac{\tilde{Q}_j}{\tilde{Q}_w}) \quad (6.1)$$

where \tilde{Q} represents the quantum catch in a given cone class, normalized for luminance (Y):

$$\tilde{Q} = \frac{Q}{Y}. \quad (6.2)$$

The subscripts j and w in eq. (6.1) indicate the input from test sample and (white) background, respectively. The latter is spectrally flat, and thus reflects the chromaticity of the illuminant. As eq. (6.1) shows, \tilde{R} is a function that incorporates two variables, representing luminance-normalized cone stimulation of the background (\tilde{Q}_w) and chromatic contrast (\tilde{Q}_j/\tilde{Q}_w). The response function

has to be applied to each of the three cone types, both for the test and match eye. The prediction of the matching sample, in terms of absolute cone inputs, is obtained by the following procedure (see Lucassen & Walraven (1992c) for more details):

1. Equate the test and match eyes' responses, $\tilde{R}^t = \tilde{R}^m$, where superscripts t and m denote test and match, respectively. These superscripts have to be applied to each element in eq. (6.1).
2. Solve for \tilde{Q}_j^m and compute its value for the L, M and S cone system:

$$\tilde{Q}_j^m = \frac{\tilde{Q}_w^m}{2.6} \left[2.6 \frac{\tilde{Q}_j^t}{\tilde{Q}_w^t} \right]^{(\tilde{Q}_w^t/\tilde{Q}_w^m)^{0.33}} \quad (6.3)$$

3. Compute the luminance of the matching sample, Y_j^m , according to

$$Y_j^m = Y_j^t \left(\frac{Y_w^m}{Y_w^t} \right)^{0.88} . \quad (6.4)$$

4. Multiply \tilde{Q}_j^m with Y_j^m to obtain the absolute values of Q_j^m for each cone class.

The above procedure for predicting Q_j^m (the observer's match), assumes that the visual system's color constant behavior can be modeled on the basis of local contrast (Q_j/Q_w). So, for the data predictions with the response function, we always used the cone input values of the *local surround* for Q_w , except for the $d = 0$ condition. In the latter condition, there is no longer a unique surround bordering on the central patch in Fig. 6.1. In that particular condition we computed the average cone input along a contour located 1 pixel outside the borders of the central patch, and substituted the average value for Q_w . This average is identical to the average cone input of patches No. 11, 17, 19 and 25 (Fig. 6.1) which are adjacent to patch No. 18 for $d = 0$.

6.3.2 Retinex algorithm (Land, 1986b)

In Land's (1986b) most recent version of the Retinex, the computation of a triplet of cone-specific lightnesses (the designators) is implemented by sampling the visual scene with a "two-aperture" light integrator. The designators, ρ^p ($p=L,M,S$),

are computed as the logarithm of the ratio of two fluxes, respectively computed for a 4' field and a 24° field, here expressed as

$$\rho^p = \log \left(\frac{F_1^p}{F_2^p} \right) \quad (6.5)$$

in which F_1^p and F_2^p represent the average flux (for cone class p) incident on the 4' field and the 24° field, respectively. The field sizes are empirical and were suggested by Land (1986b) in conjunction with a $1/r^2$ cone sensitivity profile, where r denotes the distance to the point of interest. We used a slightly modified version in which the $1/r^2$ sensitivity profile was replaced by a Gaussian type, $e^{-r^2/\sigma}$, but with σ such that 50% of the area integrated under the Gaussian is obtained within a 4° field, as was the case in Land's version (see Fig.3 in Land, 1986b). We took the Gaussian profile instead of $1/r^2$ to evade the problem of infinite sensitivity when r approaches zero. This modification does not alter the principle underlying eq. (6.5), and should hardly affect the outcome of the predictions (see Hurlbert (1986) and Moore, Allman & Goodman (1991) for a comparison of the two profiles). When using eq. (6.5) for predicting our data we proceed as follows. The designators for the test samples in the test pattern should be identical to the corresponding color matches in the match pattern. So, we may write

$$\log \left(\frac{F_1^{p,t}}{F_2^{p,t}} \right) = \log \left(\frac{F_1^{p,m}}{F_2^{p,m}} \right) . \quad (6.6)$$

Note the superscripts t and m for test and match, respectively.

In order to compute ρ^p , the Gaussian sensitivity profile was centered on the central patch of the 35-patch array (No. 18 in Fig. 1), the test sample that has to be matched by the observer. Since the 24° field is larger than our stimulus pattern, we assigned zero fluxes to the unit areas outside the stimulus. For F_1^p , the flux captured by the small aperture, we take the sample's cone input Q_j^p , since the Gaussian sensitivity profile is flat within the (small) 4' field. So, for $F_1^{p,m}$, which equals the match input $Q_j^{p,m}$ (the quantity to be predicted), we may write

$$Q_j^{p,m} = F_1^{p,m} = F_2^{p,m} \frac{F_1^{p,t}}{F_2^{p,t}} . \quad (6.7)$$

There is one slight problem with this equation. The value of $F_2^{p,m}$, the flux sampled with the large (24°) aperture, relates to the match sample; this is an unknown variable, to be set by the subject. However, $F_2^{p,m}$ can be separated into two components, $F_{2,j}^{p,m}$ and $F_{2,\text{rest}}^{p,m}$, which denote the (unknown) contribution of the matching sample (j) and the (known) remainder of the 24° field to the

integral, respectively:

$$F_2^{p,m} = F_{2,j}^{p,m} + F_{2,\text{rest}}^{p,m} = A_2^{-1} \left[\int_{\phi=0}^{2\pi} \int_{r=0}^{r_1} r e^{-r^2/\sigma} Q_j^{p,m}(r, \phi) dr d\phi + \int_{\phi=0}^{2\pi} \int_{r=r_1}^{r_2} r e^{-r^2/\sigma} Q_j^{p,m}(r, \phi) dr d\phi \right] \quad (6.8)$$

where A_2 denotes the area occupied by the 24° field. In the right-hand side of eq. (6.8), r_1 is the distance from the center of the matching sample to its boundary, and r_2 the radius (12° for the 24° field). Since $Q_j^{p,m}(r, \phi)$ is constant within the area of the matching sample, eq. (6.8) can be rewritten as

$$F_2^{p,m} = A_2^{-1} Q_j^{p,m} \int_{\phi=0}^{2\pi} \int_{r=0}^{r_1} r e^{-r^2/\sigma} dr d\phi + F_{2,\text{rest}}^{p,m}. \quad (6.9)$$

Substitution of eq. (6.9) in eq. (6.7) gives

$$Q_j^{p,m} = \left[A_2^{-1} Q_j^{p,m} \int_{\phi=0}^{2\pi} \int_{r=0}^{r_1} r e^{-r^2/\sigma} dr d\phi + F_{2,\text{rest}}^{p,m} \right] \frac{F_1^{p,t}}{F_2^{p,t}}. \quad (6.10)$$

The prediction of the cone input of the matching sample can now be made explicit by solving $Q_j^{p,m}$ from eq. (6.10):

$$Q_j^{p,m} = \frac{F_{2,\text{rest}}^{p,m} \left(\frac{F_1^{p,t}}{F_2^{p,t}} \right)}{1 - A_2^{-1} \int_{\phi=0}^{2\pi} \int_{r=0}^{r_1} r e^{-r^2/\sigma} dr d\phi \left(\frac{F_1^{p,t}}{F_2^{p,t}} \right)} \quad (6.11)$$

We implemented this algorithm in a computer program¹. In order to reduce the amount of time needed by the computer to carry out all computations, the algorithm was run on the stimuli on a 1:2 scale (576×450 pixels). This still required about 90 minutes for each stimulus condition (11 samples) on a Sun 3 workstation.

Before applying the algorithm to the stimuli that were actually used in Experiments 1 and 2, it was calibrated on a uniform grey test and match stimulus (Fig. 6.1, with all samples Munsell N 7.0/, and illuminated by D₆₅). This provided a gain factor by which F_2^p was multiplied in order to obtain equal values for F_1^p and F_2^p , and hence, zero values for the designators in this condition. Next, the algorithm was tested on the actual stimulus pattern for the “physical match”

¹The integrals in eqs. (6.9-6.11) are written here in general form in polar coordinates (r, ϕ) . Since the stimulus pattern contains rectangular patches, the computations were carried out in Cartesian coordinates (x, y) . The integrand in eqs. (6.9-6.11) then takes the form $e^{-(x^2+y^2)/\sigma} dx dy$.

condition, that is, the condition in which the test and match illuminant are identical (D_{65} in this case). The predictions of the cone inputs of the matches, as computed with eq. (6.11), were identical to the cone inputs of the test samples, as should be the case for an ideal observer making perfect color matches.

6.3.3 Data format

The main purpose of this study is to test model predictions of color constancy, for conditions that vary with respect to the spatial configuration of the stimulus pattern. The bulk of the data will therefore be presented in terms of comparisons of experimentally obtained sample matches versus predicted matches. The difference between obtained and predicted color of the matching samples can be expressed in various ways. Ideally one should like to measure color differences in a color space where the distance between two colors corresponds to their perceived difference. Such a perceptually uniform color space has yet to be constructed, but a reasonable and widely used first order approximation is the well known CIE 1976 ($L^*u^*v^*$) color space (CIE, 1971; Robertson, 1977). Therefore, we shall use the Euclidean distance unit in CIE 1976 ($L^*u^*v^*$) space, ΔE_{uv}^* , as the yardstick for measuring the error of our model predictions.

For each experimental condition (11 data points), $\bar{\Delta E}_{uv}^*$ represents the mean color difference between the experimentally obtained color matches (subscripts j, exp) and the colors predicted by the model in question (subscripts $j, pred$) :

$$\bar{\Delta E}_{uv}^* = \frac{1}{11} \sum_{j=1}^{11} \left((L_{j,exp}^* - L_{j,pred}^*)^2 + (u_{j,exp}^* - u_{j,pred}^*)^2 + (v_{j,exp}^* - v_{j,pred}^*)^2 \right)^{1/2} \quad (6.12)$$

with L_j^* , u_j^* , and v_j^* defined as

$$L_j^* = 116 \left(\frac{Y_j}{Y_n} \right)^{1/3} - 16 \quad (Y_j/Y_n > 0.01) \quad (6.13)$$

$$u_j^* = 13L_j^* (u'_j - u'_n) \quad (6.14)$$

$$v_j^* = 13L_j^* (v'_j - v'_n) . \quad (6.15)$$

The quantities u'_j, v'_j and u'_n, v'_n are given by

$$u'_j = \frac{4X_j}{X_j + 15Y_j + 3Z_j} \quad v'_j = \frac{9Y_j}{X_j + 15Y_j + 3Z_j} \quad (6.16)$$

$$u'_n = \frac{4X_n}{X_n + 15Y_n + 3Z_n} \quad v'_n = \frac{9Y_n}{X_n + 15Y_n + 3Z_n} \quad (6.17)$$

where the tristimulus values X_n, Y_n, Z_n are those of the nominally white object-color stimulus. The latter is usually given by the spectral radiant power of a standard CIE illuminant, reflected by a perfect diffuser. We used the standard D_{65} illuminant for that purpose, hence, $X_n=28.9, Y_n=30.4, Z_n=33.1$, and $u'_n=0.1979, v'_n=0.4683$.

The procedure for arriving at the final results ($\bar{\Delta}E_{uv}^*$) involves the following steps:

1. Transformation from the observers' R,G,B settings (the color matches) to CIE x, y, Y units (Lucassen & Walraven (1992a), Appendix B).
2. Computation of the model predictions in terms of L,M,S cone inputs. Transformation of those L,M,S units into CIE x, y, Y units (Lucassen & Walraven (1992a), Appendix C).
3. Transformation from x, y, Y (color matches and model predictions) to L^*, u^*, v^* according to eqns. (6.13-6.17).
4. Computation of $\bar{\Delta}E_{uv}^*$ according to eq. (6.12)

6.4 Results

6.4.1 Precision of matches and experimental variables

Before discussing the size of the errors ($\bar{\Delta}E_{uv}^*$) for the (two) model predictions, something should be said with respect to the intrinsic error of the method (the precision of haploscopic matching), and the extent to which our experimental variables affect color constancy.

An example of the (x, y) precision of the matches is given in Fig. 6.3, which shows the results for the condition in which the two eyes viewed the stimulus pattern under the same illuminant.

In Fig. 6.3 perfect matching would be indicated by coinciding open and closed circles. The deviations shown here correspond to an average error ($\bar{\Delta}E_{uv}^*$) in the order of 3.5 CIELUV units. In the experimental conditions with different test and match illuminants, the average error was somewhat higher, in the order of 4 to 5 CIELUV units.

As for the size of the effects that we measured, these are well in excess of the uncertainty range of the matches. Figure 6.4 shows an example.

The data shown in Fig. 6.4 relate to conditions that differ with respect to sample separation ($d=0$ or 1.67°) and background reflectance (N 7.0/ or N 2.5/). This results in different degrees of color constancy as indicated by the degree of

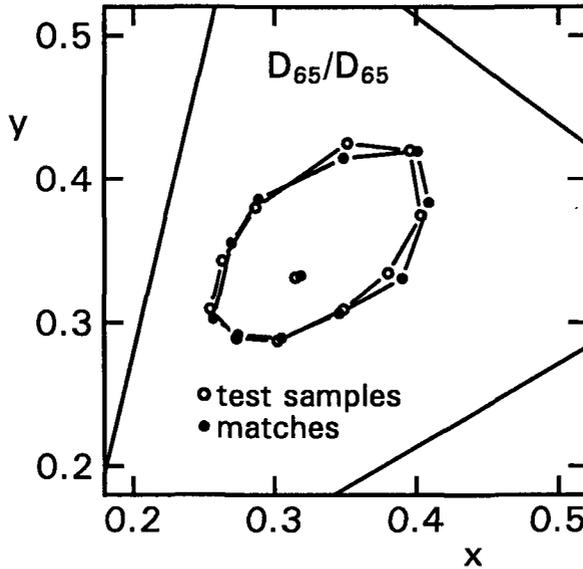


Figure 6.3: Result of a training session ($d=0.37^\circ$, N 2.5/ background). D_{65} was used for illuminating both the test and match pattern. Open circles: chromaticities of the test samples. Closed circles: chromaticities of the matches to the test samples. Perfect color matching would be indicated by coinciding closed and open circles.

overlap between filled circles (hatched area) and open circles. Note that perfect color constancy would imply coincidence of open circles and closed circles, whereas the opposite would be indicated by coincidence of closed circles and open squares (physical identity match). The matches (hatched area) fall in between those two extreme positions, indicating partial color constancy. As discussed elsewhere (Lucassen & Walraven, 1992a,b), this is the rule rather than the exception in studies on color constancy.

Comparison of the left and right panel in Fig. 6.4 shows color constancy to be better for the condition $d=0$, N 7.0/ (left panel) than for the condition $d=1.67^\circ$, N 2.5/ (right panel). We shall return to this result in the Discussion; for the moment it only serves to show that the stimulus variables we introduced, do have an effect on color constancy.

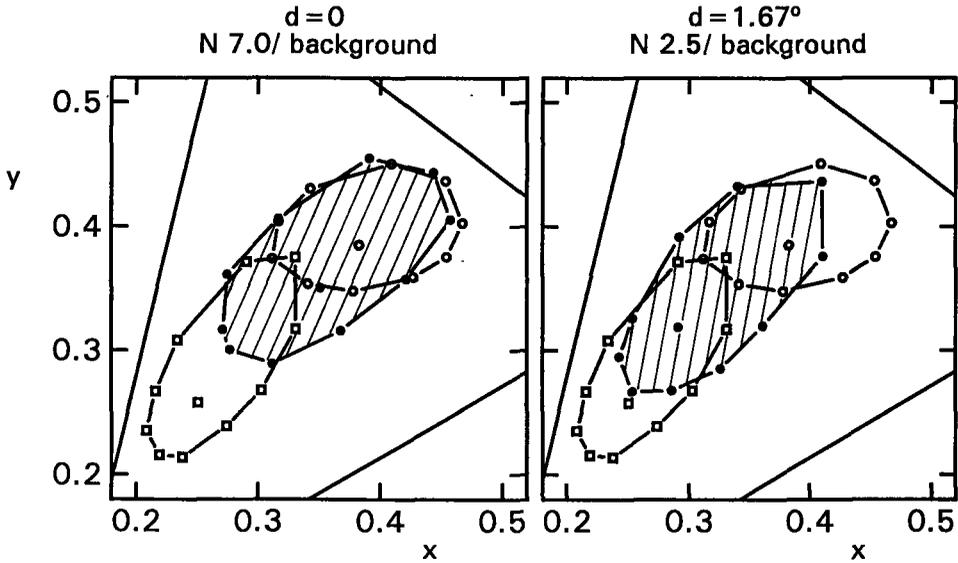


Figure 6.4: Example of how stimulus composition may effect the degree of color constancy. Left and right panel show data obtained in conditions, differing in both background reflectance (N 7.0/, N 2.5/) and sample separation (0° , 1.67°). Open symbols represent the test stimuli, as seen under blue (squares) and yellow (circles) phases of daylight, respectively. Filled symbols (hatched area) represent matches (averaged), made under yellow light, of test samples, seen under blue light. Perfect color constancy would be indicated by superimposed filled and open circles, as is more closely approximated in the panel on the left.

6.4.2 Experiment 1

In this experiment the samples were surrounded by a homogeneous background (either lighter or darker), and their mutual separation was varied from 0 to 1.67 degrees of visual angle. For each condition (eight in total) we computed the average deviation ($\bar{\Delta}E_{uv}^*$) of the model predictions, as obtained for our response function and the Retinex algorithm. The $\bar{\Delta}E_{uv}^*$ values, plotted separately for each model and each observer, are shown in Fig. 6.5. The performance of the response model (panels on the left) are about twice as accurate as those obtained with the Retinex model. The average error in the predictions of the response function are in the order of 10 CIELUV units, which is not such a bad result, when considering that this is about twice the standard error of the matches.

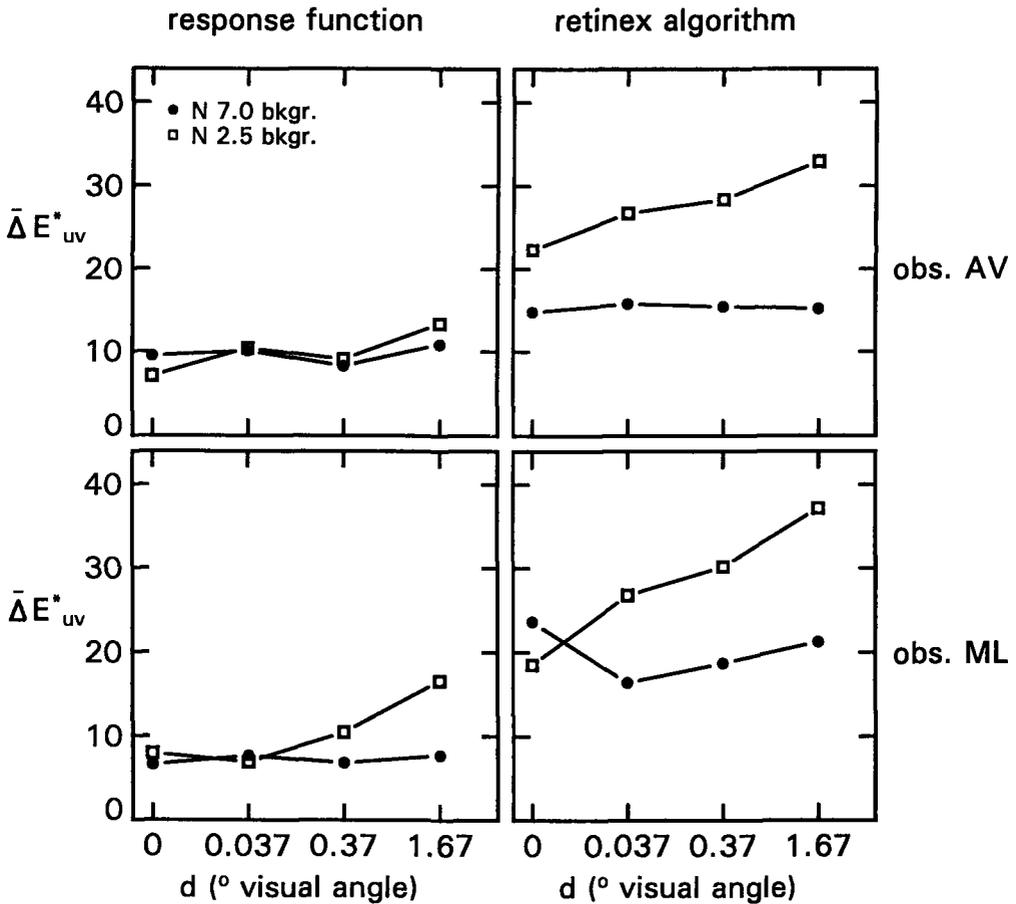


Figure 6.5: Results from Experiment 1 for obs. AV (top panels) and obs. ML (bottom panels). Mean color difference, $\bar{\Delta} E_{uv}^*$, computed according to eq. (6.12), as a function of stimulus spacing, d . Left panels: predictions with Lucassen & Walraven's (1992c) response function. Right panels: predictions with Land's (1986b) Retinex algorithm. Closed circles indicate the results for the N 7.0/ background; open squares indicate those for the N 2.5/ background conditions. Note that the abscissae are not linear.

It is of interest that $\bar{\Delta} E_{uv}^*$ is relatively low and quite constant, for the condition in which the samples are presented as decrements (filled symbols in Fig. 6.5). In

the incremental mode the predictions of the Retinex become increasingly worse with increasing sample separation, a tendency also reflected, but much less so, in the predictions of the response function. It was rather unexpected that the Retinex model, which incorporates spatial sampling, has more difficulty in dealing with the spatial variable (d), than the response function, which is sensitive only to local contrast.

6.4.3 Experiment 2

In Experiment 2 the area adjacent to the (central) test sample, the local surrounding, contrasted with the rest of the background (see Fig. 6.1). When the latter was relatively bright (N 7.0/), the local surround was dark (N 2.5/), and vice versa. We were interested to see what the influence was of the local surround as compared to that of the rest of the stimulus pattern. Note that the area of the local surround varies with the sample separation (d). For $d=0$ there is no local surround, so for that condition the data are the same as obtained in Experiment 1.

The results of Experiment 2 are shown in Fig. 6.6, which is constructed in the same way as Fig. 6.5. When comparing these results with those shown in Fig. 6.5, the values of $\bar{\Delta}E_{uv}^*$ are of about the same order of magnitude. For the response function the performance is now even better, because the slight deterioration for increments at increasing separation (obs. ML in Fig. 6.5), has now completely disappeared. A similar improvement, in the sense that the $\bar{\Delta}E_{uv}^*$ discrepancy between decrements and increments has become smaller, is observed for the Retinex model.

These results suggest that the introduction of a local surround, introducing a spatial discontinuity between test sample and the rest of the stimulus pattern, improves the performance of both models. This result is consistent with the fact that the two models are both sensitive to local contrast, either exclusively (response function), or predominantly so (Retinex model). Considering the better performance of the response function, it would seem that the Retinex model still puts a too strong weight on the contribution of distant areas in the stimulus pattern.

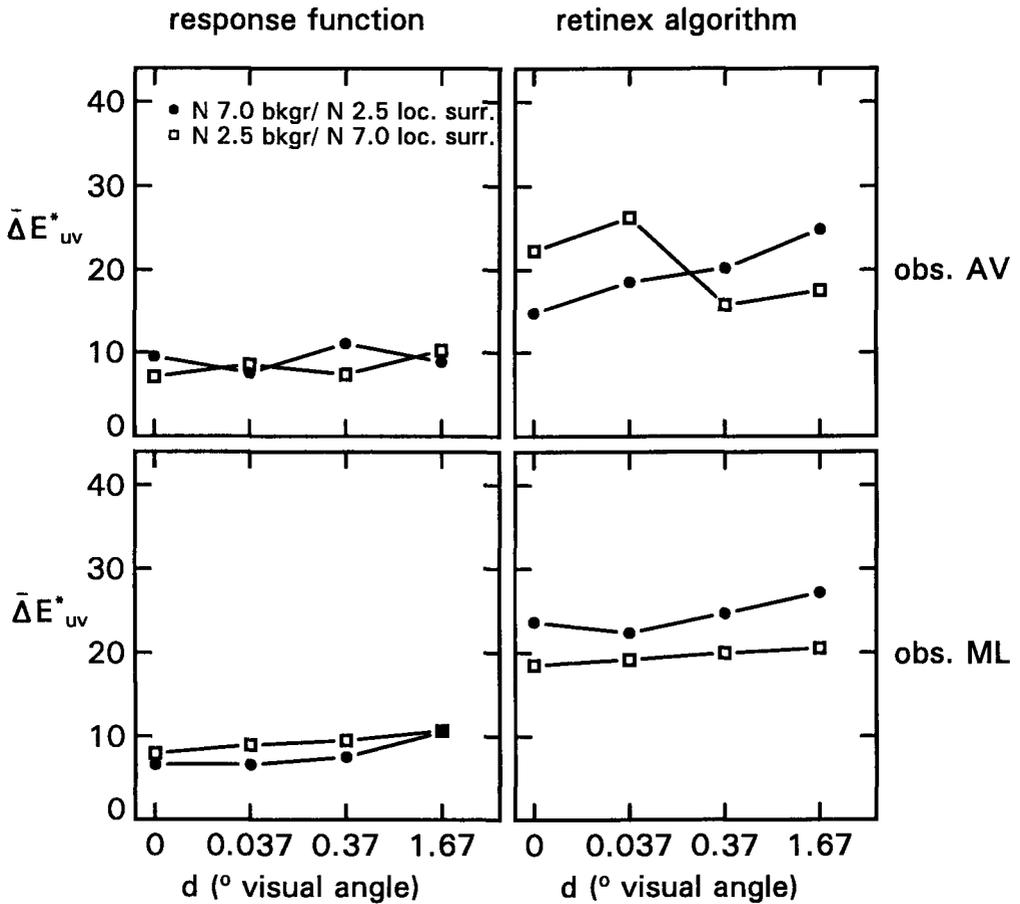


Figure 6.6: Same as Fig. 6.5, but now for the results of Experiment 2. The data points at $d = 0$ were taken from Fig. 6.5.

6.5 Discussion

The main result of the experiments reported here, is that predictions based on our response function can accommodate the effects of the quite sizable changes in spatial configuration that we introduced. That function assumes only local

processes operating on the test sample, whereas the more sophisticated Retinex algorithm is also sensitive to contributions from distant areas in the stimulus pattern. Nevertheless, the predictions of the latter model are much less accurate than those obtained with the response function. This does not mean that the principle of this most recent Retinex version is incorrect. It certainly has physiological validity, and was probably inspired by Land's collaboration with Livingstone and Hubel (1984), and Zeki (1980). The algorithm, which is also the starting point of the algorithm of Hurlbert and Poggio (1988), resembles the action of a receptive field with a narrow center and a very large surround, consistent with the "non-classical" receptive fields of V4 cells reported by Desimone *et al.* (1985). Reports of long-range (cortical) color interactions are also relevant in this respect (Land *et al.*, 1983; Pöppel, 1986; Wehrhahn, Heide & Peterson, 1989).

It is possible that the Retinex algorithm only requires "fine-tuning" of the Gaussian sensitivity profile (σ in eq. (6.11)), so as to make its predictions conform with our results. Moore *et al.* (1991), who implemented the algorithm in a real-time neural (video) network, suggested an extension by taking into account "edginess" of the image. Whatever the effect of such operations, our results suggest that the algorithm should aim at obtaining results resembling those obtained with our response function. Of course, the latter is not perfect, but the error is within reasonable bounds. If we take as just-noticeable-difference (jnd) the size of the intrinsic error of the matching task ($\Delta E_{uv}^* \approx 3.5$), then the error of the model predictions is in the order of three jnd's.

The results of this study show that the degree of color constancy is affected by spatial configuration. This is shown in Fig. 6.4 which shows that color constancy is relatively weak if the test samples are separated by a dark grid. This confirms the results of other studies (Karp, 1959; Tiplitz-Blackwell & Buchsbaum, 1988a,b; Walraven *et al.*, 1991) and is also consistent with the Helson-Judd effect (Helson, 1938; Judd, 1940), in the sense that the latter states that samples lighter than the (dark) background, take on the hue of the illuminant. This kind of result is difficult to account for by (computational) models of color constancy that sample the color content of an image (for estimating the illuminant) without reference to the spatial distribution of the color stimuli.

Since the experimental paradigm that we used does not include a simultaneous comparison of different spatial configurations (both eyes see the same pattern), it is less well suited for studying the effect of this variable on color *appearance*. Color appearance and color constancy relate to different topics in color vision, but they are clearly related, of course. The effect of spatial variables on color appearance has been frequently investigated in studies on (a)chromatic induction (e.g. Brenner & Cornelissen, 1991; Kinney, 1962; Newson, 1958; Oyama & Hsiu, 1966; Reid & Shapley, 1988; Tiplitz-Blackwell & Buchsbaum, 1988a,b; Walraven,

1973; Whipple, Wallach & Marshall, 1988; Yund & Armington, 1975). The spatial variable most frequently studied is the separation between test and inducing field and for the area of the latter. The result of such an experiment, when properly corrected for stray light (cf. Walraven, 1973), typically shows that the induction effect is almost entirely dictated by a narrow region (up to a width of 30') adjacent to the border of the test field. This suggests that the induction effect is due to the activity generated by neural structures subserving abstraction of edges. In a recent review of border effects on brightness perception (Kingdom & Moulden, 1989), edge-detectors were also considered as the most likely candidates of such effects; more so than the global spatial sampling associated with the Retinex algorithm of Land and McCann (1971).

As mentioned already, the relatively good performance of our response function is consistent with local rather than global spatial interactions. However, since the predictions are not perfect, there is still room for an additional, possibly global kind of interaction. Experiments by McCann & Savoy (1991), on lightness matching in spatially different stimulus arrangements, clearly show the presence of two mechanisms: the "normal" inducing effect, rapidly declining with separation, and a long-distance effect, reminiscent of the Gelb (1929) effect. The combined action of these two effects can also be seen in the chromatic domain (Walraven, 1973, Fig. 9). As noted above, the present experiments employing the same spatial configuration for test and match condition, are not suited for analyzing to what extent non-local spatial interactions may have contributed to our results. All we can say is that a simple response function, processing only local contrast information, may already provide a fairly accurate account of color constancy under spatially varying conditions.

Summary

Color constancy is the ability of a visual system to perceive object colors as fairly independent of the spectral composition of the illuminant. The studies comprising this thesis are directed at the quantification and modeling of human color constancy. The experimental method involves the simulation of illuminant-object interactions on a calibrated color CRT. During an experiment, a test pattern of colored “paper” samples was shown under two different illuminants, and presented successively to the left and right eye. The observer’s task was to match the color of samples seen under the “match” illuminant (seen by the “match” eye), to the corresponding samples seen under the “test” illuminant (seen by the “test” eye).

Chapter 2 addresses the problem of colorimetric calibration of the computer controlled color CRT that was used for generating the stimuli. As full calibration of such devices is time-consuming, and offers only temporary validity, a simple recalibration procedure was developed that only requires a single measurement of a reference white, just prior to an experimental session. An evaluation of the performance of the recalibration procedure showed that it enabled color control with an accuracy in the order of a just perceptible chromaticity difference.

In Chapter 3, the effect of illuminant chromaticity on the perceived color of simulated surface samples was studied. A trichromatic illuminant-object interaction was simulated, which represented a simplification of the interaction in the real world. In the latter, the light reflected from an object is characterized by the product of the spectral distribution of the light source and surface reflectance, $E(\lambda) \times R(\lambda)$. Hence, it requires the complete wavelength functions to be known. In contrast, the trichromatic simulation (using the RGB phosphors as primaries) only requires three illuminant-specific emission coefficients, and three sample-specific reflectance coefficients, which allows for a considerable freedom in the choice of illuminants and objects to be investigated. The results showed

that color constancy was never perfect, but the departures from perfect constancy were very systematic. These departures could be explained by assuming a non-linear cone-specific response function, that depends on two factors: sample to background contrast and absolute level of cone stimulation. The latter factor, which mainly shows up in the S-cone system, distinguishes the response function from other color constancy models, such as the Retinex algorithm and the von Kries adaptation.

Chapter 4 dealt with color constancy under natural and artificial illumination. The experimental method was the same as that used in Chapter 3, except that the illuminant-object interaction was now simulated in the wavelength rather than the RGB domain. That is, the stimuli now represented “real” Munsell samples, either under natural illumination (phases of daylight), or under (metameric) artificial illumination (light composed of only two wavelengths). In the latter condition, color constancy may be expected to deteriorate, due to the impoverished spectral characteristics of the illuminant. Data predictions were shown, computed on the basis of two models. One of the models was the response function derived in Chapter 3, the other was a computational model, based on an algorithm for generating smoothest (i.e. natural) reflectance functions (van Trigt, 1990). The computational model aims at the recovery of complete reflectance spectra, thereby requiring information about the (estimated) spectral distribution of the illuminant. The response function, operating in a strictly trichromatic domain (the quantum catches of the cones) is not concerned with spectral recovery. The data predictions with the response function were about a factor two more accurate than those of the computational model, confirming the general applicability of the simple response function under various types of illuminations, and questioning the physiological validity of the much more complex computational approach.

In Chapter 5 two experiments are discussed that explore the effects of the achromatic stimulus variables, luminance and luminance contrast. In the first experiment, illumination was fixed, but the reflectance of the samples and the background were varied such that the luminance contrast between samples and background - kept constant in earlier experiments - covered a two log-unit range, comprising both decrements and increments. In the second experiment, the color and intensity of the illuminants was manipulated, either to specifically stimulate the S-cone system, or all three cone types simultaneously. A mathematical equivalence was discovered - only valid over a limited stimulus range - between our response function (derived in Chapter 3) and the model of Jameson and Hurvich

(1964) for brightness contrast. The two experiments of Chapter 5 went beyond the range over which the equivalence between the two models holds, and therefore allowed a critical test between the models. We compared data predictions with these two models, showing that the model of Jameson and Hurvich (1964) has less difficulty than the response function in dealing with luminance contrast for stimuli brighter than the background (increments). However, a modified version of the response functions could be derived, that performs better than the Jameson and Hurvich model, and still accounted for the data of the experiments discussed in Chapters 3 and 4. The modification existed in a luminance-normalization of the cone signals, thus resulting in a model in which color and luminance information are separately processed.

The last chapter, Chapter 6, focuses on the spatial parameter in color constancy. A set of experiments were performed for studying the effect of varying the spacing between the samples in the stimulus pattern. Also, the effect of changing the immediate surround of the test samples was studied. The experimental data were compared with two model predictions, one of which was the improved response function derived in Chapter 5. The other model was the most recent version of the Retinex algorithm (Land, 1986b), which includes spatial sampling over the entire visual scene. The data predictions with the Retinex algorithm were about a factor two less accurate than those with the improved response function. Since the latter operates on the basis of local contrast, the results suggest that color constancy is mainly determined by the processing of local spatial information.

Samenvatting

Kleurconstantie is het vermogen van een visueel systeem om object-kleuren als vrijwel onafhankelijk waar te nemen van de spectrale samenstelling van de lichtbron. Het onderzoek dat in dit proefschrift wordt beschreven is gericht op het meten en modelleren van kleurconstantie van het menselijke visuele systeem. Bij de experimentele methode die gebruikt werd voor het kwantificeren van kleurconstantie werden illuminant-object interacties gesimuleerd op een gekleurd kleurenscherm. Tijdens een experiment werd een testpatroon van gekleurde “papieren” proefvlakjes afwisselend getoond onder twee verschillende lichtbronnen, respectievelijk aan het linker- en aan het rechteroog. De proefpersoon had tot taak om de kleuren van de proefvlakjes, zoals waargenomen onder respectievelijk “test” en “match” verlichting (door “test” en “match” oog), aan elkaar gelijk te stellen.

In Hoofdstuk 2 werd aandacht besteed aan de colorimetrische ijking van een computer-gestuurde kleurenbeeldbuis die gebruikt werd voor het genereren van de stimuli. Aangezien volledige calibratie van zo'n systeem een tijdrovende aanlegging is, en slechts een beperkte geldigheidsduur kent, werd een her-ijkingsprocedure ontwikkeld - uit te voeren voorafgaand aan een experimentele sessie -, die slechts één meting van een referentiewit vereist. In een evaluatie van de prestaties van deze procedure werd aangetoond dat hiermee kleuren gereproduceerd kunnen worden met een nauwkeurigheid in de orde van een juist waarneembaar kleurverschil.

In Hoofdstuk 3 werd de invloed bestudeerd van de kleur van de lichtbron op de waargenomen kleur van gesimuleerde kleurvlakjes. Bij de computer simulatie werd uitgegaan van een synthetische wereld waarin de lichtreflectie zich afspeelt binnen de drie afzonderlijke RGB fosfor “lichtkanalen” van het beeldscherm. Deze zgn. trichromatische illuminant-object interactie, is een vereenvoudiging van de werkelijke (fysische) interactie tussen lichtbron en object. In werkelijkheid wordt

het door een object gereflecteerde licht gekarakteriseerd door het product van de spectrale verdeling van de lichtbron en de spectrale reflectie van het object: $E(\lambda) \times R(\lambda)$. Daarvoor zijn de complete golflengte-functies vereist. De trichromatische simulatie vereist daarentegen slechts drie illuminant-specifieke emissie-coëfficiënten en drie object-specifieke reflectie-coëfficiënten, hetgeen een aanzienlijke keuzevrijheid creëert voor de te onderzoeken lichtbronnen en objecten. De resultaten laten zien dat kleurconstantie nooit perfect is, maar de afwijkingen van perfecte constantie waren zeer systematisch. Deze afwijkingen konden goed voorspeld worden door uit te gaan van een niet-lineaire, receptor-specifieke responsfunctie, die reageert op het receptor-specifieke contrast en het absolute niveau van receptor stimulatie. Deze laatste factor onderscheidt de responsfunctie van andere kleurconstantie modellen, zoals het Retinex algoritme en de von Kries adaptatie.

Hoofdstuk 4 behandelt kleurconstantie onder natuurlijke en kunstmatige lichtbronnen. Dezelfde experimentele methode werd toegepast als in Hoofdstuk 3, behalve dat de illuminant-object interactie nu niet werd gesimuleerd in het RGB domein, maar in het golflengte domein. Er werd uitgegaan van “echte” Munsell samples onder hetzij natuurlijke lichtbronnen (fasen van het daglicht), of daarop gelijkende (metamere) kunstmatige lichtbronnen (licht van slechts twee golflengten). Bij het gebruik van het kunstmatige licht mag verwacht worden dat kleurconstantie volledig faalt, tengevolge van de povere spectrale informatie in de lichtbron. Data predicties werden berekend op basis van twee modellen. Eén van de modellen was de responsfunctie zoals afgeleid in Hoofdstuk 3, het andere was een computationeel model, gebaseerd op een algoritme waarmee de meest gladde (natuurlijke) reflectiespectra worden gegenereerd (van Trigt, 1990). Het computationele model streeft naar de reconstructie van complete reflectiespectra. De responsfunctie opereert uitsluitend in het trichromatische domein (de lichtabsorptie in de kegeltjes) en heeft als zodanig geen “bemoeienis” met spectrale reconstructies. De data-predicties met de responsfunctie waren ongeveer een factor twee nauwkeuriger dan die met het computationele model. Dit resultaat bevestigt de algemene toepasbaarheid van de simpele responsfunctie onder verschillende soorten lichtbronnen, en zet vraagtekens bij de fysiologische relevantie van de veel complexere computationele modellen.

In Hoofdstuk 5 worden twee experimenten besproken waarin het effect wordt onderzocht van de achromatische stimulus variabelen, contrast en lichtniveau. In het eerste experiment lagen de lichtbronnen vast, maar werd de reflectie van

de kleurvlakjes en de achtergrond zodanig gevarieerd dat het luminantiecontrast tussen kleurvlakje en achtergrond twee log-units bestreek, met daarin zowel positief als negatief contrast. In het tweede experiment werden de kleuren en intensiteiten van de lichtbronnen gemanipuleerd om hetzij uitsluitend het S-kegel systeem, of alle drie de kegelsystemen tegelijkertijd, differentiëel te stimuleren. Er werd een mathematische equivalentie aangetoond - zij het slechts geldend voor een beperkt stimulus domein - tussen de respons functie (afgeleid in Hoofdstuk 3) en het model van Jameson en Hurvich (1964) voor helderheidscontrast. De twee experimenten van Hoofdstuk 5 speelden zich af buiten het geldigheidsgebied van de equivalentie, en vormen zodoende een kritische test voor de twee modellen. Een vergelijking van de data-predicties met deze twee modellen, laat zien dat het Jameson en Hurvich (1964) model minder problemen heeft met variaties in luminantiecontrast, van stimuli die helderder zijn dan de achtergrond (incrementen). Er kon echter een gemodificeerde versie van de respons functie worden afgeleid die beter voorspelt dan het Jameson en Hurvich model, en bovendien de experimentele data van Hoofdstukken 3 en 4 kan beschrijven. De modificatie bestaat uit een luminantie-normering van de kegelsignalen, met als resultaat een model waarin kleur- en luminantiecontrast gescheiden worden verwerkt.

Het laatste Hoofdstuk (6) is gericht op de rol van de spatiële variabele in kleurconstantie. Een aantal experimenten werd uitgevoerd met een variabele onderlinge afstand tussen de kleurvlakjes van het testpatroon. In samenhang hiermee werd ook het effect onderzocht van verandering van de directe omgeving van de kleurvlakjes. De experimentele data werden vergeleken met twee modelpredicties. Eén van de modellen is de verbeterde respons functie die werd afgeleid in Hoofdstuk 5. Het andere model is de meest recente versie van het Retinex algoritme (Land, 1986b), dat uitgaat van een bemonstering van de gehele visuele scène. De data-predicties met het Retinex algoritme bleken ongeveer een factor twee minder nauwkeurig te zijn dan die van de verbeterde respons functie. Aangezien de respons functie werkt op basis van de codering van lokaal contrast, doen de resultaten vermoeden dat kleurconstantie hoofdzakelijk wordt bepaald door locale spatiële informatie.

Epilogue

The work described in this thesis is restricted to studying the *sensory* aspect of color constancy. Undoubtedly, one also has to consider the role of what might be broadly called “cognition”. However, although this aspect has been shown to be important for our understanding of color constancy (e.g. Arend & Reeves, 1986; Troost, 1992), we feel that the potential of “hard-wired” sensory processing, has still not been fully explored. Here we discuss some possible lines for future research on color constancy, including its relevance for the application sector.

Extending the stimulus

The chromatic and achromatic variables

As a first step towards further exploration we take a look at the rather limited selection of test samples that we used for the stimulus patterns. These were (simulated) Munsell chips from the Munsell 5/4 series, so these samples were varying in hue but not in saturation. It might be useful, therefore, to also perform experiments with both more and less saturated test samples, selected for example from the Munsell 5/2 and 5/6 series.

An increase or decrease in saturation of the test samples is accompanied by an increase or decrease in the stimulus *range* of chromatic contrast. When plotted in a cone-specific contrast diagram, as shown in Fig. 3.11, we expect that the data points belonging to the set of higher or lower saturation would still be located on the same line fitted to the illumination condition in question. For example, in the “trivial” W/W condition shown in the right graph of Fig. 3.11, these data points would still fall along the dashed identity line, but with the data covering another range of contrasts.

An other stimulus variable that has not been sufficiently explored, is the absolute level of illumination. As already discussed in the Introduction, a disadvantage of using a color CRT is the restricted luminance domain that can be investigated (about two log-units). Considering the nonlinear response of the visual system

to increasing light level, it would be interesting to determine to what extent our present modeling of color constancy is capable of correctly predicting observer responses under conditions in which illumination level can be manipulated over, say, five log-units.

In its present form, the model does not predict an effect of absolute illumination level on the chromatic response, since the luminance component is filtered out by the normalization procedure ($\tilde{Q} = Q/Y$). However, the achromatic response *does* respond to absolute luminance, in the sense that perceived luminance contrast increases slightly with illumination level (Y_w), as described by eq. (5.25).

At least two problems will be encountered when investigating the effect of illumination level on color constancy. When illumination level is continuously decreased, photopic vision is gradually replaced by mesopic and scotopic vision, which implies a shift from cone vision to rod vision. Color constancy then reduces to lightness constancy. At the other extreme, the light level at which the cone response starts to saturate, color vision may break down due to the fact that an "unbalance" in the three cone systems cannot be realized anymore, all systems being driven to their respective limits. Although different mechanisms are involved at the two luminance extremes, their visual effects may nevertheless be the same: desaturation. When visualized in color space, chromaticities at both extremes will converge onto white. This implies that somewhere in between these luminance boundaries of color vision, there has to be an optimum; that might be halfway the photopic range of natural light levels in which we have evolved.

The spatial variable

The 2D scenario

Another extension of the experiments reported here would be provided by experiments in which stimulus configuration is varied. Chapter 6 dealt with varying the spacing between the samples, but still, sample size was fixed. If sample size turns out to be an important parameter in color constancy, this might be reflected in a change of the "constant" k of eq. (5.9), in the same way as it did for luminance contrast (Chapter 5).

Sample size is an important consideration with regard to the phenomenon of assimilation, which has the opposite effect of simultaneous contrast. When stimuli become small, the associated colors seem to spread out, merging with the background, rather than contrast with it. The interplay of contrast and assimilation is one of the most neglected areas in spatial (color) vision. Interested readers are referred to de Weert (1991).

The 3D scenario

In the long run, natural images are preferred above Mondrian-like patterns for studying the full potential of color constancy, even if the Mondrians may keep color scientists busy for quite a while. As a first step to a more complete visual stimulus, one may consider to introduce the third dimension. This would include stereopsis, and hence, binocular fusion, of images. Such experiments will be difficult to implement in combination with our haploscopic matching paradigm which requires successive monocular viewing. However, dynamic (monocular) parallax as a depth cue could be used instead of binocular parallax. The next step could be, then, to compare the results from such experiments with those obtained with true stereoscopic vision.

With respect to the way in which the visual system uses three-dimensional information, there is still much to learn from the studies that have been conducted in the context of lightness and brightness constancy (Gilchrist, 1977; Schirillo, Reeves & Arend, 1990; Schirillo & Shevell, 1992). The most important conclusion that can be drawn from such studies is that the third dimension is an important cue for separating light from matter. This can be demonstrated, for example, by monocularly viewing of a folded piece of white cardboard. In the absence of sufficient monocular depth cues, the fold in the cardboard may appear to be inverted in depth. The side of the cardboard that is turned away from the predominant direction of illumination will reflect less of the incident light, and will be perceived as such (the perception of a shadow) if seen in the proper three-dimensional context. However, when the fold appears to be inverted in depth, the shadowed side, which now seems to face the light from which it is actually shielded, will not be perceived as a shadow, but rather as if the material has been painted black. So, a change in incident light is now interpreted as a change in surface reflectance. This "illusion", which is quite compelling, shows how a perceived change in the (achromatic) color domain is evoked by a perceptual change in the spatial domain.

Practical considerations

Fidelity of color coding

The quantification of color constancy is not only important for its understanding, but also has consequences for practical applications. An important application filed in this respect, is the color coding of information on electronic displays, by now probably the most important man-machine interface. Designers and users of color-coded displays may be confronted with perceptual "artifacts", that often

find their origin in the way the visual system processes color information (e.g. Walraven, 1985, 1991). For example, a green symbol on a blue background may be perceived as yellow, a “mistake” that can be attributed to the visual system’s misdirected attempt at color constancy (Walraven & Lucassen, 1991). Surround induced color changes, like the aforementioned, should be avoided, of course, if color coding is to be effective as such. It should be possible, in principle, to use the response function we derived for developing an algorithm that adjusts the RGB specification of a color coded symbol, so as to retain its proper color when presented against differently colored backgrounds.

The feasibility of the above application has already been shown while playing with a computer program that demonstrates the predictive power of the response function. The program generates two center-surround patterns (test and match), shown side by side on the (calibrated) color monitor. The colors and luminances of the test pattern can be adjusted at will, as well as the color and luminance of the surround in the match pattern. The program generates a color in the center of the match pattern, that according to the model, matches color and luminance in the test pattern (under the restriction that the prediction falls within the color gamut of the monitor). Ideally, the centers in the test and match pattern would have the same color appearance. We found that the model is quite successful in predicting the center in the match pattern, over a fairly large portion of the color space covered by the monitor. Deviations from perfect prediction that we noted could be attributed to the fact that the two patterns were displayed simultaneously, instead of successively (as in the experiments).

Screen to printer color reproduction

Another field of applied color science where the response function could be useful, is that of matching hard-copy output of color printers to colored images on electronic displays. This problem is not so easy to solve, since it involves different techniques of creating color, and depends on environmental conditions (room lighting, for instance). The colors on an electronic display are generated by additive mixing of the light emitted by the RGB primaries, whereas the colors on the hard-copy are produced by subtractive mixing of dyes. The light reflected from a color print not only depends on the choice and amounts of toner primaries that are printed on the paper, but also on the ambient light, which may vary because of changes in natural lighting (phases of daylight, clouds) or the addition of artificial light (fluorescent tube). This is where color constancy becomes an important consideration, as has also been recognized by the Commission Internationale d’Eclairage (CIE), which has led to the formation of a CIE Technical Committee (TC 1-27) that specifically addresses the problem of maintaining color

fidelity in the process of “soft to hard-copy” color reproduction. What is needed for solving this kind of problem, is device independent color processing, which means that for both the electronic display and the printer, the input-output relations have to be known in terms of standardized output quantities (CIE *XYZ* units, for example). This is a technical (calibration) problem, which can be solved, of course. In Chapter 2 we described an acceptable solution to the problem of CRT calibration. What remains is the selection of printer paper, toner and room lighting. In as far as lighting is concerned, we have shown in Chapter 4 that the interaction between light and selected materials, can be simulated (and studied) on a color CRT. Simulating printed matter on a color monitor may well be one of the more promising roads to solving the matching problem. When the quantum catches (the effective inputs to the visual system) are known that result from the combination of a certain light source and (toner) reflectance spectra, the response function may be used to predict the associated visual effect.

Machine vision

Much of the current interest in color constancy can be traced to its obvious relevance for machine vision. Machines that can see are confronted by the same problem as biological visual systems: they should be capable of discriminating changes in surface reflectance from changes in illumination.

Depending on the application, different principles may be used. Often there is no need to resort to look for the biological solution. As a matter of fact, non-biological approaches may often be preferable, in particular when the machine operates in conditions where it can use stored information (a reference reflectance, for example). For less predictable visual scenarios, however, the successful color constancy algorithm still has to be developed.

Several neural networks have been proposed, that, like the model derived in this thesis, encode color information in terms of contrast at boundaries (e.g. Grossberg, 1987), or some kind of spatial operator that relates locally recorded input to all other inputs in the image. The retinex algorithm (Land, 1964, 1986) was the precursor of this approach, of which there is even already a hard-wired implementation (Moore, Allman & Goodman, 1991). Still, we have a long way to go, before such algorithms will be capable of handling natural 3D images in which the light is not homogeneously distributed, but is modulated by a complex of spatial variables.

References

1. Arend, L. E. (1973). Spatial differential integral operations in human vision: implications of stabilized retinal image fading. *Psychological Review*, 80, 374-395. *Journal of the Optical Society of America A*, 8, 661-672.
2. Arend, L. E. & Goldstein, R. (1987). Simultaneous constancy, lightness, and brightness. *Journal of the Optical Society of America A*, 4, 2281-2285.
3. Arend, L. E. & Goldstein, R. (1990). Lightness and brightness over spatial illumination gradients. *Journal of the Optical Society of America A*, 7, 1929-1936.
4. Arend, L. E. & Reeves, A. (1986). Simultaneous color constancy. *Journal of the Optical Society of America A*, 3, 1743-1751.
5. Arend, L. E., Reeves, A., Schirillo, J. A. & Goldstein, R. (1991). Simultaneous color constancy: papers with diverse Munsell values. *Journal of the Optical Society of America A*, 8, 661-672.
6. Benzschawel, T. (1992). Colorimetry of self-luminous displays. In: Widdel, H. & Post, D. L. (Eds.), *Color in Electronic Displays*. Plenum press, New York and London, 86-87.
7. Benzschawel, T., Walraven, J. & Rogowitz, B. E. (1987). Studies of color constancy. *Investigative Ophthalmology and Visual Science (Suppl.)*, 28, 92.
8. Borges, C. F. (1991). Trichromatic approximation method for surface illumination. *Journal of the Optical Society of America A*, 8, 1319-1323.
9. Boynton, R. M., (1979). *Human Color Vision*. Holt, Rinehart & Winston, New York, 211-215.
10. Boynton, R. M. & Whitten, D. N. (1972). Selective chromatic adaptation in primate photoreceptors. *Vision Research*, 12, 855-874.
11. Brainard, D. H. (1989). Calibration of a computer controlled color monitor. *Color Research and Application*, 14, 23-34.
12. 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.

13. Brainard, D. H. & Wandell, B. A. (1991). A bilinear model of the illuminant's effect of color appearance. In: Movshon, J. A. & Landy, M. S. (Eds), *Computational models of visual processing*, MIT Press, Cambridge, Mass.
14. Brainard, D. H. & Wandell, B. A. (1992). Asymmetric color matching: how color appearance depends on the illuminant. *Journal of the Optical Society of America A*, *9*, 1433-1448.
15. Brenner, E. & Cornelissen, F. W. (1991). Spatial interactions in color vision depend on distances between boundaries. *Naturwissenschaften*, *78*, 70-73.
16. Brill, M. H. (1978). A device performing illuminant-invariant assessment of chromatic relations. *Journal of Theoretical Biology*, *71*, 473-478.
17. Brill, M. H. & West, G. (1986). Chromatic adaptation and color constancy: a possible dichotomy. *Color Research and Application*, *11*, 196-204.
18. Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute*, *310*, 1-26.
19. Buchsbaum, G. & Gottschalk, A. (1983). Trichromacy, opponent colours coding and optimum colour information transmission in the retina. *Proceedings of the Royal Society of London B*, *220*, 89-113.
20. Burkhardt, D. A., Gottesman, J., Kersten, D. & Legge, G. E. (1984). Symmetry and constancy in the perception of negative and positive luminance contrast. *Journal of the Optical Society of America A*, *1*, 309-316.
21. CIE (1971). Colorimetry. Publication No. 15, Bureau Central de la CIE, Paris.
22. Cohen, J. (1964). Dependency of the spectral reflectance curves of the Munsell color chips. *Psychonomic Science*, *1*, 369-370.
23. Cornsweet, T. N. & Pinsker, H. M. (1965). Luminance discrimination of brief flashes under various conditions of adaptation. *Journal of Physiology (London)*, *176*, 294-310.
24. Cowan, W. B. (1983). An inexpensive scheme for calibration of a colour monitor in terms of CIE standard coordinates. *Computer Graphics*, *17*, 315-321.
25. Cowan W. B. (1986). CIE calibration of video monitors. (available from National Research Council of Canada, Ottawa, Ontario, K1A 0R6, Canada.)
26. 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.
27. Creutzfeldt, O., Lange-Malecki, B. & Wortmann, K. (1987). Darkness induction, retinex algorithm and cooperative mechanism in vision. *Experimental Brain Research*, *63*, 270-283.

28. Creutzfeldt, O. D., Lee, B. B. & Elepfandt, A. (1979). A quantitative study of chromatic organization and receptive fields of cells in the lateral geniculate body of the rhesus monkey. *Experimental Brain Research*, *35*, 527-545.
29. Dannemiller, J. L. (1989). Computational approaches to color constancy: adaptive and ontogenetic considerations. *Psychological Review*, *96*, 255-266.
30. Dannemiller, J. L. (1991). Lightness is not illuminant invariant: reply to Troost & de Weert (1990). *Psychological Review*, *98*, 146-148.
31. Dannemiller, J. L. (1992). Spectral reflectance of natural objects: how many basis functions are necessary? *Journal of the Optical Society of America A*, *9*, 507-515.
32. Desimone, R., Schein, S. J., Moran, J. & Ungerleider, L. G. (1985). Contour, color, and shape analysis beyond the striate cortex. *Vision Research*, *25*, 441-452.
33. Dixon, E. R. (1978). Spectral distribution of Australian daylight. *Journal of the Optical Society of America*, *68*, 437-450.
34. D'Zmura, M. (1992). Color constancy: surface color from changing illumination. *Journal of the Optical Society of America A*, *9*, 490-493.
35. D'Zmura, M. & Lennie, P. (1986). Mechanisms of color constancy. *Journal of the Optical Society of America A*, *3*, 1662-1672.
36. Eastman, A. A. & Brecker, S. A. (1972). The subjective measurements of color shifts with and without chromatic adaptation. *Journal of Illuminating Engineering Society*, *2*, 239.
37. 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.
38. Enroth-Cugell, C. & Robson, J. G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology*, *187*, 517-522.
39. Estévez, O. (1979). On the fundamental data-base of normal and dichromatic color vision, Ph.D. dissertation, University of Amsterdam.
40. Fairchild, M. D. (1991). Formulation and testing of an incomplete chromatic adaptation model. *Color Research and Application*, *16*, 243-250.
41. Fairchild, M. D. & Lennie, P. (1992). Chromatic adaptation to natural and incandescent illuminants. *Vision Research*, *32*, 2077-2085.
42. Forsyth, D. A. (1990). A novel algorithm for color constancy. *International Journal of Computer Vision*, *5*, 5-36.
43. Foster, D. H., Craven, B. J. & Sale, R. H. (1992). Immediate colour constancy. *Ophthalmic and Physiological Optics*, *12*, 157-160.
44. Gelb, A. (1929). Die "Farbenkonstanz" der Sehdinge. In: Bethe, A. (Ed.), *Handbuch der Normalen und Pathologischen Physiologie*, Vol. 12. Berlin: Springer.

45. Gershon, R. & Jepson, A. D. (1989). The computation of color constant descriptors in chromatic images. *Color Research and Application*, *14*, 325-334.
46. Georgeson, M. A. (1984). Contrast overconstancy. *Journal of the Optical Society of America A*, *8*, 579-586.
47. Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. *Science*, *195*, 185-187.
48. Gilchrist, A. L. (1988). Lightness contrast and failures of constancy: a common explanation. *Perception & Psychophysics*, *43*, 415-424.
49. Gilchrist, A. L., Delman, S. & Jacobsen, A. (1983). The classification and integration of edges as critical to the perception of reflectance and illumination. *Perception & Psychophysics*, *33*, 425-436.
50. Grossberg, S. (1987). Cortical dynamics of three-dimensional form, color, and brightness perception. *Perception & Psychophysics*, *41*, 87-158.
51. Guth, S. L. (1991). Model for color vision and light adaptation. *Journal of the Optical Society of America A*, *8*, 976-993.
52. Hayhoe, M. M., Benimoff, N. I. & Hood, D. C. (1987). The time-course of multiplicative and subtractive adaptation process. *Vision Research*, *27*, 1981-1996.
53. Helmholtz, H. L. F. von (1911). *Handbuch der physiologischen Optik* (3rd ed., 1st ed. 1866). W. Nagel, A. Gullstrand, and J. von Kries (Eds.). Voss, Hamburg and Leipzig.
54. Helson, H. (1938). Fundamental problems in color vision. I. The principle governing changes in non-selective samples in chromatic illumination. *Journal of Experimental Psychology*, *23*, 439-476.
55. Helson, H. (1943). Some factors and implications of color constancy. *Journal of the Optical Society of America*, *33*, 555-567.
56. Helson, H. & Jeffers, V. B. (1940). Fundamental problems in color vision. II. Hue, lightness and saturation of selective samples in chromatic illumination. *Journal of Experimental Psychology*, *26*, 1-29.
57. Hering, E. (1878). *Grundzüge der Lehre vom Lichtsinne*. Gerolds Sohn, Vienna.
58. Hess, C. & Pretori, H. (1894). Messende Untersuchungen über die Gesetzmässigkeit des simultanen Helligkeits-contrastes. *Archives of Ophthalmology*, *40*, 1-24.
59. Ho, J., Funt, B. V. & Drew, M. S. (1990). Separating a color signal into illumination and surface reflectance components: theory and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*, 966-977.
60. Hunt, R. W. G. (1991). Revised Colour-appearance model for related and unrelated colours. *Color Research and Application*, *16*, 146-165.

61. Hurlbert, A. C. (1986). Formal connections between lightness algorithms, *Journal of the Optical Society of America A*, *3*, 1684-1693.
62. Hurlbert, A. C. (1991). Deciphering the colour code. *Nature*, *349*, 191.
63. Hurlbert, A. C. & Poggio, T. A. (1988). Synthesizing a color algorithm from examples. *Science*, *239*, 482-485.
64. Jacobsen, A. & Gilchrist, A. L. (1988). The ratio principle holds over a million-to-one range of illumination. *Perception & Psychophysics*, *43*, 1-6.
65. Jameson, D. & Hurvich, L. M. (1955). Some quantitative aspects of an opponent-colors theory. I. Chromatic responses and spectral saturation, *Journal of the Optical Society of America*, *45*, 546-552.
66. Jameson, D. & Hurvich, L. M. (1961). Complexities of perceived brightness. *Science*, *133*, 174-179.
67. Jameson, D. & Hurvich, L. M. (1964). Theory of brightness and color contrast in human vision, *Vision Research*, *4*, 135-154.
68. Jameson, D. & Hurvich, L. M. (1989). Essay concerning color constancy, *Annual Review of Psychology*, *40*, 1-22.
69. Judd, D. B. (1940). Hue, saturation and lightness of surface colors with chromatic illumination. *Journal of the Optical Society of America*, *30*, 2-32.
70. Judd, D. B. (1951). Colorimetry and artificial daylight. *Proceedings 12th Session CIE, Stockholm, Vol. 1 (TC7) Bureau Central de la CIE, Paris*, 11.
71. Judd, D. B. (1960). Appraisal of Land's work on two-primary color projections. *Journal of the Optical Society of America*, *50*, 254-268.
72. Judd, D. B., MacAdam, D. L. & Wyszecki, G. (1964). Spectral distribution of typical daylight as a function of correlated color temperature. *Journal of the Optical Society of America*, *54*, 1031-1040.
73. Kaplan, E. & Shapley, R. M. (1986). The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proceedings of the National Academy of Science, USA*, *83*, 2755-2757.
74. Karp, A. (1959). Colour-image synthesis with two unorthodox primaries. *Nature*, *184*, 710-712.
75. Katz, D. (1911). Die Erscheinungsweisen der Farben und ihrer Beeinflussung durch die individuelle Erfahrung. *Zeitschrift für Psychologie, Ergbd.*, *7*.
76. Kelly, D. H. (1969). Flickering patterns and lateral inhibition. *Journal of the Optical Society of America*, *59*, 1361-1365.
77. Kingdom, F. & Moulden, B. (1989). Corner effect in induced hue: evidence for chromatic band-pass filters. *Spatial Vision*, *4*, 253-266.
78. Kingdom, F. & Moulden, B. (1991). A model for contrast discrimination with incremental and decremental test patches. *Vision Research*, *31*, 851-858.

79. Kinney, J. A. S. (1962). Factors affecting induced color. *Vision Research*, *2*, 503-525.
80. Koenderink, J. J., van de Grind, W. A. & Bouman, M. A. (1971). Foveal information processing at photopic luminances. *Kybernetik*, *8*, 128-144.
81. 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.
82. Krauss, S. (1926). Das Farbensehen in bunter Beleuchtung und die Farbenkonstanz der Sehdinge. *Zeitschrift für Psychologie*, *96*, 58.
83. Kries, J. von (1905). Die Gesichtsempfindungen. In: *Handbuch der Physiologie des Menschen*, Vol 3, 109-282 (Ed. W. Nagel). Vieweg, Braunschweig.
84. Land, E. H. (1959a). Color vision and the natural image, Part I. *Proceedings of the National Academy of Science, USA*, *45*, 115-129.
85. Land, E. H. (1959b). Color vision and the natural image, Part II. *Proceedings of the National Academy of Science, USA*, *45*, 636-644.
86. Land, E. H. (1964). The retinex. *American Scientist*, *52*, 247-264.
87. Land, E. H. (1974). The retinex theory of colour vision. *Proceedings of the Royal Institution, Gr. Br.*, *47*, 23-58.
88. Land, E. H. (1977). The retinex theory of color vision. *Scientific American*, *237*, 108-128.
89. Land, E. H. (1983). Recent advances in retinex theory and some implications for cortical computations : colour vision and the natural image. *Proceedings of the National Academy of Science, USA*, *80*, 5163-5169.
90. Land, E. H. (1986a). Recent advances in retinex theory. *Vision Research*, *26*, 7-21.
91. 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, USA*, *83*, 3078-3080.
92. Land, E. H., Hubel, D. H., Livingstone, M. S., Perry, S. H. & Burns, M. M. (1983). Colour-generating interactions across the corpus callosum. *Nature, Lond.*, *303*, 616-618.
93. Land, E. H. & McCann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America A*, *61*, 1-11.
94. 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.
95. Livingstone, M. S. & Hubel, D. H. (1984). Anatomy and physiology of a color system in primate primary visual cortex. *Journal of Neuroscience*, *4*, 309-356.

96. Lucassen, M. P. & Walraven, J. (1990). Evaluation of a simple method for color monitor recalibration. *Color Research and Application*, 15, 321-326.
97. Lucassen, M. P. & Walraven, J. (1992a). Quantifying color constancy: evidence for nonlinear processing of cone-specific contrast. *Vision Research* (in press).
98. Lucassen, M. P. & Walraven, J. (1992b). Color constancy under natural and artificial illumination (submitted to *Vision Research*).
99. Lucassen, M. P. & Walraven, J. (1992c). Separate processing of chromatic and achromatic contrast in color constancy (submitted to *Vision Research*).
100. Maloney, L. T. (1986). Evaluation of linear models of surface spectral reflectance with small number of parameters. *Journal of the Optical Society of America A*, 3, 1673-1683.
101. Maloney, L. T. (1992). A mathematical framework for biological vision. *Behavioral and Brain Sciences*, 15, 45.
102. 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.
103. McCann, J. J. (1987). Local/global mechanisms for color constancy. *Die Farbe*, 34, 275-283 (1987).
104. McCann, J. J. (1989). The role of simple nonlinear operations in modeling human lightness and color sensations. *Human vision, visual processing, and digital display*, SPIE Vol. 1077, 355-363.
105. McCann, J. J. (1992). Rules for color constancy. *Ophthalmic and Physiological Optics*, 12, 175-177.
106. 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.
107. 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.
108. McCann, J. J. & Savoy, R. L. (1991). Measurements of lightness: dependence on the position of a white in the field of view. *Human vision, visual processing, and digital display II*, SPIE Vol. 1453, 402-411.
109. McManus, P. & Hoffman, G. (1985). A method for matching hard copy color to display color, *SID 85 Digest*, 204.
110. Monasterio, de F. M. & Gouras, P. (1975). Functional properties of ganglion cells of the rhesus monkey retina. *Journal of Physiology*, 251, 167-195.
111. Moore, A., Allman, J. & Goodman, R. M. (1991). A real-time neural system for color constancy. *IEEE Transactions on Neural Networks*, 2, 237-247.

112. Naka, K. I. & Rushton, W. A. (1966). S-potentials from colour units in the retina of fish. *Journal of Physiology*, *185*, 587-599.
113. Nayatani, Y., Takahama, K., Sobagaki, H. & Hashimoto, K. (1990). Color appearance model and chromatic-adaptation transform. *Color Research and Application*, *15*, 210-221.
114. Newson, L. J. (1958). Some principles governing changes in the apparent lightness of test surfaces isolated from their normal backgrounds. *Quarterly Journal of Experimental Psychology*, *10*, 82-95.
115. Oyama, T. & Hsia, Y. (1966). Compensatory hue shift in simultaneous color contrast as a function of separation between inducing and test fields. *Journal of Experimental Psychology*, *71*, 405-413.
116. Parkkinen, J. P. S., Hallikainen, J. & Jaaskelainen, T. (1989). Characteristic spectra of Munsell colors. *Journal of the Optical Society of America A*, *6*, 318-322.
117. Pöppel, E. (1986). Long-range colour-generating interactions across the retina. *Nature, Lond.*, *320*, 523-525.
118. Post, D. L. & Calhoun, C. S. (1987). An evaluation of methods for producing specific colors on CRTs. *Proceedings of the Human Factors Society 31st Annual Meeting, Human Factors Society, Santa Monica, CA*, 1276-1280.
119. Post, D. L. & Calhoun, C. S. (1989). An evaluation of methods for producing desired colors on CRT monitors. *Color Research and Application*, *14*, 172-186.
120. Reeves, A., Arend, L. E. & Schirillo, J. A. (1989). Color constancy in isolated displays. *Perception*, *18*, 1743-1751.
121. Reid, R. C. & Shapley, R. W. (1988). Brightness induction by local contrast and the spatial dependence of assimilation. *Vision Research*, *28*, 115-132.
122. Reid, R. C. & Shapley, R. W. (1992). Spatial structure of cone inputs to receptive fields in primate lateral geniculate nucleus. *Nature*, *356*, 716-717.
123. Robertson, A. R. (1977). The CIE 1976 color-difference formulae. *Color Research and Application*, *2*, 7-11.
124. Rubin, J. M. & Richards, W. A. (1982). Color vision and image intensities: when are changes material? *Biological Cybernetics*, *45*, 215-226.
125. Rushton, W. A. H. (1965). Visual adaptation. *Proceedings of the Royal Society (B)*, *162*, 20-46.
126. Rushton, W. A. H. (1972). Visual pigments in man. In: *Handbook of Sensory Physiology, Vol. II/1. Photochemistry of Vision* (Ed. H.J.A. Dartnall), 364-394, Springer-Verlag, New York.

127. Sällström, P. (1973). Colour and physics: some remarks concerning the physical aspects of human colour vision. Report 73-09. Stockholm, University of Stockholm, Institute of Physics.
128. Schirillo, J. A. & Shevell, S. K. (1992). Brightness perception depends on the perceived depth plane of a surface relative to the depth planes of other surfaces. *Investigative Ophthalmology and Visual Science (Suppl.)*, *33*, 1258.
129. Schirillo, J. A., Reeves, A. & Arend, L. E. (1990). Perceived lightness, but not brightness, of achromatic surfaces depends on perceived depth information. *Perception & Psychophysics*, *48*, 82-90.
130. Semmelroth, C. C. (1970). Prediction of lightness and brightness on different backgrounds. *Journal of the Optical Society of America*, *60*, 1685.
131. Shapley, R. (1986). The importance of contrast for the activity of single neurons, the VEP and perception. *Vision Research*, *26*, 45-61.
132. Shapley, R., Caelli, T., Grossberg, S., Morgan, M. & Rentschler, I. (1990). Computational theories of visual perception. In: Spillmann, L. & Werner, J. S. (Eds), *Visual Perception: the Neurophysiological Foundations* (pp. 53-101). San Diego: Academic Press.
133. Shapley, R. & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain control. In: Osborne, N. & Chader, G. (Eds.), *Progress in Retinal Research*, *3*. Oxford: Pergamon.
134. Sproson, W. N. (1983). *Colour science in television and display systems*, Bristol: Hilger.
135. Stevens, J. C. & Stevens, S. S. (1963). Brightness function: effects of adaptation. *Journal of the Optical Society of America*, *53*, 375-385.
136. Thompson, E., Palacios, A. & Varela, F. J. (1992). Ways of coloring: comparative color vision as a case study for cognitive science. *Behavioral and Brain Sciences*, *15*, 1-26.
137. Tiplitz-Blackwell, K. & Buchsbaum, G. (1988a). Quantitative studies of color constancy. *Journal of the Optical Society of America A*, *5*, 1772-1780.
138. Tiplitz-Blackwell, K. & Buchsbaum, G. (1988b). The effect of spatial and chromatic parameters on chromatic induction. *Color Research and Application*, *13*, 166-173.
139. Todorovic, D. (1987). The Craick-O'Brien-Cornsweet effect: new varieties and their theoretical implications. *Perception & Psychophysics*, *42*, 545-560.
140. Trigt, C. van (1990). Smoothest reflectance functions. I. Definition and main results. *Journal of the Optical Society of America A*, *7*, 1891-1904.
141. Troost, J. M. (1992). Perceptual and computational aspects of color constancy. Ph.D. thesis, Nijmegen Institute for Cognition and Information.

142. Troost, J. M. & de Weert, C. M. M. (1991a). Surface reflectance and human color constancy: comment on Dannemiller (1989). *Psychological Review*, *98*, 143-145.
143. Troost, J. M. & de Weert, C. M. M. (1991b). Naming vs matching in color constancy. *Perception & Psychophysics*, *50*, 591-602.
144. Troost, J. M. & de Weert, C. M. M. (1992). Techniques for simulating object color under changing illuminant conditions on electronic displays. *Color Research and Application*, *17*, 316-327.
145. Troost, J. M., Wei, L. & de Weert, C. M. M. (1992). Binocular measurements of chromatic adaptation. *Vision Research*, *10*, 1987-1997.
146. Ullman, S. & Schechtman, G. (1982). Adaptation and gain normalization. *Proceedings of the Royal Society of London, B*, *216*, 299-313.
147. 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.
148. Vos, J. J. (1978). Colorimetric and photometric properties of a 2° fundamental observer. *Color Research and Application*, *3*, 125-128.
149. Vos, J. J. & Walraven, P. L. (1971). On the derivation of the foveal receptor primaries, *Vision Research*, *11*, 799-818.
150. Wallach, H. (1948). Brightness constancy and the nature of achromatic colors, *Journal of Experimental Psychology*, *38*, 310-324.
151. Walls, G. L. (1960). "Land! Land!". *Psychological Bulletin*, *57*, 29-48.
152. Walraven, J. (1973). Spatial characteristics of chromatic induction; the segregation of lateral effects from straylight artifacts. *Vision Research*, *11*, 1739-1753.
153. Walraven, J. (1976). Discounting the background - the missing link in the explanation of chromatic induction. *Vision Research*, *16*, 289-219.
154. Walraven, J. (1977). Color signals from incremental and decremental light stimuli. *Vision Research*, *17*, 71-76.
155. Walraven, J. (1980). The derivation of nerve signals from contrast flash data; a re-analysis. *Biological Cybernetics*, *38*, 23-29.
156. Walraven, J. (1981). Perceived colour under conditions of chromatic adaptation: evidence for gain control by π -mechanisms. *Vision Research*, *21*, 611-620.
157. Walraven, J. (1985). The colours are not on the display: a survey of non-veridical perceptions that may turn up on a colour display. *Displays*, *6*, 35-42.
158. Walraven, J. (1988). Colour coding for ATC displays. *ECC Report No 112, Eurocontrol, Brussels*.
159. Walraven, J., Benzschawel, T. & Rogowitz, B. E. (1989). Color-constancy interpretation of chromatic induction. *Die Farbe*, *34*, 269-273.

160. 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.
161. 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: Academic Press.
162. Walraven, J. (1992). Color basics for the display designer. In: Widdel, H. & Post, D. L. (Eds.), *Color in Electronic Displays*. Plenum press, New York and London, 3-38.
163. Walraven, J. & Lucassen, M. P. (1991). Color constancy in an RGB world. *Proceedings Eurodisplay '90*, 112-115.
164. 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.
165. Walraven, J. & Werner, J. S. (1991). The invariance of unique white; possible implications for normalizing cone action spectra. *Vision Research*, *31*, 2185-2193.
166. Walraven, P. L. (1962). On the mechanisms of colour vision. Thesis, University of Utrecht.
167. Weert, de C. M. M. (1991). Assimilation versus contrast. In: Valberg, A. & Lee, B. B. (Eds). *From Pigments to Perception* (pp. 305-311). New York: Plenum Press.
168. Wehrhahn, C., Heide, W. & Petersen, D. (1990). Long-range colour interactions in human visual cortex. *Clinical Vision Sciences*, *4*, 401-406.
169. 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.
170. Whipple, W. R., Wallach, H. & Marshall, F. J. (1988). The effect of area, separation, and dichoptic presentation on the perception of achromatic color. *Perception & Psychophysics*, *43*, 367-372.
171. Whittle, P. (1986). Increments and decrements: luminance discrimination. *Vision Research*, *26*, 1677-1691.
172. Whittle, P. & Challands, P. D. C. (1969). The effect of background luminance on the brightness of flashes. *Vision Research*, *9*, 1095-1110.
173. Wiesel, T. N. & Hubel, D. H. (1966). Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *Journal of Neurophysiology*, *29*, 1115-1156.

174. Worthey, J. A. (1985). Limitations of color constancy. *Journal of the Optical Society of America A*, *2*, 1014-1026.
175. 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.
176. Wyszecki, G. & Stiles, W. S. (1982). *Color Science, Concepts and Methods, Quantitative Data and Formulae*. 2nd edition, New York: Wiley.
177. Young, R. A. (1987). Color vision and the Retinex theory. *Science*, *238*, 1731-1732
178. Yuille, A. (1987). A method for computing spectral reflectance. *Biological Cybernetics*, *56*, 195-201.
179. Yund, E. W. & Armington, J. C. (1975). Color and brightness contrast effects as a function of spatial variables. *Vision Research*, *15*, 917-929.
180. Zeki, S. M. (1980). The representation of colours in the cerebral cortex. *Nature*, *284*, 412-418.

Bibliography

1. Lucassen, M. P. & Walraven, J. (1990). Evaluation of a simple method for color monitor recalibration. *Color Research and Application*, 15, 321-326.
2. Lucassen, M. P. & Walraven, J. (1992a). Quantifying color constancy: evidence for nonlinear processing of cone-specific contrast. *Vision Research* (in press).
3. Lucassen, M. P. & Walraven, J. (1992b). Color constancy under natural and artificial illumination. (submitted for publication in *Vision Research*).
4. Lucassen, M. P. & Walraven, J. (1992c). Separate processing of chromatic and achromatic contrast in color constancy. (submitted for publication in *Vision Research*).
5. Lucassen, M. P. & Walraven, J. (1992d). Color constancy under conditions varying in spatial configuration. (submitted for publication in *Vision Research*).
6. Theeuwes, J. & Lucassen, M. P. (1992). An adaptation-induced pop-out in visual search. (submitted for publication in *Vision Research*).
7. 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.
8. Walraven, J. & Lucassen, M. P. (1991). Kleurtabellen voor de elektronische zeekaart (ECDIS); aanpassing van de RGB specificaties aan de fosfor- en lumiantiekaracteristieken van het kleurenbeeldscherm. TNO report, No. 1991-M8.
9. Walraven, J. & Lucassen, M. P. (1992). Optimalisering van de kleurweergave bij een false-colour helderheidsversterker. TNO report, No. 1992-M9.

Curriculum Vitae

Marcel Lucassen werd op 15 juli 1964 te Oosterbeek geboren. Na verhuisd te zijn naar Doetinchem werd aldaar in 1982 het VWO-diploma aan het St. Ludgercollege behaald. In datzelfde jaar begon hij zijn studie Technische Natuurkunde aan de Universiteit Twente te Enschede. In de zes jaar die daarop volgden leerde hij (o.a.) dat men "door meten" nog steeds "tot weten" komt, onder meer door een stage aan de Loughborough University of Technology (Physics Department) in Engeland. Het onderzoek dat tijdens die stage werd verricht betrof de toepassing van een piëzo-electrische transducer als detector voor laser-gegenereerde acoustische emissie. Het laatste jaar van zijn studie was hij werkzaam bij de taakgroep Biofysische Techniek van de afdeling Technische Natuurkunde. Daar hielp hij mee aan de ontwikkeling van een multi-parameter fluorescentie-systeem voor de cytodisk-opstelling, een apparaat waarmee (o.m.) bloedcellen optisch bemeten kunnen worden. In augustus 1988 ontving hij het Ingenieurs diploma Technische Natuurkunde. Eind 1988 begon hij als NWO Onderzoeker In Opleiding aan dit promotieonderzoek op het Instituut voor Zintuigfysiologie TNO te Soesterberg. In de daarop volgende vier jaar verrichtte hij psychofysisch onderzoek naar het mechanisme van de kleurconstantie, resulterend in dit proefschrift. Daarnaast werd ervaring opgedaan in het toegepaste kleurenonderzoek.