Mechanisms of Color Constancy

in Trichromats and Dichromats

Dissertation

zur Erlangung des
Doktorgrades (Dr. phil.)

vorgelegt

der Philosophischen Fakultät I
der Martin-Luther-Universität Halle-Wittenberg,

von Sven Nicklas
geb. am 20.05.1975 in Wittenberg

Gutachter: Prof. Dr. Dieter Heyer
Prof. Dr. Laurence T. Maloney

Tag der Verteidigung: 18.07.2008

urn:nbn:de:gbv:3-000014355
[http://nbn-resolving.de/urn/resolver.pl?urn=nbn%3Ade%3Agbv%3A3-000014355]

Acknowledgements

This work would not have been possible without the encouragement and support from a number of significant people. The least I can do is to express my gratitude to them.

First of all I am indebted to my advisor Dieter Heyer who introduced me patiently to the field of color science and sharpened my understanding of theoretical concepts. I heartily thank Larry Maloney for sharing his knowledge and for the opportunity to conduct some of the experiments in his lab. Gisela Müller–Plath contributed her patience and time as a member of my committee. I am grateful to Eike Richter and Katja Doerschner for fruitful discussions on color issues and for helping me with the experimental setup. I would also like to thank Franz Faul for providing his extensive C++ color library and Katrin Heier for running the experiments on increment-decrement asymmetries. I am indebted to my dear friend Fidelindo Lim who helped me improving the language and style of the present thesis. On the long way that eventually led to this thesis my parents have always supported me. At this point I would like to express my love and gratitude to them. My beloved wife Stefanie encouraged me with her cheerfulness and her creative way of thinking. I am grateful to have her on my side.

This work was supported by a doctoral scholarship from the federal state of Saxony-Anhalt and in part by a grant from DAAD.

Contents

Chapter 1

Introduction

The perception of color is an important feature of our visual system which maintains our orientation in the world. As a matter of course, we assign colors to objects assuming implicitly that color is a stable attribute of the object. However, the pattern of light that reaches the eye from an object is the result of a complex interaction between the incident illumination and the object's surface properties. The physical light signal from a given object can vary dramatically across different illuminant conditions, yet the color percept linked with the object remains roughly the same. Mostly we are not aware of this phenomenon of *color constancy* but it enables us to recognize objects across a great variety of different illumination conditions.

In addition, our visual system is also able to deal with two different classes of illuminant changes: changes across time and changes across space. In natural environments we are confronted with both types of changes, if we imagine for example slow gradual changes of daylight over time or a scene with one part lit by direct sunlight while another part is situated in the shade. It is assumed that two different classes of mechanisms, sensory and perceptual processes, which are presumably located at different stages in the visual path enable us to optimally tune to the present illumination conditions (Bäuml, 1997). In the investigations presented here I focused on the ability of observers to adjust to illuminant changes across time which is sometimes referred to as *successive color constancy*.

The investigation of the phenomenon of color constancy has a long and rich tradition in vision science. Hermann von Helmholtz (1867) and Ewald Hering (1920) both examined the problem and—like on other aspects of human color vision—came to different conclusions. Helmholtz assumed perceptual cognitive mechanisms to underlie color constancy and emphasized the role of experience whereas Hering suggested that the phenomenon is mainly due to sensory low-level processes (see also Gelb, 1929). It is now a well established fact that both cognitive (Hansen & Gegenfurtner, 2006) and sensory processes (Smithson & Zaidi, 2004) contribute to color constancy.

A new perspective on the phenomenon of color constancy was outlined by

David Katz (1911/1935). He introduced the notion that our color perception is structured in different modes of appearance. According to Katz, two distinct modes of appearance, surface color and film color, are related with the encoding of surfaces' properties and characteristics of the illumination respectively. Arend and Reeves (1986) demonstrated that observers have in fact access to these two different modes of appearance at the same time. Recently, it was hypothesized that the distinction between incremental and decremental stimuli may trigger two different modes of appearance that refer to illumination (film) color and object (surface) color (Mausfeld, 1998; Bäuml, 2001). This assumption will be investigated in the experiments presented in Chapter 4.

The preceding remarks on different facets of color constancy indicate that it is not plausible to assume only one single mechanism underlying the phenomenon. In fact, research in the last two decades has led to the conclusion that color constancy is maintained rather by a number of different principles depending on environmental constraints (Kraft & Brainard, 1999; Maloney, 2002). Some of the principles discussed in the literature are chromatic adaptation, consideration of the brightest surface in the scene, adaptation to the (virtual) average light signal across all surfaces, local contrast due to adjacent surfaces and cognitive factors based on experience (for review see Smithson, 2005). In the experimental part of this work effects of two of these mechanisms, the mean light signal and chromatic adaptation, will be examined.

The present work consists of two experimental investigations focusing on different aspects of color constancy. In the first study, which is described in detail in Chapter 4 the role of increments and decrements in color constancy were examined systematically using traditional stimuli and measures. It was found that the strength of asymmetric processing of incremental and decremental stimuli is correlated with the color of daylights. Furthermore, in this study color constancy performance of normal observers was compared with performance of a subpopulation of observers who are generally characterized by the absence of one type of cone photoreceptors. This class of color deficient observers is termed *dichromats*. Results from recent studies bring into question the assumption that a simple reduction model describes dichromacy adequately (Wachtler, Dohrmann & Hertel, 2004). In the present study this issue was investigated in terms of color constancy. My results indicate that color constancy of dichromats does not necessarily break down in conditions where this would be predicted by the reduction model. Accordingly, the present data does not strictly argue in favor of this hypothesis.

In a recent essay on color constancy, David Foster (2003) raised the question of whether traditional quantitative measures of color constancy provide an adequate access to the performance of observers. The second study which is presented in Chapter 5 introduces hue scaling as an alternative method to investigate color constancy. This technique promises to treat the problem of color constancy in a more phenomenological sense and yet allows comparisons with common quantita-

tive measures. In addition, stimuli used in the corresponding experiments are simulated according to a more complex three dimensional model of the environment. It was found that the degree of observers color constancy obtained by the hue scaling technique was comparable to quantitative measurements. Furthermore, the results demonstrate the isolated effect of temporal chromatic adaptation on color constancy.

Before I present the experimental work, I will give an introduction to theoretical concepts of color science and related empirical findings. In Chapter 2, two fundamental theories of color vision: trichromatic theory and the theory of opponent colors will be discussed. Within this theoretical framework specific characteristics of dichromatic observers will also be covered. In Chapter 3, I will give a comprehensive introduction to the phenomenon of color constancy. In this context influential models of color constancy and new perspectives on the phenomenon will be discussed.

Chapter 2

Basic Color Theory

In this chapter I will give an introduction to two classical theories of color vision. Probably the most influential theory is the theory of trichromatic color vision which will be discussed first. In the second part I will examine characteristics of a subpopulation of observers called *dichromats*. The existence and the characteristics of these color deficient observers are well explained within the framework of trichromatic theory. The chapter will be concluded with an introduction to the second classical theory of color vision, the theory of opponent colors. This theory had been standing in opposition to trichromatic theory for many years. Now it is commonly accepted that the two theories refer to different stages of color processing in the visual system.

2.1 Primary Color Coding

The theory of trichromatic color vision was mainly developed in the 19th century and is closely connected with the names of Thomas Young, Hermann von Helmholtz, James Clerk Maxwell and Hermann Grassmann. The enormous success of the theory results especially from the fact that the theory of trichromatic color vision provides an adequate description of processing of light at retinal level. Furthermore, conclusions derived from the theory play an important role in most applications that deal with color vision. In particular, quantitative aspects of the theory constitute the foundations of the field of *colorimetry* that these applications are based on.

The main goal of the theory of trichromatic color vision is to connect characteristics of the physical stimulus of light, that is electromagnetic radiation within the visible spectrum, with the matching behavior of observers which is determined by psychological variables. The physical stimulus—light—can be characterized by its spectral power distribution within the visible spectrum (Figure 2.1). In the following I will describe how the relation between the physical stimulus and the psychological percept is conceptualized and 'measured' within the framework

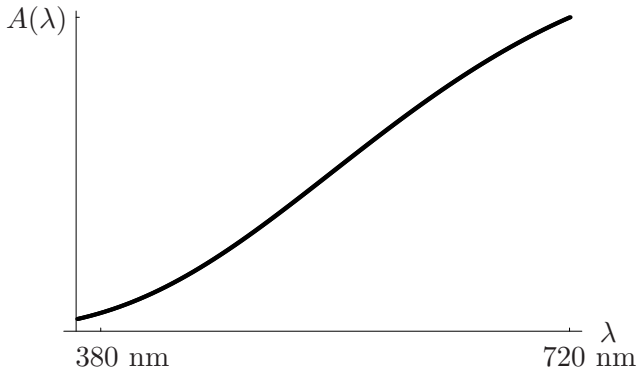

Figure 2.1: Spectral power distribution of a tungsten light source A.

of trichromatic theory. This relation can be understood as a mapping from the physical world to psychological sensations.

2.1.1 The Experimental Paradigm

In classical experiments of trichromatic theory subjects are presented with two lights a and b on the two halves of a bipartite disk (Figure 2.2a).[1] The task of the subject is to change light b in color appearance until it is perceptually indistinguishable to light a on the other half of the disk. This criterion of perceptual indistinguishability is called *metamerism* and is written $a \sim b$. Properties of the relation of metamerism are reflexivity, symmetry and transitivity. Therefore metamerism is an equivalence relation which is only defined by a psychological criterion. In other words, two lights that are metamers generally differ in their spectral power distributions. The classical color matching experiments are based on two *physical* operations, namely the additive mixture of lights ($\oplus$) and scalar multiplication ($*$). In practice additive mixture is realized as superposition of lights whereas scalar multiplication can be understood as change of the intensity of lights. These operations allow us to define linear combinations of lights m_1, m_2, m_3 with intensities t_1, t_2, t_3:

$$t_1 * m_1 \oplus t_2 * m_2 \oplus t_3 * m_3. \tag{2.1}$$

In the experiment, the light b is replaced with such a linear combination of three lights on the right half of the disk (Figure 2.2b). The subject is asked to adjust the intensities t_1, t_2, t_3 of the three *independent* basis lights until the

[1] Two restrictions on the experimental conditions are usually made. First, the stimulus is presented against a black background. Second, the size of the disk is limited to two degrees of visual angle to yield only foveal stimulation.

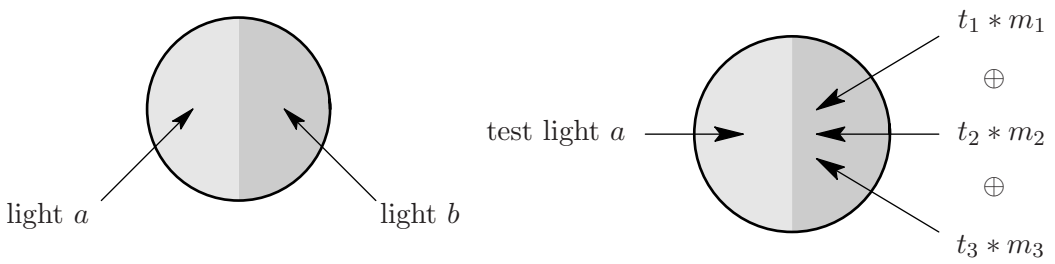

Figure 2.2: (a) Presentation of two lights a and b on the two halves of a bipartite disk (left). (b) Configuration in the classical color matching experiment. Light a is matched with a mixture of lights m_1, m_2, m_3 (right).

mixture is metamer to a given test light a on the left half of the disk.[2] This procedure is called *color matching*. In general, subjects need just three basis lights to find a color match for any given test light a. Each test light can be characterized by the three numbers t_1, t_2, t_3, which we will later refer to as *color codes*.[3] This fundamental characteristic of human color vision is called *trichromacy*. In contrast, some observers need only two basis lights to find a color match for any given test light. These observers are called *dichromats*. We will come back to this point later in this chapter.

2.1.2 The Grassmann Laws

Grassmann (1853) put results from color matching experiments in a formal theoretical framework. He expressed three characteristics of metamerism that define the space of lights as an infinite vector space. These fundamental conclusions are known as *Grassmann laws*. Given three lights a, b, c and scalars $t \geq 0$, the following equations hold for the metamerism relation:

$$a \sim b \Leftrightarrow a \oplus c \sim b \oplus c, \tag{2.2}$$

$$a \sim b \Leftrightarrow t * a \sim t * b. \tag{2.3}$$

The third of Grassmann's laws is the already mentioned trichromacy of normal human observers. Empirical work has shown that the Grassmann laws generally hold. An axiomatic formulation of the Grassmann theory is given in Krantz (1975).

[2]The term *independent basis lights* refers to the claim that each of the basis lights can not be represented as a mixture of the two other basis lights.

[3]For some test lights it is impossible to find a satisfying match using the method just described. In this case one of the three basis light is presented on the half of the test light. Using this indirect color matching procedure the subject is asked to adjust the three intensities t_1, t_2, t_3 until both halves are indistinguishable in color appearance.

2.1.3 Primary Color Codes

If we choose a fixed set of three basis lights, we can define each test light a by the three numbers t_1, t_2, t_3 that are needed to match a with the mixture of the three basis lights m_1, m_2, m_3. This triple of numbers is called *primary color code* or *tristimulus values* and we will denote it with the vector $\mathbf{t} := (t_1, t_2, t_3)'$. Lights that are metamers have identical primary color codes and lights with identical color codes are perceptually indistinguishable:

$$a \sim b \Leftrightarrow \mathbf{t}^a = \mathbf{t}^b. \tag{2.4}$$

The primary color code t_1^a, t_2^a, t_3^a of a given light a is a linear function of the spectral power distribution of that light. Together with (2.2) and (2.3) the space of color codes is therefore defined as a three dimensional vector space. This means that

$$\mathbf{t}^{a \oplus b} = \mathbf{t}^a + \mathbf{t}^b \tag{2.5}$$

and

$$\mathbf{t}^{s*a} = s \cdot \mathbf{t}^a \tag{2.6}$$

hold for all lights a, b and scalars $s \geq 0$. Imagine we were able to carry out the color matching experiment with a fixed set of basis lights and determine primary color codes for all possible lights as test lights. If we represent all obtained color codes in three dimensional vector space the result would be a geometric structure that is called *color cone* (Figure 2.3). The color cone can also be interpreted as the mapping of all realizable lights (which are a subset of infinite vector space) onto a three dimensional vector space which we will refer to as *color space*.[4]

Usually the basis lights that are used in the color matching experiment are narrow band lights with a bandwidth of only one nanometer. These lights are called *spectral lights* and they are characterized by the respective wavelength number. The set of spectral lights with unit radiant power $\{e_{300}, \ldots, e_{700}\}$ is called *equal energy spectrum*. For each of these spectral lights e_λ the spectral power distribution $e_\lambda(x)$ is 1 if $x \in [\lambda - \frac{1}{2}; \lambda + \frac{1}{2}]$ and else 0. Later the spectral power distribution of a given light will be approximated using the equal energy spectrum.

A common set of basis lights is the *RGB-system*. The three basis lights of this system appear reddish ('R', 700 nm), greenish ('G', 546 nm) and bluish ('B', 436 nm) to the normal observer.[5] If we want to determine the primary color code of a given light a regarding to the three basis vectors R, G, B we could carry out the color matching experiment. Now a more convenient method to obtain the intensities t_1^a, t_2^a, t_3^a for a given light a without conducting the color matching experiment will be described. First, we determine experimentally the

[4]Note that the color cone is a subspace of color space which contains tristimulus values of all physically realizable lights.

[5]Note that the color codes of the three basis lights must lie within the color cone

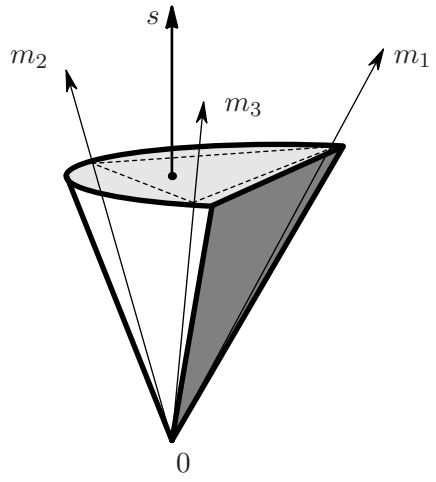

Figure 2.3: The color cone and the three basis lights m_1, m_2 and m_3.

tristimulus values of all spectral lights e_λ of the equal energy spectrum using e_{700}, e_{546} and e_{436} as basis lights. As a result we get mean tristimulus values $\bar{t}_1^\lambda$, $\bar{t}_2^\lambda$, $\bar{t}_3^\lambda$ for each spectral light λ across several observers. These mean intensities can be denoted as three functions $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ of the wavelength λ and are called *color matching functions* (Figure 2.4). In the next step we can approximate the spectral power distribution of a given light a as linear combination of the equal energy spectrum with $a(\lambda)$ as coefficients:

$$a \sim \sum_{\lambda=300}^{700} a(\lambda) * e_\lambda. \tag{2.7}$$

Given this approximation and the relations (2.5) and (2.6) we get the tristimulus values of the light a with

$$\mathbf{t}(a) = \sum_{\lambda=300}^{700} a(\lambda) \cdot \mathbf{t}^{e_\lambda}. \tag{2.8}$$

As we used the color matching functions $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ to denote the tristimulus values of all spectral lights e_λ we can also write

$$\begin{aligned} t_1^a &= \sum_{\lambda=300}^{700} a(\lambda) \cdot \bar{r}(\lambda), \\ t_2^a &= \sum_{\lambda=300}^{700} a(\lambda) \cdot \bar{g}(\lambda), \\ t_3^a &= \sum_{\lambda=300}^{700} a(\lambda) \cdot \bar{b}(\lambda). \end{aligned} \tag{2.9}$$

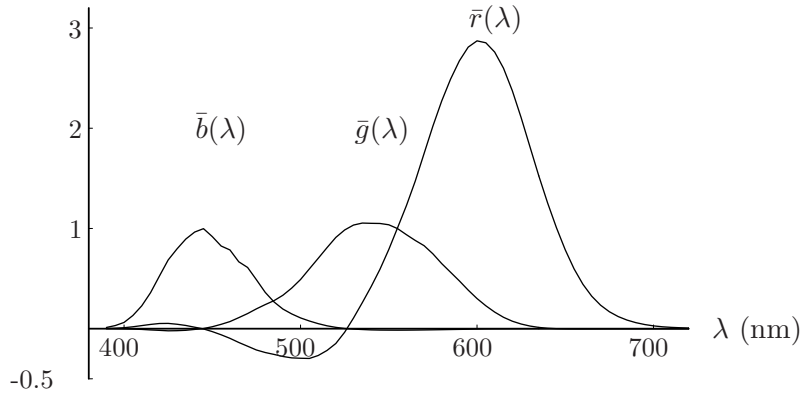

Figure 2.4: *Color matching functions* $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ and $\bar{b}(\lambda)$ of the CIE-1931 standard observer.

If the color matching functions for a given set of basis lights are known, the color codes of any light a can be computed easily.

One disadvantage of the RGB-System is that—since the color codes of the basis lights are part of the color cone—for some lights, negative coefficients result. If we prefer a set of basis lights which provides only positive coefficients the color codes of these 'basis lights' must lie necessarily outside the color cone. It is evident that in this case the basis of such a system is constituted by not realizable 'lights'.

An alternative system with this characteristic is the *CIE-XYZ-system* of the *Commission Internationale de l'Eclairage (CIE)*. In this coordinatization the basis vectors X, Y and Z are not part of the color cone and cannot be related with physical lights. As a consequence, for any given light only positive XYZ-tristimulus values result (Figure 2.5).

Another advantage of the XYZ-system is that the Y-coordinate is defined as the *luminance* of the light. Before we continue I will introduce the concept of luminance in brief. Roughly speaking, luminance corresponds to perceived brightness. The luminance $L(a)$ is defined as a linear function of the light a so that

$$L(a \oplus b) = L(a) + L(b) \tag{2.10}$$

and

$$L(s * a) = s \cdot L(a) \tag{2.11}$$

hold for all lights a, b and and scalars $s \geq 0$. As a consequence, all lights with identical luminance lie on a plane in color space. Analogous to the three color matching functions there is function of λ which is related to luminance. It is called the *luminous efficiency function* $V(\lambda)$ (Figure 2.6). The function $V(\lambda)$ can be also interpreted as a sensitivity function of human daylight vision. The luminance of a light is simply the integral of the product of $a(\lambda)$ and $V(\lambda)$ over

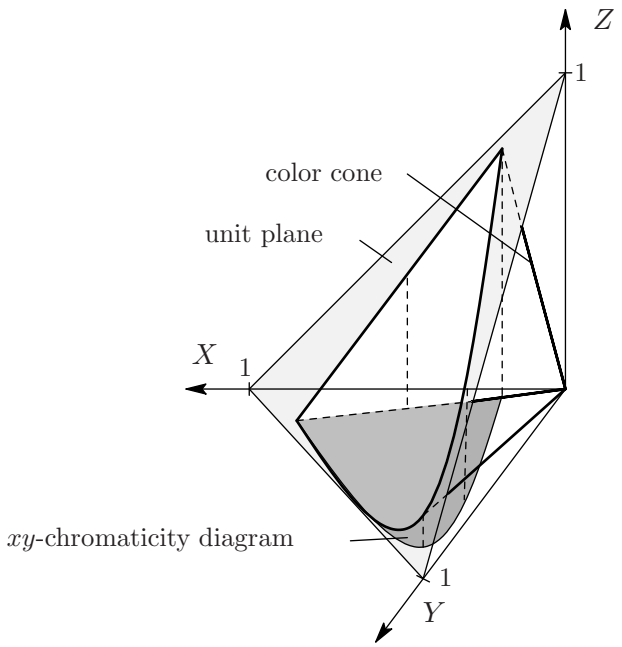

Figure 2.5: The CIE-XYZ-system and the color cone. The xy-chromaticity diagram can be interpreted as the projection of the unit plane onto the XY-plane.

the visible spectrum:

$$L(a) = \int_{\lambda} a(\lambda)V(\lambda)\, d\lambda. \tag{2.12}$$

Experimentally, $V(\lambda)$ is determined using a technique called *heterochromatic flicker photometry*.[6] Two spectral lights are presented to the observer in rapid interchange at the same location. At a certain frequency of presentation differences in luminance of both stimuli are perceived as flicker whereas chromatic differences are not detectable for the visual system. This phenomenon has to do with the fact that the temporal resolution of the luminance path is finer than those of the chromatic paths. In the flicker experiment, one of the stimuli serves as a fixed standard which is usually the spectral light e_{555}. The task of the observer is to adjust the intensity of the other stimulus, the test light, until perceived flicker is canceled or minimized. If this procedure is done with all spectral lights as test lights, the resulting intensities m_λ can be expressed in relation to the intensity m_{555} of the standard. This yields to the luminous efficiency function $V(\lambda)$:

$$V(\lambda) = m_{555}/m_\lambda. \tag{2.13}$$

[6] As an alternative one might suggest to determine equally bright stimuli using equal brightness judgments. Unfortunately this method does not lead to a measure which is a linear function of the physical stimulus (Wyszecki & Stiles, 1982).

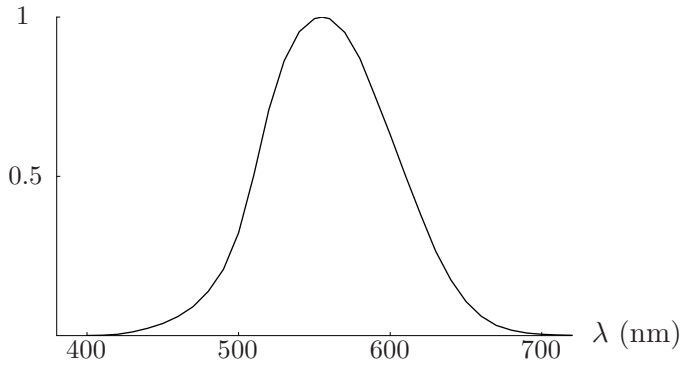

Figure 2.6: The luminous efficiency function $V(\lambda)$.

The RGB- and the XYZ-system can be both interpreted as coordinatizations of the three dimensional color space with different basis vectors. Therefore the transformation from one system to the other is simply given by multiplication with the respective 3×3-transformation matrix. A two dimensional representation of chromaticities within the *CIE-XYZ-system* is the *xy-chromaticity diagram* (Figure 2.5 and Figure 2.7). The xy-chromaticity coordinates can be easily transformed from given $XYZ-$coordinates:

$$x = \frac{X}{X+Y+Z} \quad \text{and} \quad y = \frac{Y}{X+Y+Z} \quad . \tag{2.14}$$

Several efforts have been made to construct a transformation of color space which is based on perceptual distances. The *CIE* has suggested two systems that claim this purpose, the *CIE-Lab-system* and the *CIE-Luv-system*. Both coordinatizations are given by non-linear transformations of color space and depend on a reference light that refers to the adaptational state of the observer. However, none of the two systems provides an adequate representation of perceptual differences in terms of Euclidean distances. The chromaticities of the *CIE-Luv*-system are represented in the $u'v'$-diagram which is commonly used for data analysis and representation (Figure 2.7).[7] An important feature of the $u'v'$-diagram is that straight lines in the $xy-$chromaticity diagram remain straight in the $u'v'$-diagram. The non-linear transformation from xy- to $u'v'$-coordinates is given by:

$$u' = \frac{4x}{-2x+12y+3} \quad \text{and} \quad v' = \frac{9y}{-2x+12y+3} \quad . \tag{2.15}$$

[7]The $u'v'$-diagram is often referred to as the CIE 1976 UCS (*uniform-chromaticity-scale*) diagram.

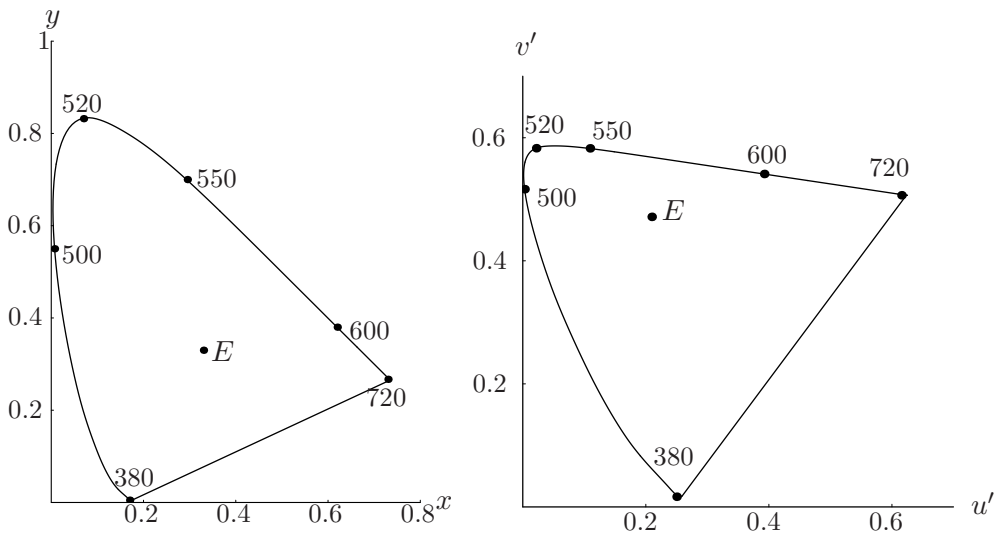

Figure 2.7: The xy-chromaticity diagram (left) and the $u'v'$-chromaticity diagram (right) with the chromaticity of the equal energy spectrum (E) and selected spectral lights.

2.1.4 Receptor Codes

The finding from color matching experiments that most human observers need exactly three basis lights to find a color match to any given test light led to the conclusion that the matching behavior of observers is mediated by three different classes of photoreceptors that are often called *cone fundamentals*.[8] These classes of photoreceptors differ in the spectral region of their maximal sensitivity. The first photoreceptor ('L') has its maximal sensitivity in the longwavelength region of the spectrum, the second photoreceptor ('M') is maximally sensitive in the middlewavelength region and the third photoreceptor ('S') has its maximal sensitivity in the shortwavelength region of the spectrum.

At the time the theory of trichromacy was formulated first, the sensitivities of the photoreceptors were unknown. Therefore, one of the main challenges of trichromatic theory was to obtain functions that represent sensitivities of these assumed photoreceptors. The psychophysical determination of these sensitivity functions was first done by König and Dieterici (1893). Their work was based on two main assumptions. First, König and Dieterici assumed that a coordinatization of color space exists in which axes are represented by the three classes of photoreceptors. Second, they supposed that three classes of dichromats exist which differ in their matching behavior. Each of these classes of dichromats can

[8]The human retina contains a fourth class of photoreceptors that are called *rods*. The rods are extremely sensitive to light so that only one quantum of light is sufficient to excite a single rod (Hecht, Schlaer & Pirenne, 1942). Hence, rods mediate vision at very low light levels which is called *scotopic vision*. It is assumed that rods do not contribute to *photopic* daylight vision (Wyszecki & Stiles, 1982).

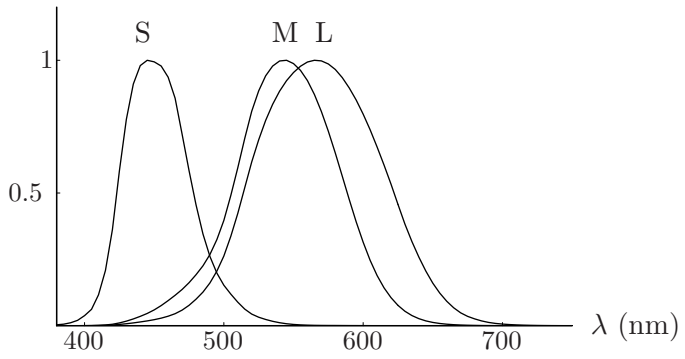

Figure 2.8: Sensitivity functions of the three photoreceptor classes L, M and S after Smith and Pokorny (1975). In this graph the functions are normalized to maximum value 1.

be characterized by the lack of one class of photoreceptors.

As a consequence of these assumptions, two lights that are perceptually indistinguishable for a given type of dichromat but in fact do not match for a normal trichromat can only differ in the excitation of the missing fundamental. If we present a given dichromat with a fixed test light in the color matching experiment, the color codes of the lights which are indistinguishable from the test light for that dichromat lie on a line in color space. Such a line is called *confusion line* of the respective dichromat. The lights which color codes are aligned on a confusion line of a given dichromat differ only in the excitation of the missing photoreceptor of that dichromat. Hence confusion lines of a dichromat are parallel to the axis of the missing fundamental. König and Dieterici determined several confusion lines for each of the three classes of dichromats and shifted these lines to the origin of color space. The three resulting axes can be interpreted as a basis of color space which is related to cone excitations. The sensitivity functions of the three classes of photoreceptors can be obtained as a linear transformation of the color matching functions. It is worth noting that König and Dietericis estimates of the sensitivity functions based on psychophysical experiments are in good agreement with physiological measurements that were done in the second half of the 20th century (e.g. Schnapf, Kraft & Baylor, 1987). Figure 2.8 shows sensitivity functions that were determined by Smith and Pokorny (1975). These estimates of the cone fundamentals will be used throughout this work.

Now we can characterize each light a by a triplet of numbers $\phi(a) = (\phi_1(a), \phi_2(a), \phi_3(a))$ that gives the photoreceptor excitations for this light. We will refer to this triplet of cone excitations as *LMS-* or *receptor codes*. If we denote the three sensitivity functions with $R_1(\lambda)$, $R_2(\lambda)$, $R_3(\lambda)$, the receptor codes of any given light a can be simply determined by

$$\phi_i(a) = \int_\lambda a(\lambda) \cdot R_i(\lambda)\, d\lambda, \qquad i = 1, 2, 3. \tag{2.16}$$

The function ϕ can be interpreted as a linear mapping from the infinite space

of lights onto the three dimensional vector space which basis is given by the sensitivities of the three classes of photoreceptors [9]. This space is called LMS- or *photoreceptor excitation space*. Transformations to and from other coordinatizations of color space (RGB, XYZ) are simply given by respective 3×3 transformation matrices. In Equation (2.4) we already noted that two metameric lights a and b will be mapped to identical tristimulus values and vice versa. The same is true for receptor codes:

$$a \sim b \Leftrightarrow \phi(a) = \phi(b). \tag{2.17}$$

Here the psychophysical relation of metamerism is linked with the neurophysiological structure of cone photoreceptors. Therefore will refer to Equation 2.17 as *linking proposition*.

2.2 Dichromacy

In the following section we will focus on certain characteristics of dichromats. This subpopulation of color deficient observers differs in several respects from normal observers.

2.2.1 Dichromatic Color Spaces

We already defined dichromats as observers who need exactly two basis lights to find a color match for any given test light. As a consequence of this matching behavior, it was supposed that one class of photoreceptors is missing in dichromats. Given this assumption, three classes of dichromats can be distinguished: *protanopes*, which lack the L-cone, *deuteranopes*, where the M-cone is missing and *tritanopes* which lack the S-cone. The three classes of dichromats differ in their adjustments in the color matching experiment. Therefore we will denote matches of each class by corresponding relations of metamerism $\sim_P$, $\sim_D$ and $\sim_T$. Trichomatic matches are accepted by dichromats (König & Dieterici, 1893). As a consequence, the relation $\sim$ is a subset of each and every dichromatic relation of metamerism:

$$\sim \subset \ (\sim_P \cap \sim_D \cap \sim_T). \tag{2.18}$$

In the last section I introduced a representation of color space that is given by photoreceptor excitations (ϕ_1, ϕ_2, ϕ_3). As dichromats lack one type of cone photoreceptor, it is sufficient for them to characterize each light only by a pair of numbers which gives excitations of the two remaining cones. For example, in the protanope case we can denote this pair of numbers ϕ^P as

[9] The physically realizable lights constitute a *convex cone* within infinite vector space. Roughly speaking, a convex cone is a subset C of a vector space V where scalar multiplication is only allowed for positive scalars. The function ϕ can be understood as a linear mapping from this convex cone onto a convex cone in photoreceptor excitation space.

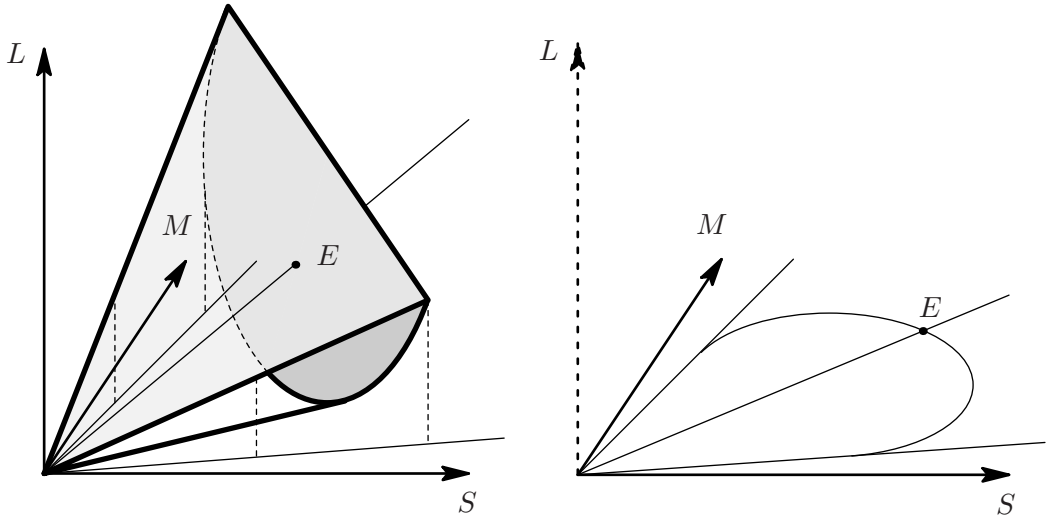

Figure 2.9: Cone excitation space of protanopes as a projection of the trichromatic color cone onto the MS-plane.

$$\phi^P(a) = (\phi_2(a), \phi_3(a)), \tag{2.19}$$

where $\phi_2(a)$ and $\phi_3(a)$ are receptor codes of the remaining cones. Hence dichromatic color spaces can be represented as two dimensional vector spaces.[10] Analogous to the trichromatic case we can state a dichromatic linking proposition, which connects metamerism and receptor codes:

$$a \sim_P b \Leftrightarrow \phi^P(a) = \phi^P(b) \Leftrightarrow (\phi_2(a), \phi_3(a)) = (\phi_2(b), \phi_3(b)). \tag{2.20}$$

If in addition the lights a and b are distinguishable for normal trichromats so that $\phi_1(a) \neq \phi_1(b)$ the receptor codes of these two lights lie on a confusion line in trichromatic color space. This confusion line is parallel to the axis of the missing cone. I already mentioned that König and Dieterici (1893) used this fact when they estimated sensitivity functions of the three cone photoreceptors. If we visualize the two dimensional dichromatic cone excitation space, this space is defined as an orthogonal projection of trichromatic LMS-space. This projection is parallel to the missing receptor axis onto the plane which is spanned by the two remaining fundamentals (Figure 2.9). In this sense, the protanopic cone excitation space is given by the MS-plane, the deuteranopic space by the LS-plane and the tritanopic space by the LM-plane.

2.2.2 Classes of Dichromats

We already introduced the three different classes of dichromats protanopes, deuteranopes and tritanopes that are characterized by the absence of the L-, M- or

[10]The function ϕ^P can be interpreted as the linear mapping from the infinite space of lights to the two dimensional color space of protanopes.

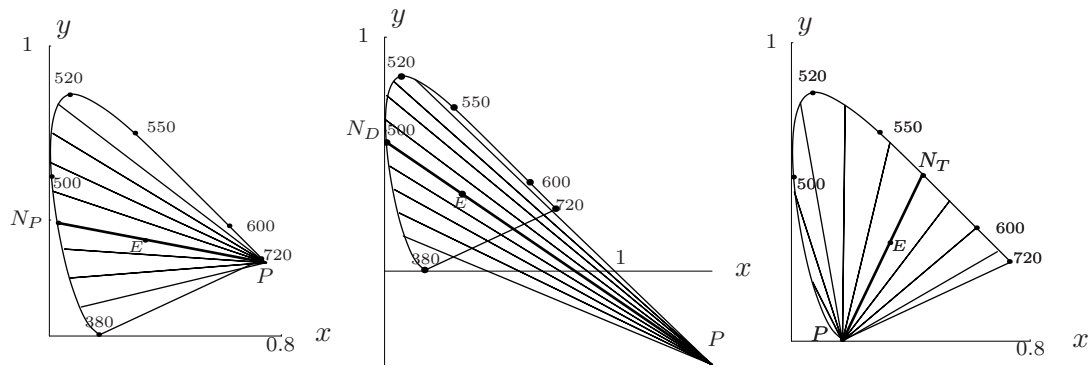

Figure 2.10: Pattern of confusion lines of protanopes (left), deuteranopes (center) and tritanopes (right) in $xy-$chromaticity diagram with confusion points ('P') and neutral points ('N'). 'E' denotes the chromaticity of an equal energy spectrum.

S-photoreceptor respectively. Roughly speaking, protanopes and deuteranopes can be identified with red-green color blinds and tritanopes with blue-yellow color blinds. In the preceding section we discussed differences between types of dichromats that referred to confusion lines, relations of metamerism and color space. In this section I will introduce further distinctive characteristics of the three classes of dichromats.

We can define a plane in color space which contains tristimulus values of all lights b with:

$$a \sim_P t * b, \tag{2.21}$$

for scalars $t \geq 0$ and any fixed light a. These lights are metamers for protanopes except for intensity changes. Additionally, each of such planes contains all colors with receptor codes $t \cdot (\phi_1, 0, 0)'$. Hence, in the case of protanopes the intersection line of all these planes is the L-cone axis. If we picture these planes in xy-chromaticity diagram, each of the planes is represented by a confusion line. These patterns of confusion lines are distinctive for each class of dichromats. Figure 2.10 shows the patterns of confusion lines of the three types of dichromats in xy-chromaticity diagram. The confusion lines intersect at a specific point, the *confusion point*. This point has a clear geometrical interpretation. It is the point where the missing fundamental crosses the XYZ-unit plane. Protanopes, deuteranopes and tritanopes differ in their specific confusion points which are given in Table 2.1. Another distinctive feature of different classes of dichromats are *neutral points*. If we present an isolated stimulus which corresponds with an equal energy spectrum ('E' in Figure 2.10) to normal observers it will appear achromatic to them. For a given dichromat all stimuli that are metamer to this stimulus lie on a confusion line that crosses 'E'. Consequently, there is also a spectral light that can be matched with white light by this dichromat and hence appears achromatic to him. The wavelength of this spectral light is the neutral

Table 2.1: Wavelengths of neutral points and xy-coordinates of confusion points of protanopes, deuteranopes and tritanopes (after Wyszecki & Stiles, 1982).

dichromat	confusion point	neutral point
protanope	x = 0.747 y = 0.253	490 - 495 nm
deuteranope	x = 1.400 y = -0.400	495 - 505 nm
tritanope	x = 0.171 y = 0	568 - 570 nm

point of this dichromat. The three types of dichromats can be also characterized by their respective neutral points which are given in Table 2.1.

Protanopes, deuteranopes and tritanopes differ also in their respective luminance functions. As explained above, only the L- and the M-cones contribute to the luminance function $V(\lambda)$ of color normal observers. Hence the luminance function of tritanopes $V_T(\lambda)$ is identical to that of normal trichromats. For protanopes only the M-cones contribute to luminance. The protanopic luminous efficiency function $V_P(\lambda)$ coincides with the sensitivity function of the M-cone photopigment. In the same way the luminosity function of deuteranopes is identical with the sensitivity function of the L-photoreceptor (Figure 2.11).

2.2.3 Diagnostics of Color Vision Deficiencies

Several diagnostic methods have been developed to detect color vision deficiencies. The most common instrument are pseudoisochromatic plates such as the Ishihara plates (Ishihara, 1997). These plates consist of a pattern of colored dots. Subjects are asked to detect a target which differs in color appearance from the background. The targets are normally letters or numbers. While tests based on pseudoisochromatic plates claim to discriminate reliably between different kinds of color deficiencies, inaccuracies in construction and in printing make them a choice for a first screening only.

The *anomaloscope* is a device which is used to classify different types of another group of color deficient observers, *anomalous trichromats*. In color matching experiments, anomalous trichromats need three basis lights to find a match for any given test light a. Hence the retina of those observers contains three different types of cones but the sensitivity function of one type of photoreceptors is shifted along the visible spectrum. The main forms of anomalous trichromacy are *protanomaly*, where the L-cone is modified and *deuteranomaly* where the M-cone is different from normal. For both classes of observers the ability to discriminate between spectral lights in the middle and long wavelength region is limited. Using the anomaloscope, the observer is presented with a mixture of two spectral lights, namely e_{535} and e_{670} which appear green and red to normal observers re-

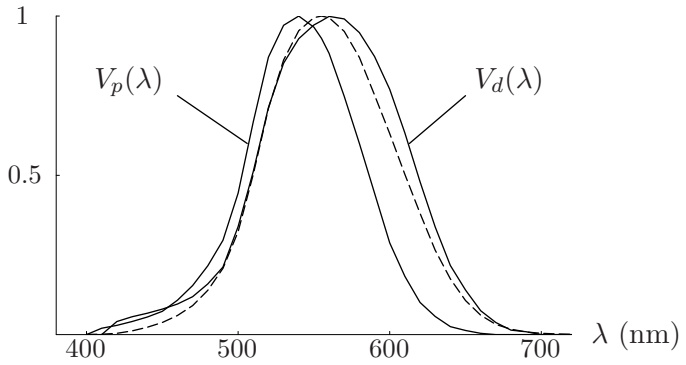

Figure 2.11: Luminosity functions of protanopes $V_p(\lambda)$ and deuteranopes $V_d(\lambda)$. The luminosity function of normal trichromats $V(\lambda)$ is shown in dashed line.

spectively. Then he is asked to match this mixture to a spectral light e_{589} that appears yellow to normal observers:[11]

$$e_{589} \sim \alpha * e_{670} \oplus \beta * e_{535}. \tag{2.22}$$

The intensities α and β that provide a satisfying match constitute the *Rayleigh-ratio* R:

$$R = \frac{\alpha}{\beta}. \tag{2.23}$$

The test light e_{589} is chosen so that for normal observers $\alpha = \beta$ and hence $R = 1$ results. Protanomalous observers need a greater amount of the long wavelength light e_{589} for a satisfying match which means that for them $R > 1$. The Rayleigh ratio of deuteranomalous is smaller than one as they need a higher intensity β for their matches. The Rayleigh ratios of dichromats vary unsystematically over a wide range as they generally find a match when they are allowed to change only one of the two intensities.

A psychophysical approach to classify dichromatic observers is to obtain corresponding confusion lines. The most prominent example of this method is the pioneer work of König and Dieterici (1893) which was already reported in the previous section. In Chapter 4 I will describe a variation of this approach in detail that was used in the present study to classify dichromatic subjects.

2.2.4 Genetics of Color Vision Deficiencies

In former times it had already been noticed that color vision deficiencies are congenital and handed from one generation to the next. Now it is a well established fact that color vision defects have a molecular genetics basis (Nathans,

[11]Note that the three spectral lights e_{535}, e_{670} and e_{589} are *not* independent mixture lights as it is generally claimed in color matching experiments. Their chromaticities lie on a line in xy-chromaticity diagram (see Figure 2.5) which indicates that all spectral lights on that line can be expressed as a mixture of only two of these spectral lights.

Table 2.2: Incident rates and genetics of color vision deficiencies (after Pokorny, Smith, Verriest & Pinckers, 1979; Wyszecki & Stiles, 1982).

defect	incident rate (%)		genetics
	men	women	
protanopia	1.0	0.02	recessive X-linked
deuteranopia	1.1	0.01	recessive X-linked
tritanopia	0.002	0.001	dominant autosomal
protanomaly	1.0	0.02	recessive X-linked
deuteranomaly	4.9	0.38	recessive X-linked

Piantanida, Eddy, Shows & Hogness, 1986). More precisely, genetical analyses have shown that the L- and M-cone photopigments are encoded by genes that lie on the X-chromosome. As color vision deficiencies related to these cones are much more frequently observed in men than in women, it was supposed that protan and deutan defects are inherited as sex-linked recessive characteristics. In contrast, inherited tritan defects are rare and are not sex-linked. Table 2.2 gives an overview of the incident rates and genetics of the color vision deficiencies discussed here. Results from recent studies have shown that the patterns and variations of the underlying genetics are more complex than expected given only an examination of the phenotypes (Neitz, Neitz, He & Shevell, 1999; Neitz & Neitz, 2000).

2.2.5 Recent Results

Until this point we discussed dichromacy from the classical perspective which considers dichromatic vision as a reduced form of normal trichromatic vision. We will refer to this view as the *reduction hypothesis* of dichromatic color vision. According to this hypothesis, dichromatic color spaces are understood as two dimensional subspaces of the three dimensional trichromatic color space. However, several results from a number of recent studies are contradictory to the reduction hypothesis. A typical finding is that dichromats use the whole range of color terms when asked to describe their color percepts (Scheibner & Boynton, 1968; Montag, 1994; Paramei, Bimler & Cavonius, 1998; Wachtler, Dohrmann & Hertel, 2004). Moreover, under certain conditions, such as long stimulus presentation and large size of stimuli, dichromats are able to name the color of presented stimuli almost as accurately as normal trichromats (Montag & Boynton, 1987). The precision of dichromatic color naming depends on the size of the stimulus (Smith & Pokorny, 1977), the duration of stimulus presentation (Montag, 1994) and the level of luminance (Paramei et al., 1998).

Not only the color naming behavior of dichromats differs from what is predicted by the reduction hypothesis, contradictory findings come also from other areas of color vision. Smith and Pokorny (1977) report that dichromats do not accept classical dichromatic matches when the size of the stimulus extends to 8° of visual angle. Wavelength discrimination of dichromats is improved at high light levels (McMahon & MacLeod, 1998). Some authors report lower absolute thresholds of dichromats compared to trichromats under scotopic conditions (Verhulst & Maes, 1998) whereas others found no differences (Simunovic, Regan & Mollon, 2001). Morgan, Adam and Mollon (1992) report that dichromats detected color-camouflaged objects that were not detected by normal trichromats. Using a recognition memory task, Gegenfurtner, Wichmann and Sharpe (1998) found that dichromats recognized color images of natural scenes with the same accuracy as normal trichromats.

Several explanations of these findings contradictory to classical reduction hypothesis have been suggested. First, one assumption explaining large field trichromacy of dichromats is the contribution of rods to dichromatic color vision (Smith & Pokorny, 1977; Montag & Boynton, 1987). Second, the improved discrimination and color naming abilities of dichromats may result from the participation of a novel unknown photopigment (Montag, 1994). Third, at very high light levels, S-cones may contribute to discrimination in the long wavelength region as they receive enough light to exceed their threshold (McMahon & MacLeod, 1998).

These three explanations were tested in a study incorporating psychophysical and genetic methods by Crognale, Teller, Yamaguchi, Motulsky and Deeb (1999). When controlling for all three possible mechanisms, the range of large field Rayleigh matches of dichromats was still much smaller than expected. Crognale et al. suggest that this unexpected performance of dichromatic observers might be due to variations in optical density, that means differences in photopigment concentration which may lead to broader sensitivity functions, or due to variations in photopigment concentration across the retina.

An alternative explanation of the extraordinary color naming abilities of dichromatic observers is simply the assumption that dichromats may use learned strategies that increase the probability of correct assignments (Jameson & Hurvich, 1978; Bonnardel, 2006). This does not mean that dichromats necessarily perceive the same distinctive color categories as normal trichromats. We only know that dichromats are able to label presented stimuli more or less correctly but we do not know if different color categories of dichromats are actually related with qualitatively different percepts. However, this explanation can not account for many of the findings discussed above.

A different mechanism that may account for deviations from classical reduction hypothesis is suggested by Wachtler, Dohrmann and Hertel (2004). They propose a non-linear processing of signals from the two remaining cone types which results in a novel spectral response $\tilde{Q}$. If we consider for example the protanopic case, the $\tilde{Q}$-response is given by a non-linear transformation of the

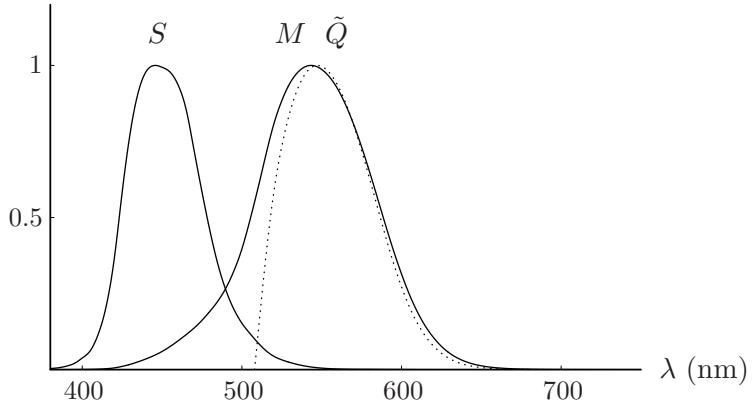

Figure 2.12: Pseudo-trichromatic signals of a protanope as suggested by the model of Wachtler et al. (2004). The solid lines show the sensitivity functions of the two remaining receptors M and S. The dotted line shows the response function of the virtual receptor $\tilde{Q}$ which is given by a non-linear transformation of the M- and S-signals with parameters $\gamma = 1.15$ and $\alpha = 5$.

M-signal minus the weighted signal from the S-cone:

$$\tilde{Q} = M^{\gamma} - \alpha S, \tag{2.24}$$

where γ is the parameter determining the degree of non–linearity and α the weight of the S-cone signal. Figure 2.12 shows the estimated response function of the virtual receptor $\tilde{Q}$ using parameter values of $\gamma = 1.15$ and $\alpha = 5$ as suggested by Wachtler et al. The authors were able to predict dichromatic hue scalings which were contradictory to the classical theory of dichromacy.[12]

Nonetheless, the model of Wachtler et al. raises several questions. First, it does not seem plausible that a non-linear processing of color signals must be assumed in dichromats but not in trichromats. Second, several well established aspects of dichromatic vision such as confusion lines, the existence of a neutral point or wavelength discrimination are not explained by the model.

To summarize, recent studies have shown deviations from the classical reduction hypothesis of dichromacy in several respects. Explanations that may account for these findings are variations of optical density across the retina and across observers, learned strategies and non–linear processing of signals from remaining cones in order to derive pseudo–trichromatic signals.

2.3 Opponent Colors Theory

2.3.1 Introduction

The trichromatic theory provided a framework within which many phenomena of color vision could be explained. However, critics of this theoretical approach

[12] An introduction to the hue scaling technique is given in Chapter 5.

argued that some phenomena such as afterimages and colored shadows lay beyond the scope of the theory and could not be explained adequately. The most prominent critic of trichromatic theory was Ewald Hering, who developed an alternative theory, the *opponent colors theory* (Hering, 1920). The main idea of this theory is the organization of color perception in pairs of antagonistic colors red-green, blue-yellow and white-black. The two basic colors of each chromatic pair are mutually exclusive which means that color impressions cannot be, for example, bluish and yellowish at the same time. Another important feature of the theory is the assumption that the four basic colors blue, yellow, red and green are conceptualized as *unique hues*. For example, a stimulus which is characterized as uniquely blue appears neither reddish nor greenish to the observer.

At first glance it seems as if the two theories—trichromatic theory and theory of opponent colors—are contradictory. In the late 19th century Helmholtz and Hering were involved in a long debate to decide which one of the two theories could describe human color vision more adequately. Now it is a well established fact that both theories complement one another and refer to two different stages of color processing in the visual system. This idea was already suggested by Johannes von Kries (1905) who integrated both theories within a so called *zone theory*. The main idea of the zone theory is that opponent mechanisms are located at higher stages of color processing and that these mechanisms get their input from the three different types of cones. In neurophysiological studies, neurons which response patterns are in good agreement with opponent colors theory have been found at several levels of the visual pathway (Svaetichin, 1956; De Valois, Abramov & Jacobs, 1966; see Gegenfurtner, 2003 for review).[13]

2.3.2 Quantitative Aspects of the Theory

Jameson and Hurvich (1955) formulated a quantitative version of the opponent colors theory based on psychophysical measurements (see also Hurvich, 1981). They determined experimentally response functions of the two chromatic opponent channels. Based on these measurements of opponent responses they were able to explain several phenomena, such as the Bezold-Brücke shift, wavelength discrimination and color vision deficiencies within the framework of opponent colors theory (Hurvich & Jameson, 1955). In their classical experiment subjects were presented with the mixture of two spectral lights a_λ and b_λ against a black background (Figure 2.13). The wavelength of light a_λ is fixed, but its intensity can be adjusted by the observer. For example, a_λ is chosen as unique blue with $\lambda = 475\,nm$, which appears neither reddish nor greenish to the observer. The wavelength of the second spectral light b_λ is varied throughout the experiment.

[13] As a matter of fact, the response patterns of opponent cells in the visual pathway do not exactly correspond with the opponent channels proposed by Hering (Krauskopf, Williams & Heeley, 1982; Derrington, Krauskopf & Lennie, 1984). Thus, the neural substrate of unique hues is still to be identified (Valberg, 2001).

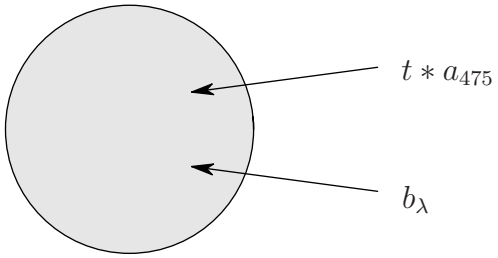

Figure 2.13: Hue cancellation experiment with the spectral light a_{475} as standard.

The task of the subject is to adjust the intensity t of light a_{475} until the mixture c appears neither bluish nor yellowish:

$$c = t * a_{475} \oplus b_\lambda. \tag{2.25}$$

In other words, the subject is asked to cancel the blue and the yellow hue from the mixture c. Therefore, this method is also called *hue cancellation*. In the terminology of opponent colors theory the mixture c is in *equilibrium* and the related percepts are called *BY-equilibrium colors*. We can now determine the intensities which are needed to cancel the blue and the yellow hue from the mixture c for all different b_λ. As a result we get a response function of the yellow process of the blue-yellow mechanism. If we replace a_{475} by unique yellow a_{580} and measure the amount of yellow across the visible spectrum that is needed to let the mixture appear neither bluish nor yellowish, the result will be the response function of the blue process. Both processes, the blue and the yellow process, complement one another and constitute the response function of the blue-yellow mechanism.[14] By convention the responses of one of the two processes are represented with negative signs.

In the same manner, the response function of the red-green process can be determined by replacing only a_{475} by appropriate spectral lights. In this case, the task of the subject is to adjust the intensity t of the standard a until the mixture c appears neither reddish nor greenish. The response function of the black-white mechanism gives only the 'whiteness' response and is identical with the luminous efficiency function (see Figure 2.6). The response functions of the three opponent mechanisms are shown in Figure 2.14.

[14]To see this recall that a stimulus can appear only bluish or yellowish. If we use a_{475} as standard and b_λ appears also bluish to the observer it is impossible to cancel the blue hue from the mixture. We can only measure a response from the yellow process if the spectral light b_λ appears yellowish but not bluish to the observer.

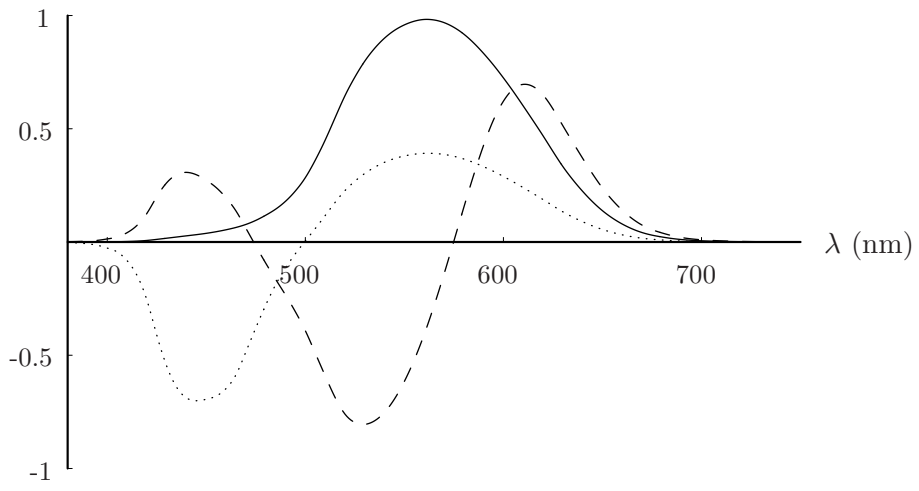

Figure 2.14: Response functions of the three opponent channels blue-yellow (dotted line), red-green (dashed line) and white-black (solid line).

2.3.3 Opponent Color Codes and Relation to Photoreceptor Excitations

Jameson and Hurvich supposed that the responses of the opponent mechanisms are linear functions of photoreceptor excitations. In this case, the BY- and the RG-equilibria, that is the set of all respective equilibrium colors, can be represented by two different planes in color space (Figure 2.15). The intersection of both planes contains all achromatic colors, the set of colors for which both chromatic mechanisms are in equilibrium. Now I want to introduce opponent color codes $\psi(a) = (\psi_1(a),\ \psi_2(a), \psi_3(a))$ which give responses of the three opponent mechanisms for each light a. The opponent color codes are defined so that positive and negative codes are possible for the two chromatic mechanisms. Positive and negative signs indicate the activity of one or the other process within each chromatic opponent channel respectively. If we assume the opponent color codes to be linear functions of the stimulus, the following relation must hold for all lights a, b and scalars $s, t \geq 0$:

$$\psi_i(s * a \oplus t * b) = s \cdot \psi_i(a) + t \cdot \psi_i(b); \qquad i = 1, 2, 3. \tag{2.26}$$

If we assume (2.26) to be true, we can express the responses of the opponent channels as a linear transformation of photoreceptor excitations. If we denote the response of the BY-mechanism for a given light a with $\psi_1(a)$, we can write $\psi_1(a)$ as a linear combination of photoreceptor codes $\phi(a)$ with fixed weights l_{BY}, m_{BY}, s_{BY}:

$$\psi_1(a) = l_{BY} \cdot \phi_1(a) + m_{BY} \cdot \phi_2(a) + s_{BY} \cdot \phi_3(a). \tag{2.27}$$

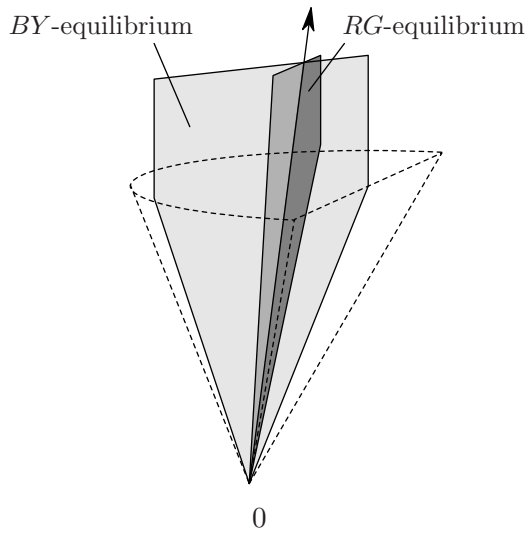

Figure 2.15: *RG*- and *BY*-equilibrium planes in color space. The intersection of both planes (arrow) defines the locus of achromatic colors.

For all BY-equilibrium colors the following equation must hold:

$$0 = l_{BY} \cdot \phi_1(a) + m_{BY} \cdot \phi_2(a) + s_{BY} \cdot \phi_3(a). \tag{2.28}$$

The three opponent mechanisms provide a new set of basis vectors of color space. In matrix notation the transformation of receptor codes to opponent color codes can be written as:

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} = \begin{pmatrix} l_{BY} & m_{BY} & s_{BY} \\ l_{RG} & m_{RG} & s_{RG} \\ l_{WBk} & m_{WBk} & s_{WBk} \end{pmatrix} \cdot \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix}. \tag{2.29}$$

Equation (2.29) can be also interpreted as the linkage between cone receptor output and response of the three opponent mechanisms. This linkage is shown schematically in Figure 2.16. The signs of the coefficients of the transformation matrix in (2.29) indicate if there is an excitatory ('+') or an inhibitory ('−') relation between one type of cone and an opponent channel. For example, the BY-response is given by a weighted sum of the L- and M-cone signals minus the weighted signal from the S-cone as seen in Figure 2.16. The weighting factors of the luminance (or white-black) channel are all positive except s_{WBk}. It is now generally accepted that the S-cones do not contribute to luminance and hence $s_{WBk} = 0$ (Eisner & MacLeod, 1980).

In the color cancellation experiment of Jameson and Hurvich (1955) the linearity of the opponent mechanisms as it is stated in (2.26) was not explicitly tested. However, Jameson and Hurvich defined the response functions of the

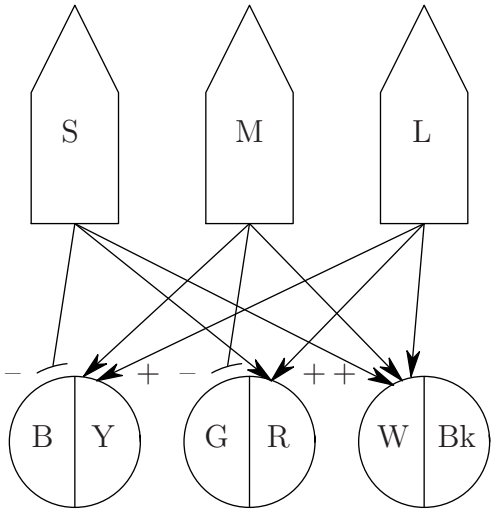

Figure 2.16: Schematic representation of the linkage between cone receptor output and opponent channels. The three photoreceptors are denoted with L, M and S. The signs '+' and '−' indicate excitatory and inhibitory effects respectively.

opponent channels as linear transformations of the color matching functions. Linearity of both chromatic opponent mechanisms was tested experimentally by Larimer, Krantz and Cicerone (1974; 1975). They found that (2.26) holds for the RG-mechanism but not for the BY-mechanism. Further evidence for the nonlinearity of the blue-yellow channel comes from experimental studies by Werner and Wooten (1979) and Webster, Miyahara, Malkoc and Raker (2000).

We shall now discuss implications of the opponent theory for red-green deficient dichromats. Exemplarily, we will focus on protanopic vision. Assuming the reduction hypothesis of dichromacy to be true these observers are lacking the L-cone photoreceptors. Therefore, the transformation of receptor codes to opponent signals (see Equation 2.29) reduces to

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} = \begin{pmatrix} m_{BY} & s_{BY} \\ m_{RG} & s_{RG} \\ m_{WBk} & s_{WBk} \end{pmatrix} \cdot \begin{pmatrix} \phi_2 \\ \phi_3 \end{pmatrix}. \tag{2.30}$$

We note that the transformation matrix has only rank 2. Therefore, the three opponent signals ψ_1, ψ_2, ψ_3 contain redundant information. It would be sufficient for the protanope to code signals only in two different opponent channels. One can verify easily that the same is true for deuteranopic observers.

Results from psychophysical studies indicate that protanopes and deuteranopes possess only two different opponent channels which can be characterized as white-black and blue-yellow channel (Hurvich & Jameson, 1955; Knoblauch, Sirovich & Wooten, 1985). The red-green channel is missing in these two classes of observers. The response functions of the protanopic opponent channels are

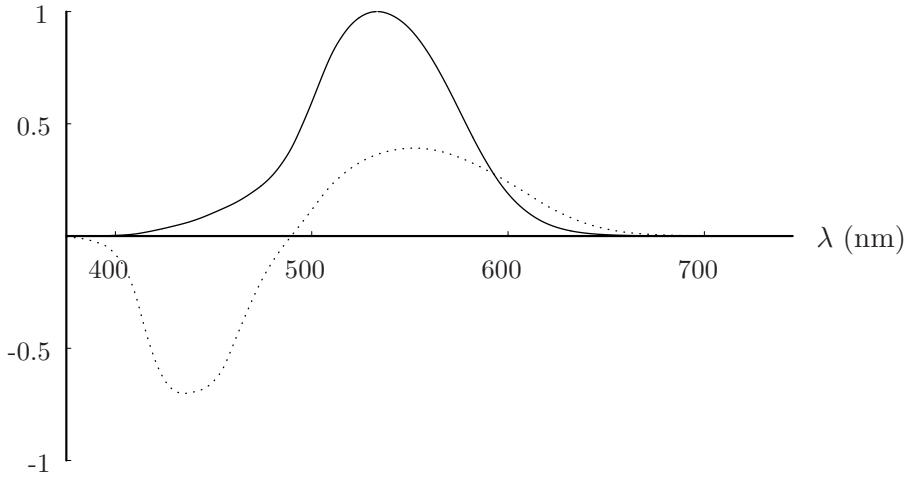

Figure 2.17: Response functions of the two protanopic opponent channels blue-yellow (dotted line), and white-black (solid line). The equilibrium of the BY-channel can be interpreted as the protanopic neutral point. Note that in comparison with Figure 2.14 the response functions are slightly shifted toward the shortwavelength region.

shown in Figure 2.17. Note that due to the lack of L-cone signals response functions are slightly shifted toward the shortwavelength region. The equilibrium of the BY-channel is perceived as achromatic in appearance by the protanope. It can be interpreted as the protanopic neutral point (see Table 2.1).

In Equation (2.30) we assumed silently that opponent signals are a linear function of cone excitations. Results from a study by Knoblauch, Sirovich and Wooten (1985) indicate that linearity of the chromatic blue-yellow channel holds for both protanopes and deuteranopes. This finding is not consistent with detected non-linearities of the blue-yellow channel in normal observers reported earlier.

Chapter 3

Color Constancy

In the previous chapter I gave an introduction to fundamental theories of color vision. These theories referred to stimuli that were viewed in isolation. A central issue within this framework was the assignment of primary color codes to isolated light stimuli. Hitherto we discussed the color appearance of such stimuli only marginally. In this chapter we will extend the examined stimulus to complex spatial patterns that contain more than one light. We will see that the color appearance of light stimuli in this context is not determined by corresponding primary color codes. Rather, the color appearance of a given light depends on temporal and spatial characteristics of the stimulus. We can study the influence of the context on the color appearance of a stimulus when we present two different lights either sequentially or spatially (as a center-surround configuration) to the observer. The corresponding phenomena are called *successive contrast* and *simultaneous contrast*. In this chapter we will focus on a related phenomenon that refers to color appearance of fixed objects under varying lighting conditions. This phenomenon is known as *color constancy*. Loosely speaking, color constancy is described as the ability of visual systems to assign stable color appearance to a fixed object under changing illuminant conditions.

3.1 The Problem of Color Constancy

3.1.1 Introduction

Probably the most important function of human color vision is object recognition. In this sense, the color appearance of an object should persist with changes of the illumination. We can experience this phenomenon every day but we rarely notice it. For example, a leaf appears green to us under bright sunny daylight and inside under tungsten light as well. This phenomenon which is called *color constancy* is an ability of our visual system. Research over the last few decades has shown that the human visual system achieves a high degree of color constancy over a

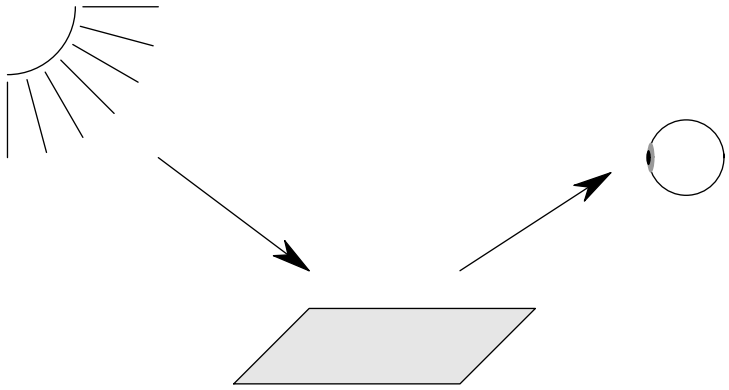

Figure 3.1: Formation of the color signal that reaches the eye of the observer.

wide range of illumination conditions. In fact, this feature of our visual system is very impressive for at least two reasons. First, the physical stimulus can change dramatically with changes of the illuminant. Second, the information that our visual system uses is very little and noisy.

We will now characterize the physical stimulus which will be referred to as the *color signal* that reaches the eye from an object (Figure 3.1).[1] This stimulus consists of two components, the incident illumination and characteristics of the object's surface. For reasons of simplicity we will assume that the light source is spatially uniform and that the objects are flat and matte. This environment is often referred to as the *Mondrian-* or *Flat World* (Maloney, 1999; Brainard, 2004). The illumination can be described by its spectral power distribution $E(\lambda)$. In studies on color constancy illuminants are often characterized as natural daylights. Appendix A gives an introduction to properties of natural daylights and describes a method to approximate their spectral power distributions. Objects in the scene differ in the way they reflect incident light at different wavelengths. This characteristic of a surface is represented by the spectral reflectance function $S(\lambda)$. For each wavelength, $S(\lambda)$ gives the fraction of incident light that is reflected from the surface. The color signal $C(\lambda)$, that is the spectral power distribution of the light which is reflected from the surface to the eye of the observer, can then be characterized by:

$$C(\lambda) = E(\lambda)S(\lambda). \tag{3.1}$$

Clearly, in the color signal informations about the illumination and the reflectance characteristics of the surface are already confounded. We are only interested in the latter when we wish to maintain persistent object recognition.

[1]The term *color signal* might be misleading as it is defined physically but not psychologically. I use this term here because its use is common in the color constancy literature (see for example Maloney, 1999).

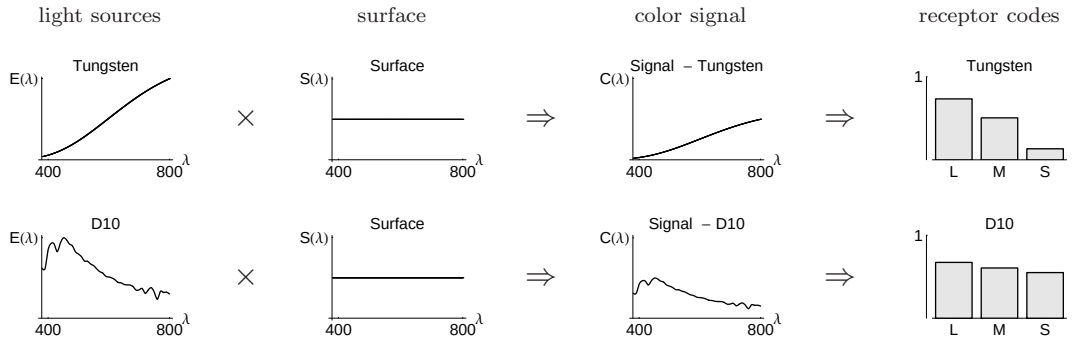

Figure 3.2: A fixed surface which appears gray under white daylight is rendered under two different illuminations: tungsten and bluish daylight D10. The first column shows the spectral power distributions of the two light sources. The second column gives the reflectance function of the surface. The resulting color signals and receptor codes (arbitrary units) are shown in columns three and four respectively. A comparison of both lines reveals that the color signals and the receptor codes differ strongly.

Hence, our visual system has to somehow discount the effect of the illuminant from the color signal. A second problem arises if we consider the limitations of our visual system. When the color signal reaches the cone photoreceptors a dramatic reduction of information results. In (2.16) we denoted the sensitivity functions of the three types of cones with $R_1(\lambda)$, $R_2(\lambda)$, $R_3(\lambda)$. The receptor codes $\phi_i(C)$ of the color signal are given with:

$$\phi_i(C) = \int_\lambda E(\lambda)\, S(\lambda)\, R_i(\lambda)\, d\lambda, \quad i = 1, 2, 3. \tag{3.2}$$

In general, for a fixed surface S and two different illuminants E_1 and E_2, the color signals and corresponding receptor codes differ.

In Figure 3.2, the effects of two different light sources on the color signal from a surface which appears achromatic to the observer under white daylight are shown. When seen under a bluish daylight $D10$ the three cone photoreceptors are excited almost equally. Under a yellowish tungsten illuminant the L-cone is much more strongly excited than M- and S-cones. If the surface is seen only locally, these different patterns of cone excitations induce very different color impressions. If the two surfaces are embedded in larger scenes instead, we are probably able to recognize that we are observing one and the same surface under two different illuminants.

On the other hand, identical color signals do not necessarily lead to identical color impressions. For example, imagine two different surfaces S_1 and S_2 which appear yellowish and bluish under white daylight respectively. The surface S_1 is illuminated with a bluish and surface S_2 with a yellowish light source so that identical receptor excitations from both surfaces result. Edwin Land showed in his impressive demonstrations of color constancy that when surfaces S_1 and S_2 are seen in corresponding contexts very different color appearances may result

(Land, 1977; see also Lotto & Purves, 2004). But if we present both surfaces in isolation to an observer they are perceptually indistinguishable for him. Hence it is impossible for the observer to decide if he sees a yellow surface under blue illumination or a blue surface under yellow illumination. A demonstration of this phenomenon is shown in Color Plate D.2 (Appendix D). From this example we can see that the confounding of information from the illuminant and the surface in the color signal can not be resolved without further information. Several suggestions have been made as to what kind of information the visual system uses to arrive at constant estimations of surface characteristics with changes of illumination. We will discuss some of these suggestions in the next section.

At this point I will try to define the concept of color constancy more precisely. In the simple lighting model that was just introduced we considered only effects of the illuminant on the color signal but ignored other variables that may influence stable perception of surface colors such as interreflections between surfaces in the scene or atmospheric haze. In the literature color constancy is often defined as the ability of a visual system to discount only the effect of the illuminant from the color signal (e.g. Bäuml, 1995). Maloney (1999) has pointed out correctly that a definition of color constancy should take into account the aforementioned additional factors as well. According to Maloney (1999)

> ... an observer has (perfect) color constancy precisely when the color appearance assigned to a small surface patch by the visual system is completely determined by that surface's local spectral properties (Maloney, 1999, p. 389).

It should be noted that human performance in this sense is not perfect and that color constancy is achieved only to a certain degree. Throughout this work we will mainly concentrate on the ability of the visual system to discount the effect of the illuminant from the color signal and ignore other potential variables. In this sense, the objective of a color constant system is to provide stable estimates of reflectance characteristics of surfaces across different illuminations. Often this estimation problem is broken down into two subproblems. First, the system has to get an estimate of the illumination using different cues from the viewed scene.[2] Second, this estimate is used to determine certain intrinsic properties of the surface that are invariant with changes of the illuminant. In the next section I will describe this estimation process in more detail.

3.1.2 Simultaneous and Successive Color Constancy

In recent years, the distinction between *simultaneous* and *successive color constancy* has evoked wider interest (Bäuml, 1997; Brainard, 1998). Let us first

[2]The term 'estimate' refers to the estimation of photoreceptor excitations that are related with the illuminant. In Section 3.2.3 we will discuss models where it is assumed that the illuminant is estimated only up to a scaling factor.

illustrate both phenomena on two examples. When we take a look at Figure D.2 (Appendix D) we see one and the same simulated scene under yellowish and bluish illumination. We note that we can easily assign a fixed surface on the checkerboard under yellowish illumination to the corresponding surface under bluish illumination. We are immediately able to recognize that it is one and the same surface under two different illuminants. Also, we are instantly able to tell the color of the surface, but we have the impression that in both cases the 'true color' of the surface is veiled with the colors of the respective illuminants. It seems to us as if the color appearance of both surfaces is different, but the 'true' color of the surface is the same. In other words, in such situations we are also somehow able to get an impression of the color of the illuminant. When the observer is asked to tell the 'true' color of the surface which is often termed *surface color* we find that observers solve this task quite reliably across different simultaneously presented illuminations. This phenomenon is often referred to as *simultaneous color constancy*.

Let us now imagine a situation where an observer is watching a real object (e. g. a blue traffic sign) under daylight conditions in the morning and at noon. At both occasions the observer will be able to recognize the object as blue traffic sign although the illumination conditions may have changed dramatically. Moreover, the observer may note that the color appearance of the sign, which means its hue, saturation and brightness, did not change substantially from morning until noon. The phenomenon of perceiving the color of a fixed object as unchanged across temporal changes in illumination is often referred to as *successive color constancy*.

Under natural viewing conditions both phenomena, simultaneous and successive color constancy, do rarely occur in isolated form. Often we are confronted with situations where both modes of color constancy are involved. We could for example imagine a scene illuminated by gradually changing daylight with one part of the scene exposed to direct sunlight whereas an adjacent part situated in the shade. Therefore, it is difficult to extract the mechanisms underlying each of the two phenomena.

One approach to investigate simultaneous color constancy is to present subjects with one and the same (simulated) scene under two different illuminations simultaneously side by side (Arend & Reeves, 1986). The task of the subject is then to match surface colors. More precisely, the subject is asked to adjust a test surface in the first scene so that it looked as if it was cut from the same piece of paper as the corresponding surface in the second scene. Successive color constancy is often investigated using achromatic settings. The subject is presented with one (simulated) scene under a fixed illuminant (Bäuml, 1997; Brainard, 1998). Before making the settings, the subject adapts to the stimulus for a sufficient time. The task of the subject is to set a given test surface so that it is perceived as neither bluish nor yellowish and neither reddish nor greenish.

It is assumed that two different visual processes are associated with the two

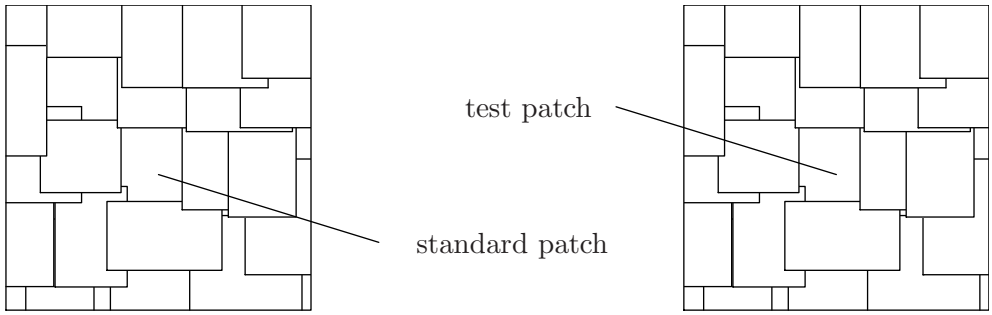

Figure 3.3: Two *Mondrian* stimuli in asymmetric matching. Both Mondrians are usually rendered under different illuminations. Subjects are asked to match a surface from the test *Mondrian* (right) according to given criteria with a patch from a standard *Mondrian* (left).

classes of color constancy (Bäuml, 1997). Both processes differ in their time courses and are presumably located at different stages of the visual system. The *perceptual processes* that are supposed to underlie simultaneous color constancy are acting very fast. In contrast, the *sensory processes* that are related with successive color constancy work very slowly. The former class of processes is assumed to be located at higher levels of visual processing whereas the latter is presumably located at low levels of the visual system.

In summary, the phenomenon of color constancy can be understood as an interplay of both simultaneous and successive constancy and their underlying mechanisms. This interaction of both classes of processes maintains constant perception of object colors over a wide range of different situations.

3.1.3 Methods for Investigating Color Constancy

Typical stimuli

The classical stimulus in the field of color constancy is called *Mondrian* stimulus.[3] It was introduced by Edwin Land (Land & McCann, 1971) and consists of a number of matte homogeneous surfaces that differ in color appearance (Figure 3.3). The planar *Mondrian* is rendered under a spatially uniform illumination. The intention of using this sort of stimuli is to study color constancy in an environment that provides on one hand no meaning to the observer but on the other hand is more complex than simple center-surround configurations. In recent times, mostly simulations of *Mondrians*, rendered under chromatic illuminations are presented on CRT-displays to the observer.

Planar matte stimuli such as *Mondrians* represent a simplified model of natural environments which we already referred to as the Flat World model. It has been criticized that this model is not sufficient to describe the conditions in nat-

[3]The term *Mondrian* is a reference to the Dutch painter Piet Mondriaan who created similar paintings.

ural environments adequately (Maloney, 1999). Hence, in recent years there is a tendency to use enriched stimuli that contain additional cues to the illuminant. One group of stimuli in this context are three dimensional representations of simulated scenes (Boyaci, Maloney & Hersh, 2003; Delahunt & Brainard, 2004b). Examples of such simulated scenes are given in Color Plates D.1–D.3 (Appendix D). The model underlying these representations is more complex than the Flat World model and is often referred to as the *Shape World model*. We will come back to this model in Section 3.3. Another approach is the use of real stimuli that are arranged in experimental chambers (Brainard, Brunt & Speigle, 1997; Kraft & Brainard, 1999). These authors promise to investigate color constancy under nearly natural viewing conditions. A third class of stimuli that has evoked interest in recent years are known meaningful objects such as fruits and vegetables. Studies using these stimuli have shown that knowledge about the meaning of objects can enhance color constancy (Hansen & Gegenfurtner, 2006; Hansen, Olkkonen, Walter & Gegenfurtner, 2006).

Asymmetric matches

The most popular method to investigate color constancy are asymmetric matches. Subjects are presented simultaneously with two scenes (preferably two *Mondrians*) that are usually rendered under different illuminations (Figure 3.3). The task of the subject is to match a surface in one scene according to a certain criterion with a standard surface in the other scene. The instruction given to the subject is crucial in asymmetric matching as emphasized by Arend and Reeves (1986). In the previous section I already mentioned that settings of subjects can differ enormously if they are either asked to set paper matches or appearance matches.

There are several problems related with asymmetric matches. First, as outlined above, the instruction plays a crucial role in the matching procedure and it remains unclear what the observer actually matches. Furthermore, observers often report that they cannot find a satisfying setting (see Ekroll, Faul & Niederée, 2004). One possible explanation for these difficulties of subjects in finding exact matches is provided by Logvinenko and Maloney (2006). They investigated the proximity structure of achromatic surfaces under neutral illuminations of different intensities. The authors found that achromatic colors constitute a two-dimensional manifold with dimensions interpretable as perceived albedo and intensity of the illuminant. They conclude that 'an exact match in asymmetric lightness matching is impossible in principal' (Logvinenko & Maloney, 2006, p. 80). We will return to this issue in Chapter 5. A third problem related with asymmetric matches may occur when presenting two unchanged stimuli side by side while only the illumination changes. If the standard patch is identical in appearance to a second surface in the standard *Mondrian*, the observer only needs to match the test patch with the corresponding surface in the test *Mondrian*.

In this case, observers are able to produce color constant matches even without having color constancy in reality. Finally, when presenting two stimuli under different illuminations side by side, this procedure raises questions concerning the adaptational state of the observer.

A variant of classical asymmetric matches are successive asymmetric matches (Brainard & Wandell, 1992). In this case, the two scenes are presented successively to the observer. The task of the subject is to adjust a given test surface from the second scene in appearance with the corresponding surface from the standard scene. Naturally, this matching procedure requires the subjects to recall the appearance of the standard surface from memory and is hence also called *memory matching*. Although this rationale may raise questions about the influence of color memory on the matches, subjects are generally setting quite reliable matches. The advantage of this procedure is the control of the adaptational state of the observer which is normally not given in simultaneous asymmetric matches.

Achromatic settings

Another technique to measure the degree of color constancy exhibited by human observers are achromatic settings. The subject is presented with only one scene rendered under a certain uniform illumination. The scene and the illuminant are often simulated on CRT-displays. In the experiment the subject is asked to adjust a given test surface in the scene so that it appears neither bluish nor yellowish and neither reddish nor greenish to him. An advantage of this method is that it is easy to explain and that observers produce more reliable settings compared to asymmetric matches. The main problem related with achromatic settings is that illuminant adjustments are measured essentially only for achromatic surfaces. Nevertheless, Speigle and Brainard (1999) were able to predict accurately asymmetric matches from achromatic settings.

Other methods

Asymmetric matches and achromatic settings focus on quantitative aspects of color constancy. The problems related with these techniques arise doubt if these two methods effectively measure color constancy (Foster, 2003). A more direct approach to determine qualitatively the degree of color constancy is color naming. Subjects are simply asked to label a given test surface using basic color terms (Troost & de Weert, 1991). Problems of this technique are that the number of categories is limited and that it is difficult to compare results with quantitative measures of color constancy. In Chapter 5 hue scalings will be introduced as an alternative method which combines aspects of quantitative and qualitative measures of color constancy.

An interesting approach has been suggested by Foster and colleagues (Craven & Foster, 1992; Nascimento & Foster, 1997). They focus on the question if

subjects are able to attribute changes in color appearance correctly to either illumination or surface changes. This method has proven useful when investigating conditions under which human color constancy exists and those under which color constancy fails. A non-experimental group of methods is the theoretical analysis of light-surface interactions and its relation to color constancy (Foster & Nascimento, 1994). Such analysis focuses on properties of scenes that are invariant with changes of the illuminant and might underlie color constancy mechanisms of our visual system. Recently, the analysis of scene statistics of natural scenes has raised interest within this context (Mausfeld & Andres, 2002; Nascimento, Ferreira & Foster, 2002; Golz & MacLeod, 2002).

3.2 Models of Color Constancy

3.2.1 Introduction

In the previous section I discussed the problem of color constancy. One approach for a visual system to achieve color constancy could be to somehow segregate information about the illumination and the surface that are confounded in the color signal. In fact, a group of models of color constancy is following this idea. These models are generally organized in two steps: first, the illumination in the scene is estimated using certain available cues and second, based on an estimate of the illuminant, designators of the surfaces' reflectance properties are derived. Such models are often referred to as *computational models* of color constancy.

A second group of models focuses on relations between different surfaces within the scene across illuminant changes. Models of this type which we will refer to as *relational models* of color constancy try to derive codes from such relations that promise to be invariant across different lighting conditions. The basic principle inherent in these models is the so called *von Kries receptor scaling* which goes back to von Kries (1905; see also MacAdam, 1970). The original idea of von Kries was that the sensitivities of cone photoreceptors are rescaled after temporal adaptation to a given stimulus. However, the principle of von Kries receptor scaling has been influential not only in modeling temporal context effects (as originally proposed by von Kries), but also in modeling spatial context effects (e.g. Walraven, 1976) and color constancy (e.g. Land, 1977). The proposed linear rescaling is performed independently for each class of cones. As a consequence, the rescaling can be described as multiplication of the cone sensitivity $R(\lambda)$ with

a scalar ρ which depends on the context a:[4]

$$R_i^a(\lambda) = \rho_i^a \cdot R_i(\lambda), \qquad i = 1, 2, 3. \tag{3.3}$$

Equation (3.3) implies that in case of no adaptation $\rho_i = 1$ for all i. Two adaptational stimuli a, b that are metamers are linked with identical von Kries coefficients and vice versa:

$$a \sim b \Leftrightarrow \rho_i^a = \rho_i^b, \qquad i = 1, 2, 3. \tag{3.4}$$

If we denote a stimulus c after adaptation to context a with (c, a), we can use a new psychophysical relation $\approx$ which describes matches of stimuli in different contexts, for example $(c, a) \approx (d, b)$. If the von Kries coefficient law from (3.3) holds, the following relations must also hold for lights a, b, c, d and scalars $t > 0$:

$$\begin{aligned} (c, a) \approx (d, a) \quad &\Rightarrow \quad (c, b) \approx (d, b) \qquad (persistency) \\ (c, a) \approx (d, b) \quad &\Rightarrow \quad (t * c, a) \approx (t * d, b) \quad (proportionality). \end{aligned} \tag{3.5}$$

These rules of persistency and proportionality provide empirical testable assumptions of the von Kries coefficient law. In the next section we will discuss models of color constancy which are based on this elementary principle of adaptational processes.

3.2.2 Relational Models of Color Constancy

The idea that the visual system might analyze relations between surfaces within a scene in order to maintain color constancy is not new and goes back to Schrödinger (1920). This notion is based on the assumption that relations between surfaces in terms of ratios of color codes of the respective color signals remain quite stable with changes of illumination. If we denote the color signal from surfaces s under illuminant e with $e \cdot s$ then for all surfaces s, s' and illuminants e, e' we can write this assumption in terms of receptor codes ϕ_i as:

$$\frac{\phi_i(e \cdot s)}{\phi_i(e \cdot s')} = \frac{\phi_i(e' \cdot s)}{\phi_i(e' \cdot s')}, \qquad i = 1, 2, 3. \tag{3.6}$$

In a simulation study, Foster and Nascimento (1994) analyzed theoretical receptor codes of color signals from a large set of different natural and artificial surfaces and illuminants. They showed that for natural illuminations and surfaces Equation (3.6) approximately holds. In contrast, this principle was violated if random

[4]The term 'context' can have different meanings depending on the examined phenomenon. In the original von Kries coefficient law the context referred to temporal adaptation to a previous stimulus. In case of simultaneous contrast, the context is defined by the spatial surrounding. When investigating color constancy the context refers to characteristics of the scene that are related with the global illumination conditions.

power distributions of lights and random reflectance spectra of surfaces were used. Similarly, in a simulation study Dannemiller (1993) analyzed rank orderings of photoreceptor excitations from natural objects under different illuminations. He found that such rank orderings persist with changes of the illuminant. This conjecture is weaker than the assumption of invariant cone-excitation ratios from Equation (3.6). Taken together these two findings suggest that information of invariant properties of natural scenes across illuminant changes is available to the visual system.

The crucial question is now whether our visual system in fact uses this sort of information to achieve color constancy. Craven and Foster (1992) showed that when presented with two successive Mondrian patterns observers are able to discriminate reliably between illuminant and surface changes. The two conditions differed in the stability of cone excitation ratios. When presenting illuminant changes cone excitation ratios were approximately preserved whereas this was not the case for surface changes. In another experiment, Nascimento and Foster (1997) found strong evidence for the assumption that observers analyze effectively cone excitation ratios. Again, they presented their subjects with two successive Mondrian patterns and asked them to detect illuminant changes. The changes were either natural illuminant changes with approximately preserved cone excitation ratios or artificial illuminant changes with perfectly preserved ratios. The subjects reported more often to have seen an illuminant change when cone excitation ratios were constant compared to natural illuminant changes. However, results from a recent study by Rinner and Gegenfurtner (2002) suggest that constant cone ratios of signals from adjacent surfaces are not necessary for color constancy. They found that observers adjusted equally well to the incident illumination regardless of whether cone excitation ratios across illuminant changes were preserved or decorrelated.

The preceding analyses of color signals in natural scenes showed that simple von Kries receptor scaling could be a useful mechanism to provide approximately invariant percepts across illuminant changes. We will now discuss specifications of the von Kries model in the context of color constancy. First we will focus on the conjecture that the visual system adjusts color signals from surfaces in relation to the most intense region in the scene, an idea which was already proposed by Schrödinger (1920). This assumption implies that in general, the most intense surface in the scene is interpreted as 'white' by the visual system. Furthermore, it is assumed that this surface reflects the incident light almost completely and hence provides the observer directly with an estimate of the illumination. If we weight the color signal from any surface s under fixed illumination e with the color signal from a white surface w, this may yield a code which is invariant with changes of the illuminant. We can denote the receptor codes of the color signal

from the most intense surface with $\phi_i(e \cdot w)$ and then define a code $\tau_i(s, e)$ with

$$\tau_i(s, e) = \frac{1}{\phi_i(e \cdot w)} \cdot \phi_i(e \cdot s) = \frac{1}{\phi_i(e)} \cdot \phi_i(e \cdot s), \qquad i = 1, 2, 3. \tag{3.7}$$

Obviously, the codes $\tau_i(s, e)$ are all one if $s = w$. Moreover, Equation (3.7) bears resemblance to (3.3) which expressed the unspecified von Kries model. Clearly the model (3.7) is a specification of the von Kries coefficient law with $\rho_i = 1/\phi_i(e)$. We will discuss empirical evidence for model (3.7) further below.

Another relational model of color constancy is based on the idea that the visual system is taking the mean color signal across all surfaces in the scene into account. Furthermore, in this context it is assumed that the mean color signal (which is in fact virtual) appears achromatic to observers under neutral illumination. Hence, this notion has been discussed in the literature as the *grayworld assumption* (Buchsbaum, 1980). Similarly to model (3.7) it is assumed that color signals from surfaces are weighted with coefficients to derive codes that are invariant across changing illumination conditions. In this case the weighting coefficients are related with the receptor codes of the average color signal in the scene. Thus we can define codes $\kappa_i(s, e)$ as

$$\kappa_i(s, e) = \frac{1}{\frac{1}{n}\sum_j^n \phi_i(e \cdot s_j)} \cdot \phi_i(e \cdot s), \qquad i = 1, 2, 3. \tag{3.8}$$

An influential model of color constancy which shares some structural similarities with model (3.8) is the Retinex-model of Land and colleagues (Land & McCann, 1971; Land, 1986).[5] Several versions of the model exist which suggest algorithms that are supposed to provide color constancy for artificial systems. Brainard and Wandell (1986) showed in a detailed analysis of the Retinex-model that one of the algorithms can be rewritten as normalization of receptor codes based on the geometric mean of cone excitations. Thus, this algorithm is equivalent with model (3.8) when replacing the arithmetic mean with $\sqrt[n]{\Pi_j^n \phi_i(e \cdot s_j)}$. Obviously both versions of the model can be interpreted as specifications of von Kries receptor scaling with inverse mean cone excitations as coefficients. Now we return to the grayworld assumption and derive an even more specific version of model (3.8) that is closely related to model (3.7). Taking the grayworld assumption into account, we can rewrite the virtual mean surface as $t * w$ with $0 < t < 1$. The model becomes then

$$\kappa_i(s, e) = \frac{1}{\phi_i(e \cdot (t * w))} \cdot \phi_i(e \cdot s) = \frac{1}{t \cdot \phi_i(e)} \cdot \phi_i(e \cdot s), \qquad i = 1, 2, 3. \tag{3.9}$$

Thus, assuming that the visual system is able to determine mean cone excitations, the system gets a direct estimate of the chromaticity of the illuminant. The codes τ from model (3.7) and codes κ are identical up to a multiplicative constant.

[5]The neologism *Retinex* reflects the assumption that the suggested mechanisms underlying color constancy are located at retinal and cortical levels.

We will now discuss empirical evidence for models (3.7) and (3.8). McCann, McKee and Taylor (1976) analyzed theoretical cone excitations of color signals from Munsell surfaces under different illumination conditions. In a first simulation study they tested if normalization by the most intense region yields constant designators of surfaces (which they called *integrated reflectances*) for a fixed surface under different illuminants. They found that in general such designators were invariant across illumination changes. However, McCann et al. reported systematic deviations from constant designators in the sense that integrated reflectances became larger with increasing intensity of the illuminant. In the second part of the study, McCann et al. analyzed whether normalization by the geometric mean of the scene provides invariant codes (which they called *computed lightness*) with changes of the illumination. They found that for their two dimensional matte *Mondrian* stimuli such codes were preserved across different illumination conditions.

In a more recent study, Kraft and Brainard (1999) compared experimentally the effectiveness of different mechanisms of color constancy. These experiments were carried out in an illuminated chamber under nearly natural viewing conditions. Subjects were asked to make achromatic settings for a test patch under two different illuminants. In one condition the color signal from the most intense region was manipulated so that it was identical under both illuminations. In the second condition, the mean cone excitations were equaled in the same way. In both cases subjects still showed moderate residual color constancy. In contrast, when provided with all available cues subjects showed very good constancy. The results demonstrate that neither of the two mechanisms alone accounts for color constancy. Rather, the visual system might analyze different cues from the scene and combine them in order to estimate invariant designators (Maloney, 2002). We will come back to this idea later in this chapter.

Evidence for the hypothesis that the spatial mean alone can not account for color constancy also comes from Brown and MacLeod (1997). They report that color appearance of fixed surfaces changed dramatically when the variance of the background was altered while the spatial mean cone excitations were kept constant. This finding leads to the hypothesis that the visual system might analyze further characteristics of the scene that are not evident at first glance. Such characteristics are sometimes referred to as *second order statistics* of the scene. Further evidence for the importance of second order statistics as possible cues to the illuminant comes from Mausfeld and Andres (2002) and Golz and MacLeod (2003).

The preceding remarks make clear that the models (3.7) and (3.8) cannot be taken as exclusive explanations of the phenomenon of color constancy. Another problem of both models has to do with the additional assumptions of white and gray appearance of the most intense and average surface respectively. If these assumptions are violated it is questionable if the visual system is able to arrive at constant designators (Brainard & Wandell, 1986). Moreover, it is unclear if

the grayworld assumption generally holds for natural scenes. A further problem occurs if the grayworld assumption is violated or not made. This problem can be referred to as the missing neutral point or missing anchor. It becomes evident if we imagine two scenes which contain for example only reddish and only greenish surfaces respectively. If we determine designators for a surface in each scene that may represent the respective mean, model (3.8) leads to identical designators in both cases although the two surfaces differ in color appearance.

Another problem of color constancy models of von Kries type is that the influence of simultaneous contrast effects is largely ignored. Such models predict identical designators for a fixed surface regardless of the arrangement within the scene. However, the local surround plays undoubtedly an important role as a possible mechanism that is underlying color constancy (Valberg & Lange-Malecki, 1990; Kraft & Brainard, 1999). Finally, the described models explicitly assume only integration of information over spatial characteristics of a scene. In Chapter 5 I will report evidence for the conjecture that the visual system is also able to integrate information from isolated surfaces over time.

3.2.3 Computational Models of Color Constancy

We already discussed the problem of color constancy at the beginning of this chapter. In this context I outlined that in the color signal which reaches the eye of the observer from a surface, information about the illuminant and reflectance properties of the surface are already confounded. Without further restrictions on illuminations and surfaces, this confounding in the color signal cannot be resolved. The analysis of spectral power distributions of natural daylights has shown that these distributions follow certain regularities (Judd, MacAdam & Wyszecki, 1964). Judd et al. measured spectral power distributions of a large number of daylights at different times of the day and at different weather conditions. They found that these distributions were fairly smooth and showed similar patterns of variation. Furthermore, the chromaticities of these daylights were located in CIE-xy-diagram on a curve close to the Planckian locus (see Figure A.1 in Appendix A). Due to these regularities it is possible to approximate the spectral power distributions of daylights using a linear model with a small number of basis functions. Judd et al. performed a principal components analysis on their data and were able to recover the spectral power distributions of daylights extremely well using only three basis functions. The estimated spectral power distribution of a daylight $\hat{E}(\lambda)$ can then be written as a linear combination of the $M = 3$ basis functions $E_j(\lambda)$ with weights ε_j:

$$\hat{E}(\lambda) = \sum_{j=1}^{3} \varepsilon_j E_j(\lambda). \tag{3.10}$$

Hence, given the three basis functions each daylight can be represented by only three numbers ε_1, ε_2 and ε_3.

Similarly, the reflectance functions of natural surfaces show regularities that make it possible to describe them using a linear model of small dimension. Several researchers were able to approximate reflectance functions of Munsell papers or natural objects using only three to eight basis functions (Cohen, 1964; Maloney, 1986; Parkkinen, Hallikainen & Jaaskelainen, 1989). When the limitations of the human visual system in terms of trichromacy are taken into account, three basis functions provide already a good approximation (Maloney, 1986). We can denote estimated reflectance functions $\hat{S}(\lambda)$ as linear combinations of $N = 3$ basis functions $S_k(\lambda)$ with corresponding weights σ_k:

$$\hat{S}(\lambda) = \sum_{k=1}^{3} \sigma_k S_k(\lambda). \tag{3.11}$$

Using the linear models from (3.10) and (3.11) we can approximate the color signal $C(\lambda)$ (see Equation 3.1) with

$$\hat{C}(\lambda) = \sum_{j=1}^{3} \sum_{k=1}^{3} \varepsilon_j \; \sigma_k \; E_j(\lambda) \; S_k(\lambda). \tag{3.12}$$

Substituting this equation into Equation (3.2) we get:

$$\phi_i(\hat{C}) = \int_\lambda \sum_{j=1}^{3} \sum_{k=1}^{3} \varepsilon_j \; \sigma_k \; E_j(\lambda) \; S_k(\lambda) \, R_i(\lambda) \; d\lambda, \quad i = 1, 2, 3, \tag{3.13}$$

which is equivalent to

$$\phi_i(\hat{C}) = \sum_{k=1}^{3} \sigma_k \int_\lambda \sum_{j=1}^{3} \varepsilon_j \; E_j(\lambda) \; S_k(\lambda) \, R_i(\lambda) \; d\lambda, \quad i = 1, 2, 3. \tag{3.14}$$

The representation in (3.14) has the advantage that components which depend on the illuminant and those which are independent of the illuminant are disentangled. The integral $\int_\lambda \sum_{j=1}^{3} \varepsilon_j \; E_j(\lambda) \; S_k(\lambda) \, R_i(\lambda) \; d\lambda$ varies with the illuminant whereas the weighting factors of the surface reflectance linear model σ_k are constant across different illuminations. These coefficients represent invariant surface characteristics and are sometimes called *intrinsic colors*. The main goal of most computational models is to find reliable estimates of these intrinsic colors. The idea of disentangling the color signal becomes more evident when we write (3.14) in compact matrix notation:

$$\boldsymbol{\phi} = \boldsymbol{\Lambda}_\varepsilon \boldsymbol{\sigma}, \tag{3.15}$$

where $\boldsymbol{\phi}$ is the vector of $P = 3$ cone excitations, $\boldsymbol{\sigma}$ is the vector of intrinsic colors ($N = 3$) and $\boldsymbol{\Lambda}_\varepsilon$ is a 3×3 matrix. This matrix is often referred to as the *lighting matrix*.[6] The ik-th entry of $\boldsymbol{\Lambda}_\varepsilon$ is $\int_\lambda \hat{E}(\lambda) \; S_k(\lambda) \, R_i(\lambda) \; d\lambda$ which

[6] In general, the lighting matrix $\boldsymbol{\Lambda}_\varepsilon$ is a $P \times N$ matrix.

can be interpreted as the excitation in the i-th photoreceptor that results from observing the virtual basis surface S_k under the estimated illuminant $\hat{E}$. The lighting matrix $\mathbf{\Lambda}_\varepsilon$ provides a linear mapping of a surface S that is characterized by the coefficients $\boldsymbol{\sigma}$ to cone excitations $\boldsymbol{\phi}$. Obviously, the intrinsic colors of a fixed surface are mapped onto different receptor codes when illumination varies. Suppose the lighting matrix $\mathbf{\Lambda}_\varepsilon$ is known. In this case the intrinsic colors of a given surface can be easily determined from receptor codes by inverting $\mathbf{\Lambda}_\varepsilon$ and solving the system of equations:

$$\boldsymbol{\sigma} = \mathbf{\Lambda}_\varepsilon^{-1}\boldsymbol{\phi}. \tag{3.16}$$

Unfortunately the illumination is usually not known. Hence, the first goal of computational models of color constancy is the estimation of $\mathbf{\Lambda}_\varepsilon$. In a second step, Equation (3.16) is then solved for the intrinsic colors $\boldsymbol{\sigma}$. The models differ particularly in the algorithm of estimating the illuminant and in underlying assumptions. I will now briefly introduce three influential algorithms: the reference surface algorithm of Brill (1978), the grayworld algorithm of Buchsbaum (1980) and the subspace algorithm of Maloney and Wandell (1986).

The reference surface algorithm of Brill

Brill (1978) assumes that three linearly independent reference surfaces S_1, S_2, S_3 are given in the scene and $P = N = 3$. Furthermore, the three reference surfaces are taken as the basis of the three dimensional space of surfaces. In contrast, there are no restrictions on the illuminant. This leads to a system of nine equations:

$$\begin{aligned} \boldsymbol{\phi}_1 &= \mathbf{\Lambda}_\varepsilon\boldsymbol{\sigma}_1 \\ \boldsymbol{\phi}_2 &= \mathbf{\Lambda}_\varepsilon\boldsymbol{\sigma}_2 \\ \boldsymbol{\phi}_3 &= \mathbf{\Lambda}_\varepsilon\boldsymbol{\sigma}_3. \end{aligned} \tag{3.17}$$

This system of linear equations can be solved for the nine unknowns which are the elements of $\mathbf{\Lambda}_\varepsilon$. Once the lighting matrix is known, Equation (3.16) can be solved for all the other surfaces in the scene. The resulting intrinsic colors refer to the basis formed by the three reference surfaces. The Brill algorithm has the advantage that no restrictions on the illuminant are needed. The disadvantage of this algorithm is that at least three reference surfaces have to be available in the scene.

The grayworld algorithm of Buchsbaum

Buchsbaum (1980) assumes that only one reference surface is needed and that $M = N = 3$. He suggests that this reference surface might be just the mean of all surfaces in the scene and he assumes that this average surface appears gray

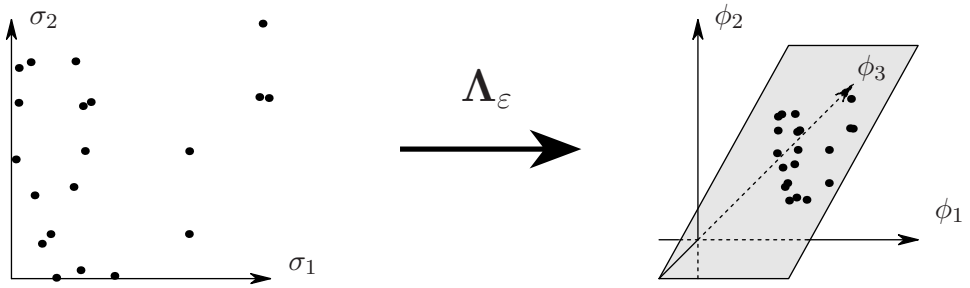

Figure 3.4: Schematic representation of the subspace algorithm. The lighting matrix $\mathbf{\Lambda}_\varepsilon$ provides a mapping from the two dimensional space of intrinsic colors onto a plane in cone excitation space (after Maloney & Wandell, 1986).

to the observer. We already referred to this assumption earlier as the *grayworld assumption*. Interestingly, this assumption is not necessary for Buchsbaums algorithm to work. The idea of his algorithm is the following: The only unknown quantities in $\mathbf{\Lambda}_\varepsilon$ are the three coefficients ε_1, ε_2 and ε_3. Initially, receptor codes $\boldsymbol{\phi}_1$, $\boldsymbol{\phi}_2$, $\boldsymbol{\phi}_3$ of the reference surface S_0 under the three basis lights E_1, E_2, E_3 are determined. This is possible when we assume that the codes $\boldsymbol{\sigma}_0$ of the reference surface are known. For example, the receptor codes $\boldsymbol{\phi}_1$ of the reference surface S_0 under the first basis light E_1 are given with:

$$\boldsymbol{\phi}_1 = \mathbf{\Lambda}_1 \boldsymbol{\sigma}_0, \tag{3.18}$$

where $\mathbf{\Lambda}_1$ is the lighting matrix with $\varepsilon_1 = 1$ and $\varepsilon_2 = \varepsilon_3 = 0$. After we got the three vectors of cone absorptions $\boldsymbol{\phi}_1$, $\boldsymbol{\phi}_2$, $\boldsymbol{\phi}_3$ Equation (3.15) can be rewritten as

$$\boldsymbol{\phi}_0 = \mathbf{\Lambda}_\varepsilon \boldsymbol{\sigma}_0 = \varepsilon_1 \boldsymbol{\phi}_1 + \varepsilon_2 \boldsymbol{\phi}_2 + \varepsilon_3 \boldsymbol{\phi}_3 \tag{3.19}$$

and solved for the ε_j. Once the ε_j are known, Equation (3.16) can be solved for the intrinsic colors of surfaces in the scene. The advantage of Buchsbaums algorithm is that it only requires a single reference surface.

The subspace algorithm of Maloney and Wandell

The main assumption of Maloney and Wandell (1986) is $P > N$ which means that the number of classes of photoreceptors is larger than the dimension of the surface reflectance model. For reasons of simplicity and practical relevance we will assume that $P = 3$ and $N = 2$. Furthermore, the algorithm works without presupposing the existence of intrinsic colors of reference surfaces. The only additional assumption is that the receptor codes of at least two surfaces under a given illuminant E are available.

The lighting matrix $\mathbf{\Lambda}_\varepsilon$, which is in this case a 3×2 matrix, provides a mapping of the two dimensional space of intrinsic colors onto the three dimensional cone excitation space. Therefore, receptor codes of surfaces under a fixed illumination

must lie in a plane in photoreceptor space. Different illuminations are related with different planes of cone excitations. Each plane and therefore also each illuminant can be characterized by its normal $\boldsymbol{\varepsilon}$ through the origin. The first step of the algorithm is the estimation of this plane in cone excitation space. This leads to an estimate of the normal $\hat{\boldsymbol{\varepsilon}}$ up to a multiplicative scaling factor c. As the $\hat{\boldsymbol{\varepsilon}}$ are the coefficients of the linear illumination model, with the normal $\hat{\boldsymbol{\varepsilon}}$ also the lighting matrix is given. In the second step, the estimated lighting matrix is used to determine intrinsic colors of surfaces. Roughly speaking, in analogy to (3.16) an inverse of $\boldsymbol{\Lambda}_\varepsilon$ is computed and used to recover the intrinsic colors. However, as $\boldsymbol{\Lambda}_\varepsilon$ is not quadratic the inverse $\boldsymbol{\Lambda}_\varepsilon^{-1}$ is not defined. In this case a pseudo-inverse can be found which provides the inverse mapping.[7] The idea of the algorithm is shown schematically in Figure 3.4.

The advantage of the subspace algorithm is that no reference surfaces are needed. Nevertheless, it is questionable if—in case of human color vision—two dimensional models of surfaces are sufficient to describe natural viewing conditions adequately. When we imagine an even more reduced system, such as dichromatic vision, the assumption of $P > N$ becomes highly improbable. Another problem of all computational models is that the arrangement of surfaces within the scene is not taken into account. The algorithms lead to identical estimates of the intrinsic colors of a given surface regardless of the local surround. On the other hand, the main purpose of computational models is not to provide an adequate explanation of human color constancy, but rather to develop algorithms that can be implemented in artificial systems.

3.3 Color Constancy in the Shape World

The models we discussed in the previous section referred to a two dimensional representation of environments that were called earlier the *Flat World*. We described corresponding stimuli as flat matte Mondrians (see Figure 3.3). The restrictions made on these stimuli and on the incident illumination allowed us to characterize these stimuli physically in a very simple way (see Equation 3.1). However, a number of typical characteristics of real world stimuli are ignored in this simplistic view. Additional features of the physical stimulus in three dimensional environments that should be taken into account are specular highlights, mutual reflections, shades and shadows.

Important for our considerations is the observation that the percepts related with flat Mondrians and three dimensional stimuli differ qualitatively. Let us for example take a look at the simulated scene in Figure D.1. First, let us examine only the left image. The color checker board on the back wall is an example of a flat Mondrian. The objects in the foreground possess additional physical properties such as shading or specular highlights. These features enable us to

[7] Details of the computations are given in Maloney (1984).

interpret them as three dimensional objects. The percept changes when we fuse the left and the right image of the scene. Now the three dimensional impression is more vivid. However, we can imagine that this impression is still different from the situation where we are sitting in a chamber containing real objects. We note that in each of these situations a given stimulus evokes a qualitatively different percept. The motivation to enrich stimuli with additional features is not only to provide more realistic impressions. Experimental studies have shown that observers in fact use some of these additional features as cues to the illuminant (Yang & Maloney, 2001; Delahunt & Brainard, 2004a; Snyder, Doerschner & Maloney, 2005; Boyaci, Doerschner & Maloney, 2006).

At this point we will introduce and discuss aspects of more complex models of the environment. We will refer to these three dimensional environments as *Shape World*. Compared with the simplistic *Flat World*, surface characteristics and lighting conditions in *Shape World* environments are much more complex. We will also report empirical work that considers the rich structure of such environments.

3.3.1 Bidirectional Reflectance Density Functions

In the previous sections we characterized reflectance properties of surfaces with reflectance functions $S(\lambda)$. In the Shape World the amount of light reaching the eye of the observer from a surface also depends on geometric viewing conditions. These viewing conditions refer to the direction of the incident light and to the direction of the observer. Let us first assume that each point on the retina can be characterized by two coordinates x and y. Each visible point on a surface can be identified with a corresponding location xy on the retina.[8] The reflectance of a surface at location xy can be understood as a function of wavelength λ and of geometric viewing conditions g. For the moment we will write this function as $S^{xy}(\lambda, g)$. If we denote the normal vector of the surface at location xy with $\mathbf{n}^{xy}$, we can define θ^{xy} as the angle between $\mathbf{n}^{xy}$ and the vector of incident light $\mathbf{e}^{xy}$ and ν^{xy} as the angle between $\mathbf{n}^{xy}$ and the direction of the observer $\mathbf{v}^{xy}$ (Figure 3.5).

An extension of surfaces reflectance functions are *bidirectional reflectance density functions* (BRDFs) that we will denote with $S^{xy}(\lambda, \nu^{xy}, \theta^{xy})$. The concept of BRDFs takes geometric viewing conditions into account. Obviously, in the BRDF informations about reflectance properties of a surface and the geometric viewing conditions are confounded.[9] For some surfaces it is possible to separate reflectance characteristics from geometric viewing conditions. This property of surfaces is called *geometry-reflectance separability*. For these surfaces the BRDF can be decomposed in order to get separated spectral and geometric components (Shafer, 1985). The resulting function $S^{xy}(\lambda, \nu^{xy}, \theta^{xy})$ can be understood as a

[8] For reasons of simplicity we will exclude the possibility of transparent surfaces from our considerations.

[9] It should be clear that BRDFs provide physical descriptions of stimuli.

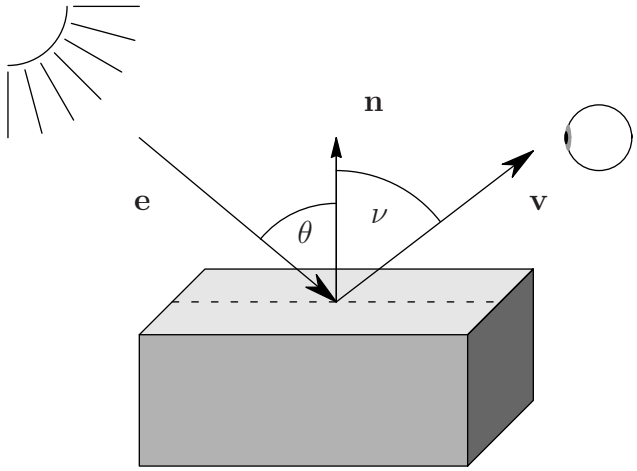

Figure 3.5: Viewing conditions in the Shape World. The amount of light that reaches the eye of the observer from a surface depends on the light source, the reflectance characteristics of the surface, the angle θ between incident light and surface normal and the angle ν between the direction of the observer and the surface normal.

rescaling of the reflectance function of the surface depending on geometric viewing conditions that are represented by the function $G(\nu^{xy}, \theta^{xy})$:

$$S^{xy}(\lambda, \nu^{xy}, \theta^{xy}) = S^{xy}(\lambda) \cdot G(\nu^{xy}, \theta^{xy}). \tag{3.20}$$

The reflectance component $S^{xy}(\lambda)$ can be interpreted analogously to the reflectance function $S(\lambda)$ in the Flat World model: For each λ it simply gives the relative amount of reflected light at location xy. When the surface normal $\mathbf{n}^{xy}$ is identical with the vector of incident light $\mathbf{e}^{xy}$ and the viewing direction of the observer $\mathbf{v}^{xy}$ (so that $\nu^{xy} = \theta^{xy} = 0$) the geometric component is $G(\nu^{xy}, \theta^{xy}) = 1$.

Surfaces often do not satisfy geometry-reflectance separability. However, for a wide range of surfaces that are called *dielectric surfaces* the BRDF can be represented as a weighted sum of two BRDFs that refer to two different components of the surface.[10] The BRDFs of both components satisfy geometry-reflectance separability:

$$S^{xy}(\lambda, \nu^{xy}, \theta^{xy}) = m_d \cdot S^{xy}(\lambda) \cdot G_d(\nu^{xy}, \theta^{xy}) + m_s \cdot Spec(\lambda) \cdot G_s(\nu^{xy}, \theta^{xy}). \tag{3.21}$$

The first component is the *diffuse* component which can be understood roughly as diffuse reflection of incident light from colorant particles embedded in the surface. It depends on the spectral reflectance function of the surface and on the geometric function $G_d(\nu^{xy}, \theta^{xy})$. The second *specular* component refers to the reflection of incident light at the interface of the surface. It can be understood as a mirror-like surface with reflectance function $Spec(\lambda) = 1$ for all λ which is

[10]The term *dielectric surfaces* refers to non-conducting surfaces such as plastics.

rescaled with the geometric function $G_s(\nu^{xy}, \theta^{xy})$. The numbers m_d and m_s are weights that specify the proportion of both components. The property of surfaces given in Equation (3.21) is called *diffuse-specular superposition* (Shafer, 1985). Analyses of natural surfaces have shown that some of them, such as 'plastics', 'green leaves' and 'orange peel', satisfy diffuse-specular superposition whereas others, such as 'blue paper' do not (Lee, Breeneman & Schulte, 1990). The property of diffuse-specular superposition plays an important role in modeling natural surfaces in computer graphics.

3.3.2 The Lambertian Model

In the preceding section we discussed properties of surfaces in the Shape World and approaches that claim to model such properties adequately. In the following a model will be described that introduces further constraints on surfaces. Specifically, we will consider a group of matte surfaces that we will refer to as *Lambertian surfaces*. A typical example of a Lambertian surface is white paper. These surfaces satisfy the assumption of geometry-reflectance separability. We can write the BRDF of a Lambertian surface at location xy as

$$S^{xy}(\lambda, \theta^{xy}) = S^{xy}(\lambda) \cdot \cos\theta^{xy}, \tag{3.22}$$

where θ is the angle between the surface normal and the direction of incident light and $S^{xy}(\lambda)$ denotes the spectral properties of the surface independent of geometric factors. We see that in this case $S^{xy}(\lambda)$ is scaled by the cosine of θ. When the light source is diffuse the BRDF of a Lambertian surface is simply its reflectance function $S(\lambda)$.

Let us now consider a scene with two light sources, a spatially uniform illuminant E_D which we will refer to as the *diffuse source* and a *punctate source* E_P (Figure 3.6). The total power distribution $E(\lambda)$ is then

$$E(\lambda) = t_1 \cdot E_P(\lambda) + t_2 \cdot E_D(\lambda). \tag{3.23}$$

For reasons of simplicity we will assume that $t_1 = t_2 = 1$. Then, the color signal from a Lambertian surface is given by

$$C^{xy}(\lambda) = S^{xy}(\lambda) \cdot (E_P(\lambda) \cdot \cos^{xy}\theta + E_D(\lambda)). \tag{3.24}$$

In the Lambertian model, the color signal depends on the reflectance properties of the surface, the spectral power distributions of the light sources and the angle θ between the surface normal and incident light from the punctate source. In contrast, $C(\lambda)$ is independent of the position of the observer (as long as $-90° \leq \nu \leq 90°$). Surfaces which can be approximately considered as Lambertian surfaces are matte Munsell papers from the Munsell book of colors. In the experiments described in Chapter 5 computer rendered simulations of three dimensional scenes with simulated Munsell papers as test stimuli were presented to the subjects.

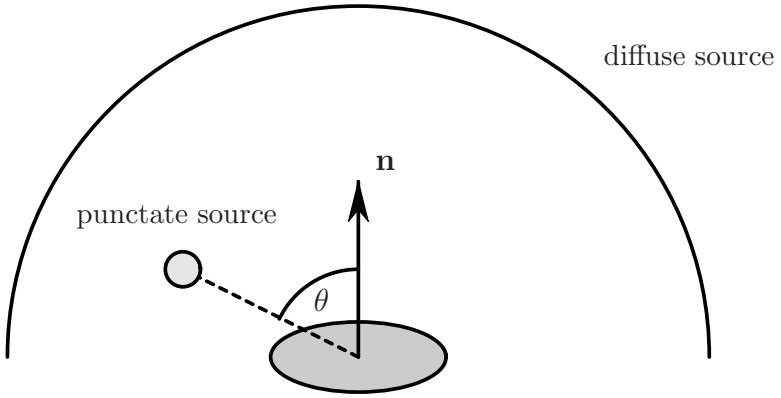

Figure 3.6: The Lambertian model. The amount of light that reaches the eye of the observer from a surface depends on the angle θ between the incident light from the punctate source and the surface normal.

3.3.3 Cues to the Illuminant

In Section 3.2 we discussed prominent models of color constancy that focused on cues and mechanisms available in the restricted two dimensional Flat World. In principle, these models are also applicable to the more complex three dimensional Shape World (Kraft & Brainard, 1999). In addition, further cues to the illuminant are available to the visual system in this environment, such as specular highlights, full surface specularity, cast shadows and surface shading. In order to evaluate the validity of these cues we have to address two questions. First, the question arises if it is possible to arrive at a reliable estimate of the illumination in the scene using a given cue. Second, we have to investigate if such cues are in fact used by the visual system in order to estimate the lighting conditions.

An influential approach that refers to the first question is the specularity algorithm of Lee (1986) and D'Zmura and Lennie (1986). This algorithm focuses on surfaces which satisfy diffuse-specular superposition (see Equation 3.21). When we consider a scene with two different surfaces T_1 and T_2, these surfaces differ in the diffuse component but share the same specular component:

$$\begin{aligned} T_1(\lambda,\nu,\theta) &= m_{d_1} \cdot T_1(\lambda) \cdot G_{d_1}(\nu,\theta) + m_{s_1} \cdot Spec(\lambda) \cdot G_{s_1}(\nu,\theta) \\ T_2(\lambda,\nu,\theta) &= m_{d_2} \cdot T_2(\lambda) \cdot G_{d_2}(\nu,\theta) + m_{s_2} \cdot Spec(\lambda) \cdot G_{s_2}(\nu,\theta). \end{aligned} \quad (3.25)$$

From (3.25) we can derive corresponding color signals C_1, C_2 so that the color signal from a given point on the surface can be understood as a linear combination of two lights, the color signal from the diffuse component of the surface and the illuminant E. The resulting color signals C_1, C_2 of points on surfaces T_1, T_2 are given by:

$$\begin{aligned} C_1(\lambda) &= k_{d_1}^{\nu,\theta} \cdot T_1(\lambda) \cdot E(\lambda) + k_{s_1}^{\nu,\theta} \cdot E(\lambda), \\ C_2(\lambda) &= k_{d_2}^{\nu,\theta} \cdot T_2(\lambda) \cdot E(\lambda) + k_{s_2}^{\nu,\theta} \cdot E(\lambda), \end{aligned} \quad (3.26)$$

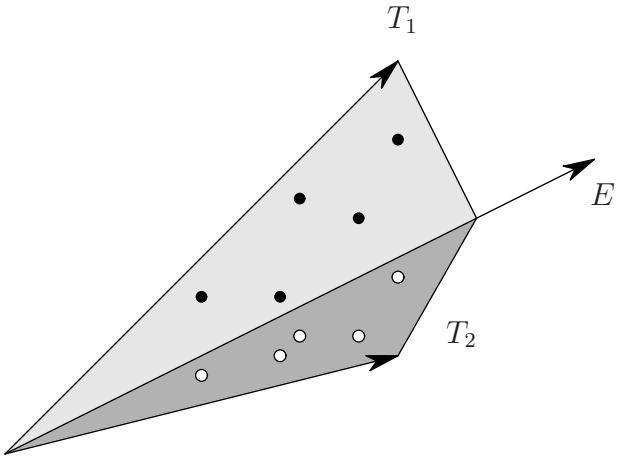

Figure 3.7: The specularity algorithm of Lee (1986) and D'Zmura and Lennie (1986). The receptor codes of color signals from different points on surface T_1 (filled circles) lie in a plane in cone excitation space that is given by the color signal from the diffuse component (T_1) and the illuminant (E). The same is true for points on surface T_2 (open circles). Both planes intersect in a line that is given by the chromaticity of the illuminant E.

where the $k^{\nu,\theta}$ are weights that represent the proportions of both components and the geometric viewing conditions, for example:

$$k_{d_1}^{\nu,\theta} = m_{d_1} \cdot G_{d_1}(\nu, \theta). \tag{3.27}$$

We can then write the photoreceptor excitations of color signals C_1, C_2 as linear combinations of the receptor codes of the respective diffuse and specular components:

$$\begin{aligned} \boldsymbol{\phi}_{C_1} &= k_{d_1}^{\nu,\theta} \cdot \boldsymbol{\phi}_{d_1} + k_{s_1}^{\nu,\theta} \cdot \boldsymbol{\phi}_E, \\ \boldsymbol{\phi}_{C_2} &= k_{d_2}^{\nu,\theta} \cdot \boldsymbol{\phi}_{d_2} + k_{s_2}^{\nu,\theta} \cdot \boldsymbol{\phi}_E, \end{aligned} \tag{3.28}$$

where $\boldsymbol{\phi}_{d_i}$ and $\boldsymbol{\phi}_E$ denote receptor codes of the diffuse components and the illuminant respectively.

From (3.28) it becomes clear that the photoreceptor excitations of color signals from all different points on C_1 must lie in a plane in cone excitation space that is given by the receptor codes of the signal from the diffuse component and the receptor codes of the illuminant. As a consequence, the planes in color space related with T_1 and T_2 intersect in a line that can be identified with the specular component and hence gives the chromaticity of the illuminant (Figure 3.7).

In the specularity algorithm, it is only assumed that surfaces satisfy diffuse-specular superposition but the specular component does not necessarily consist of bright specular highlights. Rather, the algorithm also works for surfaces that are only slightly specular such as human skin. Limitations of the algorithm are that at least two surfaces are needed for both of which Equation (3.21) must

hold and that mutual reflections among surfaces or scenes with more than one illuminant are not considered.

Does the human visual system in fact make use of the specularity cue in order to estimate the illumination as suggested by Lee (1986) and D'Zmura and Lennie (1986)? This question was addressed in an experimental study conducted by Yang and Maloney (2001). They investigated if bright specular highlights and slight specularities according to Equation (3.21) which they referred to as full surface specularity influenced achromatic settings of subjects under different illuminations. The full surface specularity cue represented a direct test of the specularity algorithm. Yang and Maloney found that observers used specular highlights as a cue to the illuminant, but they did not use the full surface specularity cue as suggested by the specularity algorithm.

In natural three dimensional environments usually several cues to the illuminant are available to the visual system. In the *cue combination* framework suggested by Maloney (2002) it is assumed that the information from different cues is used and combined by the visual system in order to estimate the illuminant. This approach is related with similar models from the depth and shape perception literature (Maloney & Landy, 1989; Landy, Maloney, Johnson & Young, 1995).

The estimation of the illuminant within the cue combination approach can be described in two discrete steps. The first step is to get independent estimates of the illuminant in terms of receptor codes based on each of the cues available in the scene. The information from several cues is then combined in order to get an aggregated estimate of the illuminant. A simple model of the cue combination process assumes the final estimate of the illuminant $\hat{\boldsymbol{\phi}}^e$ to be a weighted average of the independent estimates from the cues:

$$\hat{\boldsymbol{\phi}}^e = \alpha_{DV}\hat{\boldsymbol{\phi}}^{DV} + \alpha_{SH}\hat{\boldsymbol{\phi}}^{SH} + \alpha_{GW}\hat{\boldsymbol{\phi}}^{GW}, \tag{3.29}$$

where examples of cues are direct viewing (DV), specular highlights (SH) and mean cone excitations (GW). The αs are corresponding non-negative weights which sum to one. They can be understood as an indicator of the importance of a given cue. A central issue within the cue combination framework is the question how the visual system assigns weights to potential cues. One plausible solution of this problem could be to put available cues in hierarchical order. Such cues to the illuminant differ in the precision of information they provide to the observer. They could give direct access to the illuminant (e.g. if the light source is viewed directly), provide a complete estimate of the illuminant (e.g. the mean cone excitations cue) or give an estimate to the illuminant up to a scaling factor (e.g. the full surface specularity cue). One possible rule of weight assignment could be to use the first cue in the hierarchy exclusively which means to set the corresponding weight to one whereas all other weights are zero. In this sense, most of the models introduced in this chapter could be understood as special cases of the cue combination approach with only one cue used exclusively.

Alternatively, all available cues could contribute to the estimate of the illuminant so that corresponding weights of cues are set to values between zero and one. In this case estimates of weights could be determined based on the reliability of a given cue.

In general, the weights assigned to a given cue are not fixed. The availability of a given cue may vary across different scenes. Hence an additional step termed *dynamic reweighting* allows weights assigned to a fixed cue to vary from scene to scene. The cue combination approach provides rather a framework than an explicit model. Nonetheless, empirical work from related fields, e.g. depth and shape vision, has shown that the cue combination approach might be a useful tool to integrate information from relevant attributes of the environment in near-optimal fashion (Ernst & Banks, 2002).

The cue combination hypothesis with respect to color constancy was tested in a study by Boyaci, Doerschner and Maloney (2006). They investigated whether the cues specular highlights, cast shadows and surface shading are used in isolation when only one cue type is present and compared estimates of the illuminant with those from scenes where all cues were available. Boyaci et al. found that specular highlights and cast shadows were used by several observers when presented in isolated form whereas the shading cue was used only by one subject. When presented with scenes containing all three cues, subjects combined the different cues and got estimates of the illuminant that were more reliable than any of the cues in isolation. This finding is not consistent with the assumption that one of the available cues is used exclusively. Additionally, Boyaci et al. reported large differences in the use of specific cue types across observers. This result can be explained within the cue combination framework in the sense that different observers assign different weights to available cues. One problem related with the cue combination approach is that as the conditions of the cue combination process are not exactly specified it is difficult to derive empirical testable hypotheses.

Chapter 4

Experiments I: Color Constancy of Normal and Dichromatic Observers

In the present chapter I will report results from a series of experiments that were conducted under the constraints of the simple Flat World. This study focuses on the color constancy performance of two classes of observers—normal trichromats and color deficient dichromats—under different illuminant conditions. First, an overview of the questions addressed in this investigation will be given.

4.1 Motivation

4.1.1 Overview

In the last three decades the color constancy performance of normal observers has been studied extensively under a wide range of viewing conditions using a number of different experimental paradigms. This considerable interest in the phenomenon of color constancy has led to the identification of several factors that are assumed to maintain color constancy. Many of these ideas are reflected by the models that we discussed earlier in Chapter 3.

In contrast, only little is known about the color constancy performance of color deficient observers. Rüttiger, Mayser, Sérey and Sharpe (2001) investigated the color constancy of dichromatic observers for simulated illuminant shifts along a blue-yellow and a red-green axis. They found that color constancy of dichromats was comparable to that of normal trichromats under all illuminant conditions. However, dichromats showed higher variability in their settings under red and green illumination compared to normal observers. Using an operational approach, Foster and Linnell (1995) found that dichromats can discriminate between illuminant changes and surface changes as reliably as normal observers when these

changes occur along the daylight locus. Some dichromats were able to perform this task even when illuminants changed along a red-green axis in color space.

According to the reduction hypothesis of dichromatic color vision, dichromacy is understood as a submodel of trichromatic vision which is characterized by the lack of one type of cone photoreceptors (see Section 2.2). As a consequence, dichromatic color spaces can be described algebraically as subspaces of trichromatic color space. However, results from several studies on different aspects of dichromatic vision are not in agreement with this view (Montag, 1994; Wachtler, Dohrmann & Hertel 2004). Similarly, the results from studies on color constancy of dichromatic observers suggest that their color constancy is better than predicted by the classical reduction hypothesis.

The main objective of the work reported in the present chapter is to study the color constancy performance of a class of dichromatic observers, protanopes, under several illuminant conditions and to compare it with predictions derived from the reduction hypothesis. We already noted that—according to the reduction hypothesis—protanopes are lacking L-cone photoreceptors. Of particular interest are those conditions under which color constancy of protanopes is expected to break down given the reduction hypothesis to be true. In addition, elaborated models of color constancy that have been proven useful for normal trichromats will be evaluated for the protanopic case.

A second goal of the research presented in this chapter is to analyze color constancy performance of observers for two distinct classes of stimuli, namely increments and decrements. Roughly speaking, incremental lights refer to stimuli that induce stronger cone excitations than some reference (e.g. the mean color signal across the scene) whereas decrements correspond with lower cone excitations. Bäuml (1997; 2001) showed that increments and decrements in complex scenes are processed differently and that color constancy of observers is better for decremental than for incremental stimuli. But, as reanalysis of his data suggests, the size of this effect seems to depend on the chromaticity of the illumination. In the experiments reported here, the relationship between the correlated color temperature of daylights and the strength of increment-decrement asymmetries was analyzed systematically.[1]

Three main experiments were conducted in order to investigate these questions. The visual display used in the experiments always consisted of simulations of flat matte Mondrian stimuli that were rendered under different illuminants and presented on CRT-screens. Such simulations have been used frequently to study color constancy (Arend & Reeves, 1986; Brainard & Wandell, 1992; Bäuml, 2001).[2] The simple lighting model underlying the rendering of stimuli is given in Equation (3.1). An important presupposition of this model is that the distribu-

[1] An introduction to the concept of correlated color temperature is given in Appendix A.

[2] Several issues on the use of this kind of stimulus configuration will be discussed in the general methods section.

tion of the illumination is spatially uniform. Illuminant adjustment of observers was determined using achromatic settings (see Section 3.1.3). In each of the experiments, the adaptational state of the subjects was controlled. In this sense, the present work refers to successive color constancy. In the following, the questions of the present study will be discussed in more detail and hypotheses will be derived accordingly.

4.1.2 Questions Addressed in this Study

Color constancy of protanopes

The first objective of the present study is to obtain the color constancy performance of normal and protanopic observers under different illuminant conditions and to evaluate it in the light of the reduction hypothesis of dichromacy. At this point, we shall discuss theoretical aspects of protanopic color constancy from the viewpoint of the reduction hypothesis. Further, we will derive expectations for protanopic constancy judgments under different illuminant conditions. Let us denote the light signal from surface s under illuminant e with (s, e). First, we will define a protanopic psychophysical relation '$\approx_P$' so that

$$(s, e) \approx_P (s', e') \tag{4.1}$$

means that surface s under illuminant e is indistinguishable for the protanope from surface s' under illuminant e'. In Section 3.2.2 we introduced relational models of color constancy. In this context, we discussed a model where it is assumed that weighting the color signal from surface s under illuminant e with the spatial mean signal in terms of receptor codes ϕ yield an illuminant invariant code $\kappa(s, e)$. We can denote a protanopic version of this model with:

$$(s, e) \approx_P (s', e') \Leftrightarrow \kappa_i(s, e) = \kappa_i(s', e'), \qquad i = 2, 3. \tag{4.2}$$

In the model it is assumed that codes $\kappa_i(s, e)$ are given with

$$\kappa_i(s, e) = \frac{1}{\frac{1}{N}\sum_j^N \phi_i(e \cdot s_j)} \cdot \phi_i(e \cdot s), \qquad i = 2, 3, \tag{4.3}$$

where s_j denotes the j-th surface and N the number of surfaces in the scene. If we assume that the virtual mean surface appears gray to the observer, we can express constancy codes $\kappa(s, e)$ as

$$\kappa_i(s, e) = \frac{1}{\phi_i(e \cdot (t * w))} \cdot \phi_i(e \cdot s) = \frac{1}{t \cdot \phi_i(e)} \cdot \phi_i(e \cdot s), \qquad i = 2, 3. \tag{4.4}$$

Now we shall investigate exemplarily predictions from this model for the protanopic case. In particular, we will discuss two special cases: one for which

illuminant shifts affect only L-cone excitations and not M- and S-cone signals, that is illuminant changes which occur along protanopic confusion lines. In the second case, we will investigate illuminant shifts that mainly occur along the S-cone axis but do not affect L-cone signals. An example of this situation are changes of natural daylights.

First, recall that we understand color constancy as the ability of a visual system to assign stable color appearance to a fixed surface under changing illuminant conditions. Imagine now a surface s that is observed by a protanopic observer under illuminants e and e'. Let us assume that the receptor codes of the corresponding color signals differ only in the L-cone (which is missing in the protanope) but not in the M- and S-cone:[3]

$$\begin{aligned} \phi^P(e \cdot s) &= \phi^P(e' \cdot s) \quad \text{and} \\ \phi_1(e \cdot s) &\neq \phi_1(e' \cdot s). \end{aligned} \tag{4.5}$$

In addition, let us assume that $\phi^P(e) = \phi^P(e')$ so that with the linking proposition noted in Equation (2.20) illumination e is indistinguishable for the protanope from illumination e'. Further, we will denote a protanopic constancy code with $\kappa^P(s, e) = (\kappa_2(s, e), \kappa_3(s, e))$. As in the given example

$$\frac{\phi_i(e \cdot s)}{\phi_i(e)} = \frac{\phi_i(e' \cdot s)}{\phi_i(e')}, \qquad i = 2, 3, \tag{4.6}$$

model (4.2) predicts that in this case $\kappa^P(s, e) = \kappa^P(s, e')$. Thus, such a protanopic observer may come to judgments that we would classify as color constant without having any color constancy at all. In this case, a protanope can simply not distinguish stimulus (s, e) from (s, e') due to the lack of L-cone signals. Therefore we can treat a situation where illuminant invariant judgments of protanopes are based on identical photoreceptor signals so that

$$\begin{aligned} \phi^P(e \cdot s) &= \phi^P(e' \cdot s) \quad \text{and} \\ \phi^P(e) &= \phi^P(e') \end{aligned} \tag{4.7}$$

as a trivial form of color constancy which we shall separate from the phenomenon of color constancy. Rather, we are in favor of saying that in this situation the color constancy of a protanope breaks down.

We shall now investigate whether pairs of receptor codes of any given surface s under illuminants e and e' are located on one and the same protanopic confusion line if $\phi^P(e) = \phi^P(e')$. In particular, we are interested if

$$\phi^P(e) = \phi^P(e') \Rightarrow \phi^P(e \cdot s) = \phi^P(e' \cdot s) \tag{4.8}$$

[3]In Section 2.2 we denoted protanopic photoreceptor excitations of any light a with $\phi^P(a) = (\phi_2(a), \phi_3(a))$.

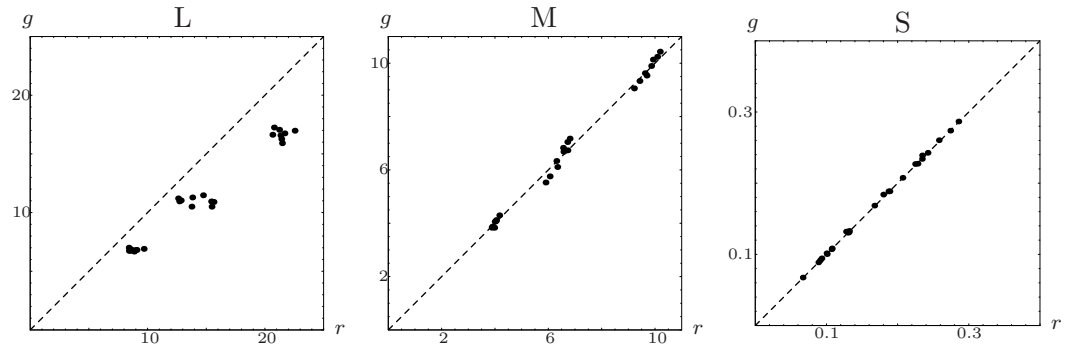

Figure 4.1: Receptor codes of 25 Munsell surfaces under reddish illuminant r and greenish illuminant g with $\phi^P(r) = \phi^P(g)$. The data are shown separately for L-, M- and S-cone. Data points on the dashed line indicate surfaces for which corresponding receptor codes under illuminant r and g are identical.

holds for all surfaces s. This would be the case if it were possible to decompose the receptor code of the color signal $(e \cdot s)$ in the following way:

$$\phi_i(e \cdot s) = \phi_i(e) \cdot c_i^s, \qquad i = 2, 3, \tag{4.9}$$

where c_i^s is a constant which depends on the surface s but not on the illuminant. A special case of Equation (4.9) would be $c_i^s = \phi_i(s)$, where $\phi_i(s)$ are photoreceptor excitations of surface s under an illumination with constant, unit spectral power density. The trichromatic version of this assumption is often referred to as *RGB heuristic* (see Yang & Maloney, 2001):[4]

$$\phi_i(e \cdot s) = \phi_i(e) \cdot \phi_i(s), \qquad i = 1, 2, 3. \tag{4.10}$$

It can be shown mathematically that this assumption does *not* hold in general.

However, if we use only illuminants with smooth spectral power distributions, Equation (4.9) may provide a useful heuristic. Let us for example investigate a fixed set of 25 Munsell surfaces under reddish light r and under greenish light g with $\phi^P(r) = \phi^P(g)$. Figure 4.1 shows L-, M- and S-cone excitations of color signals from these surfaces under illuminants r and g. For this set of surfaces—which are in fact stimuli from Experiment 2 of the present study—Equation (4.8) approximately holds. As a consequence, we assume that protanopic observers are in this case not able to provide reliable color constancy judgments.

Let us now consider illuminations that are distinguishable for the protanopic observer. In particular, we want to investigate the non-trivial case where

$$\begin{aligned} \phi^P(e) &\neq \phi^P(e') \quad \text{and} \\ \phi^P(e \cdot s) &\neq \phi^P(e' \cdot s). \end{aligned} \tag{4.11}$$

[4]If we assume that constancy codes $\kappa_i(s, e)$ in model (4.2) are in fact independent of the illuminant e, we could replace c_i^s in Equation (4.9) with $t \cdot \kappa_i(s, e)$. In this case, Equation (4.8) would be a necessary condition for model (4.2).

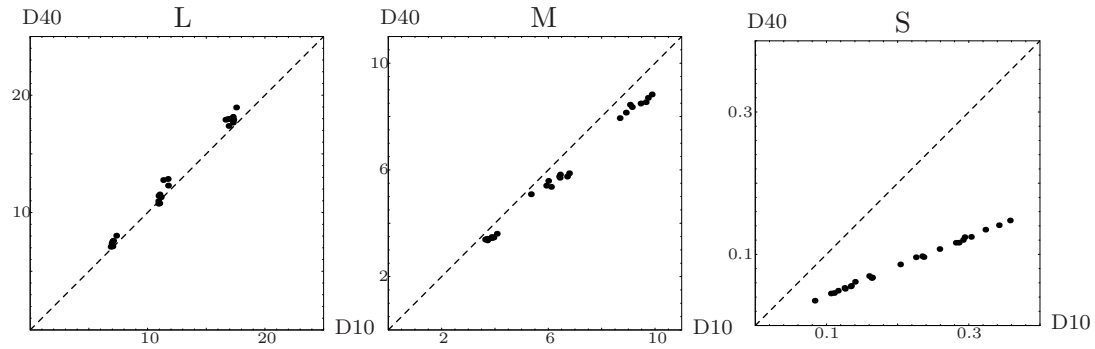

Figure 4.2: Receptor codes of 25 Munsell surfaces under daylights D10 and D40. The data are shown separately for L-, M- and S-cone. Data points on the dashed line indicate surfaces for which corresponding receptor codes under both illuminants are identical.

If we assume that model (4.2) holds for the protanopic observer, such an observer should be able to make more or less reliable constancy judgments. However, if we compared constancy judgments of a protanopic observer with those of a normal trichromat we would expect that the judgments of the latter are in general more accurate as additional information from L-cone photoreceptors is provided.

Let us now consider a special case of Equation (4.11) with two illuminants e, e' and surfaces s for which

$$\begin{aligned} \phi_1(e) &= \phi_1(e') \quad \text{and} \\ \phi_1(e \cdot s) &= \phi_1(e' \cdot s) \end{aligned} \tag{4.12}$$

hold. In this situation trichromatic observers should not have an advantage over protanopic observers. Thus, we would expect that under these conditions protanopes should be able to make reliable constancy judgments comparable to those of normal trichromats. When we consider typical changes of daylights with identical intensities, we find that Equations (4.11) and (4.12) approximately hold for these daylights and a large number of surfaces. Figure 4.2 shows L-, M- and S-cone excitations of color signals from a fixed set of 25 Munsell surfaces under daylights D10 and D40. Simulations of these surfaces are stimuli in Experiments 1 and 3 of the present study. We see that the color signals from a fixed surface s under the two daylights differ mainly in S-cone excitation whereas L-cone signals are roughly identical under both illuminants. A similar pattern of cone excitations occurs for other daylight changes. Therefore, we would expect that under daylight conditions constancy judgments of protanopes should be comparable quantitatively and in precision to those of normal trichromats.

The color constancy of protanopic observers will be investigated in Experiments 1 and 2. The first experiment represents a replication of the study by Rüttiger et al. (2001) but goes beyond the scope of this study in several respect. In this experiment two different classes of simulated illuminations will be used. One class of lighting conditions are simulations of natural daylights. A second

class of illuminations comprises simulations of red and green illuminants off the daylight locus that are slightly displaced from protanopic confusion lines. We expect that location and precision of protanopes' achromatic settings under daylight conditions will be comparable to those of normal trichromats. Protanopes are presumably not able to distinguish between the stimulus configuration under simulated reddish and greenish light. Therefore, we expect a wide range of protanopic achromatic settings along a corresponding confusion line.

In Experiment 2, the observers are presented with red and green illuminants which are aligned with protanopic confusion lines. The settings of protanopes under these conditions will be compared with predictions from the reduction model. Given this hypothesis to be true, we expect that protanopic observers accept any setting on the confusion line through the reddish and the greenish illuminant as achromatic setting.

The role of daylights

The class of illuminations that we are exposed to most frequently in our natural environment are daylights. In Chapter 3 it was mentioned that daylights show some regularities which might be used by the visual system to achieve color constancy (Judd, MacAdam & Wyszecki, 1964; Shepard, 1994). In this sense, color constancy performance of observers should be better under daylight conditions rather than under illuminations off the daylight locus.

Evidence supporting this assumption comes from a study by Worthey (1985). He reanalyzed the data from McCann, McKee and Taylor (1976) and showed that adjustment to the illuminant is better for illumination shifts along the opponent blue-yellow axis, which is roughly aligned with the daylight locus, than for illuminant shifts along the opponent red-green axis. In contrast, results from more recent studies suggest that adjustment to illuminants off the daylights locus can be comparable to performance under daylight conditions (Brainard, 1998; Delahunt & Brainard, 2004b). In Experiment 1, the hypothesis that color constancy performance of normal observers is better under daylight conditions than under non-daylight illumination will also be tested.

The increment-decrement hypothesis

Increment-decrement asymmetries—that is different processing of incremental and decremental lights—have been investigated mainly for the phenomenon of simultaneous contrast. In this case, mostly center-surround configurations are used as stimuli so that increments and decrements are simply defined in relation to the homogeneous surround. Several recent studies have shown that—in contrast to results from previous studies (Burnham, Evans & Newhall, 1957; Walraven, 1976; Chichilnisky & Wandell, 1995)—simple von-Kries receptor scaling is not sufficient to predict the color appearance of the center (Chichilnisky & Wandell,

1996; Heyer, 1997; Richter, 2002). Rather, results from these studies suggest that the degree of rescaling is different for incremental and decremental stimuli. A model which considers this increment-decrement asymmetry is the Octant-model by Mausfeld and Niederee (1993). In the model it is assumed that six scaling factors, one related with increments and one with decrements within each cone class, are needed to describe the color appearance of the test light appropriately. Results from a study by Schiller, Sandell and Maunsell (1986) suggest that a neurophysiological substrate of this asymmetry might be the segregation of retinal signals into ON and OFF pathways.

A plausible explanation for the observed increment-decrement asymmetry is the notion that incremental stimuli could be interpreted by the visual system as light sources whereas decrements may be perceived as surfaces (Mausfeld, 1998). Taken into account the assumption that the phenomena of simultaneous contrast and color constancy are maintained by similar underlying principles (Valberg & Lange-Malecki, 1990; Heyer, 1997; Kraft & Brainard, 1999), it would not be surprising to find differences in the processing of incremental and decremental lights also in complex environments.

Recently, Bäuml (2001) reported increment-decrement asymmetries also for more complex Mondrian stimuli. Moreover, he found that color constancy performance was better for decremental than for incremental stimuli. This result is in agreement with the assumption that decrements are encoded rather as surfaces than as lights. On the other hand, less adjustment to the illuminant in case of incremental stimuli may provide the visual system with information about the color of the present illumination (Walraven, Enroth-Cugell, Hood, MacLeod & Schnapf, 1990).

In the first experiment I investigated under which simulated illumination conditions increment-decrement asymmetries occur using complex Mondrian stimuli. The relationship between the correlated color temperature of daylights and the strength of increment-decrement asymmetries is studied systematically in Experiment 3. In simple center-surround configurations increments and decrements are defined relative to the background. In more complex scenes the question arises which reference the visual system might use to separate increments and decrements. As a starting point and in agreement with Bäuml (2001) increments and decrements were defined relative to the mean cone signal of the entire stimulus. This assumption is also to be tested in the present study.

Models of color constancy

In the previous chapter we briefly introduced models of color constancy. Now it will be examined if subjects' achromatic settings are well predicted by three selected models. The three models are: a version of the Retinex model, an unspecified model of von-Kries type and a model that takes increment-decrement asymmetries into account. In the following, formal descriptions of these three

models will be given.

First, we shall introduce a global notation. Let us denote the light signal from surface s under illuminant e with (s, e). We can write the corresponding local receptor code as column vector $\boldsymbol{\phi}_{s,e}$. Let us then define a relation of metamerism '$\approx$' where

$$(s, e) \approx (s', e') \tag{4.13}$$

denotes that the surface s under illuminant e is perceptually indistinguishable from the surface s' under illuminant e'. In each of the constancy models introduced in the following it is assumed that $(s, e) \approx (s', e')$ if and only if some constancy codes $\boldsymbol{\kappa}$ (denoted as vectors) are identical:[5]

$$(s, e) \approx (s', e') \Leftrightarrow \boldsymbol{\kappa}_{s,e} = \boldsymbol{\kappa}_{s',e'}. \tag{4.14}$$

The models differ in the specification of these constancy codes. However, in each of these models the constancy codes $\boldsymbol{\kappa}$ are given by a linear transformation of receptor codes $\boldsymbol{\phi}$ which can be written as:

$$\boldsymbol{\kappa}_{s,e} = \mathbf{M}_e \boldsymbol{\phi}_{s,e}, \tag{4.15}$$

where $\mathbf{M}_e$ is a 3×3 matrix specific to illuminant e. This matrix expresses the rescaling of receptor codes under fixed illuminant conditions e. Results from several studies on color constancy suggest that—according to von Kries' (1905) original hypothesis—this rescaling is performed independently within each class of cones (Brainard & Wandell, 1992; Bäuml, 1995; 1999a). This result has at least two implications for models of color constancy. First, transformation matrices $\mathbf{M}_e$ can be expressed as diagonal matrices $\mathbf{D}_e$. Second, the transformation of receptor codes into constancy codes is presumably located before the opponent stage in the visual path.

In Section 2.3 we denoted the opponent code of a given light a with $\psi(a)$. Let us now summarize the opponent signal of surface s under illuminant e in the vector $\boldsymbol{\psi}_{s,e}$. Due to the empirical finding of independent photoreceptor rescaling it seems plausible to locate the coding of illuminant invariant signals prior to opponent processing. We can derive opponent signals $\boldsymbol{\psi}$ as a linear transformation of constancy codes $\boldsymbol{\kappa}$. This transformation is given by a 3×3 matrix $\mathbf{T}$ (see Equation 2.29):

$$\boldsymbol{\psi}_{s,e} = \mathbf{T}(\mathbf{M}_e \boldsymbol{\phi}_{s,e}). \tag{4.16}$$

We shall denote a surface a under illuminant e which appears achromatic to the normal trichromatic observer with $A(a, e)$. In this case the chromatic channels BY and RG are in equilibrium:

$$A(a, e) \Leftrightarrow \psi_1(a, e) = \psi_2(a, e) = 0. \tag{4.17}$$

[5]Equation (4.14) can be understood as linking proposition. The psychophysical relation '$\approx$' is linked with assumed neurophysiological constancy codes (see also Equation 2.17).

We can then reformulate the general model from Equation (4.14) so that achromatic settings of observers at any luminance level k with $k = \psi_3(a, e)$ are related with opponent codes predicted by the model:

$$A(a, e) \Leftrightarrow \mathbf{T}(\mathbf{M}_e \boldsymbol{\phi}_{a,e}) = (0, 0, k)'. \tag{4.18}$$

Let us now consider the three models we wish to test. The most general model takes increment decrement asymmetries into account. We will refer to this model as the *ID model*. The second model is an unspecific model of von-Kries type which can be understood as a special case of the ID model. We will refer to this model as the *diagonal model*. The third model is a special case of the diagonal model. In this model the rescaling coefficients are defined as the reciprocals of corresponding mean cone excitations across the whole scene. It resembles a version of the Retinex-model of Land and coworkers (Land, 1977). We will refer to this model as the *spatial mean model*. The three models are nested models, that is the ID model contains the other two models as special cases and the diagonal model contains the spatial mean model as special case. We will use this property when testing the models against each other.

The ID model can be understood as an extension of the octant model by Mausfeld and Niederée (1993) to more complex stimulus configurations. Incremental and decremental stimuli are defined in relation to a reference r.[6] In the following we will only consider the case of pure increments and decrements so that in this model two diagonal matrices, $\mathbf{D}_e^+$ for increments and $\mathbf{D}_e^-$ for decrements, describe the rescaling of receptor codes under illuminant e. If we denote the difference $\phi_i(a, e) - \phi_i(r)$ with $\Delta_i(e, r)$ and corresponding vectors of differences with $\boldsymbol{\Delta}_{e,r}$, we can write the affine ID model as follows:

$$\begin{aligned} A(a, e) &\Leftrightarrow \mathbf{T}\mathbf{D}_e^+ \boldsymbol{\Delta}_{e,r} = (0, 0, k^+)', \quad \text{if } \Delta_i(e, r) \geq 0; \\ A(a, e) &\Leftrightarrow \mathbf{T}\mathbf{D}_e^- \boldsymbol{\Delta}_{e,r} = (0, 0, k^-)', \quad \text{if } \Delta_i(e, r) < 0; \quad \text{for } i = 1, 2, 3. \end{aligned} \tag{4.19}$$

As noted above, the expression $A(a, e)$ refers to the observation that surface a appears achromatic under illuminant e. The matrix $\mathbf{T}$ translates constancy codes into opponent codes. The numbers k^+ and k^- represent luminance levels of incremental and decremental stimuli respectively where the sign indicates that luminances are larger ('+') or smaller ('–') than the luminance of reference r. The model has 12 parameters, the six rescaling coefficients that are simply the diagonal elements of matrices $\mathbf{D}_e^+$ and $\mathbf{D}_e^-$ and six additive constants of corresponding linear regressions.

The diagonal model can be derived from Equation (4.18) by simply replacing matrix $\mathbf{M}_e$ with a corresponding diagonal matrix $\mathbf{D}_e$:

$$A(a, e) \Leftrightarrow \mathbf{T}\mathbf{D}_e \boldsymbol{\phi}_{a,e} = (0, 0, k)'. \tag{4.20}$$

[6] As a first approximation we assume that the reference r is identical with mean cone excitations across the whole scene (Bäuml, 2001). This heuristic was considered when constructing the stimuli. Later the reference r will be estimated from the data.

The diagonal model has three parameters which are simply the rescaling coefficients $d_i(e)$ in the diagonal of matrix $\mathbf{D}_e$.

A special case of model (4.20) is the spatial mean model where the diagonal elements $d_i(e)$ of $\mathbf{D}_e$ are specified as the reciprocals of corresponding mean cone excitations across the whole scene. We will denote this matrix with $\mathbf{D}_e^*$ and its diagonal elements with $d_i^*(e)$. For the i-th class of cone photoreceptors the diagonal element $d_i^*(e)$ can be written as:

$$d_i^*(e) = \frac{1}{\frac{1}{N}\sum_j^N \phi_i(s_j, e)}, \qquad i = 1, 2, 3, \tag{4.21}$$

where s_j denotes the j-th surface and N the number of surfaces in the scene. We can then write the model analogous to model (4.20) as:

$$A(a, e) \Leftrightarrow \mathbf{T}\mathbf{D}_e^*\boldsymbol{\phi}_{a,e} = (0, 0, k)'. \tag{4.22}$$

For a given illuminant, this model has no free parameters. The nested design of the three models allows us to test the models against each other.

We shall now derive empirically testable conditions of models (4.19), (4.20) and (4.22). First, let us consider a second surface a' under a different illumination e' which appears achromatic to the normal observer. For convenience we will assume that $\psi_3(a, e) = \psi_3(a', e')$ which means that both stimuli provide identical luminance signals. Therefore we can note that in this case

$$\boldsymbol{\psi}_{a,e} = \boldsymbol{\psi}_{a',e'} \tag{4.23}$$

holds for these stimuli which is equivalent to

$$\mathbf{T}(\mathbf{M}_e\boldsymbol{\phi}_{a,e}) = \mathbf{T}(\mathbf{M}_{e'}\boldsymbol{\phi}_{a',e'}). \tag{4.24}$$

For the moment, we choose unspecified 3×3 transformation matrices $\mathbf{M}_e$, $\mathbf{M}_{e'}$ which will be replaced later by matrices specific to the respective model. The matrix $\mathbf{T}$ which translates constancy codes into opponent codes is independent of the illuminant and has full rank. Therefore we can multiply each side with the inverse $\mathbf{T}^{-1}$ and see that in this case constancy codes as defined in Equation (4.15) are identical:

$$\mathbf{M}_e\boldsymbol{\phi}_{a,e} = \mathbf{M}_{e'}\boldsymbol{\phi}_{a',e'}. \tag{4.25}$$

Taken together this result and the linking proposition from Equation (4.14) leads to

$$(a, e) \approx (a', e') \Leftrightarrow \mathbf{M}_e\boldsymbol{\phi}_{a,e} = \mathbf{M}_{e'}\boldsymbol{\phi}_{a',e'}, \tag{4.26}$$

which is equivalent to

$$(a, e) \approx (a', e') \Leftrightarrow \boldsymbol{\psi}_{a,e} = \boldsymbol{\psi}_{a',e'}. \tag{4.27}$$

From Equation (4.25) we can derive a necessary condition that must hold for all models we consider. If we assume that $\psi_3(a, e) = \psi_3(a', e')$, then photoreceptor excitations of achromatic settings a under illuminant e can be expressed as a linear transformation of receptor codes of achromatic settings a' under illuminant e':

$$(a, e) \approx (a', e') \Leftrightarrow \boldsymbol{\phi}_{a,e} = \mathbf{M}_e^{-1}\mathbf{M}_{e'}\boldsymbol{\phi}_{a',e'}. \tag{4.28}$$

In the experiments, achromatic settings of subjects are determined under different illuminants. It should be noted that when dealing with achromatic settings we are testing only the special case of $\psi_1 = \psi_2 = 0$. However, results from experimental studies indicate that it is possible to predict asymmetric matches from achromatic settings (Speigle & Brainard, 1999) and observers' hue scalings from achromatic settings (see Chaper 5 of the present work).

In the model test, D65 will be used as reference illuminant. In particular, we want to predict the subjects' achromatic settings under test illuminant e from ideal achromatic settings a' under D65. If we consider for example the diagonal model (4.20), achromatic settings a under test illuminant e can be predicted by

$$\boldsymbol{\phi}_{a,e} = \mathbf{D}_e^{-1}\mathbf{D}_{D65}\,\boldsymbol{\phi}_{a',D65}. \tag{4.29}$$

The matrix $\mathbf{D}_e^{-1}\mathbf{D}_{D65}$ is a diagonal matrix which contains rescaling coefficients $d_i(D65)/d_i(e)$ of each photoreceptor class i as diagonal elements. For each subject and illuminant condition, the set of corresponding coefficients will be estimated for each of the three models.[7]

First, the estimation of these rescaling parameters will be described. Under the assumption of independent rescaling within each class of photoreceptors, estimations can be understood as weights from univariate regressions of the general form

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \boldsymbol{\varepsilon}. \tag{4.30}$$

Using this notation we can write corresponding predictions by the ID model for the i-th receptor class as

$$\begin{pmatrix} \phi_i^-(a_1, e) \\ \vdots \\ \phi_i^-(a_{\frac{N}{2}}, e) \\ \phi_i^+(a_{\frac{N}{2}+1}, e) \\ \vdots \\ \phi_i^+(a_N, e) \end{pmatrix} = \begin{pmatrix} \phi_i^-(a'_1, D65) & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \phi_i^-(a'_{\frac{N}{2}}, D65) & 1 & 0 & 0 \\ 0 & 0 & \phi_i^+(a'_{\frac{N}{2}+1}, D65) & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \phi_i^+(a'_N, D65) & 1 \end{pmatrix} \begin{pmatrix} b_i^- \\ c_i^- \\ b_i^+ \\ c_i^+ \end{pmatrix} + \boldsymbol{\varepsilon}, \tag{4.31}$$

[7] It should be clear that when dealing with the ID model a pair of coefficients, one for incremental and one for decremental stimuli, needs to be estimated for each receptor channel. Furthermore, corresponding coefficients of the spatial mean model are identical across subjects.

where $\phi_i(a_j, e)$ denotes the i-th receptor code of the j-th achromatic setting under test illuminant e. N is the number of settings for each illuminant condition and the vector $\boldsymbol{\varepsilon}$ is the vector of prediction errors. Cone excitations of corresponding settings under D65 are written as $\phi_i(a'_j, D65)$. The symbols '$-$' and '$+$' indicate decremental and incremental stimuli respectively. The parameters of the model are the rescaling coefficients b_i^-, b_i^+ and the additive constants c_i^-, c_i^+. More precisely, the rescaling coefficients are ratios of corresponding illuminant dependent scalars so that for example $b_i^+ = d_i^+(D65)/d_i^+(e)$. The additive constant c_i^+ can be written as

$$c_i^+ = \phi_i(r, e) - b_i^+ \phi_i(r', D65), \tag{4.32}$$

where r and r' denote reference stimuli that separate increments and decrements under illuminants e and D65 respectively. Using the parameters c_i^+ and c_i^- we can estimate the reference stimulus r for each test illuminant e. Geometrically, the i-th receptor code of reference stimulus r is simply the intersection point of the incremental and the decremental regression lines for that channel.

Using the notation from Equation (4.30) predictions by the diagonal model for the i-th photoreceptor channel can be written as

$$\begin{pmatrix} \phi_i(a_1, e) \\ \vdots \\ \phi_i(a_N, e) \end{pmatrix} = \begin{pmatrix} \phi_i(a'_1, D65) \\ \vdots \\ \phi_i(a'_N, D65) \end{pmatrix} \cdot b_i + \boldsymbol{\varepsilon}. \tag{4.33}$$

In this case we only have to estimate one parameter per channel, the rescaling coefficient b_i. Predictions of subjects' achromatic settings by the spatial mean model can be written analogously with specified coefficients b_i^* so that

$$b_i^* = \frac{\sum_j^K \phi_i(s_j, e)}{\sum_j^K \phi_i(s_j, D65)}, \tag{4.34}$$

where the s_j denotes the j-th surface in the scene and K is the number of surfaces.

We already noted that the three models can be understood as nested models with the ID-model as the most general and the spatial mean model as the most restricted model. Therefore, we can test the models pairwise against each other. If we test the ID-model against the diagonal model, the null hypothesis for the i-th channel is

$$H_0: \quad \mathbf{b}_i^{ID} = (b_i^D, 0, b_i^D, 0)', \tag{4.35}$$

where $\mathbf{b}_i^{ID}$ denotes the vector of parameters of the ID-model and b_i^D the corresponding rescaling coefficient of the diagonal model. The null hypothesis for a test of the diagonal model against the spatial mean model in the i-th channel is then

$$H_0: \quad b_i^D = \frac{\sum_j^K \phi_i(s_j, e)}{\sum_j^K \phi_i(s_j, D65)}. \tag{4.36}$$

The model tests follow a simple rationale. For a given photoreceptor channel the sum of squared prediction errors of two models, a general and a restricted model, will be compared. If we denote the sum of squares of the general model with SS_G and that of the restricted model with SS_R, the following F-test can be applied to test the hypotheses (4.35) and (4.36) (see for example Maxwell and Delaney, 2003):

$$F = \frac{(SS_R - SS_G)/(q-p)}{SS_G/(N-q)}, \tag{4.37}$$

where q is the number of parameters of the general model, p is the number of parameters of the restricted model and N is the number of observations. This statistic is $F_{q-p,N-q}$-distributed under the null hypothesis.

4.1.3 Hypotheses

From the discussion above we can extract the following hypotheses of the present study:

1. Given the reduction hypothesis of dichromacy to be true, the location and precision of protanopes' achromatic settings should be comparable to those of normal observers under natural daylight conditions. Under illumination shifts aligned with a protanopic confusion line it should be impossible for these dichromats to make precise achromatic settings. It is expected that they accept all settings along the confusion line of the illuminant shift.

2. Illuminant adjustment of normal observers under natural daylights should not be better than under red and green illuminants.

3. It is assumed that the segregation of incremental and decremental stimuli is a substantial principle of the visual system. Therefore, increment-decrement asymmetries should also occur when complex Mondrian stimuli are used.

4. If increment-decrement asymmetries occur, adjustment to the illuminant should be better for decremental than for incremental stimuli.

5. The strength of increment-decrement asymmetries varies for illuminations along the daylight locus but not for illuminants along a red-green axis.

6. The strength of the increment-decrement asymmetry is correlated with the color temperature of natural daylights.

7. It is assumed that the ID model describes achromatic settings of subjects substantially better than the diagonal and the spatial mean model.

4.2 General Methods

In the following section general methods that were used in the experiments described in the present chapter will be introduced. First, the procedure of classifying dichromatic observers will be explained. This technique is essentially based on the determination of dichromatic confusion lines.

4.2.1 Diagnostics of Dichromatic Observers

Pseudoisochromatic plates such as Ishihara plates are useful instruments to distinguish between color normal and color deficient observers. However, they fail in the task to separate dichromats from anomalous trichromats. Furthermore, it is not possible to classify dichromatic observers correctly using only pseudoisochromatic plates. Therefore, a different approach was used in order to identify and classify dichromatic observers. This method is based on the determination of dichromatic confusion lines. In the following I will summarize this method in brief.

The method consisted of two parts. First, for each color deficient observer an individual isoluminance plane at a luminance of approximately 41.6 cd/m^2 was estimated using the heterochromatic flicker technique described in Section 2.1.3. The stimulus at the center of the monitors gamut served as fixed standard ($R = G = B = 0.5$). Test lights were 30 different lights that were drawn at random from the gamut of the screen. Second, if possible, dichromatic confusion lines were determined using the following rationale. If a given subject is protanope or deuteranope the remaining M- or L-cones are excited equally by corresponding isoluminant stimuli. Perceptual differences of stimuli taken from the dichromatic isoluminance plane are only due to different S-cone excitations. Within the isoluminance plane, lines of colors that excite the S-cones in equal manner can be identified. These lines are confusion lines of the respective dichromat.

In the experiment, subjects virtually rotate lines of isoluminant stimuli around a fixed standard P until all stimuli from the given line are perceptually indistinguishable (Figure 4.3a). The observers were presented with two stimuli on the two halves of a bipartite disk. One stimulus was the fixed standard P. The color of the second stimulus was sinusoidally modulated along a line within the given isoluminance plane. The line was defined by starting point U and ending point V and went through P. The subject perceived this stimulus as changing permanently in color. This impression can be described as pulsing. In the experiment, subjects were asked to rotate the line $\overline{UV}$ around P until the pulsing disappeared. At this point the subject perceived a homogeneous disk. In this case, P and the resulting stimuli U' and V' are indistinguishable for the subject and hence lie on a confusion line. In other words, P and all points on the line $\overline{U'V'}$ are projected onto one point of that subject's color space (see Figure 4.3a).

It should be clear that this task can only be completed by dichromatic ob-

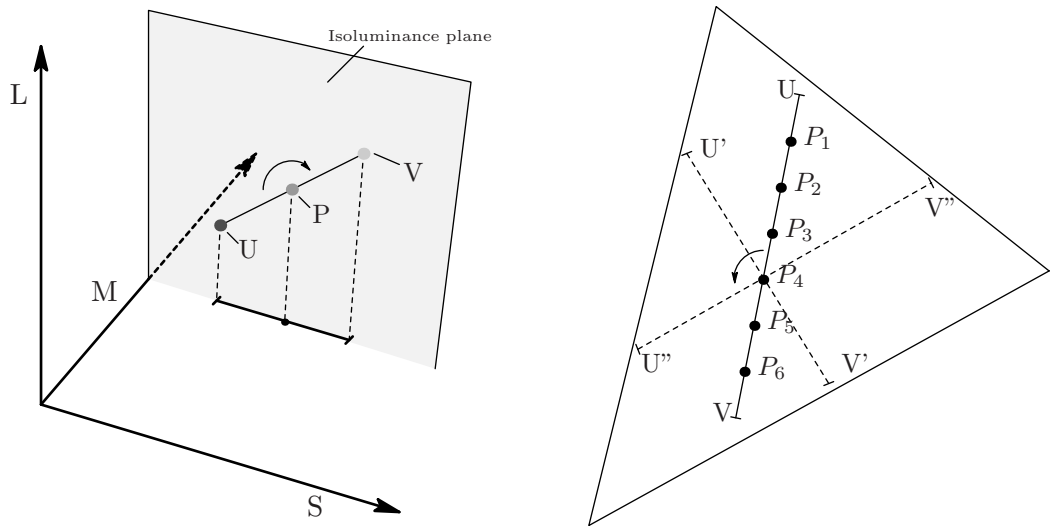

Figure 4.3: Determination of confusion lines. (a) In the experiment subjects rotate isoluminant stimuli on the line $\overline{UV}$ around the standard P. The rationale of the experiment is shown in LMS-space. A protanopic confusion line is found when the projection of the rotated line onto the MS-plane can be described as one point (left). (b) The experiment is carried out for six different standard stimuli $P_1 - P_6$. Starting and ending points of the rotated line are chosen so that they lie close to the border of the monitors gamut (right).

servers. The experiment is carried out with six different standard stimuli so that six different confusion lines can be determined (Figure 4.3b). The different standards are chosen from a line within the isoluminance plane that represents S-cone excitations. The mean direction of confusion lines in receptor space and the estimated intersection point of these lines in xy-chromaticity diagram are then used to classify the dichromatic observer.

4.2.2 Stimuli

Visual display

The stimuli that were used in the experiments consisted of an array of 25 flat matte surfaces. The array of surfaces was presented in front of a large homogeneous background. The background represented a surface with mean cone excitations of all surfaces in the array. This visual display has been used in previous studies on color constancy (Brainard & Wandell, 1992; Bäuml, 1999b; 2001). A control experiment revealed that this configuration increased the reliability of the subjects' settings compared to a stimulus with surfaces arranged side by side. An example of the visual display is shown in Figure 4.4. The surfaces represented simulations of Munsell papers rendered under a simulated spatially uniform illumination. The square in the center of the stimulus was the test patch which could be manipulated in chromaticity by the subject. The fixed set of surfaces

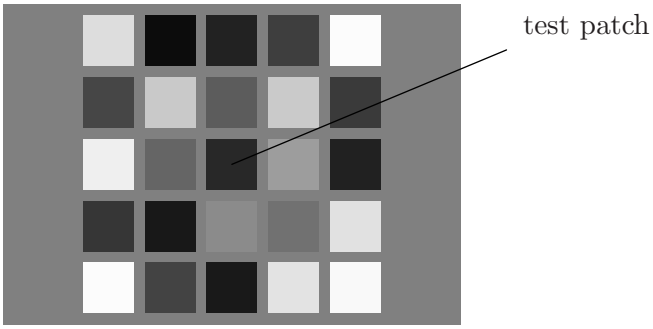

Figure 4.4: Typical stimulus configuration in the experiments. The surface in the center of the stimulus is the test patch.

was chosen carefully and contained surfaces of different hues and luminances. The surfaces were selected so that their mean receptor codes represented a surface which appeared achromatic to the observer under neutral illumination. The Munsell notation, chromaticity coordinates and luminances of the set of surfaces under neutral illuminant $D65$ are given in Table B.2 (Appendix B).

Reflectance functions of Munsell surfaces and approximated spectral power distributions of illuminants were used in order to simulate the stimuli under the given illumination, according to the simple lighting model introduced in Equation (3.1). The reflectance functions of the surfaces are based on spectral reflectance measurements made with a spectrophotometer on 1,269 color chips from the 1976 *Munsell Book of Color* at a 1-nm resolution from 380 to 800 nm. This data set was obtained from `http://spectral.joensuu.fi/`. The range and resolution of approximated illuminant spectra was identical to those of the surfaces' reflectance functions. From the given reflectance functions of surfaces and spectral power distributions of illuminants $CIE\text{-}XYZ$-coordinates of stimuli were calculated using corresponding color matching functions. For example, the $CIE\ X$-coordinate of surface s under illuminant e is given by:

$$X_{s,e} = \sum_{\lambda=380}^{800} s(\lambda) \cdot e(\lambda) \cdot \bar{x}(\lambda), \tag{4.38}$$

where $\bar{x}(\lambda)$ denotes the corresponding color matching function. The xy-chromaticity coordinates of the set of surfaces used in the present study under neutral daylight $D65$ are shown in Figure 4.5.

The size of each surface in the array was 4.7cm × 4.7cm which is equivalent to a visual angle of 1°46' × 1°46'. The background extended over the whole screen and subtended 40.5cm × 30cm (15° × 11° of visual angle).

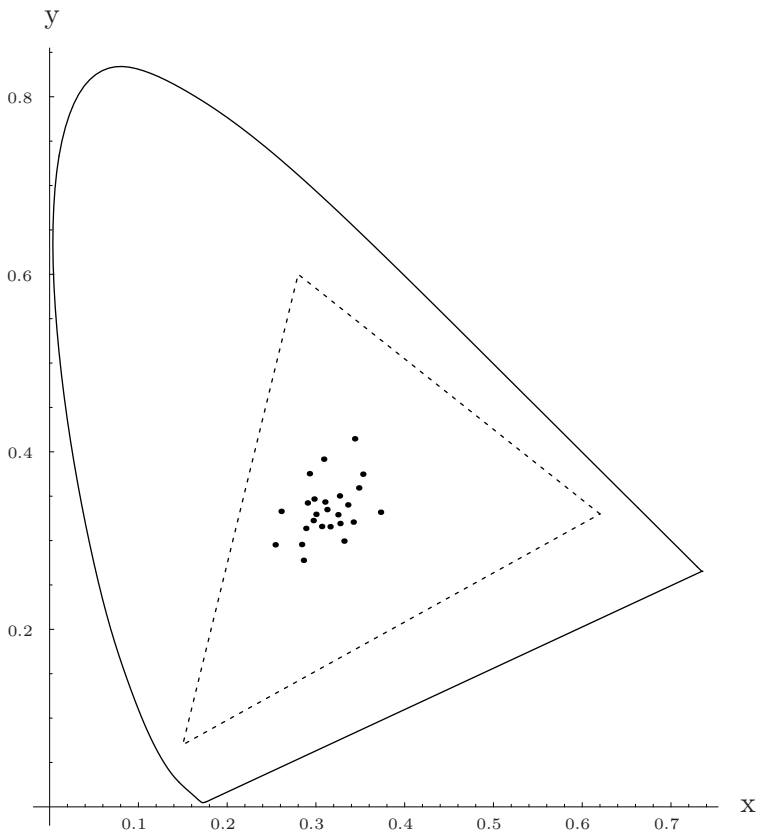

Figure 4.5: $xy-$chromaticity coordinates of stimuli under $D65$. The dashed triangle indicates the gamut of the monitor.

Simulated light sources

The light sources were simulations of either natural daylights or red and green illuminants. The spectral power distributions of these illuminants were approximated using the CIE-method described in Appendix A. The spectral power distribution of each simulated light was normalized to luminance $100\,cd/m^2$. Using illuminations with luminance of $100\,cd/m^2$ provided optimal stimuli within the gamut of the CRT-screen. In Experiments 1 and 3, simulations of daylights $D65$ (6500 K), $D40$ (4000 K) and $D10$ (10000 K) were used. These illuminants appear achromatic, yellowish and bluish to normal observers respectively.

Additionally, I used a red and a green illuminant with chromaticities far from the CIE-daylight locus in the first experiment. Two restrictions were made on these illuminants so that chromaticities of both lights could be determined uniquely. First, the distance of the respective $u'v'$-chromaticities of both lights from $D65$ equaled the mean distance of $D40$ and $D10$ $u'v'$-chromaticities from $D65$. Next the $u'v'$-chromaticities of the red and the green illuminant were selected so that they were located on the bisecting line of protanopic and deuter-

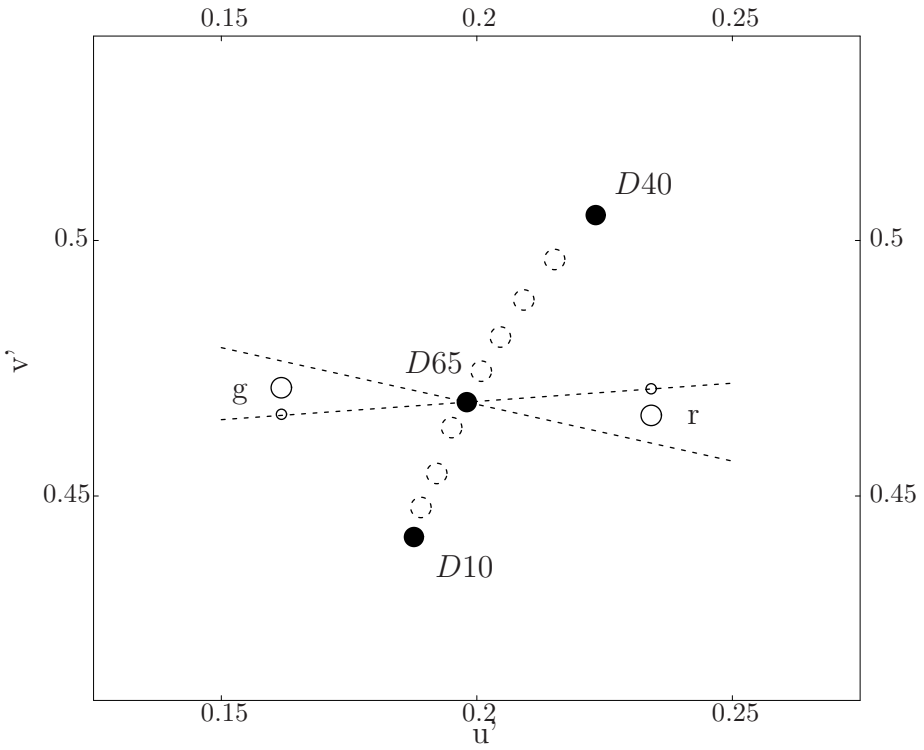

Figure 4.6: Illuminants used in Experiments 1, 2 and 3 in $u'v'$-chromaticity diagram. The filled circles indicate the illuminants D40 (4000 K), D65 (6500 K) and D10 (10000 K) that were used in Experiments 1 and 3. The large open solid circles show the red ('r') and the green ('g') illuminant of Experiment 1. The small open circles indicate red and green illuminants on a protanopic confusion line that were used in Experiment 2. The open dashed circles refer to daylights that were used only in Experiment 3. The dashed lines show protanopic and deuteranopic confusion lines.

anopic confusion lines through $D65$. In Experiment 2, light sources were represented by simulations of a red and a green illuminant with similar characteristics. For these lights the restriction was made that their chromaticities were aligned with a protanopic confusion line. In Experiment 3, seven additional daylights were used whose chromaticities spanned across the CIE-daylight locus. Figure 4.6 shows the $u'v'$-chromaticity coordinates of the illuminants used in the experiments. The xy-chromaticities and $u'v'$-chromaticities of these lights are given in Table B.1 (Appendix B).

On the use of CRT-simulations

Certainly it would be desirable to study the phenomenon of color constancy under natural viewing conditions. However, several objections argue against this approach. First, many factors that are assumed to affect the degree of color constancy are not under experimental control, such as temporal and spatial char-

acteristics of stimulus presentation and the adaptational state of the observer. Second, randomization of stimuli is almost impossible when real illuminants and surfaces are used. Third, under natural viewing conditions reliable settings of subjects are difficult to obtain. Taken together these objections make the use of simulations on CRT-screens a valuable alternative to experimental studies under natural viewing conditions. Most of the research on color constancy in recent times are based on simulations of surfaces and illuminants on CRT-monitors.

Nevertheless, some questions arise when simulations of surfaces and illuminants on CRT-screens are used. First of all, it is not clear if such simulations represent stimuli that are adequate to surfaces and illuminations in natural scenes. We only know that under certain assumptions the light signal from a simulated surface to the eye of the observer is identical to that from the corresponding physical surface under natural illumination.[8]

A related problem arises when we consider that natural surfaces only reflect the incoming light from the illuminant whereas simulated surfaces on monitors can be understood as light sources themselves. In many studies on color constancy only simulated surfaces that were incremental in relation to a global background were presented to the observer. It is more likely to interpret such incremental stimuli rather as self-luminous objects than as reflecting surfaces. In this sense, it has been argued that the distinction between increments and decrements can be considered as an important factor that affects the degree of adjustment to the illuminant (Bäuml, 2001).

In recent times several efforts have been made to extend simulations on CRT-displays to three dimensional scenes that contain additional cues to the illuminant. These configurations promise to be more comparable with conditions in natural scenes. We discussed the models that these displays are based upon already in Chapter 3. Another restriction of simulations refers to the fact that the range of stimuli realizable on CRT-screens is limited. For example, the intensity of the light reflected from a surface under daylight illumination is generally much higher than intensities of stimuli that can be realized on CRT-screens.

Despite these limitations, simulations of surfaces and illuminants on CRT-displays are an useful approach in studying color constancy under fixed conditions. They represent a compromise in the sense that the main features of natural viewing conditions are modeled in the controlled setting of the laboratory. Main principles that are assumed to underlie color constancy have been revealed both in studies using CRT-simulations (Valberg & Lange-Malecki, 1990; Brainard & Wandell, 1992) and in studies under nearly natural viewing conditions (Brainard, 1998; Kraft & Brainard, 1999).

[8]Important assumptions that are usually made when presenting simulations were discussed in Chapter 3. These assumptions referred to uniformity of illumination and to simulated surfaces that are matte and flat.

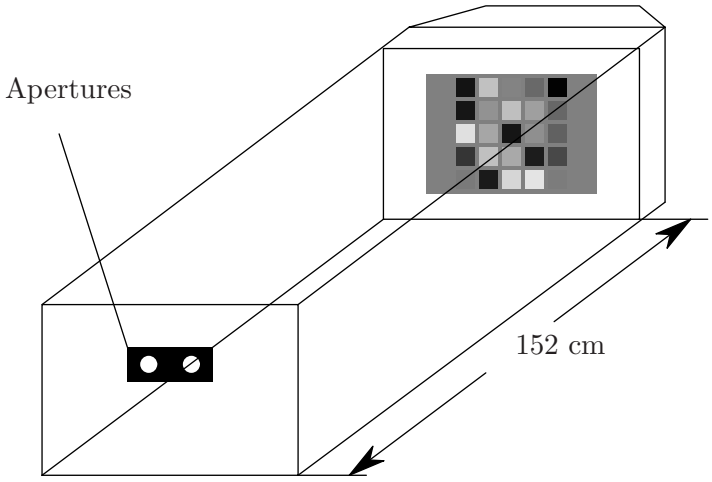

Figure 4.7: Apparatus: In the experiments, simulated stimuli were presented on a CRT-screen. A wooden box which was painted black inside to avoid disturbing reflections was attached to the screen. Two small apertures of 25mm diameter each were mounted at the end of the box. The subjects observed the stimuli presented on the screen through these apertures.

4.2.3 Apparatus

The stimuli were presented to the subjects on an EIZO FlexScan F980 CRT screen with resolution 1.024×768 pixel and frame rate of 80Hz. The screen was controlled by a PC with a VSG 5/2 graphics card. The monitor was calibrated in order to correct for nonlinearities of the gun responses. The calibration procedure was carried out using the VSG Desktop software and a Colorcal colorimeter (Cambridge Research Systems Ltd.). This procedure followed the guidelines suggested by Brainard (1989).

A wooden box which was painted black inside to avoid disturbing reflections was attached to the screen. Two small apertures of 25mm diameter each were mounted at the end of the box. The observers viewed the stimuli presented on the screen through these apertures. The distance between the observers eyes and the screen was 152 cm. The apparatus is shown in Figure 4.7. The experiments were programmed by myself in the C++ language. Subjects made their settings using a game pad.

4.2.4 Task

In the experiments subjects made achromatic settings for the test patch in the center of the visual display. They were asked to adjust the chromaticity of the test patch until it appeared neither bluish nor yellowish and/or neither reddish nor greenish to them. The luminance of the test patch remained unchanged in each trial. Dichromats made their settings always on one line that was defined as the line from $D65$ to the test illuminant. This restriction was made in order

to avoid large deviations in the settings. Under illuminant $D65$ protanopes made their adjustments on a line that represented the opponent BY-axis.

For reasons of comparison trichromatic subjects made their settings in Experiments 1 and 2 on the same lines as protanopes except for $D65$-conditions. Even under this restriction trichromats were always able to find satisfying settings. In Experiment 3 and in all $D65$-conditions, trichromatic subjects used two lines to make their adjustments. In addition, settings of two normal observers were also obtained on two lines for all conditions of Experiment 1. These lines were always defined as opponent axes that corresponded to the two opponent channels RG and BY.

4.2.5 Procedure

In the following the general procedure that was used throughout all experiments will be described. In general, different illuminant conditions were blocked into different sessions. The number of different illuminant conditions was five in Experiment 1, two in Experiment 2 and ten in Experiment 3. Within each illuminant block test surfaces at ten different luminance levels were presented to the subject. Half of the test patches had luminances below that of the mean luminance of the whole array (decrements) and the other half had luminances above it (increments). Subjects repeated settings for each of the ten different test surfaces four times resulting in 40 settings per block. The order of illuminant blocks was randomized and differed across subjects. Within one illuminant block the test surface order was randomized and different subjects saw different randomizations.

Before starting a session the subject dark adapted for about 10 min. Then the subject was presented with the first stimulus. Studies on the time course of chromatic adaptation have revealed that chromatic adaptation is essentially complete after 2 min (Fairchild & Reniff, 1995; Rinner & Gegenfurtner, 2000). In order to control the adaptational state of the subject, settings on the first stimulus could be confirmed by the subject only after 2 min of presentation.

Subjects made their settings on either one or two chromatic lines using a game pad. After confirming the achromatic setting the next stimulus appeared on the screen. In order to avoid local contrast effects the positions of all surfaces except the test patch were randomly changed every three seconds. The initial chromaticity of the test surface was chosen randomly from an interval around the chromaticity of the test illuminant.

4.2.6 Data Analysis

Color constancy indices

First, we shall describe in some detail what kind of achromatic settings we would expect if an observer had perfect or no color constancy. In case of perfect color

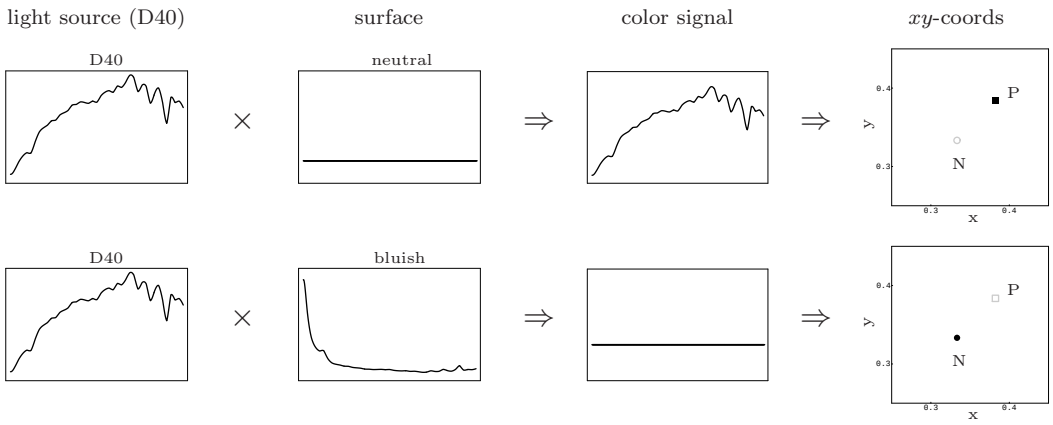

Figure 4.8: Examples of achromatic settings in case of perfect and no color constancy. Two surfaces, one that appears achromatic and one that appears bluish under neutral light are illuminated by yellowish daylight D40. The first column shows the spectral power density of D40 and the second column shows the reflectance functions of the surfaces. In the third column, color signals from these surfaces under D40 are depicted. In the last column, xy-chromaticities of the respective color signals are shown as black filled symbols. The expected chromaticity in case of perfect constancy is shown as square and denoted with 'P'. The expected chromaticity in case of no constancy is shown as circle and denoted with 'N'. We see that the expected setting in case of perfect constancy is identical with the chromaticity of the illuminant (top row). In contrast, the expected setting in case of no constancy is identical with the neutral point 'N' (bottom row).

constancy we expect the achromatic settings of the subject to coincide with the chromaticity of the illuminant. An observer without any color constancy would choose always one and the same chromaticities of a neutral standard as achromatic settings under all different illuminant conditions. In the color constancy literature this neutral standard of 'zero' constancy is often set to the chromaticity of neutral daylight D65 (e.g. Arend, Reeves, Schirillo & Goldstein; 1991)).

Initially, it may not be intuitive for us that an observer who constantly chooses one and the same chromaticities across different illuminants as achromatic in appearance should have no color constancy at all. To make this point clear we should keep in mind that the observer is not asked to choose the chromaticity of the surface but the chromaticity of the color signal from that surface under test illumination. Lets take a look at Figure 4.8 where color signals from two surfaces under yellowish daylight D40 are analyzed. The first surface reflects incident light equally at all wavelengths. This surface would appear achromatic to the observer under neutral illumination. When seen under D40, the xy-coordinates of the color signal from that surface are identical with those of D40. If an observer chooses a surface with this chromaticity as achromatic in appearance under D40, we will therefore assume that he has perfect color constancy.

Let us consider now the second case. A surface which appears bluish under neutral illumination produces a color signal under D40 with constant density at

all wavelengths. The chromaticity of that color signal is identical with the neutral point in xy-diagram. An observer who chooses this chromaticity as achromatic in appearance relies only on the color signal but is completely ignorant of reflectance characteristics of the surface. Therefore, we assume that such an observer has no color constancy.

After we exemplified cases of no and perfect constancy we shall introduce an index of color constancy. In the color constancy literature it is common to express the degree of adjustment to the illuminant in terms of color constancy indices. The idea of such a constancy index goes back to Brunswick (1929). Arend, Reeves, Schirillo and Goldstein (1991) suggested a color constancy index which refers only to chromatic constancy. Deviations from color constancy are represented as Euclidean distances between settings of the subject and the point of perfect color constancy in CIE-$u'v'$-chromaticity diagram. This nonlinear transformation of CIE-xy-chromaticity diagram is roughly based on perceptual distances. Let the vector $\mathbf{m}_{D65}$ represent $u'v'$-coordinates of the mean achromatic setting of the subject under neutral daylight $D65$ and $\mathbf{m}_{data}$ corresponding mean achromatic settings under test illumination. We will denote settings that indicate perfect color constancy with $\mathbf{m}_{perf}$. This index can be written as:

$$CI = 1 - \frac{|\mathbf{m}_{data} - \mathbf{m}_{perf}|}{|\mathbf{m}_{D65} - \mathbf{m}_{perf}|}, \tag{4.39}$$

where the numerator and the denominator of the fraction can be understood as Euclidean distances. This index is 1 in case of perfect constancy and 0 in case of no constancy.

The advantage of the use of such a color constancy index is that results are directly comparable to results from previous studies. However, conclusions based on color constancy indices should be taken with care for at least two reasons. First, distances in CIE-$u'v'$ correspond only roughly with perceptual distances. Second, as deviations from perfect color constancy are weighted on the illuminant shift indices depend on the size of the illuminant shift. This often leads to small indices with small distances between the test illuminant and $D65$. We will return to this issue in Section 4.5.3.

Analysis of subjects' reliability

The reliability in the settings of a given subject was investigated using $u'v'$ coordinatization. The analysis of reliability in terms of chromaticity coordinates has the advantage that error variation is independent of luminance level. For each subject, each illuminant condition and at each luminance level variances were determined which represented the error in the subjects' settings. Total u' and v' error variances were calculated for each subject and illuminant condition by simply adding all corresponding variances at different luminance levels. Recall that in most of the conditions subjects were asked to make their settings only on

one line in color space. In this case, settings of subjects in u' and v' direction are perfectly correlated and u' and v' error variances contain redundant information so that in general only total u' error variances will be given.

4.3 Experiment 1

4.3.1 Introduction

The main objective of Experiment 1 is to study the color constancy performance of normal and protanopic observers under different illuminant conditions. In this experiment the following questions will be addressed:

1. Is the color constancy performance of protanopes under various illuminant conditions comparable with normal trichromats?

2. Is color constancy of normal observers better under daylight conditions than under non-daylights?

3. Under which conditions do increment-decrement asymmetries occur?

4. Is the observers' adjustment to the illuminant for decremental stimuli better than for incremental stimuli?

4.3.2 Methods

Stimuli and procedure

The visual display in Experiment 1 was already described in the general methods section. A fixed set of 24 simulated surfaces was used which were always presented in front of a homogeneous background (see Figure 4.5 and Table B.2, Appendix B). In Experiment 1, five different illuminants were simulated. Three of them were daylights (D65, D40 and D10) and the other two were a red and a green illuminant respectively (see Figure 4.6 and Table B.1, Appendix B). The stimuli presented to the subjects were simulated from the surfaces reflectance functions and the spectral power distributions of illuminants as described in Section 3.1.1. Test stimuli at ten different luminance levels were used. Half of the test patches were increments relative to the mean color signal and the other half were decrements. The task of the subject was always to make achromatic settings for the test stimulus. A detailed description of the procedure was already given in Section 4.2.5.

According to the study by Rüttiger, Mayser, Sérey and Sharpe (2001), the red and green test illuminations were chosen so that their chromaticities lay on a bisecting line of respective protan and deutan confusion lines. However, one might argue that the remarkably good color constancy performance of protanopes under

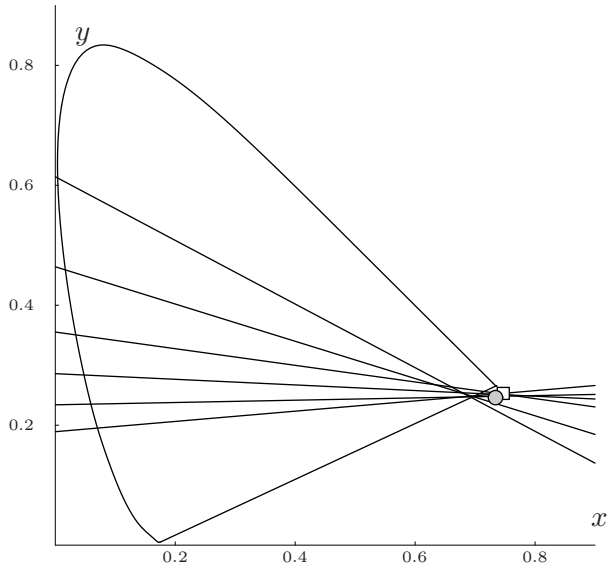

Figure 4.9: Pattern of confusion lines of protanopic subject LE in xy-chromaticity diagram. The gray filled circle shows the estimated confusion point and the open square indicates the confusion point of the standard protanope.

simulated reddish and greenish illuminants reported by Rüttiger et al. (2001) is only due to the fact that stimuli were not aligned with dichromatic confusion lines. This assumption will be tested in Experiment 2.

Observers

Seven subjects took part in Experiment 1. All subjects were students who were not aware of the purpose of the experiment, except subject SS, the author. First, color vision of all subjects was tested with Ishihara diagnostic plates (Ishihara, 1997). This screening showed that subjects KH, SA, TK and RL had normal color vision. The response pattern of subjects MB, LE and SS identified them as color deficient observers. Further analysis of confusion lines using the method described in Section 4.2.1 revealed that these observers can be classified as protanopes. Estimated confusion points and directions of missing fundamentals of these observers are given in Table C.1, Appendix C. Confusion lines of protanope LE are shown exemplarily in Figure 4.9. All protanopic subjects and the normal subjects KH, SA and TK made their settings on one line. In additional blocks, trichromats KH and RL made also settings on two chromatic lines for all illuminant conditions (see Section 4.2.4 for details).

4.3.3 Results

Color constancy

Color constancy performance of protanopic subjects was comparable to that of normal observers except for the green illuminant condition. In other words, even under red illumination performance of protanopes was surprisingly good. The analysis of color constancy of normal observers showed that performance in the two daylight conditions was not better than in the two conditions with illuminants off the daylight locus. The mean color constancy indices of all observers are shown in Figure 4.10 (see also Table C.3, Appendix C).

In addition, adjustment to the illuminant was analyzed separately for incremental and decremental stimuli. In general, color constancy indices were higher for decrements than for increments. This pattern was found consistently for conditions in which increment-decrement asymmetries occurred. Figure 4.11 shows mean color constancy indices of all observers separately for increments and decrements under illuminant D40 where strongest asymmetries were observed.

Error variation in the data

Error variation in the data of protanopic subjects was comparable to that of normal subjects in the daylight conditions $D40$ and $D10$. Due to the fact that normal observers made their settings in the $D65$ condition on two lines, their data showed larger error variation in this condition than that of the protanopes. For the red and green conditions, error variation in the data of protanopic observers was much larger compared to normal subjects. The total u' error variances for different conditions and subjects are shown in Figure 4.12.

Increment-decrement asymmetries

The analysis of increment-decrement asymmetries was based on comparisons of achromatic settings under test illuminant with ideal settings under D65. For subjects who made their settings just on one line, only the results for the S-cone channel will be reported. Recall that L-cone, M-cone and S-cone settings are pairwise perfectly correlated when settings are made on one line in isoluminant planes. For those normal subjects who used two chromatic lines for their settings M-cone results will also be reported. In general, strong increment-decrement asymmetries occurred only for the S-cone. In particular, strong asymmetries were found under D40 and weak asymmetries under D65 in the data of all subjects. Under red and green illuminations S-cone increment-decrement asymmetries did not occur. The results for the S-channel in the D10 condition are ambiguous. Asymmetries were found in the data of protanopic observers which was not the case for normal observers.

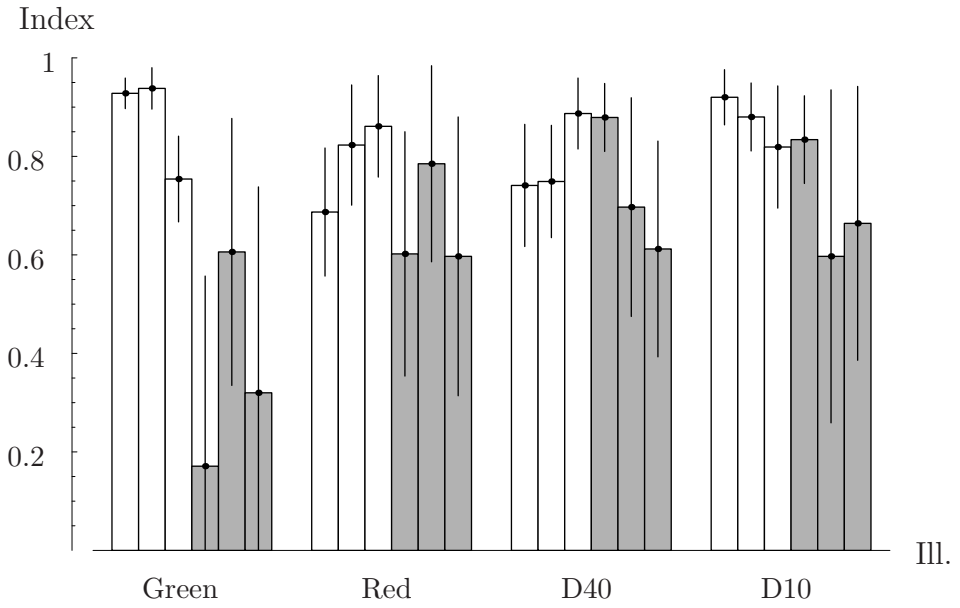

Figure 4.10: Color constancy indices of subjects in Experiment 1. The mean indices of the three normal observers KH, SA, TK (white bars) and those of the three protanopic observers SS, MB and LE (gray bars) are shown with corresponding error bars for the four different chromatic illuminant conditions.

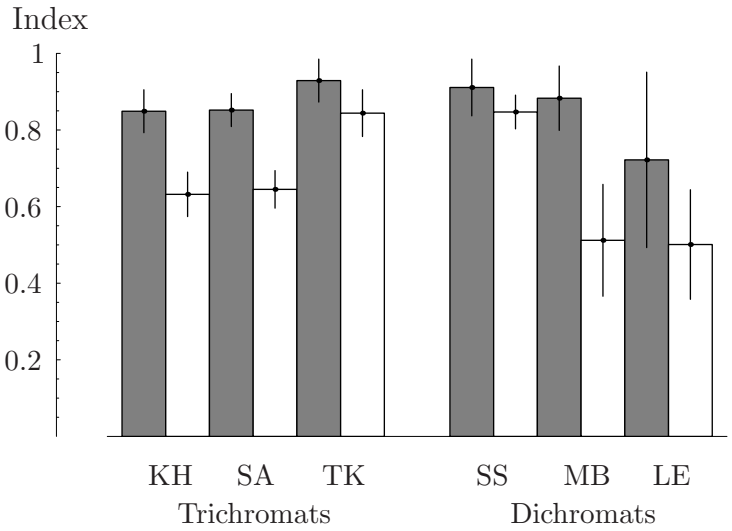

Figure 4.11: Color constancy indices of subjects under illuminant D40 separately for incremental and decremental stimuli. Gray bars refer to decrements and white bars refer to increments. Mean indices with error bars are shown separately for the three normal observers KH, SA, TK and the three protanopic observers SS, MB and LE.

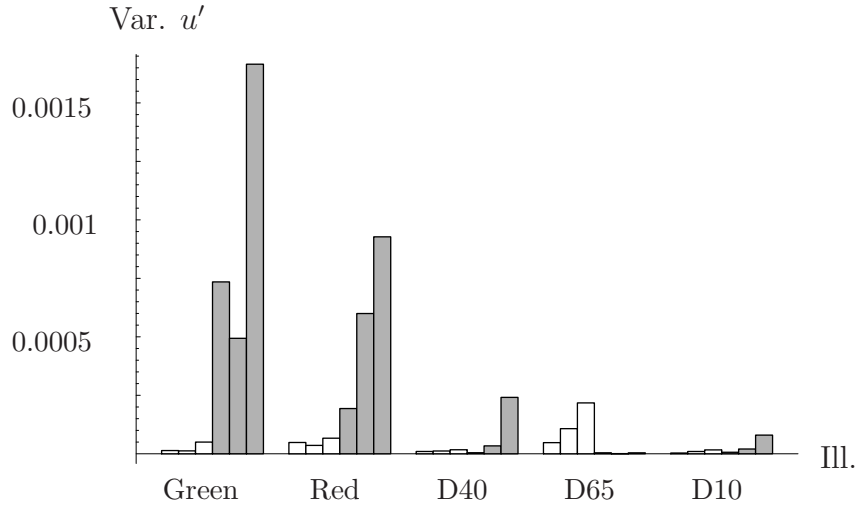

Figure 4.12: Total u' error variance of all subjects in Experiment 1 under different illuminant conditions. White bars indicate total error variances of the three normal observers KH, SA, TK and gray bars show those of the three protanopic observers SS, MB and LE.

In Figure 4.13 S-cone data of a normal and a protanopic subject are compared under three illuminant conditions. The figure shows that in the $D40$ condition, settings for incremental and decremental stimuli seem to fall on two separate straight lines that are bent at a certain point. These bends represent asymmetries in processing of increments and decrements. The intersection point can be interpreted as S-cone excitations of references r and r' which separate incremental and decremental stimuli. In conditions where increment-decrement asymmetries occurred estimates of receptor codes of the reference r are in good agreement with mean cone excitations across the whole scene (Figure 4.14).

To express the strength of increment-decrement asymmetries for a given condition c as a single number, ratios q_c of the slopes s_c^I and s_c^D of the two separate regression lines were determined:

$$q_c = \frac{s_c^I}{s_c^D}. \tag{4.40}$$

Ratios q_c considerably different from one indicate increment-decrement asymmetries. Figure 4.15 shows such ratios that refer to S-cone results for all subjects and illuminant conditions .

M-cone excitations of achromatic settings were analyzed for the two normal subjects KH and RL who made settings on two chromatic lines. Figure 4.16 shows M-cone data of these observers for selected illuminant conditions in comparison with ideal settings under D65. Inspection of Figure 4.16 indicates that increment-decrement asymmetries for the M-cone channel did not occur in the data of these

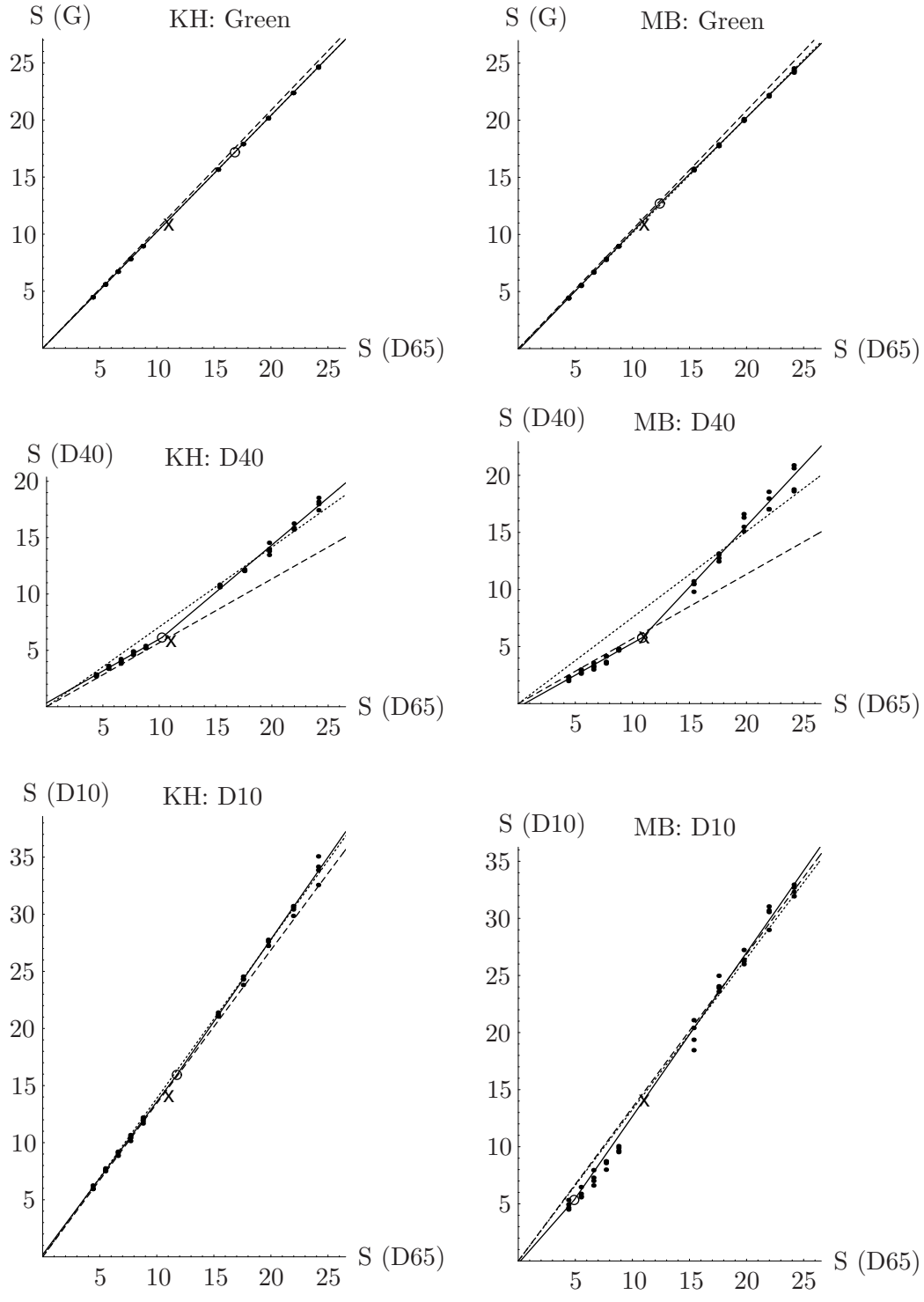

Figure 4.13: S-receptor codes of achromatic settings of the normal subject KH (left column) and the protanope MB (right column). Settings of an ideal observer under D65 are plotted against settings of the subject under test illumination. The first row shows data from the green illuminant condition, the second row settings under D40 and the third row data from the D10 condition. Solid lines represent least square fits independently for incremental and decremental stimuli. Dotted lines show predictions by a diagonal model of von Kries type and dashed lines predictions by the spatial mean model. The x indicates mean cone coordinates across the scene and the open circle S-receptor codes of the estimated reference stimulus.

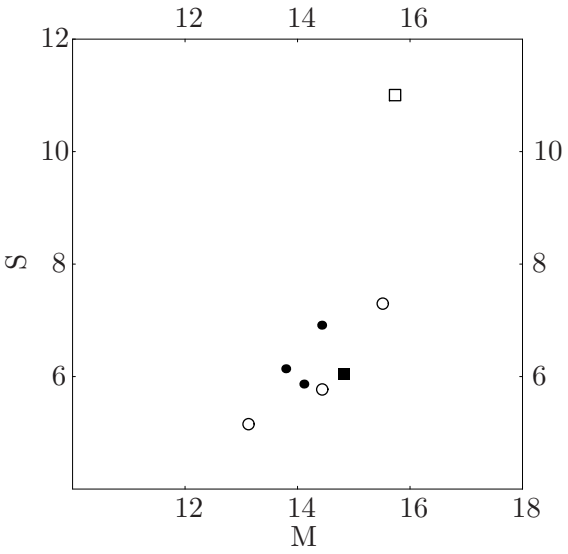

Figure 4.14: M- and S-receptor codes of estimated references r in the D40 condition for protanopic (open circles) and normal observers (filled circles). The filled square indicates mean cone excitations across the scene under D40. The open square shows these mean receptor codes under D65 for comparison.

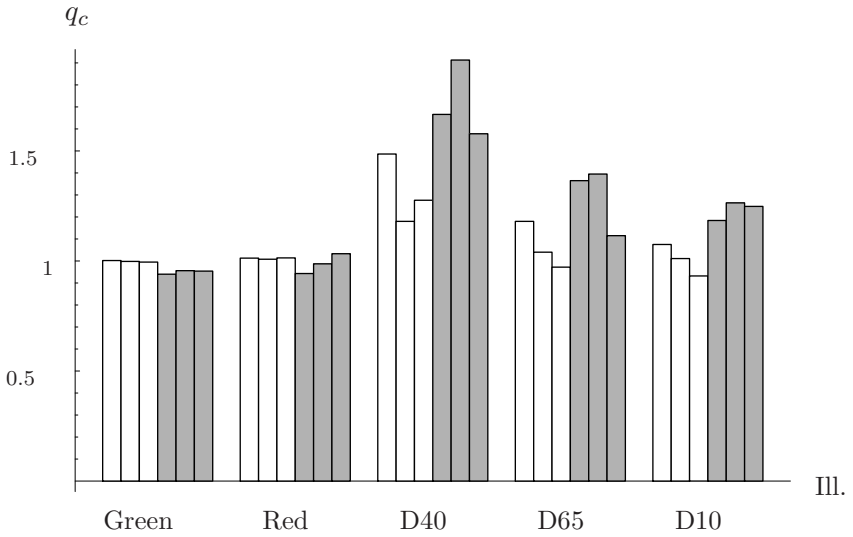

Figure 4.15: Ratios q_c of the slopes s_c^I and s_c^D of the two separate regression lines that refer to S-cone incremental and decremental stimuli respectively. White bars indicate ratios of the three normal observers KH, SA, TK and gray bars show ratios of the three protanopic observers SS, MB and LE. Ratios different from one indicate increment-decrement asymmetries.

two observers. The analysis of ratios q_c that are derived from slopes of regression lines referring to incremental and decremental stimuli confirmed this impression. These ratios are shown in Figure 4.17.

Model test

Using the rationale described in Section 4.1.2, parameters of the three models were estimated for each subject and illuminant condition. Mean estimates of S-cone rescaling coefficients across normal and protanopic observers are shown in Figure 4.18. Regarding the S-cone component of subjects' achromatic settings the three models make similar predictions in the green and red illuminant condition. This result is not surprising as the S-channel is affected only marginally by illuminant shifts along a red-green axis. Consequently, in these conditions estimated S-cone parameters of all models are close to one.

Under D40 S-cone excitations are rescaled differently depending on whether the stimulus is incremental or decremental relative to the reference. This is also true for protanopic observers in the other two daylight conditions. For normal trichromats, rescaling parameters related with increments and decrements differ slightly under D65 but they are identical in the D10 condition. If we recall that subjects' achromatic settings under test illuminant were predicted from ideal settings under D65, we would have expected coefficients of one under test illuminant D65. Surprisingly, this prediction is confirmed for decrements but not for increments. In other words, ideal achromatic settings for increments under D65 appear slightly yellowish to the subjects whereas those for decrements appear in fact achromatic.

Another interesting feature in the data becomes evident when we compare predictions of the ID model with those of the spatial mean model. In cases of strong increment-decrement asymmetries, predictions of the spatial mean model are very close to predictions of the ID model for decremental stimuli. This finding can be also observed in Figure 4.13.

In addition, M-cone rescaling coefficients of the three models for the two normal observers KH and RL who made achromatic settings on two lines were estimated. These parameters are shown in Figure 4.19. The results indicate that M-cone increment-decrement asymmetries did not occur.

I will now report the results of the statistical model tests. First, for each subject and illuminant condition, S-cone rescaling coefficients of the ID model were tested against those of the more restricted diagonal model. The number of significant results of corresponding F-tests at α-level 0.05 are shown for each condition in Figure 4.20. All F-tests were significant in the D40 condition. Five out of six tests yielded significant results for each of the red and green illuminant conditions. This finding is contradictory to what could be concluded from visual impressions of data representations (see Figures 4.13, 4.15, 4.18). Three tests became significant in the D65 and D10 conditions. Significant results essentially

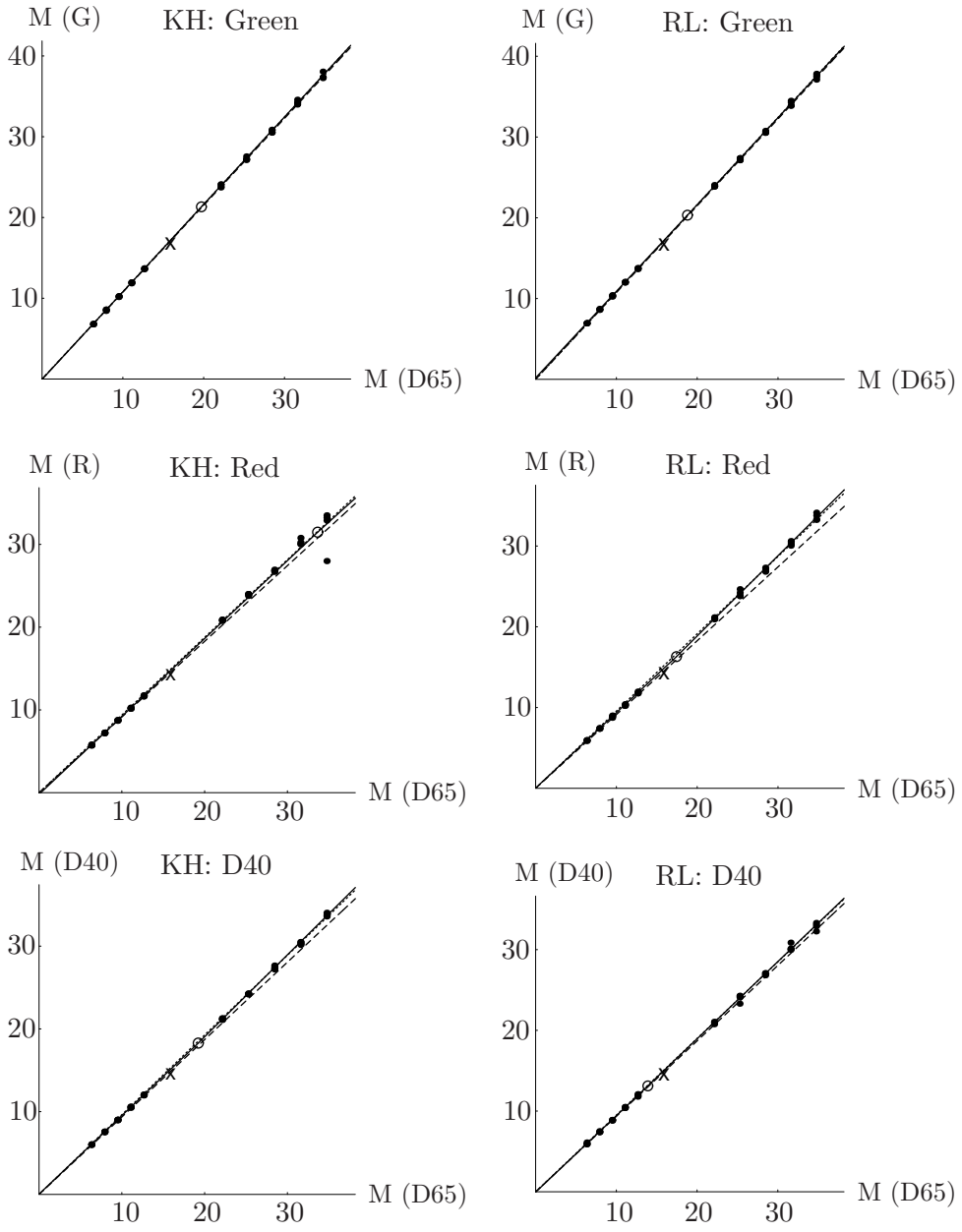

Figure 4.16: M-receptor codes of achromatic settings of the normal subjects KH (left column) and RL (right column) who made settings on two lines. Settings of an ideal observer under D65 are plotted against settings of the subject under test illumination. The first row shows data from the green illuminant condition, the second row settings under red illumination and the third row data from the D40 condition. Solid lines represent least square fits independently for incremental and decremental stimuli. Dotted lines show predictions by a diagonal model of von Kries type and dashed lines predictions by the spatial mean model. The × indicates mean cone coordinates across the scene and the open circle M-receptor codes of the estimated reference stimulus.

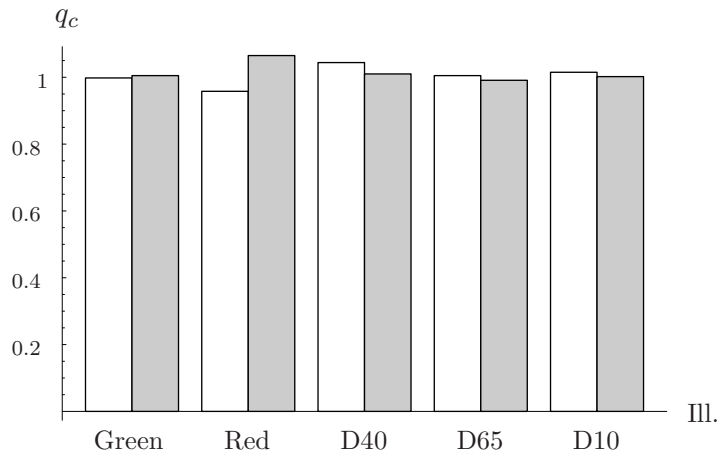

Figure 4.17: Ratios q_c of the slopes s_c^I and s_c^D of the two separate regression lines that refer to M-cone incremental and decremental stimuli respectively. White bars indicate ratios of normal observers KH and gray bars show ratios of the normal observer RL. Ratios different from one indicate increment-decrement asymmetries.

occurred for both groups of subjects.

In addition, M-cone rescaling coefficients of the ID model were tested against those of the more restricted diagonal model for subjects KH and RL. These tests yielded no consistent pattern of results. Significant results were found for subject KH in the D40 and D10 conditions but not for RL. In contrast, F-tests of RL were significant in the other three conditions whereas those of KH were not.

As a second step, parameters of the diagonal model were tested against those of the more restricted spatial mean model. In general, predictions of an unspecified diagonal model were better than predictions by the spatial mean model. Only two out of 40 tests yielded non-significant results ($\alpha = 0.05$) when contrasting both models. However, visual inspection of predictions by the spatial mean model (see Figure 4.13) suggests that this model may provide a useful heuristic, especially in the case of decremental stimuli. In general, results from statistical tests should be interpreted with care as distances in cone excitation space that the statistical analysis is based upon do not correspond with perceptual distances.

Control experiment: no background

Four normal subjects, KH, KUH, RL, SA and the protanope SS made achromatic settings under D40 using a slightly modified visual display. In this case, the 25 surfaces were presented without the homogeneous background. There was no gap between adjacenting patches. The resulting stimulus can be imagined as a 5×5 checkerboard. The other parts of stimulus and procedure were identical to those described in Section 4.2. This control condition was introduced in order to rule out the possibility that increment-decrement asymmetries are rather an artifact

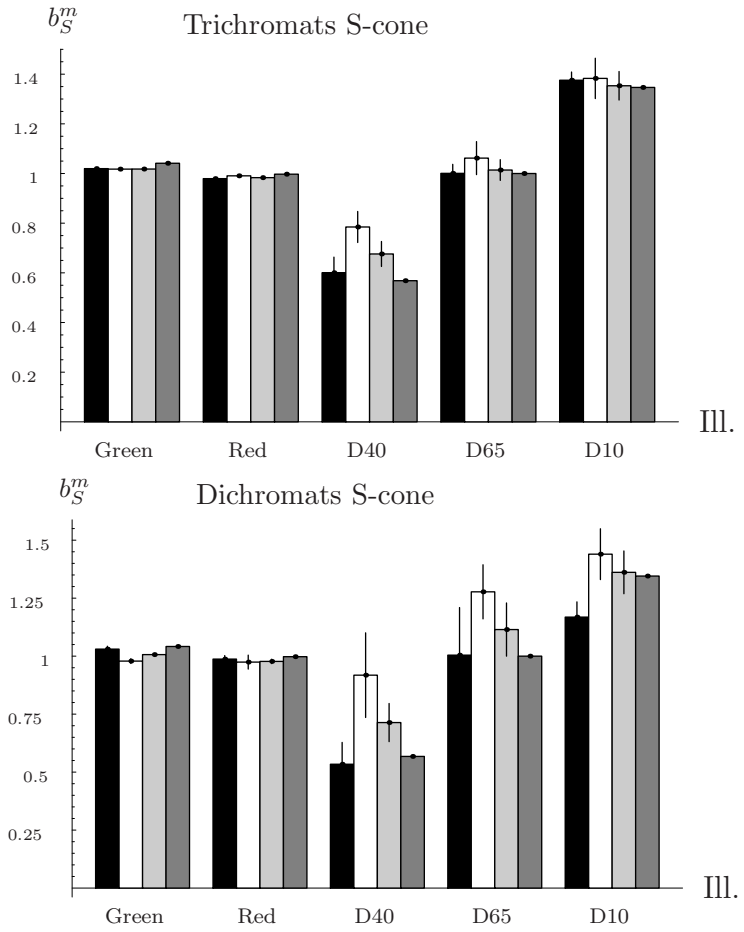

Figure 4.18: Estimated S-cone rescaling parameters of the three models. The parameters are univariate regression weights b_i^m for predicting subject's achromatic settings under test illuminant from ideal settings under D65 which refer to model m and receptor type i. Different bar shadings refer to different models: black – ID model (decrements), white – ID model (increments), light gray – diagonal model, dark gray – spatial mean model. Mean coefficients with corresponding error bars are shown for normal (top) and protanopic observers (bottom).

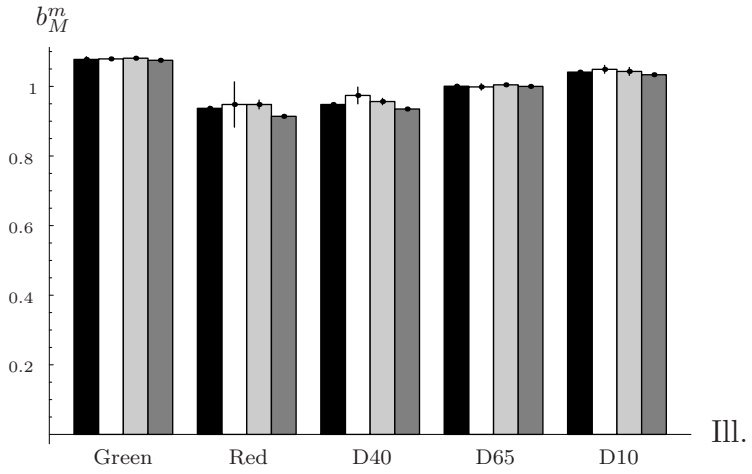

Figure 4.19: Estimated M-cone rescaling parameters of the three models. The parameters are univariate regression weights b_i^m for predicting subject's achromatic settings under test illuminant from ideal settings under D65 which refer to model m and receptor type i. Different bar shadings refer to different models: black – ID model (decrements), white – ID–model (increments), light gray – diagonal model, dark gray – spatial mean model. Mean coefficients across the two normal observers KH and RL are shown with corresponding error bars.

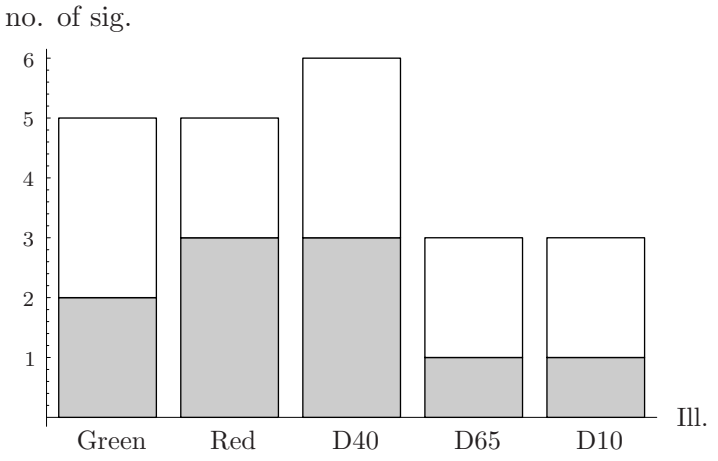

Figure 4.20: Number of significant results when testing S-cone coefficients of the ID model against the diagonal model (F-tests with $F_{3,36,.05}$). Gray bars indicate the number of significant results for normal observers and white bars the number of significant results for protanopic observers. Six tests were conducted for each illuminant condition.

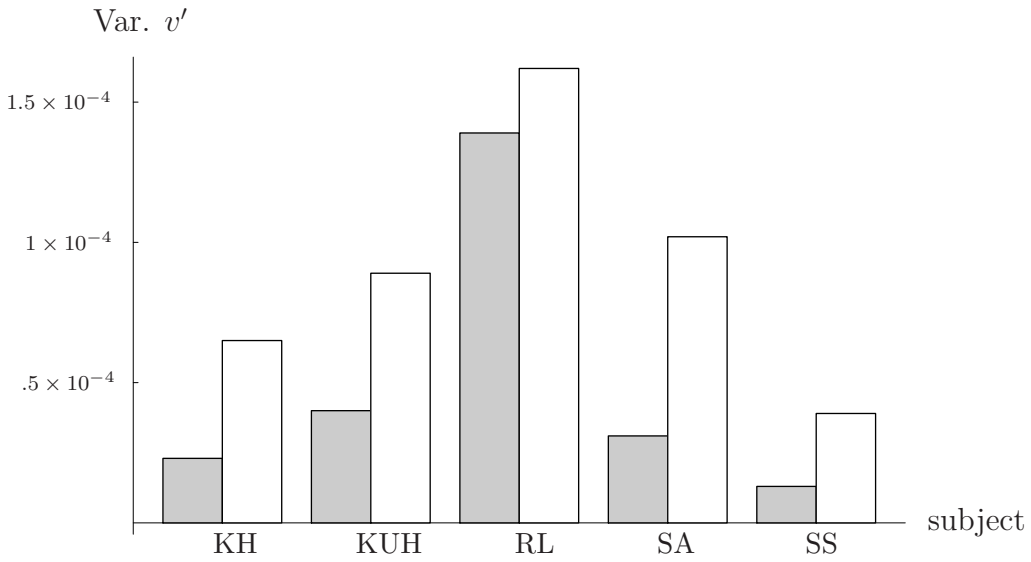

Figure 4.21: Comparison of v' total error variances under D40 in the standard condition with homogeneous background (gray bars) and in the control condition without background (white bars). Error variances are shown for normal subjects KH, KUH, RL, SA and for the protanope SS.

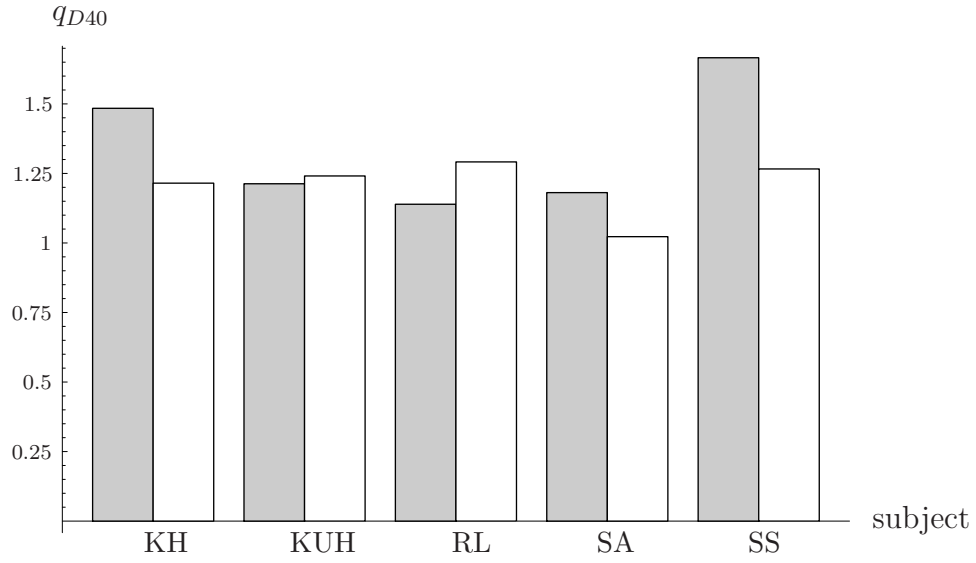

Figure 4.22: Ratios q_{D40} of the slopes s^I_{D40} and s^D_{D40} of the two separate regression lines that refer to S-cone incremental and decremental stimuli respectively under illuminant D40. Gray bars show ratios for the standard condition with homogeneous background and white bars indicate ratios for the control condition without background. Ratios are shown for normal subjects KH, KUH, RL, SA and for the protanope SS.

of the visual display, namely the homogeneous background.

All subjects showed larger variation in their settings when omitting the homogeneous background compared to the standard condition with background. Figure 4.21 depicts a comparison of v' total error variances for both conditions. Except for one subject, SA, S-cone increment-decrement asymmetries occurred also under daylight D40 when surfaces were presented without background. In this control condition, asymmetries were weaker for subjects KH and SS and larger for subject RL compared to the standard condition with homogeneous background. Ratios q_{D40} that indicate the strength of the asymmetry (see Equation 4.40) are shown for all subjects and both conditions in Figure 4.22.

4.4 Experiment 2

4.4.1 Introduction

The main purpose of Experiment 2 is to investigate whether achromatic settings of protanopic observers are in agreement with the reduction hypothesis of dichromacy. In order to do so, two important modifications of the methods used in Experiment 1 will be introduced. First, simulations of a red and a green illuminant whose chromaticities are placed on a protanopic confusion line will be used. Second, chromaticities of initial test stimuli will be taken at random from an interval around D65. In this sense, Experiment 2 serves as a test of protanopic color constancy under 'confusion conditions'. According to the reduction hypothesis, color constancy of protanopic observers should break down in this case.

4.4.2 Methods

Stimuli and procedure

The visual display in Experiment 2 was identical to that of Experiment 1. The visual display was already described in the general methods section. In Experiment 2, simulations of a red and a green illuminant with chromaticities on a protanopic confusion line were used (see also Figure 4.6 and Table B.1, Appendix B). The missing fundamentals of the protanopic subjects agreed well with the corresponding axis of the standard protanope (Table C.1, Appendix C). Hence, confusion lines of the standard observer were used. Again, the stimuli presented to the subjects were simulated from the surfaces' reflectance functions and the spectral power distributions of illuminants as described in Section 3.1.1. In contrast to Experiment 1, initial chromaticities of the test patch were chosen randomly from an interval around $D65$.

Analogous to Experiment 1, test stimuli at ten different luminance levels were used. Half of the test patches were increments relative to the mean color signal and the other half were decrements. The task of the subject was always to make achromatic settings for the test stimulus. All subjects made their settings always

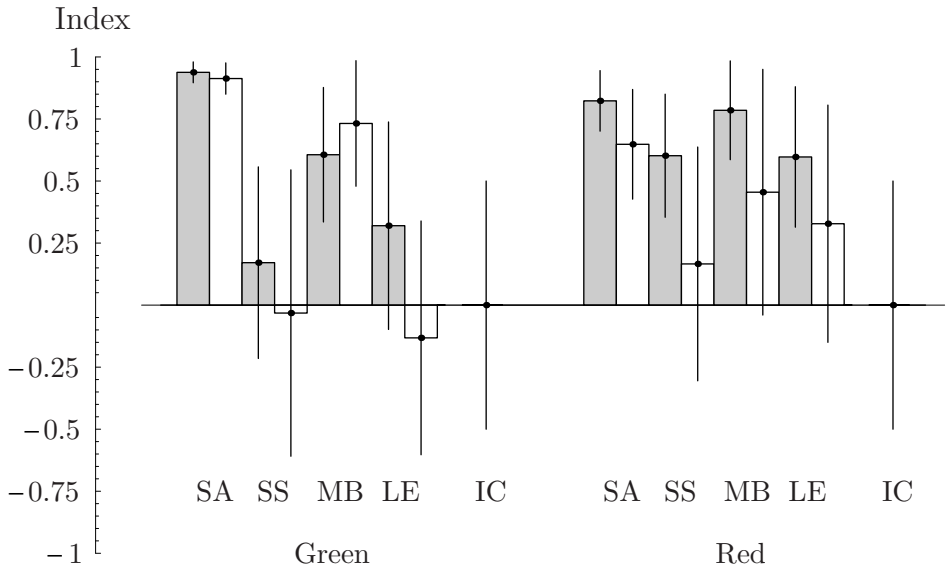

Figure 4.23: Color constancy indices of subjects under green and red illumination aligned with a protanopic confusion line (Exp. 2) in comparison with corresponding indices from Experiment 1. Gray bars refer to data from the original illuminant conditions and white bars refer to indices from the confusion conditions. Mean indices with error bars are shown separately for the normal observer SA and the protanopic observers SS, MB and LE. The zero-indices labeled 'IC' represent data of a virtual subject who always accepts the test stimulus with initial chromaticity.

on one line that was defined as the line from $D65$ to the test illuminant. A detailed description of the procedure was already given in Section 4.2.5.

Observers

The three protanopic observers LE, MB and SS and the normal observer SA took part in Experiment 2. These subjects had already made settings in Experiment 1.

4.4.3 Results

Color constancy

Color constancy performance of subjects in Experiment 2 in comparison to Experiment 1 is shown in Figure 4.23. The results suggest that color constancy of protanopes SS and LE breaks down when the green illuminant is aligned with the corresponding protanopic confusion line. The data of these observers under red illumination and the data of protanope MB under both illuminants are more ambiguous. Settings of subject MB under green illumination indicate systematic

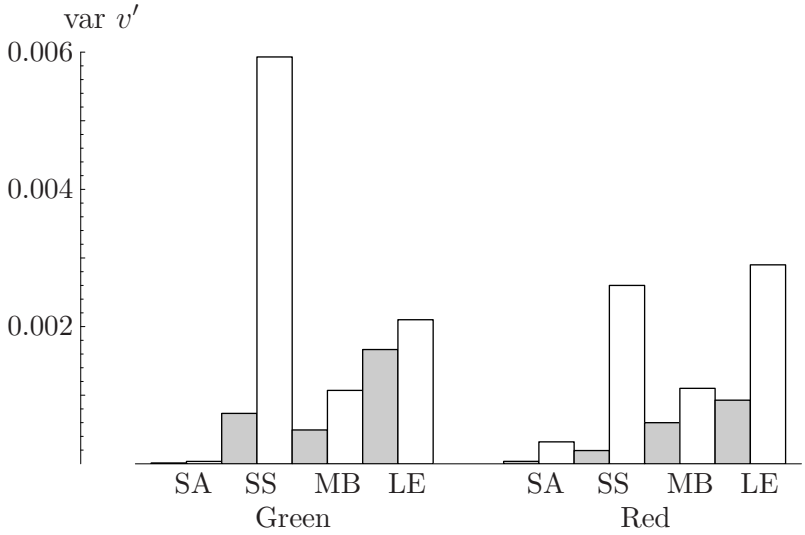

Figure 4.24: Total v' error variances of subjects under green and red illuminations that are aligned with a protanopic confusion line (Exp. 2) in comparison with corresponding variances from Experiment 1. Gray bars refer to data from Experiment 1 and white bars refer to error variances in Experiment 2. Total v' error variances are shown separately for the normal observer SA and the protanopic observers SS, MB and LE.

compensation for the ambient light and they are quite reliable. Under reddish illumination, adjustments of subjects slightly deviate from what would be expected from a perfect protanopic observer.

All subjects show the tendency to make their settings apart from the initial test stimulus toward the chromaticity of the illuminant. However, as variation in the data is large, interpretation of this finding should be taken with care. As expected, color constancy performance of the normal observer SA was comparable to respective conditions from Experiment 1. Constancy indices of all observers are also given in Table C.3 (Appendix C).

Error variation

Error variation in the data of all protanopes in the present experiment is larger than in Experiment 1 (Figure 4.24). Their settings almost span across the whole gamut of realizable colors along the corresponding confusion line. Consistent with this observation, protanopic subjects reported to perceive no chromatic differences when making their adjustments.

4.5 Experiment 3

4.5.1 Introduction

In Experiment 3, the strength of increment-decrement asymmetries under several simulated daylights along the daylight locus were analyzed systematically. A reanalysis of the data from Bäuml (2001) and my own data from Experiment 1 (see Figures 4.15 and 4.18) suggested that strong increment-decrement asymmetries occur only for the S-cone signal but not for the L- and M-cone signals. This finding is in agreement with results from studies using simple center-surround configurations (Chichilnisky & Wandell, 1996).

A comparison of the daylight locus, the BY-opponent axis and the S-cone axis in xy-chromaticity diagram shows that differences between daylights of identical intensities are mainly encoded by the S-cone photoreceptor and further on by the blue-yellow opponent channel (Figure 4.25). Therefore, the strength of increment-decrement asymmetries might also depend on the color of the daylight illuminant. This is what was found in the reanalysis of Bäumls (2001) data and in my own data from Experiment 1. Under yellowish illumination strong increment-decrement asymmetries occurred whereas asymmetries almost disappeared under bluish illuminants.

The main purpose of the present experiment is to investigate this relationship more systematically. If we find that the correlated color temperature of daylights and the strength of increment-decrement asymmetries are related in linear manner, this measure might possibly be used by the visual system as an additional cue to the color of the current illuminant.

4.5.2 Methods

Stimuli and procedure

The visual display in Experiment 3 was identical to that of Experiment 1 which was described in the general methods section. In Experiment 3, simulations of ten different daylights were used (D40, D45, D50, D55, D60, D65, D70, D80, D90, D10; see also Figure 4.6 and Table B.1, Appendix B). Again, the stimuli presented to the subjects were simulated from the surfaces' reflectance functions and the spectral power distributions of illuminants as described in Section 3.1.1.

Analogous to Experiment 1, test stimuli at ten different luminance levels were used. Half of the test patches were increments relative to the mean color signal and the other half were decrements. The task of the subject was always to make achromatic settings for the test stimulus. Unlike Experiment 1, all normal subjects except one made their settings on two lines which can be roughly characterized as opponent blue-yellow and red-green axes. Dichromats and normal subject SA made their settings always on one line that was defined as the line

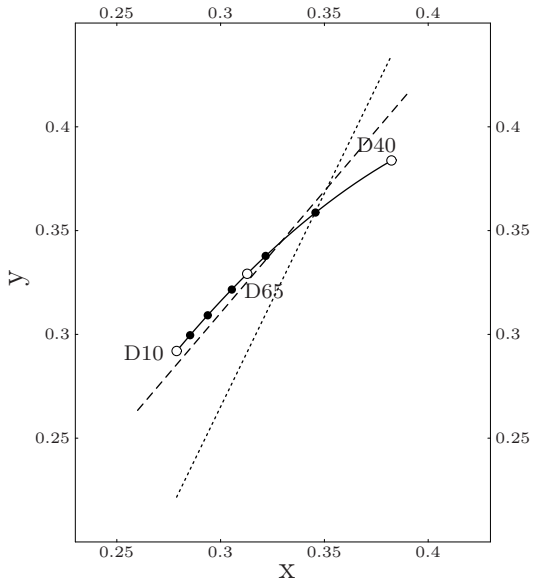

Figure 4.25: Comparison of the CIE-daylight locus (solid curve), the BY-opponent axis (dashed line) and the S-cone axis (dotted line) through the equal energy point in xy-chromaticity diagram.

from $D65$ to the test illuminant. A detailed description of the procedure was already given in Section 4.2.5.

Observers

Eight subjects took part in Experiment 3. All subjects were students who were not aware of the purpose of the experiment, except subject SS, the author. First, all subjects were tested with Ishihara diagnostic plates (Ishihara, 1997). This screening showed that subjects KH, KUH, LA, RL and SA had normal color vision. Subjects MB, LE and SS who took already part in Experiment 1 were classified as protanopes.

4.5.3 Results

Color constancy

Figure 4.26 shows mean color constancy indices of subjects for each condition. In general, the amount of adjustment to the illuminant of color deficient and normal observers was comparable. However, visual inspection of Figure 4.26 may lead to the conclusion that color constancy of subjects was poor for illuminations close to D65. As we will now see, under these illuminant conditions small indices are rather an artifact of the color constancy measure defined in Equation (4.39).

The size of this index depends on the subjects settings under test illuminant in relation to the illuminant shift. In case of small illuminant shifts and taking into account typical variation in the subjects' data, mean constancy indices are systematically biased. Figure 4.27 shows results from a simulation that clearly demonstrate this distortion.

Let us assume that color constancy indices of observers under a given test illuminant are normally distributed with expected constancy μ and σ. We can draw a sample from the population of indices with parameters μ and σ typical for the subjects in the present study. In this sense, Figure 4.27 depicts mean indices of samples each of size $n = 100,000$ drawn from populations with $\mu = 0.75$ and σ representing the mean standard deviation of normal subjects participating in Experiment 3.

The pattern of results from this simulation resembles the pattern of actual indices of subjects shown in Figure 4.26. For small illuminant shifts and with increasing σ, mean indices of subjects are more biased. Therefore, for some subjects with large variation in their data in conditions with small illuminant shifts even negative mean indices occurred.

In addition, I analyzed color constancy performance of observers separately for incremental and decremental stimuli. Like in Experiment 1, the results indicate that in general color constancy of subjects was better for decrements than for increments. Figure 4.28 shows mean constancy indices of each subject separately for increments and decrements under daylight D40.

Increment-decrement asymmetries

The main purpose of Experiment 3 was to investigate systematically the relation between correlated color temperature of daylights and the strength of S-cone increment-decrement asymmetries. Figure 4.29 shows S-cone excitations of achromatic settings for a normal and a protanopic observer under different daylight conditions. A typical pattern found in the data of all subjects is that lower correlated color temperatures of daylights are related with stronger increment-decrement asymmetries.

In order to evaluate the strength of asymmetries, I determined ratios q_c of the slopes s_c^I and s_c^D of the two separate regression lines related with incremental and decremental stimuli. Ratios q_c for all subjects and conditions are shown in Figure 4.30. In general, ratios were higher for protanopic than for normal observers. For each subject ratios q_c were predicted from Euclidean distances $\Delta u'v'$ between test illuminant and D40 in $u'v'$-chromaticity diagram. Results of these linear regressions are depicted in Figure 4.31. Linear relations between distances $\Delta u'v'$ and the strength of asymmetries in terms of ratios q_c for all subjects were negative with correlations ranging from $-.61$ to $-.94$.

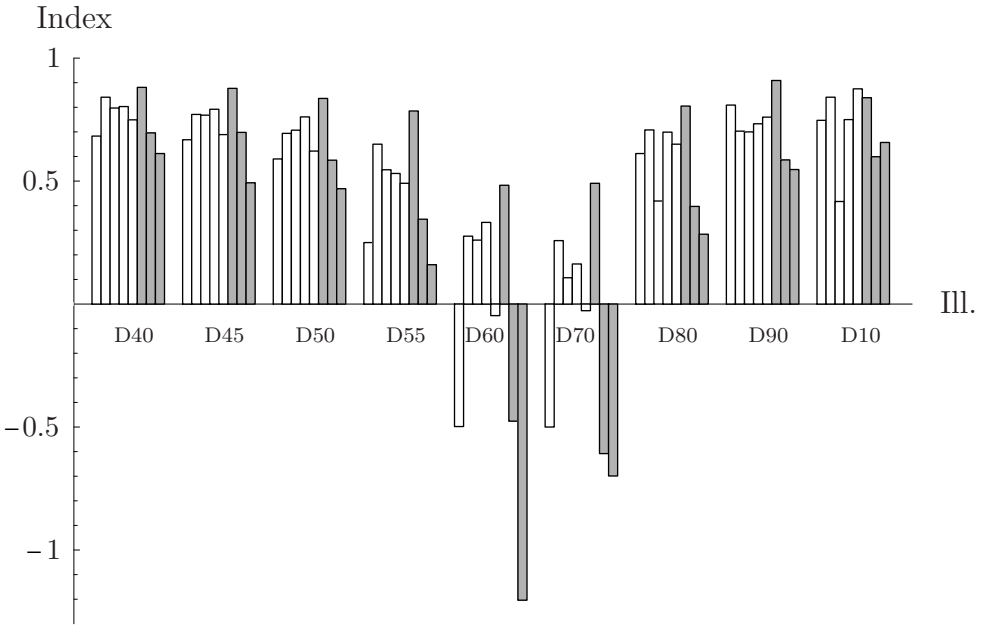

Figure 4.26: Color constancy indices of subjects in Experiment 3. The mean indices of the five normal observers KH, LA, KUH, RL and SA (white bars) and those of the three protanopic observers SS, MB and LE (gray bars) are shown for the nine different chromatic daylight conditions.

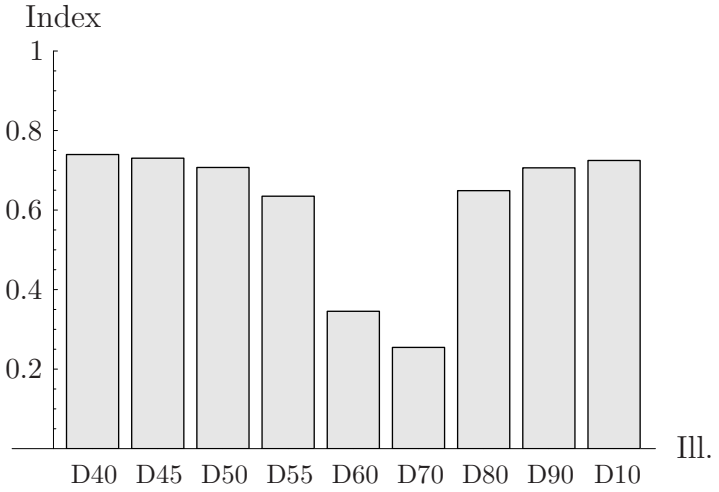

Figure 4.27: Results of a simulation that demonstrate a systematic bias of the traditional color constancy measure. The bars represent mean indices of samples each of size $n = 100,000$ drawn from populations with expected constancy $\mu = 0.75$ and σ representing the mean standard deviation of the normal subjects participating in Experiment 3 (see text for details).

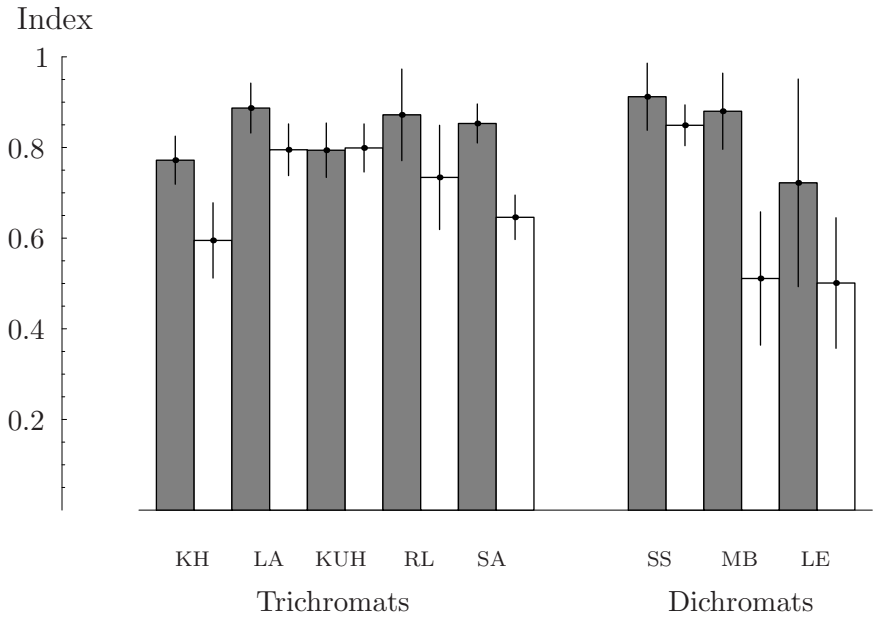

Figure 4.28: Color constancy indices of subjects under illuminant D40 separately for incremental and decremental stimuli. Gray bars refer to decrements and white bars refer to increments. Mean indices with error bars are depicted for the five normal observers KH, LA, KUH, RL, SA (labeled with 'Trichromats') and the three protanopic observers SS, MB and LE (labeled with 'Dichromats').

Model test

Parameters of the three models (ID model, diagonal model and spatial mean model) were estimated for each subject and illuminant condition. Mean estimates of S-cone rescaling coefficients across normal and protanopic observers are shown in Figure 4.32. The comparison of rescaling parameters referring to incremental and decremental stimuli shows that increment-decrement asymmetries occur and that the strength of the asymmetry weakens with increasing correlated color temperature. Furthermore, asymmetries are generally stronger for protanopes than for normal observers.

Inspection of Figure 4.32 reveals another interesting pattern in the data that was already found in Experiment 1. In cases of strong increment-decrement asymmetries, rescaling parameters for decrements are in good agreement with coefficients of the spatial mean model (see also Figure 4.29).

In addition, predictions of the three models were contrasted for the S-cone channel. According to the rationale described in Sections 4.1.2 and 4.2.6, I tested the ID model against the diagonal model. The number of significant test results decreased with increasing correlated color temperatures of daylights. Only three out of 30 tests of protanopic settings produced non-significant results which provides evidence to the aforementioned observation that increment-decrement asymmetries tend to be stronger for these color deficient observers than for nor-

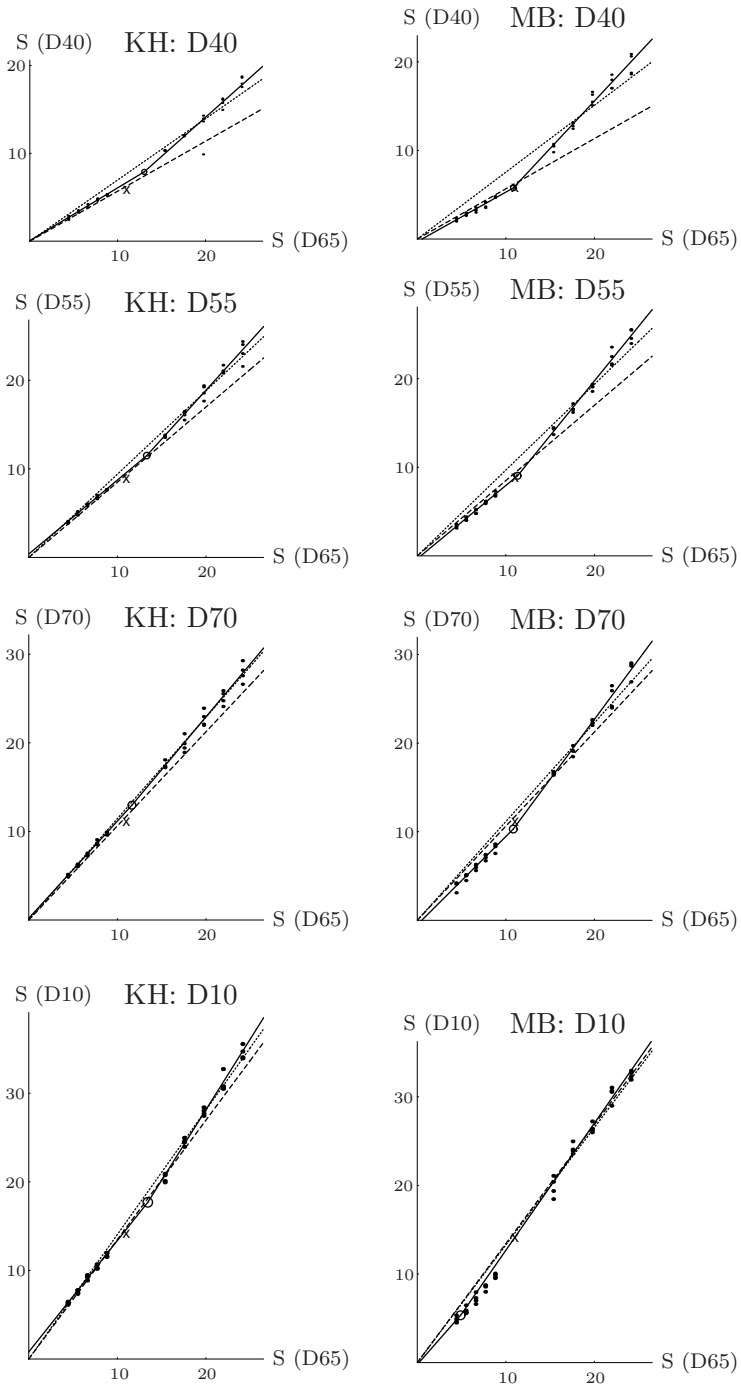

Figure 4.29: S-receptor codes of achromatic settings of the normal subject KH (left column) and the protanope MB (right column). Settings of an ideal observer under D65 are plotted against settings of the subject under test illumination. Different rows show data under test illuminant D40, D55, D70 and D10 respectively. Solid lines represent least square fits independent for incremental and decremental stimuli. Dotted lines show predictions by a diagonal model of von Kries type and dashed lines predictions by the spatial mean model. The × indicates mean cone coordinates across the scene and the open circle S-receptor codes of the estimated reference stimulus.

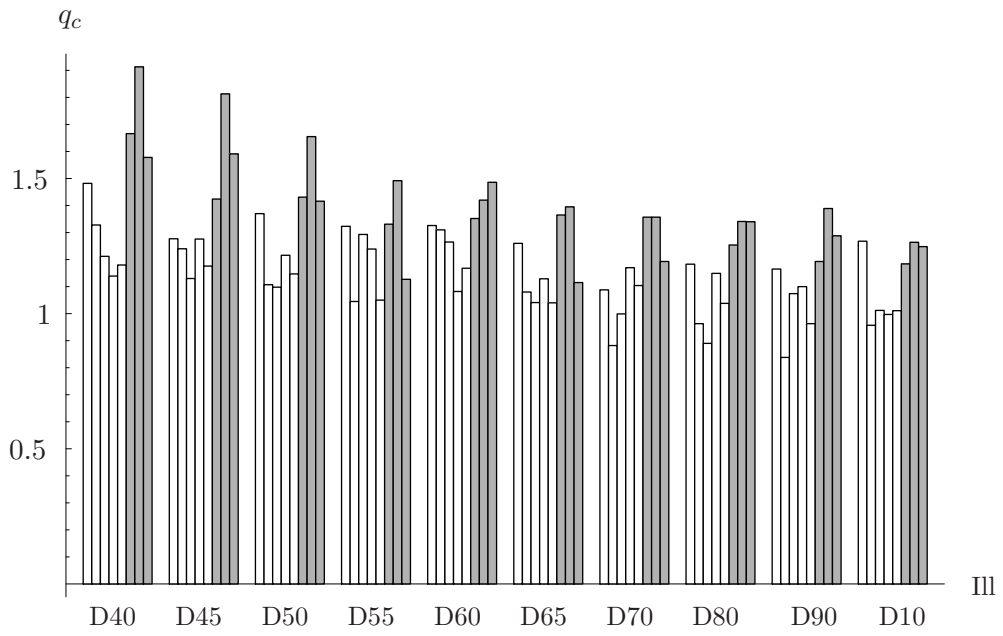

Figure 4.30: Ratios q_c of the slopes s_c^I and s_c^D of the two separate regression lines that refer to S-cone increments and decrements respectively. White bars indicate ratios of the five normal observers KH, LA, KUH, RL, SA and gray bars show ratios of the three protanopic observers SS, MB and LE.

mal trichromats. The numbers of significant results of corresponding F-tests at α-level 0.05 are shown for each condition and the two groups of observers in Figure 4.33.

The statistical tests of the diagonal model against the more restricted spatial mean model showed that predictions of an unspecified diagonal model were in most cases better than predictions by the spatial mean model. Only ten out of 160 tests yielded non-significant results ($\alpha = 0.05$) when contrasting both models. Although at first glance this result may clearly argue in favor of the diagonal model, findings from previous analyses suggest that the spatial mean model may provide a useful heuristic when dealing with decremental stimuli.

4.6 Discussion

4.6.1 Summary

We will now summarize the results of the experiments. Under daylight conditions, protanopes' adjustment to the illuminant is comparable to that of normal observers. Color constancy of two protanopic observers broke down under greenish illumination aligned with respective confusion lines. This result is in agreement with the reduction hypothesis of dichromacy. However, color constancy perfor-

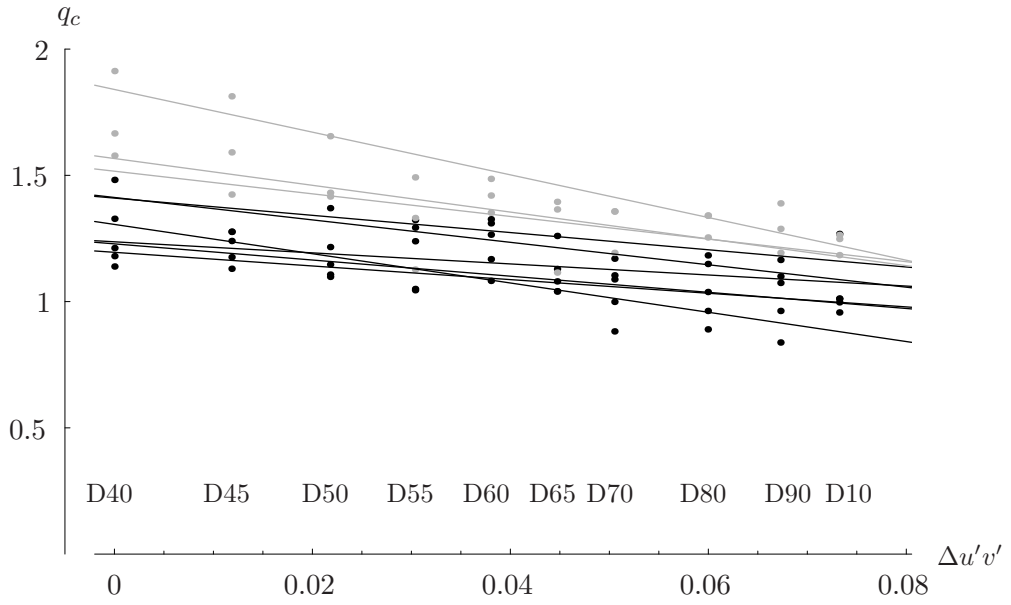

Figure 4.31: Linear regressions of distances $\Delta u'v'$ from D40 and ratios q_c plotted separately for different subjects. Black symbols indicate data and regression lines of the five normal observers KH, LA, KUH, RL, SA and gray symbols show data and regression lines of the three protanopes SS, MB and LE.

mance of all protanopic observers under corresponding reddish illumination was better than predicted by the reduction hypothesis. Furthermore, when illuminant changes were aligned with a red-green axis slightly different from a protan confusion line color constancy performance of some protanopes was remarkably good.

With respect to normal observers, adjustment to the illuminant under daylight conditions was in general not better than under red or green illuminations far off the daylight locus. This result is in agreement with findings from previous studies (Brainard, 1998; Delahunt & Brainard, 2004b).

In Experiments 1 and 3, increment-decrement asymmetries occurred. In general, observed asymmetries were stronger in the data of protanopes compared to normal subjects. Differences in the relative settings for incremental and decremental stimuli occurred essentially only for the S-photoreceptor. The systematic analysis of the relation between increment-decrement asymmetries and the correlated color temperature of daylights revealed a negative correlation between these two quantities in the sense that the more yellowish a daylight illuminant appeared to the observer the stronger the asymmetries that occurred in the S-channel. In conditions with increment-decrement asymmetries color constancy performance of subjects was better for decremental than for incremental stimuli.

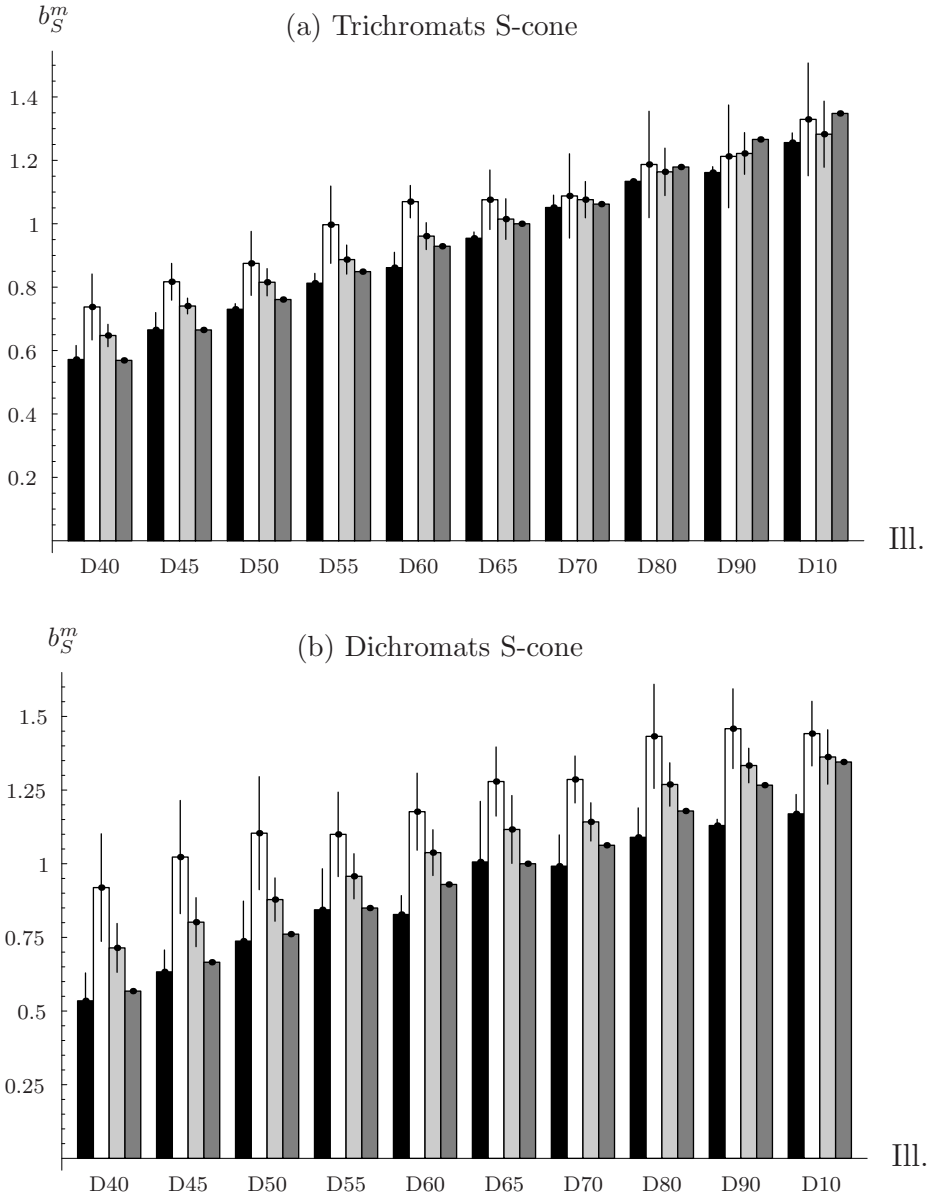

Figure 4.32: Rescaling parameters of the three models. The parameters are univariate regression weights b_i^m for predicting subject's achromatic settings under test illuminant from ideal settings under D65 which refer to model m and receptor type i. Different bar shadings refer to different models: black – ID model (decrements), white – ID model (increments), light gray – diagonal model, dark gray – spatial mean model. The two rows show mean S-cone coefficients with corresponding error bars for (a) normal and (b) protanopic observers respectively.

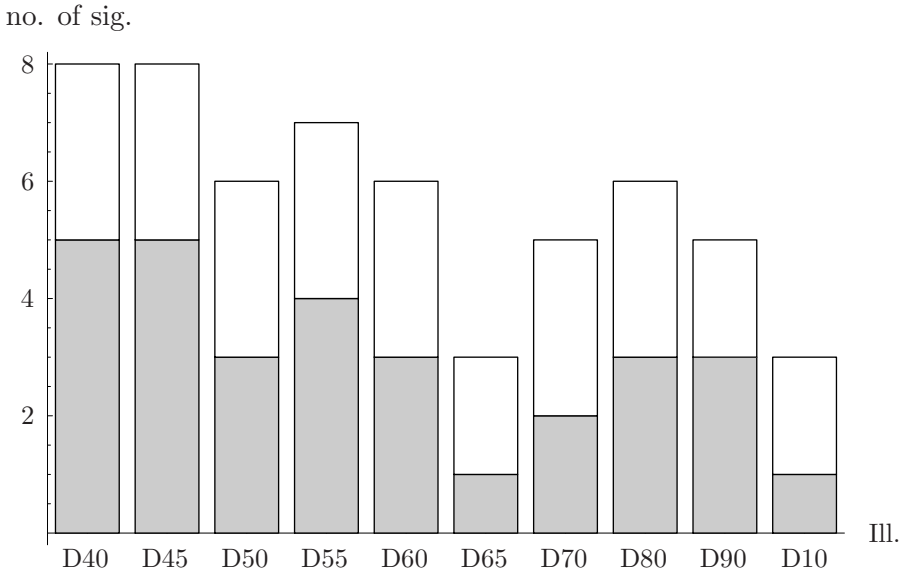

Figure 4.33: Number of significant results when testing the ID model against the diagonal model (F-tests with $F_{3,36,.05}$). Gray bars indicate the number of significant results for normal observers and white bars the number of significant results for protanopic observers. Eight tests were conducted for each illuminant condition.

4.6.2 The Reduction Hypothesis

In Experiment 1, I found that daylight adjustment of protanopic observers can be as good as that of normal observers. This result is not contradictory to the reduction hypothesis and was reported in previous studies (Foster & Linnell, 1995; Rüttiger, Mayser, Sérey & Sharpe, 2001). Protanopes and deuteranopes possess an intact blue-yellow opponent system which mainly encodes differences across daylight changes (see Figure 4.25). This residual opponent channel enables them to adjust to daylight illumination almost as good as normal observers.

The crucial question regarding the reduction hypothesis is of whether these dichromats are able to adjust to red or green illumination at least partially as suggested by Rüttiger et al. (2001) or whether color constancy of red-green deficient observers breaks down under these illuminant conditions.

At first glance, results from Experiment 1 indicate that protanopic observers can compensate for illuminant shifts which do not perfectly correspond with confusion lines. Results from Experiment 2 are more ambiguous. Color constancy of two protanopic observers broke down under greenish illumination aligned with a confusion line. This is what one would have expected if the reduction hypothesis is assumed to be true. However, the third protanope, MB, showed good adjustment to the green illuminant that was even comparable to normal observers. Under corresponding reddish illuminant all protanopes were able to at least partially adjust to the ambient light.

One protanopic subject, MB, showed remarkably good performance under all

illumination conditions. He was diagnosed as protanope based on the analysis of his confusion lines which approximately coincide with confusion lines of the standard protanope (see Appendix C, Table C.1). A genetic analysis of the protanopic observers was beyond the scope of this study. However, results from studies on the genetic basis of dichromatic and anomalous trichromatic vision indicate that for both classes of observers, a wide range of different genotypes exists which may lead to differences in the matching behavior (Neitz, Neitz & Jacobs, 1995; Sharpe, Stockman, Jägle, Knau, Klausen, Reitner & Nathans, 1999; Neitz, Neitz, He & Shevell, 1999).

Taken together these findings and the large error variation in the data indicate that it is difficult to interpret the results in the light of the reduction hypothesis. It seems as if color constancy performance of dichromats can be better than predicted from the classical view.

The partial adjustment of protanopes in Experiment 2 could be due to different sources. Observers may have used an internal standard based on lightness cues which may be more efficient under reddish illumination. However, observer MB showed better adjustment to greenish than to reddish illumination which speaks against this assumption. A second explanation has to do with the visual display that was used in the present experiments. The stimulus always extended a visual angle of $15^\circ \times 11^\circ$. Findings from a number of studies suggest that protanopes and deuteranopes show trichromatic behavior when large field stimuli are used (Smith & Pokorny, 1977; Montag & Boynton, 1987; Crognale, Teller, Yamaguchi, Motulsky & Deeb, 1999). Finally, the alignment of illuminations' chromaticities with confusion lines does not necessarily imply that light signals from one surface under these illuminants are also aligned with a confusion line. However, as Figure 4.1 indicates deviations from this assumption were small.

Two further explanations may account for the good adjustment of protanopic observers in Experiment 1 and for the results of Rüttiger et al. (2001). First, in both cases chromaticities of initial test stimuli were chosen toward the color of the illuminant which is identical with the point of perfect color constancy (see also Rüttiger, Braun, Gegenfurtner, Petersen, Schönle & Sharpe, 1999). This allows an observer to make partially color constant settings by accepting these initial values without being color constant. Second, the small deviation of the illuminants' chromaticity from the corresponding confusion line may have been sufficient for these dichromats to at least partially adjust to the ambient illumination.

4.6.3 The Daylight Hypothesis

Results from Experiment 1 suggest that for normal observers adjustment to daylights is not necessarily better than adjustment to illuminants off the daylight locus. This finding is in agreement with results from previous studies (Brainard, 1998; Rüttiger et al., 2001; Delahunt & Brainard, 2004b).

Furthermore, as reported by Delahunt and Brainard (2004b) normal observers showed in general higher color constancy for green illumination than for any other illuminant. One possible explanation for this surprising result could be that our visual system has evolved in an environment which provided mainly greenish illumination. Leaves of trees in forested areas typically produce indirect greenish light due to interreflections (Endler, 1993). However, this hypothesis is to be tested more specifically in future research.

4.6.4 Increment-Decrement Asymmetries

In Experiments 1 and 3 I found strong evidence for the assumption that the visual system processes incremental and decremental stimuli differently. This result is in agreement with findings from previous studies using different stimulus configurations and paradigms (Walraven, 1977; Chichilnisky & Wandell, 1996; Heyer, 1997; Bäuml, 2001). Therefore, the segregation of incoming light signals into increments and decrements seems to be a fundamental principle of our visual system.

In addition, analysis of the experimental data revealed two interesting features of the distinction between increments and decrements. First, strong increment-decrement asymmetries occurred essentially only for the S-cone channel but not for L- and M-cone systems. Similar patterns of results have been found in studies on simultaneous contrast effects (Chichilnisky & Wandell, 1996). These authors suggest that postreceptoral processing of S-cone signals in S-ON and S-OFF ganglion cells may be responsible for the observed increment-decrement differences. Second, the strength of increment-decrement asymmetries varied systematically with the correlated color temperature of daylights and did not occur for non-daylight illuminants. Roughly speaking, the more yellowish the illuminant the stronger the asymmetries that occurred. Both observations may be related with each other in the sense that the contribution of the S-cone system to chromatic differences of daylights is larger compared to the other two cone photoreceptor channels.

Subjects showed better adjustment to the illuminant for decremental than for incremental stimuli in conditions where strong increment-decrement asymmetries occurred. This result has been reported in previous studies (Walraven, Benzschawel & Rogowitz, 1989; Chichilnisky & Wandell, 1996; Bäuml, 2001). It provides support for the notion that the increment-decrement distinction is related with a mechanism which mediates the segregation of self-luminous light sources and reflecting surfaces (Bäuml, 1997; Mausfeld, 1998).

In general, stronger increment-decrement asymmetries were observed in the data of protanopic subjects compared with normal observers. A similar pattern of results was found in a previous study on simultaneous contrast effects (Richter, 2002). We noted earlier that asymmetries occurred essentially only for the S-cone system. Therefore, stronger asymmetries in protanopes may be related with the

fact that primarily the S-cone channel contributes to the chromatic pathways in protanopic vision. As protanopes are lacking the L-cone system, S-photoreceptor signals are only combined with input from M-cones at postreceptoral levels. In this sense, the relative importance of S-cone signals is enhanced in protanopic vision which presumably leads to stronger increment-decrement asymmetries.

4.6.5 Models of Color Constancy

In the previous section we discussed implications of the distinction between increments and decrements in some detail. The comparison of color constancy models revealed that under certain conditions a model which takes increment-decrement asymmetries into account may provide a better fit to the data than a diagonal model of von Kries type. The results indicate that the ID model is superior over simple von Kries models when S-cone excitations under neutral to yellowish illuminations are considered. In all other cases a diagonal model of von Kries type provided an adequate description of the data. However, this finding does not necessarily mean that different processing of increments and decrements asymmetries occurs only in the S-cone system. Rather, due to the experimental design the range of L- and M-cone excitations was limited and may have prevented me from detecting further cone specific asymmetries.

The results indicate that predictions by a spatial mean model are better for decremental than for incremental stimuli when substantial asymmetries occur. In this sense, the spatial mean across the scene may provide a reliable cue to illuminant invariant surface perception. This result is in agreement with findings from previous studies where it was shown that observers in fact use the spatial mean cue in order to judge the illuminant (Kraft & Brainard, 1999; Linnell & Foster, 2002). However, as demonstrated by Kraft and Brainard (1999) the spatial mean is not used as an exclusive cue to the illuminant by our visual system. Further potential cues to the illuminant such as the brightest surface in the scene, the local contrast, second order statistics or specular highlights may play a role for the estimation of the illuminant (Lee, 1986; Kraft & Brainard, 1999; Maloney, 2002; Hurlbert & Wolf, 2004). We discussed this issue earlier in Chapter 3 in some more detail. In addition, the finding that the spatial mean represents a valuable cue to the illuminant rather for decrements than for increments may account for the small effects reported in some studies using test stimuli that were incremental relative to the surround (e.g. Bäuml, 1994).

I did not investigate the question if adjustment to the illuminant is independent within each photoreceptor channel. Results from previous studies on this issue are ambiguous. Some authors report evidence for the principle of independent photoreceptor rescaling (Brainard & Wandell, 1992; Bäuml, 1995; 1999a) whereas others found violations of this principle, particularly for the S-cone (Chichilnisky & Wandell, 1996; Delahunt & Brainard, 2000). Again, the ambiguity of results could be related with the distinction between increments

and decrements. Chichilnisky and Wandell (1996) report that signals from different cone classes interact when stimuli are increments but not when stimuli are decrements.

4.6.6 Limitations of the Present Study

In this study aspects of trichromatic and dichromatic color constancy were examined using computer rendered flat Mondrian stimuli. The question arises whether the present results are generalizable to a broader range of viewing conditions. We will now discuss limitations of this study and consequences for the generalizability of results in detail.

The stimuli used throughout this study were simulations of matte surfaces presented on CRT-screens. We already discussed aspects of the use of CRT-simulations in Section 4.2.2. In general, the range of stimuli realizable on computer monitors is limited and covers only a small region of natural viewing conditions. However, main principles that are assumed to underlie color constancy have been revealed both in studies using CRT-simulations (Valberg & Lange-Malecki, 1990; Brainard & Wandell, 1992) and in studies under nearly natural viewing conditions (Brainard, 1998; Kraft & Brainard, 1999).

Another problem of CRT-simulations is related with the fact that each simulated stimulus on a CRT-screen emits radiation. Therefore, simulated surfaces may be interpreted by the visual system as self-luminous stimuli. Although I cannot definitely rule out this possibility, the results suggest that observers at least clearly distinguish between incremental and decremental stimuli which may trigger light and surface perception respectively (Mausfeld, 1998).

In the present study color constancy was investigated under viewing conditions that are often referred to as the Flat World (Maloney, 1999). Tests of color constancy models have been mainly conducted in Flat World environments (McCann, McKee & Taylor, 1976; Valberg & Lange-Malecki, 1990; Brainard & Wandell, 1992). Relevant mechanisms of color constancy have also been proven under natural viewing conditions (Kraft & Brainard, 1999). Therefore, it would not be plausible to assume that the main results of the present study, namely the characteristics of protanopic color constancy and increment-decrement asymmetries, are only an artifact of the two dimensional visual display. However, investigation of these questions in more realistic environments that take three dimensional aspects of vision into account would be desirable. We will extend our perspective to this Shape World environment in the next chapter.

In the experiments, subjects made achromatic settings for the test patch. Hence, one could argue that the results of the present study are not applicable to chromatic stimuli. Speigle and Brainard (1999) reported however, that it is possible to predict asymmetric matches from observers' achromatic settings. Further support for this assumption comes from the data that will be presented in the next chapter. Due to the achromatic setting task the present study was lim-

ited to pure incremental or decremental stimuli. The use of a different paradigm would allow us to analyze mixed stimuli that are for example increments in two photoreceptor channels but decrement in the other.

Only a small number of protanopic observers took part in the present study. The classification of dichromats was only based on the determination of confusion lines but not on genetic analysis. Hence, it is difficult to generalize the results to the class of protanopic observers. However, results from different lines of research question if protanopes constitute a homogeneous class of observers. Studies on color naming abilities of dichromats (Montag, 1994) and on hue scalings of dichromats (Scheibner & Boynton, 1968) revealed large differences between observers of one and the same dichromatic class. In addition, genetic analyses have shown a noticeable diversity of genotypes underlying each type of dichromacy (Sharpe et al., 1999; Neitz et al., 1999).

Chapter 5

Experiments II: Color constancy and Hue Scaling

In the previous chapter color constancy was investigated assuming a simplistic two-dimensional model of the environment which we referred to as the *Mondrian World*. The series of experiments presented here serves two main purposes. First, we wish to study color constancy in an enriched three-dimensional environment that resembles natural scenes. Second, hue scaling will be introduced as an alternative method to measure color constancy. This method promises to represent aspects of human color constancy performance more adequately than traditional measures (Foster, 2003).

5.1 Introduction

5.1.1 Motivation

The term *color constancy* is used to describe the phenomenon of stable color appearance of surfaces with changing illumination (Jameson & Hurvich, 1989; Kaiser & Boynton, 1996). This conception of color constancy refers directly to human perceptual experience. However, quantitative measurements of color constancy under various experimental conditions have shown that human color constancy is far from perfect (Arend & Reeves, 1986; Kuriki & Uchikawa, 1996; Kraft & Brainard, 1999). Several explanations have been suggested that may account for this imperfect performance, including effects of instruction (Arend & Reeves, 1986), inappropriateness of the stimuli used (Kraft & Brainard, 1999), increment-decrement asymmetries (Bäuml, 2001) and the size of the illuminant shift (Craven & Foster, 1992).

As pointed out by Foster (2003), it is questionable if frequently used quantitative measures are appropriate instruments to determine the degree of color constancy experienced by human observers. Therefore, he claims that alternative

methods are needed to describe the characteristics of human color constancy more adequately. In this sense, the present study is intended as a first step toward an understanding of color constancy that is more directly related with the demands on observers in natural environments.

5.1.2 Measuring Color Constancy

Psychophysical methods such as achromatic settings or asymmetric matching promise accurate, quantitative assessment of color constancy. It is not clear how well such tasks characterize the stability of observer's color appearance across the full range of surface colors. However, while spatial or temporal asymmetric matching can be used to characterize perception of colors away from the neutral point, recent results suggest that it might be inappropriate to evaluate human color constancy for a variety of reasons (Maloney, 1999; Foster, 2003; Logvinenko & Maloney, 2006). In particular, when observers are asked to set a test surface under a given illumination so that it appears perceptually indistinguishable to a reference patch under a second illumination, they sometimes report that they can not find a satisfying setting. This problem was first noted by David Katz, who reported that, when observers make a match in a lightness or color constancy experiment, there is usually a residual difference (Katz, 1911, p. 82). The following comment from a recent asymmetric color matching study may illustrate this problem:

> At this match point, however, the test and the match surfaces looked different, and the observers felt as if further adjustments of the match surface should produce a better correspondence. Yet turning any of the knobs or combinations of knobs only increased the perceptual difference (Brainard, Brunt & Speigle, 1997, p. 2098).

If asymmetric matches do not, in fact, match, then it is questionable whether they can be used to characterize color constancy.

Achromatic setting measures, for example, essentially measure only the location of the observer's neutral point and their use to characterize color appearance away from the neutral point involves assumptions. Speigle and Brainard (1999) report, for example, that achromatic settings can be used to predict asymmetric matches, but, as just noted, it is unclear whether the latter can be used to measure stability of color appearance under changes in illumination.

A task that is more directly related to the purpose of color constancy in terms of object recognition is color categorization (Boynton & Olson, 1987). The grouping of colors into a small number of discrete categories seems to be a universal feature of the visual system (Kay & Regier, 2003). Jameson and Hurvich (1989) argue that color categories of objects are preserved with changes of the illuminant, if one takes compensatory mechanisms like chromatic adaptation into account.

Troost and deWeert (1991) investigated the color constancy performance of observers in a color categorization task and asymmetric matchings. They found that observers used reliably the same color category for a test patch with changes of the illumination. Although color categorization seems to be a reliable and appropriate measure of human color constancy, a disadvantage of this task is the lacking comparability of the color categorization data and quantitative measures and the potential that the conclusions drawn depend on an arbitrary choice of categories.

In the experiments reported here, hue scalings will be introduced as an alternative method to investigate color constancy (Boynton & Gordon, 1965; Abramov & Gordon, 1994). On one hand, this task is evidently based on judgments of the appearance of chromatic surfaces under varying illumination and is therefore appropriate to study human color constancy. On the other hand, the resulting data is potentially comparable to quantitative measures of color constancy. Therefore, hue scalings of observers will also be compared with corresponding achromatic settings. Moreover, the usage of a hue scaling task promises to eliminate disadvantages that are inherent in asymmetric matchings and achromatic settings.

5.1.3 The Speigle-Brainard conjecture

Another purpose of the present study is to test a generalization of the model proposed by Speigle and Brainard (1999) to link achromatic settings and asymmetric matching. Speigle and Brainard asked subjects to make achromatic settings and asymmetric matches under different illuminations and then sought to predict the latter from the former. The question they address can be formulated as follows: Suppose that the experiment has measured the shift in color space of the achromatic point induced by a change from one illumination to a second. Can the experimenter now predict the pattern of shift of any surface (not just an achromatic surface) induced by the same change in illumination? Speigle and Brainard conclude that, although the transformations on all points in color space as measured by asymmetric matching are potentially too complex to be predicted by knowledge of the transformation of the achromatic point, in reality they can be accurately predicted. In the context of hue scaling it will be tested whether knowledge of the hue scalings of a single test patch under two lights can be used to predict scaling of all test patches under these two lights.

5.1.4 Models of the environment

Previous studies of color constancy have employed stimuli consisting of patterns of flat surfaces embedded in a fronto-parallel plane, often called *Mondrians* (Land & McCann, 1971). Although main principles assumed to underlie color constancy have been identified presupposing this model of environment (McCann, McKee & Taylor, 1976; Arend & Reeves, 1986; Brainard & Wandell, 1992),

three-dimensional aspects of vision are largely ignored. A further disadvantage of this configuration is that more complex light-surface interactions are not taken into account and reported indices of color constancy are typically low. In the experiments reported here simulated three-dimensional stimuli were presented binocularly to more closely approximate natural viewing conditions (Maloney, 1999; Boyaci, Maloney & Hersh, 2003). These scenes are viewed binocularly and contain additional cues to the illuminant such as specular highlights or shadows. Previous studies showed the relevance of these cues for color constancy performance of observers (Yang & Maloney, 2001; Boyaci, Doerschner & Maloney, 2006). The usage of these enriched stimuli was also motivated by the idea that the visual system seems to combine cues that are available in a scene to estimate the illuminant (Maloney, 2002).

5.1.5 The Role of Daylights

It has been hypothesized that the visual system might use regularities of natural daylights, like the location of the corresponding chromaticities along the so called daylight locus, to achieve color constancy (Judd, MacAdam & Wyszecki, 1964; Shepard, 1994). Recent studies have found no evidence to support this assumption (Brainard, 1998; Delahunt & Brainard, 2004b). According with these findings and the results from Experiment 1 reported in Chapter 4 it is not expected that illuminant adjustment is better for daylights compared with illuminations off the daylight locus.

5.1.6 The Role of Chromatic Adaptation

In addition to the cues presented in a scene, chromatic adaptation plays a crucial role in the adjustment of the visual system to the illuminant (Kuriki & Uchikawa, 1996). Studies of the time course of chromatic adaptation have revealed that this process consists of a fast and a slow phase of adaptation (Fairchild & Reniff, 1995; Rinner & Gegenfurtner, 2000). The slow phase of adaptation might be related to slow changes of daylights that occur in the daytime. In the second of the experiments reported in this chapter I was able to demonstrate, by accident, the strength of the isolated slow adaptation mechanism.

5.1.7 Hypotheses

From the discussion above we can derive the following hypotheses:

1. Color constancy performance of observers measured with hue scalings and achromatic settings should be comparable.

2. Observers' adjustment to natural daylights should not be better than to red and green illumination.

3. The appearance of a virtual achromatic surface under test illumination can be predicted by hue scalings of observers for a set of chromatic surfaces.

4. Color constancy of observers should break down when no spatial cues are given.

5.2 General Methods

5.2.1 Stimuli

Visual display

The stimuli in the experiments were computer rendered 3D-scenes that consisted of simple objects (such as columns and spheres) with different colors and reflectance properties (e.g. shiny, matte). These objects served as possible cues to the chromaticity and intensity of the light sources. A matte test patch was presented at the center of the scene. In Experiments 2 and 3 the scene consisted only of the isolated test patch in front of a black homogeneous background.

The scenes were illuminated by a mixture of a simulated diffuse and a simulated punctate source. In order to present the 3D-scenes stereoscopically, for each scene two images were rendered from two different viewpoints corresponding to the locations of the observer's eyes in the virtual scene (Color Plate D.3, Appendix D).[1] The scenes differed in the chromaticity of the test patch and in the color of the punctate source.

Test patch

In Experiments 1–3 a set of 16 different test patches was used that represented simulations of *Munsell* chips (Table B.4, Appendix B). All test patches had identical *Munsell* value and chroma of seven and four respectively. They differed only in hue and formed a color circle in the CIE-$u'v'$ diagram (Figure 5.1).

To increase the accuracy of stimuli, the test surfaces were rendered separately from the rest of the scene. Rendering packages such as Radiance simulate light-surface interactions as product of the respective RGB-codes. This *RGB heuristic* (Maloney, 1999) does not always lead to adequate simulation of light-surface interaction. Therefore, the light signal that reached the eye from the test patch was computed according to the Lambertian model as described in Section 3.3. Recall that the spectral power distribution of the light signal from a Lambertian

[1] An interocular distance of 6.3 cm was used as a standard in rendering. This separation has proven to be sufficient for subjects in these and previous experiments using the same apparatus. The normal human range of interocular distances is 6.0 to 7.0 cm (French, 1921). It was checked whether subjects could achieve fusion and, had any subject reported difficulty with stereo fusion, they would have been excluded from further participation in the experiment.

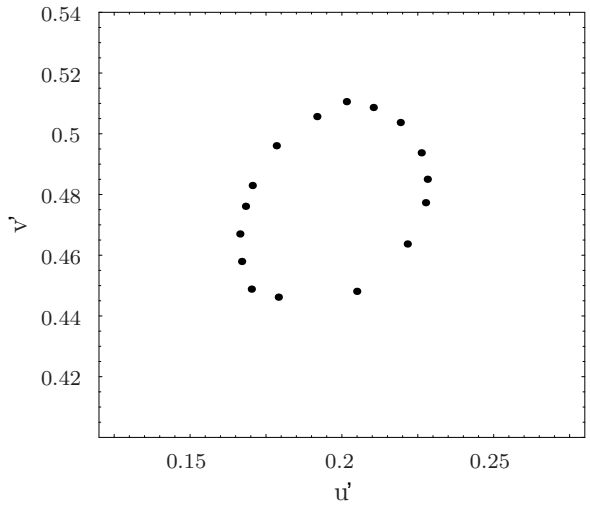

Figure 5.1: Chromaticities of the 16 Munsell surfaces under neutral illumination in u'v'. Under chromatic illumination the set of Munsell surfaces is shifted toward the illuminant.

surface that is illuminated by a diffuse source and a punctate source is given by:

$$L(\lambda) = S(\lambda) \cdot (E^P(\lambda) \cdot \cos\theta + E^D(\lambda)), \tag{5.1}$$

where λ indexes the wavelength of light in the visible spectrum and θ is the angle between the incident light from the punctate source and the surface normal. In our case θ was always set to zero. The reflectance functions of the surfaces are based on spectral reflectance measurements made with a spectrophotometer on 1,269 color chips from the 1976 *Munsell Book of Color* at a 1-nm resolution from 380 to 800 nm. This data set was obtained from `http://spectral.joensuu.fi/`.

The size of the test patch was 2.9×2.2cm which is equivalent to a visual angle of $2°22' \times 1°48'$. The distance between the test surface and the observer's viewpoint was 70 cm. The surface normal of the test patch was identical to the viewing axis and to the direction of the incident light from the punctate source. The orientation of the test patch remained constant throughout all experiments.

Light sources

As noted above, simulations of a diffuse and a punctate light source were used. The diffuse source was always a neutral daylight ($D65$). The chromaticity of the punctate source was manipulated in the experiment. Five different illuminations were used. Three of them were simulations of CIE daylights with correlated color temperatures of $6500K(D65)$, $4000K(D40)$ and $10000K(D10)$ that appeared achromatic, yellow and blue to the observer respectively.

Additionally, two illuminants off the daylight locus were used that appeared red and green to the observer (Figure 5.2). The xy- and $u'v'$-chromaticities of

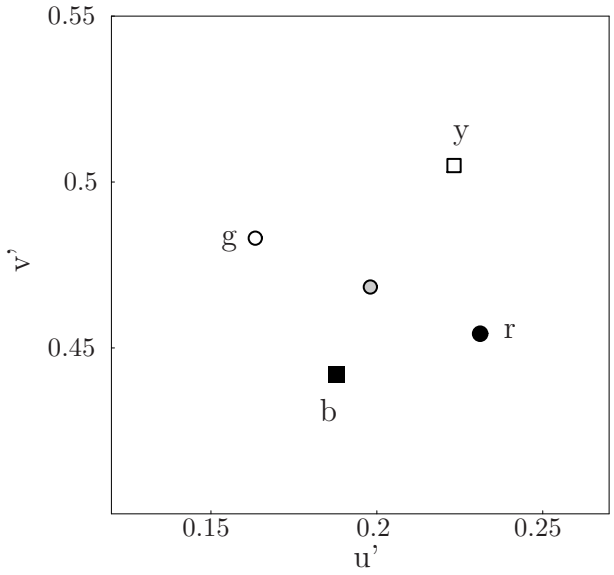

Figure 5.2: Chromaticities of the punctate illuminants in u'v'. Different symbols refer to different illuminations: filled square - blue (b), open square - yellow (y), filled circle - red (r), open circle - green (g), gray filled circle - D65).

the punctate sources are given in Table B.3 (Appendix B). For the five punctate sources spectral power distributions were computed from the CIE daylight basis functions (Wyszecki & Stiles, 1982). The calculation of the light signal that reached the eye of the observer from the test patch based on these distributions and the reflectance functions of the Munsell surfaces. The punctate source was situated behind the observer. The distance between the test patch and the punctate light source was 670cm. The position of the punctate source was held constant throughout all experiments. The punctate-total luminance ratio was always $\pi = 0.67$ (see Boyaci, Maloney & Hersh, 2003).

Software and apparatus

The stimuli were presented to the subjects stereoscopically on two Sony Trinitron Multiscan GDM-F500 21" CRT-screens (Figure 5.3). I ran the experiments under Red Hat Linux 6.1 using a Dell workstation with a NVIDIA dual-head graphics card that controlled both monitors. The two images of each scene were rendered with the Radiance package (Larson & Shakespeare, 1996). The two output files of the rendering procedure contained relative RGB triplets. These values were corrected for nonlinearities of the gun responses using measured look-up tables for each monitor. The look-up tables were based on direct measurements of the luminance values on each monitor with a Pritchard PR-650 spectrometer. Finally, the corrected RGB triplets were translated to 24-bit graphics codes. The experiments were programmed by myself in the C language using the X Window

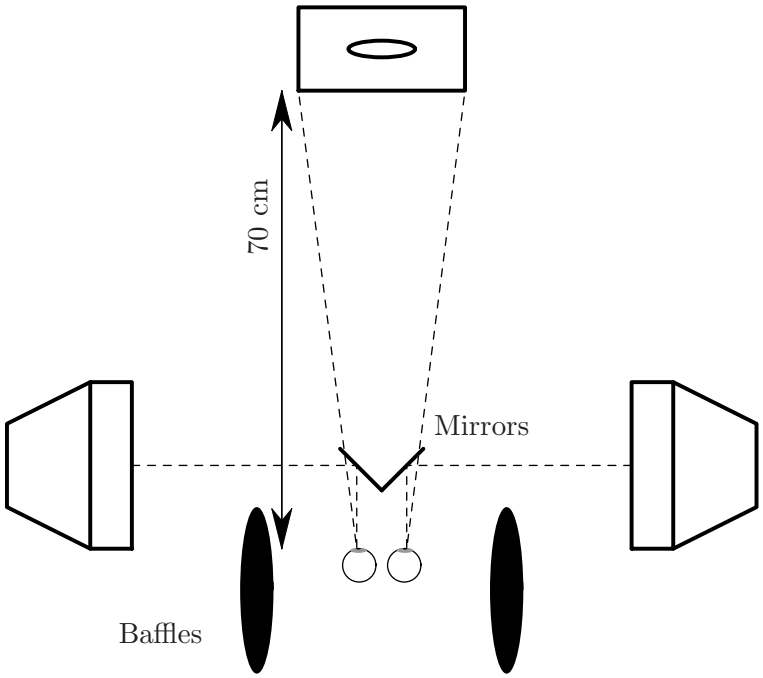

Figure 5.3: Apparatus. The observer is presented with a stereoscopic image which consists of two slightly different images displayed on two CRT-creens.

System, Version 11R6 (Scheifler & Gettys, 1996). The settings were made by the subjects using a mouse.

5.2.2 Task

The task of the subject in Experiments 1–3 was to set hue scalings for the test patch. Subjects were asked to judge how blue, yellow, red and or green a test surface appeared to them on four respective scales ranging from 0 (none) to 6 (very saturated). The active scale was presented monocularly as the number of the current value in the respective color of the scale. The subject saw this number to the right of the fused image. At any given time only one hue scale was active. The subject could increase or decrease the value of the active scale by pressing the left or the right mouse button respectively. The subject moved to a different scale by scrolling the mouse wheel up or down. The settings of one trial were confirmed by pressing the mouse wheel.

The decision to chose a rating scale with exactly six steps based on previous work. De Valois, De Valois, Switkes and Mahon (1997, p. 887) reported that:

> In preliminary trials, we found that observers preferred, and differentially used, a scale finer than the three-level scale, e.g. GGG, GGY, GYY, YYY, etc., used by Boynton and Gordon (1965), but did not require a 100-point scale as used by Abramov, Gordon and Chan (1990).

Even if subjects do not make non-zero setting on both scales of an opponent pair such as blue-yellow, then they can still make 13 distinct ratings for each opponent pair and, combining ratings from the two opponent scales, they can classify each stimulus in one of $169 = 13 \times 13$ different categories. The method presented here allows much finer comparison than color naming procedures where only 11 categories are typically allowed.

In Experiment 4, subjects were asked to make achromatic settings for the test patch.

5.2.3 Procedure

Subjects practiced hue scaling in a training session before running in the experiments. Only subjects who had reliability in their settings of 0.8 and more took part in the study. Reliability was determined as follows. In the training session subjects repeated hue scaling settings for 16 surfaces under neutral illumination five times. For each subject, all pair-wise correlations between repeated measurements were transformed using Fisher's r-to-z transformation. The subject's reliability was defined as the correlation corresponding to the mean of the transformed values. I did not measure consistency of hue scalings over time (like over a few days). However, Boynton and Gordon (1965) report that their subjects made very reliable settings with mean correlation coefficients of 0.96.

In the experiments, different illuminant conditions were generally blocked into different sessions. Each session started with 16 training trials. These training trials were prepended to the experimental trials to allow subjects to practice hue scaling and also to stabilize the adaptational state of the subject. In each trial the scene was presented to the subjects and they carried out the task described above.

In Experiment 4, subjects made achromatic settings for the test patch instead of hue scalings. Between two trials a black screen was presented to the subject for 1s to reduce the influence of afterimages. Observers repeated hue scalings for each of the 16 test surfaces four times. Each subject made 80 settings in one session. There were no time constraints. One session took about 20 minutes on the average. Each experiment consisted of five sessions that corresponded to the five different punctate sources. The order of illuminant blocks was randomized and differed across subjects. Within one illuminant block the test surface order was randomized and different subjects saw different randomizations.

5.2.4 Data analysis

First an overview of the general methods of data analysis will be given. The same analyses were carried out for Experiments 1–3.

Absolute hue scalings

As a first step, the absolute hue scalings of subjects under different illumination conditions were analyzed. For each subject and each surface mean hue scalings were calculated. These data were plotted as polar coordinates. In this representation, each of the 16 surfaces corresponds to a fixed angle. Hue scalings of a subject are plotted on the axis belonging to the respective surface. Each figure contains mean hue scalings of a subject on the four scales. The settings on a fixed scale are connected and form a geometric object that looks like a bubble. For a color constant subject it is expected that form and size of these bubbles not to change with changes of the illuminant. A subject shows no constancy if form and size of the bubbles are determined only by the local color signal of the test surface under each illumination.

Transformed hue scalings

In a second step of the analysis, mean hue scalings for each subject in each illumination condition were transformed to $u'v'$-chromaticity coordinates. This transformation makes it easier to compare the present results with those from previous studies in terms of color constancy indices. The procedure of obtaining these indices is explained next.

Subjects made non-zero settings for a given test surface only on one of the blue or the yellow scale but not both, and on only one of the red or the green scale but not both. Therefore, for each given subject and condition, blue and yellow ratings were combined into a single number a_{BY} on a BY opponent scale that ran from -6 (blue) to 6 (yellow). If, for example, the subject's rating was 2.3 on the blue scale and 0 on the yellow, the rating on the opponent scale would become $a_{BY} = -2.3$. A rating of 1.7 on the yellow scale (with rating 0 on the blue) would become $a_{BY} = 1.7$. Similarly, red and green hue ratings were combined into a rating a_{RG} on a red-green opponent scale. An observer's mean hue ratings in any given condition are then summarized by the two-dimensional column vector $\mathbf{a} = (a_{BY}, a_{RG})'$ in a two-dimensional opponent space, A.

Let $\mathbf{u}$ be the two-dimensional vector that denotes the $u'v'$-coordinates of a surface under the neutral illuminant. We assume that the mapping between the opponent space A and $u'v'$-space is affine but perturbed by judgment error. That is, we assume that there is a 2×2 matrix and a column vector $\mathbf{c}$ such that

$$\mathbf{u} = \mathbf{M}\mathbf{a} + \mathbf{c} + \varepsilon, \tag{5.2}$$

where $\varepsilon : \Phi(0, \sigma^2)$ is Gaussian judgment error with mean 0 and variance σ^2. The vector $\mathbf{c}$ captures the shift of the achromatic point.

For the neutral illumination, the chromaticity coordinates of the surfaces and mean hue scalings of the subjects are known. This data was used to estimate the six parameters of the transformation (the four elements of $\mathbf{M}$ and the two

elements of **c**) for each subject. The set of parameters was chosen so that the sum of squared distances between $u'v'$-coordinates and predictions was minimized. This set of parameters, **M** and **c**, determined by the observer's ratings under neutral illumination, was then used to compute $u'v'$-coordinates from hue scalings under other illuminations.

The Speigle-Brainard conjecture

Speigle and Brainard (1999) proposed a model to predict achromatic settings from asymmetric matching. In hue scaling terms, the conjecture of Speigle and Brainard can be stated as a claim that knowledge of the shift in hue scalings of a neutral reference surface determines all other possible hue scalings. In testing this conjecture, the centroid of all of the surface patches under a single light will be used as reference surface. It is not assumed that this surface, which the subject never sees, is precisely achromatic but, given the construction of the stimuli, it will be nearly so.

I tested the Speigle-Brainard conjecture in the following way. First, all opponent data of subject i was translated to $u'v'$-coordinates using the specific transformation that was given by $\mathbf{M}_i$ and $\mathbf{c}_i$. Then optimal ellipses to the $u'v'$-data of each subject for each illuminant were fitted (Halir & Flusser, 1998). Finally the parameters of the ellipse for neutral illumination data were compared with those of the ellipses for the chromatic illumination data.

The analysis of the parameters of the ellipses has the advantage that we can directly separate effects of color constancy from deviations from the Speigle-Brainard conjecture. If we analyze the ellipses for neutral and chromatic illumination then differences in the centroids reflect incomplete color constancy. However, differences in area, eccentricity (shape) or orientation of the ellipses indicate deviations from the Speigle-Brainard conjecture. If their conjecture were valid, then changes in illuminant chromaticity would lead to simple translations of the ellipses in $u'v'$-space. They will not rotate, shrink or grow, or change shape.

I next describe how area, eccentricity and orientation were measured. Let a be the semimajor axis and b be the semiminor axis of an ellipse. Then the eccentricity e of the ellipse is given by:

$$e = \sqrt{1 - \frac{b^2}{a^2}}. \tag{5.3}$$

The orientation of the ellipse is given by the radian angle θ between the major axis of the ellipse and the x-axis. The area A of an ellipse is defined as $A = \pi ab$.

Color constancy indices

Color constancy indices were calculated in order to compare performance of observers across different tasks. The rationale behind the determination of indices differed for the hue scaling and the achromatic setting task.

First, the derivation of indices from transformed hue scalings of subjects will be described. In case of perfect constancy the $u'v'$-coordinates of the scalings under test illumination should completely overlap with those under neutral illumination. If a subject has no constancy, the $u'v'$-coordinates of the scalings under test illumination coincide with the coordinates of the test surfaces under this illumination.

The centroid of $u'v'$-chromaticities of all surfaces under neutral illumination was almost identical to the centroid of the predicted coordinates. Therefore, the centroids of test surfaces and settings in $u'v'$ were used to calculate color constancy indices. The indices have the form:

$$CI = 1 - \frac{|\mathbf{m}_{D65} - \mathbf{m}_{data}|}{|\mathbf{m}_{D65} - \mathbf{m}_{test}|}, \tag{5.4}$$

where $\mathbf{m}_{D65}$ is the centroid of the surfaces under neutral illumination, $\mathbf{m}_{test}$ is the centroid of the surfaces under test illumination and $\mathbf{m}_{data}$ is the centroid of the transformed hue scalings of the subject. This index is 1 in case of perfect constancy and 0 in case of no constancy. By design, it is directly comparable to Brunswik ratios typically reported in studies of color constancy (Arend, Reeves, Schirillo & Goldstein, 1991; Brunswik, 1929).

In Section 4.2.6 it was described how to derive an index of color constancy performance from subject's achromatic settings (see also Arend, Reeves, Schirillo & Goldstein, 1991). In case of perfect color constancy, we expect the achromatic setting of the subject to coincide with the chromaticity of the test illuminant. A subject has no color constancy if the achromatic setting under chromatic test illumination is identical with the chromaticity of neutral daylight $D65$. Roughly speaking, the color constancy index sets the shift of the subject's achromatic setting in relation to the illuminant shift. The index is 1 in case of perfect constancy and 0 in case of no constancy.

5.3 Experiment 1: 3D Scenes – Hue Scalings

5.3.1 Methods

Stimuli and apparatus

In Experiment 1, the observers were presented with simulations of 3D scenes rendered under five different illuminations. The apparatus and stimuli that were used in Experiment 1, namely the visual display, the set of test patches and the simulated light sources, were already described in the General Methods Section.

Task and procedure

In Experiment 1, the observers were asked to set hue scalings for the test patch in the center of the scene. The details of task and procedure were already described

in the General Methods Section. Subjects were told that they are viewing a scene under a certain illumination.

Observers

Five subjects took part in Experiment 1. All subjects were paid undergraduate students who were not aware of the purpose of the study. All had normal color vision as tested with Ishihara color plates (Ishihara, 1997).

5.3.2 Results

Expected results

In the 3D-experiment subjects are enabled to use various cues to the illuminant. Therefore it is expected that hue scalings are constant with changes of the illuminant. This constancy might not be perfect, but rather lead to settings different from zero on one scale for a surface under different illuminations. Consistent with previous empirical findings (Brainard, 1998; Delahunt & Brainard, 2004b) it is not expected that color constancy performance will be better for natural daylights than for artificial red and green illuminations displaced from the daylight locus.

Experiment 1: results

Subjects consistently used the same scale settings to judge the color appearance of a fixed surface under different illuminants. For example, a test surface that appeared blue under neutral illumination was also rated mainly blue under the chromatic illuminations. The magnitude of the scalings usually differed with changes of the illuminant. For example, the 'blue' surface was rated to appear more saturated blue under blue illumination and less saturated blue under yellow illumination. In this sense, subjects showed stable hue scalings for a given test surface with changes of the illuminant (Figure 5.4). The transformed hue scaling settings for a given test patch under a chromatic test illumination fell closer to the chromaticity of this test patch under neutral illumination (Figure 5.5A).

For each subject, the fitted ellipses of the neutral illuminant data were compared with those of each chromatic illuminant condition to test the Speigle-Brainard conjecture. In general, the shapes and the sizes of ellipses of the chromatic illuminant conditions did not deviate systematically from those of the neutral illuminant condition (Figures 5.6–5.7). The analysis of the orientation parameter θ should be interpreted with care as the estimated ellipses of one subject (XH) were almost circular which resulted in unreliable estimates of θ.

In general, small systematic deviations occurred for the orientation of the ellipse in the yellow and green illuminant conditions (Figure 5.8). The analysis of the ellipses indicates that the Speigle-Brainard conjecture holds also for hue

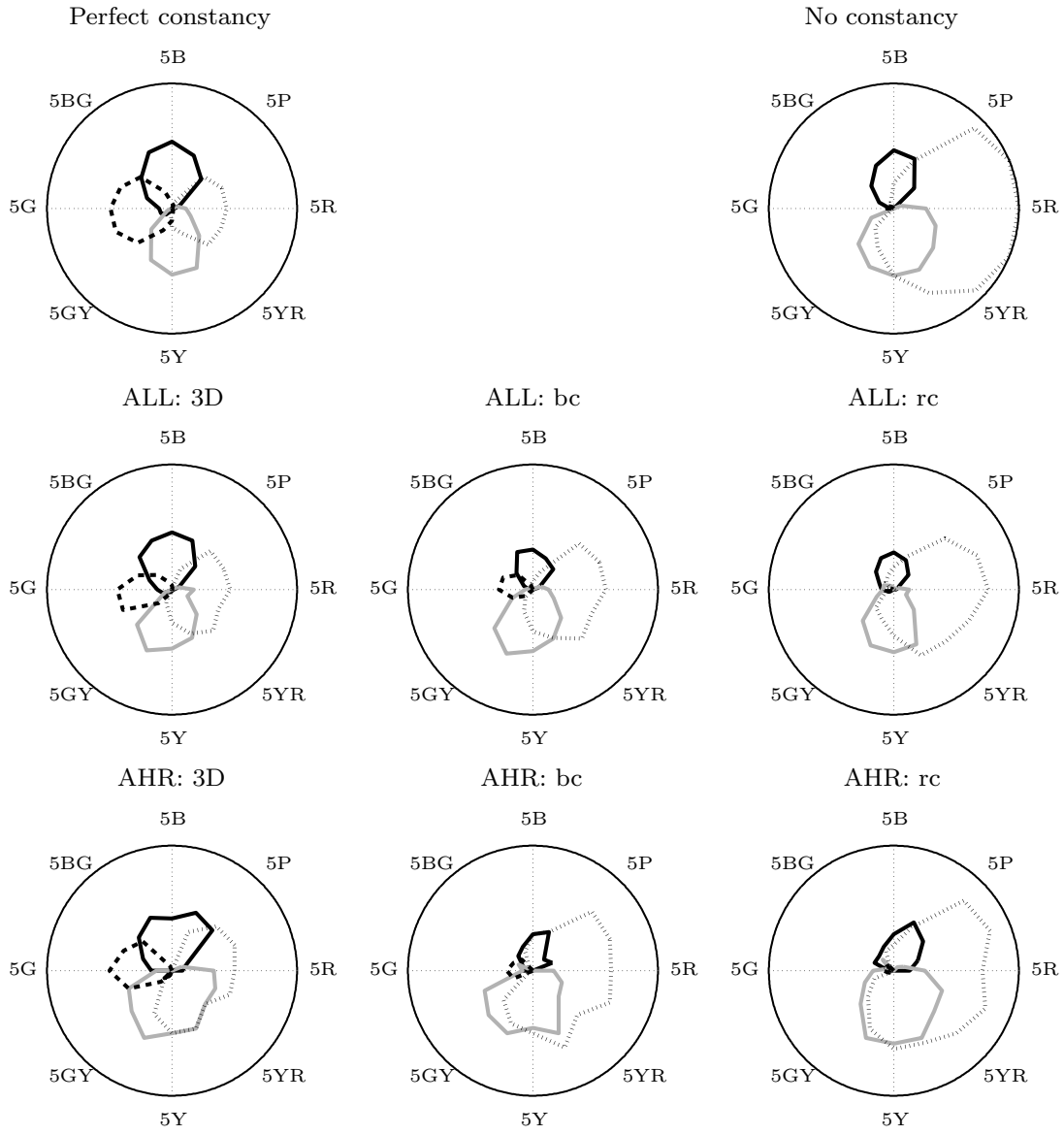

Figure 5.4: Mean hue scalings of all subjects and of subject AHR under red punctate illumination as polar plots. Surfaces are given in Munsell notation. Different line styles and colors indicate the four scales: solid black – blue, solid gray – yellow, bold dashed – red, dotted – green. The first row gives results corresponding to perfect color constancy, i.e. scalings under neutral illumination (left) and hypothetical results corresponding to complete failure of constancy (right). The plots in the three columns show data from the 3D–experiment (3D), the blocked control (bc) and the randomized control (rc) respectively. The second row shows mean scalings of all subjects.

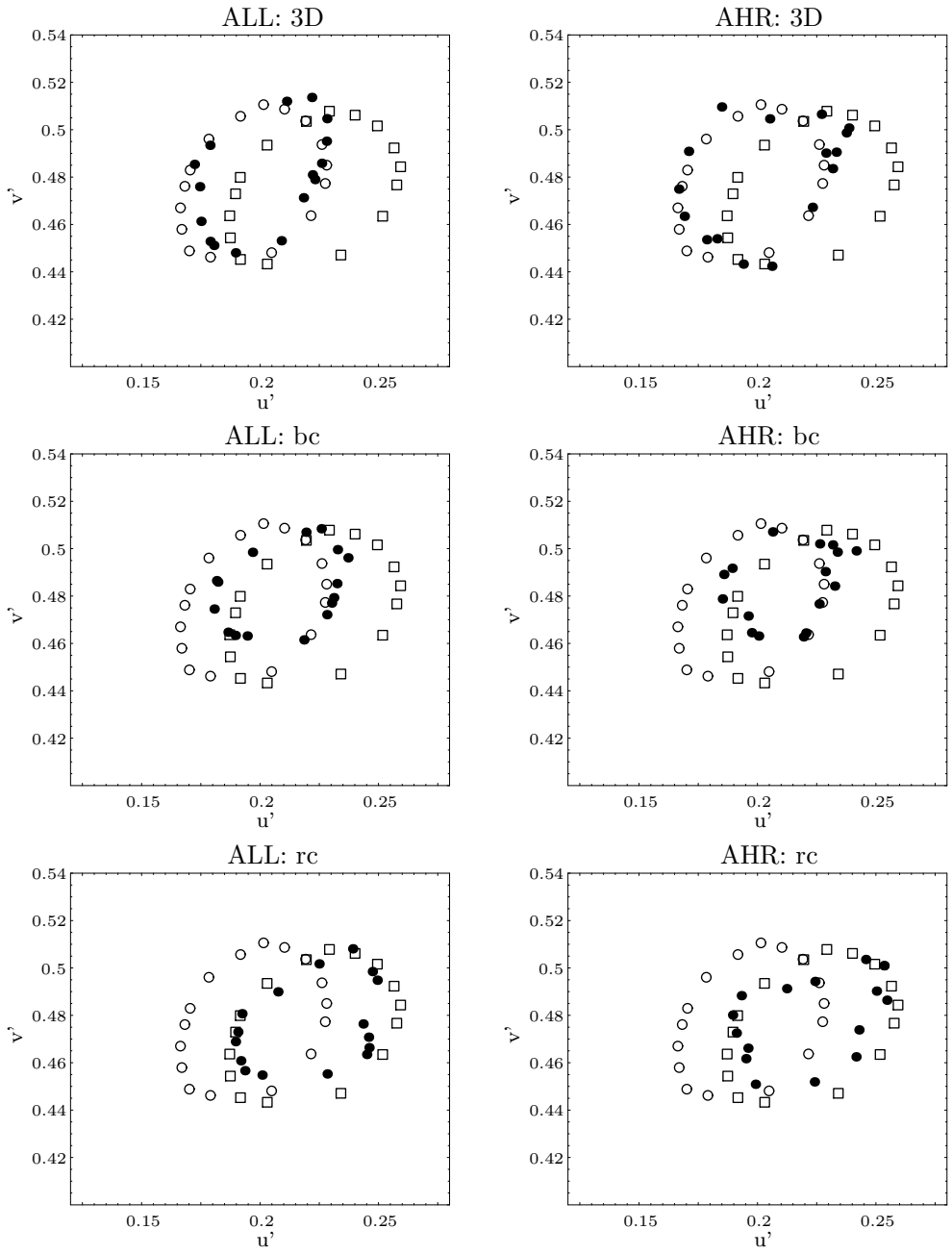

Figure 5.5: Transformed mean hue scalings of all subjects and of one subject (AHR) under red punctate illumination. The plots in the three rows show data from A. Experiment 1 (3D), B. Experiment 2 (bc) and C. Experiment 3 (rc). Open circles – surfaces under neutral illumination, open squares – surfaces under red illumination, filled circles – transformed mean hue scalings. For an observer with perfect color constancy we expect the pattern of mean hue scalings to coincide with the ellipse of color signals under neutral illumination (open circles). If the observer has no color constancy the pattern of mean hue scalings will fall together with the ellipse of color signals under chromatic illumination (open squares).

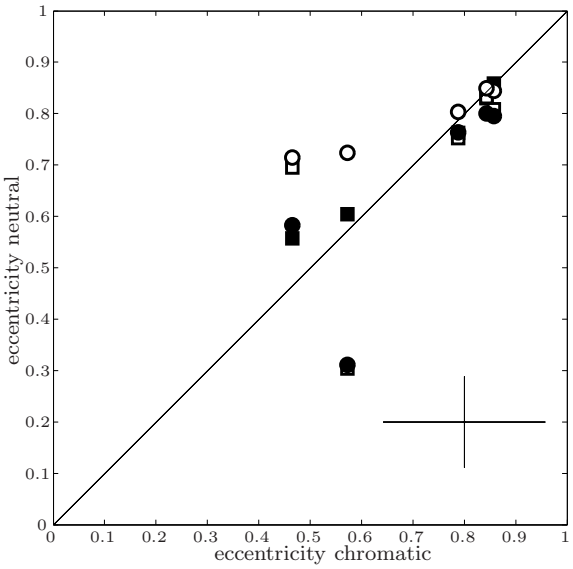

Figure 5.6: Eccentricities of the estimated ellipses for neutral and chromatic illuminant conditions. Each symbol refers to a different chromatic illuminant (blue – filled square, yellow – open square, red – filled circle, green – open circle). The eccentricities of a given subject lie on a line parallel to the ordinate (from left to right: AHR, XH, ALL, LGS, AST). The standard deviations for the conditions are shown in the cross at the right part of the figure.

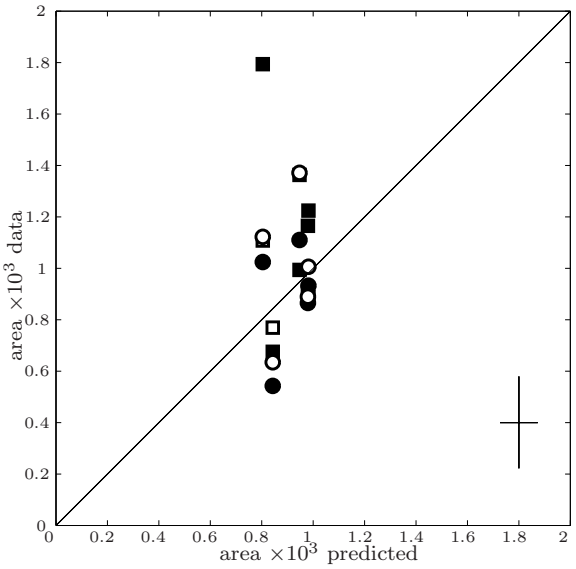

Figure 5.7: Area of the estimated ellipses for neutral and chromatic illuminant conditions. Each symbol refers to a different chromatic illuminant (blue – filled square, yellow – open square, red – filled circle, green – open circle). The area values of a given subject lie on a line parallel to the ordinate (from left to right: AHR, XH, ALL, LGS, AST). The standard deviations for the conditions are shown in the cross at the right part of the figure.

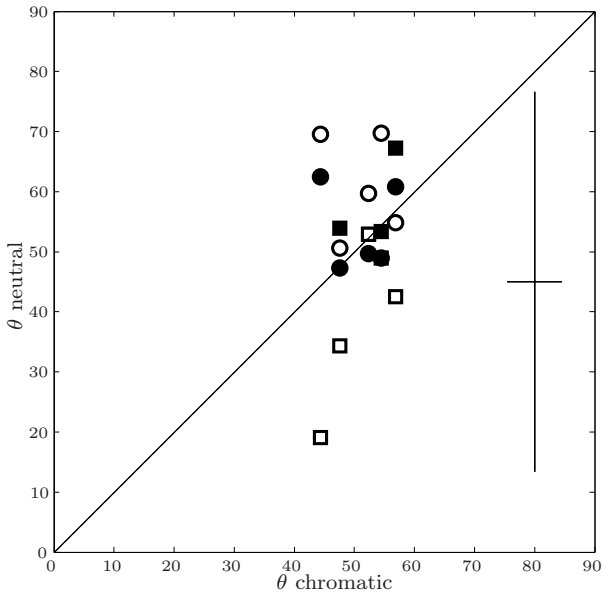

Figure 5.8: Orientation (θ) of the estimated ellipses for neutral and chromatic illuminant conditions. Each symbol refers to a different chromatic illuminant (blue – filled square, yellow – open square, red – filled circle, green – open circle). The θ values of a given subject lie on a line parallel to the ordinate (from left to right: XH, AST, LGS, AHR, ALL). The standard deviations for the conditions are shown in the cross at the right part of the figure.

scaling data. The constancy indices that were calculated from these transformed data lay between 0.58 and 0.91 with median 0.705 (Table C.5, Appendix C). This indicates good to excellent color constancy (Figure 5.9). The results from Experiment 1 also indicate that color constancy performance is not better for daylights than for artificial red and green illuminations.

5.4 Experiment 2: Blocked Control

5.4.1 Methods

Stimuli and apparatus

In Experiment 2, the isolated test patch was presented in front of a black background. All stimuli of one illumination condition were shown as a block. The apparatus, the set of test patches and the simulated light sources in Experiment 2 were as described in the General Methods Section. The same rationale as in Experiment 1 was used to calculate the light signals of the test patches.

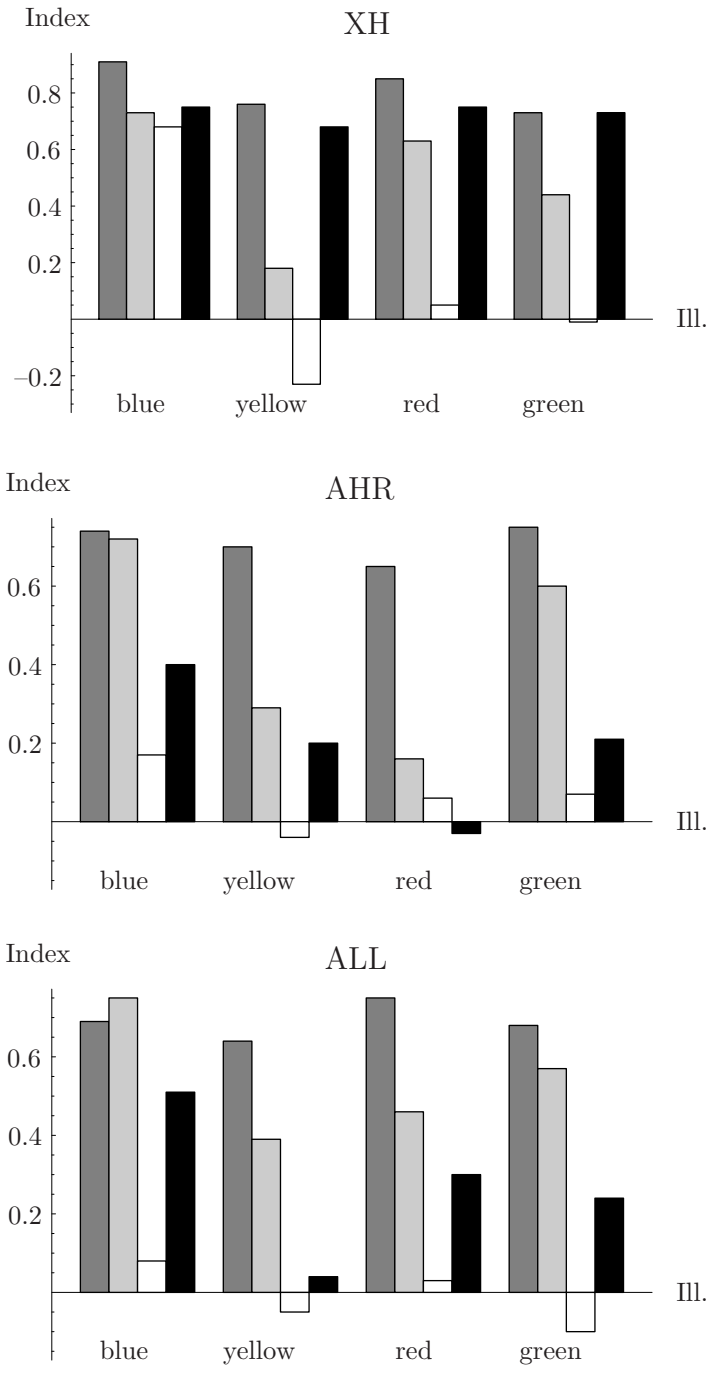

Figure 5.9: Color constancy indices for subjects XH (top), AHR (center) and the mean data of all subjects (bottom) under four chromatic illuminant conditions and each of the Experiments 1–4. Different shadings refer to different experiments: Dark gray bars – hue scalings 3D (Exp. 1), light gray bars – blocked control (Exp. 2), white bars – randomized control (Exp. 3), black bars - achromatic settings (Exp. 4). Note that the rationale the determination of indices is based upon differs for the hue scaling and the achromatic setting task. Therefore, interpretation and comparison of indices across different tasks should be done carefully.

Task and procedure

Subjects set hue scalings for the test patch as in Experiment 1. In the present experiment, subjects were not told that the stimuli might be interpreted as a surface under a certain illumination. As in Experiment 1, the order of illuminant blocks was randomized and differed across subjects. Within one illuminant block the test surface order was randomized and different subjects saw different randomizations. The randomizations used were not the same as those used in the previous experiment.

Observers

The same as in Experiment 1.

5.4.2 Results

Expected results

In designing Experiment 2, it was originally assumed that no cue to the illumination had been given to the subject. Therefore, mean hue scalings of the subjects were expected to depend only on the color codes of the surface under each illumination. In this sense, subjects should show no constancy in their settings.

Experiment 2: results

In the blocked control, all test surfaces of one illumination condition were presented to the subjects within one session in front of a black background. There were neither cues to interpret the stimulus configuration as a surface under a illumination nor was this possible interpretation mentioned explicitly.

Surprisingly, the scalings of the subjects in this experiment were not simply determined by the color codes of the test surface (Figure 5.4, Figure 5.5). For example, when a surface was used that appeared green under neutral illumination and a test patch that appeared red under neutral illumination. The green surface under the red illuminant and the red surface under green illumination produced identical color codes, but, even in this case subjects rated the former green whereas the latter was judged to appear red. The constancy indices suggest that it appears as if subjects had developed a kind of partial constancy (Table C.5, Appendix C; Figure 5.9). This result held for all subjects and under all illumination conditions.

Smithson and Zaidi (2004) found evidence of a similar dependence of color appearance on the chromaticities of test surfaces presented sequentially. They presented test stimuli one at a time against a fixed background. The task of the observer was to judge the color appearance of the test patch using only the terms ‘blue’, ‘yellow’, ‘red’ and ‘green’. Smithson and Zaidi estimated the neutral

points of the two opponent channels under a given illumination from observers' data. In one experiment, they rendered test patches and background under different illuminants. They found that, in this case, the judgments of observers were significantly affected by the chromaticity of the illumination of the test patches, implying that the visual system somehow estimated this chromaticity across successive presentations in time. Their results are consistent with those we can draw from Experiment 2. It should be noted, however, that in Experiment 2, there was no visible background. Consequently, in the present experiment there is no possibility that the effect depended in some way on the presence of a background. The most plausible explanation for my results and those of Smithson and Zaidi is that the visual system is averaging chromaticities from multiple surfaces across time and using this average as an estimate of the chromaticity of the illuminant in estimating surface colors. We return to this point in the discussion.

5.5 Experiment 3: Random Control

5.5.1 Methods

Stimuli and apparatus

As in Experiment 2, only the isolated test patch was presented in front of a black background. In this experiment, however, test surfaces from all illumination conditions were shown in random order. Subjects made hue scalings for 80 different test patches within one session. The apparatus, the set of test patches and the simulated light sources in the present experiment were as described in the General Methods Section. The same rationale as in Experiment 1 was used to calculate the light signals of the test patches.

Task and procedure

Subjects set hue scalings for the test patch as in Experiment 1. In this experiment, subjects were not told that the stimuli might be interpreted as a surface under a certain illumination. The procedure of Experiment 3 differed in the following respect from those of Experiments 1 and 2. Each session included all 16 different surfaces rendered under all five illuminants. The 80 different stimuli of each session were presented in random order and different subjects saw different randomizations. Each subject completed five sessions.

Observers Four of the five subjects from the previous experiments took part in Experiment 3.

5.5.2 Results

Expected results

In Experiment 3, no cues to the illumination were given. The sequential presentation of the test surface × illumination combinations was fully randomized. Therefore, mean hue scalings of the subjects are expected to depend only on the color codes of the surface under test illumination. This means that subjects should show no constancy in their settings.

Experiment 3: results

The hue scalings in this control were completely determined by the color codes of the test surface (Figure 5.4). For example, subjects made identical scalings for the green surface under red illumination and the red test patch under green illumination. The transformed hue scalings of the subjects under a chromatic test illumination fell together with the color codes of the surfaces under this illuminant (Figure 5.5). According to this pattern, most of the constancy indices in this experiment are close to zero and indices are always smaller than those from Experiments 1 and 2 (Table C.5, Appendix C; Figure 5.9).

5.6 Experiment 4: 3D Scenes – Achromatic Settings

5.6.1 Methods

Stimuli and apparatus

In Experiment 4, the observers were presented with simulations of 3D scenes rendered under either one of the five different illuminations. The visual display was identical to Experiment 1 except for the color of the test patches. Achromatic settings for the test patches were obtained at four fixed luminance levels (8, 12, 18, 27 cd/m^2). The initial chromaticity of a given test patch was chosen at random from a line D65 – test illuminant. The apparatus was identical to the one used in the previous experiments which was already described in the General Methods Section.

Task and procedure

In Experiment 4, observers made achromatic settings for the test patch. They were asked to adjust the chromaticity of the test patch until it appeared neither bluish nor yellowish and/or neither reddish nor greenish to them. Results from previous studies have shown that achromatic settings of subjects are essentially aligned on a line that is determined by the direction of the adaptational stimulus

in chromaticity space (Bäuml, 1994; Rinner & Gegenfurtner, 2002). Therefore, the subjects made their settings on only one line which was defined as the line D65 – test illuminant in xy-chromaticity diagram. This modification considerably reduced the complexity of the task. The subjects reported being able to always find satisfying settings. Under $D65$ punctate illumination observers used two axes to make their adjustments that corresponded to the two opponent channels RG and BY. Subjects practiced achromatic settings in a preliminary training session before running in the experiment. Like in Experiment 1, they were told that they are viewing a scene under a certain illumination.

Different illuminant conditions were blocked into different sessions. Subjects repeated settings for each of the four different luminance levels of the test surface five times resulting in 20 settings per block. In order to control the adaptational state of the subject, the first test patch appeared only after 2 min of presentation of the simulated 3D-scene. There were no time constraints. One session took about 15 minutes on the average. The order of illuminant blocks was randomized and differed across subjects. Within one illuminant block the test surface order was randomized and different subjects saw different randomizations.

Observers

Four of the five subjects from the previous experiments took part in Experiment 4.

5.6.2 Results

Expected results

Like in Experiment 1, subjects are provided with various cues to the illumination from the simulated 3D-scene. Therefore, achromatic settings of subjects are expected to be located closely at the chromaticity of the punctate light source. The degree of observers' color constancy should be comparable to results from previous studies using a similar paradigm (Kraft & Brainard, 1999; Rinner & Gegenfurtner, 2002). The degree of color constancy reported in these studies typically lies in the range 60–90%. According to the outcome of the experiments reported earlier in this work, color constancy performance is not expected to be better for natural daylights than for red and green illuminations displaced from the daylight locus.

Experiment 4: results

The interpretation of results in the achromatic setting task in terms of color constancy follows a rationale which differs from the hue scaling task. In the present case, observer has perfect color constancy if chromaticities of his achromatic settings coincide with the chromaticity of the test illuminant. In contrast, it is

expected that chromaticities of the observers achromatic settings do not shift with the illuminant in case of missing color constancy.

Figure 5.10 shows achromatic settings of two observers and the pooled data of all observers. The results indicate strong differences across subjects which are probably due to the experience with the achromatic setting task. This interpretation emerges from comparison with corresponding hue scaling data of these observers and the different degree of experience with the achromatic setting task. Only one subject, XH, who had participated in comparable psychophysical experiments before showed a pattern of results that one would expect for an observer with good color constancy. Achromatic settings of subjects AHR, AST and LGS who were not experienced with the task varied only little across different illuminants.

As a consequence, observers' adjustment to the illuminant measured with achromatic settings was generally much smaller compared with the hue scaling data obtained in Experiment 1. The degree of color constancy was expressed in terms of a constancy index. In Figure 5.9 constancy indices of observers XH, AHR and indices determined from the pooled data of all observers are compared across the different experiments. Constancy indices of individual observers are given in Table C.5 (Appendix C). Only the indices of subject XH are comparable across the two different tasks. There is no evidence in the data for a systematic difference between the adjustment to daylights and to lights off the daylight locus.

5.7 Discussion

In the present study, color constancy performance was measured using two different paradigms: a hue scaling task and achromatic settings. In the first experiment, subjects showed constant hue scalings for a given test patch with changes of the illuminant. The degree of color constancy observed in the first experiment was comparable to that found in previous studies (Brainard, 1998; Foster, Amano & Nascimento, 2001a). In Experiment 2, subjects were presented only with a sequence of isolated surfaces that were related to a given illumination condition. Nevertheless, hue scalings of the subjects still indicated a moderate degree of constancy. It seemed as if subjects had adapted to the whole sequence of previously presented stimuli. In contrast, the constancy disappeared, when the surfaces were presented completely randomized to the subjects (Exp. 3). In Experiment 4, color constancy performance of observers measured with achromatic settings was generally much smaller in comparison with the hue scaling data from Experiment 1. This finding was probably due to the lack of experience of most subjects with the achromatic setting task.

Three results stand out. First, the results from Experiment 1 indicate that hue scaling is an appropriate technique to investigate color constancy in a more

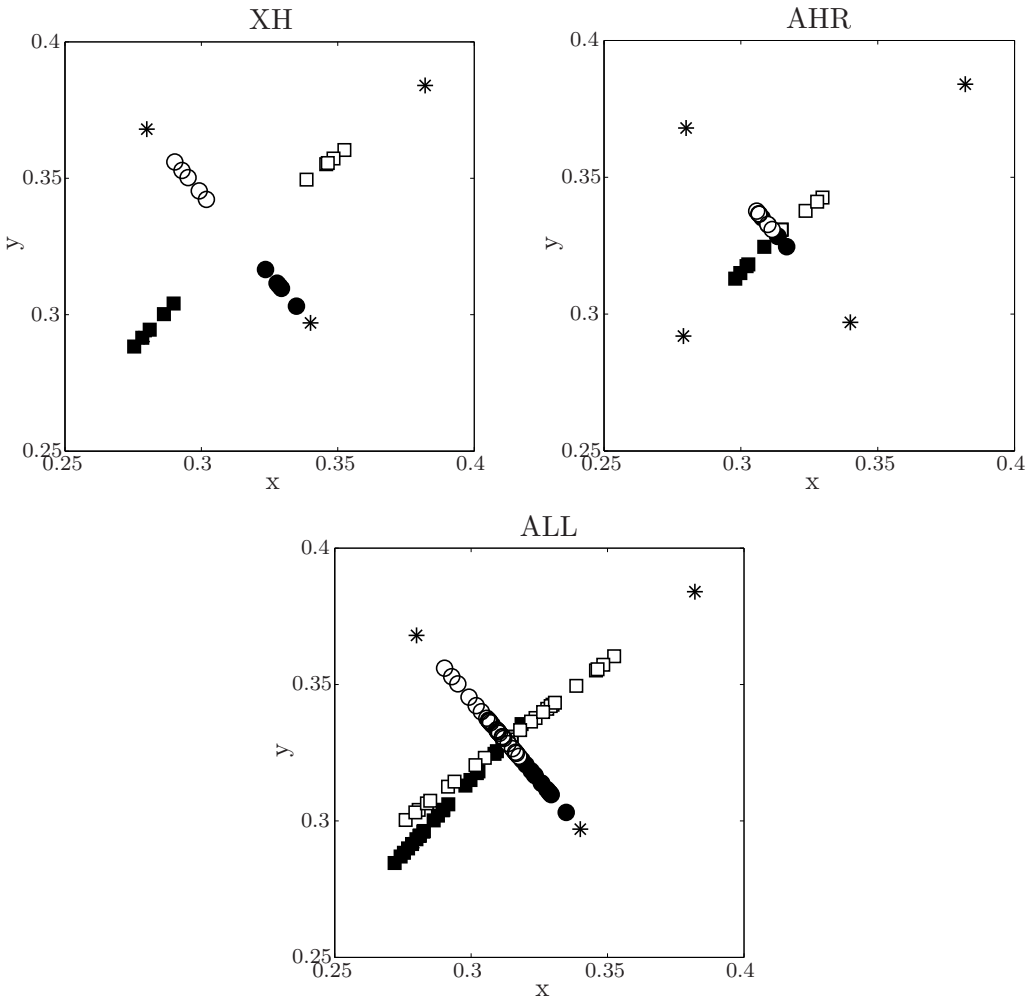

Figure 5.10: Achromatic settings of subjects under chromatic punctate light sources in xy-chromaticity diagram. The top row shows data of an experienced subject (XH, top left) and an inexperienced subject (AHR, top right). Achromatic settings of all subjects are shown below (ALL, bottom). Different symbols refer to different illuminant conditions: filled square - blue, open square - yellow, filled circle - red, open circle - green. The asterisks denote chromaticities of the illuminations which correspond with settings of an perfectly color constant observer.

phenomenological sense. Furthermore, the data shows that a measure of color constancy based on hue scaling is comparable to quantitative measures of color constancy performance.

In a recent essay on color constancy, David Foster pointed out that achromatic settings and asymmetric matchings are problematic methods to measure color constancy and that more direct approaches are needed that serve this purpose (Foster, 2003). The present study is a first step toward developing and validating measures of color constancy that on one hand are more closely related to the phenomenon of color constancy and on the other hand can be used to quantify the degree of color constancy.

Second, over the range of stimuli considered, the results indicate that the Speigle-Brainard conjecture (generalized to hue scaling) held. Speigle and Brainard (1999) were able to predict asymmetric matchings from achromatic settings. In the analysis of the data ellipses it was shown that the centroid of the ellipse under a given chromatic illumination (that is the achromatic point) is needed to reconstruct an approximation to the locations of the respective chromatic test surfaces. This result, based on hue scaling, supports the view of Speigle and Brainard (1999).

Third, the surprising results from Experiment 2 in comparison to those from Experiment 3 showed the strong dependence of settings from previously presented surfaces. A comparison across all panels of Figure 5.4 illustrates the differences between the three experimental hue scaling conditions. This finding underlines the strong contribution of slow chromatic adaptation to the illuminant adjustment of the visual system.

In the literature the distinction between successive and simultaneous color constancy has been made (Bäuml 1999; Brainard, 1998; Brainard, Brunt & Speigle, 1997). Both terms can be related to two different situations in the natural environment. Successive color constancy refers to gradual changes of one uniform illumination as it happens when daylight changes from dawn until noon. Simultaneous color constancy can be related to a situation where one part of a scene is exposed to direct sunlight whereas another part is situated in the shade. In this terminology the present study refers to successive color constancy. My results are consistent with the claim that successive color constancy is simply due to chromatic adaptation (Kuriki & Uchikawa, 1996). However, further investigation of the factors that contribute to changes in adaptational state across time is needed before drawing conclusions.

Recent studies on the time course of chromatic adaptation revealed that chromatic adaptation consists of at least two processes, a fast and a slow phase of adaptation, and that chromatic adaptation is complete after about two minutes (Fairchild & Reniff, 1995; Rinner & Gegenfurtner, 2000; Shevell, 2001; Werner, Sharpe & Zrenner, 2000). In these studies, test stimuli were presented to the subjects on larger adaptation backgrounds. As in this stimulus configuration simultaneous and successive contrast are confounded, the two suggested processes

of chromatic adaptation might be related to these two classes of phenomena (Rinner & Gegenfurtner, 2000). In Experiment 2, only isolated test patches were presented to the subjects. The results from Experiment 2 can be interpreted as the isolated effect of slow (sensory) chromatic adaptation. The observer, given samples of surfaces from a scene, one at a time, is able to develop an estimate of the illuminant and correct surface color appearance for this illuminant estimate, at least in part. More research is needed to clarify the relations between the phenomena of successive and simultaneous contrast, the two processes of chromatic adaptation, and stability of color appearance in ongoing scenes.

Another interesting outcome of Experiment 2 is the fact that the visual system is able to integrate information from isolated stimuli temporally. As noted above, Smithson and Zaidi (2004) found a similar effect in scenes where single test surfaces were presented in sequence against a fixed background. It is interesting to speculate what role this temporal integration of surface chromaticities might serve in surface color perception. D'Zmura and Lennie (1986) conjectured that the visual system estimates illuminant chromaticity by integration of chromaticities across successive eye movements, in effect developing an estimate of mean chromaticity of the scene by sampling across time. The present results and those of Smithson and Zaidi suggest that a similar averaging process occurs even when the change from one surface to the next is not due to an eye movement but is under experimenters control. There is considerable evidence for slow adaptation processes to exchange of very intense backgrounds differing in chromaticity (Augenstein & Pugh, 1977; Reeves, 1982). It remains to be seen how and when such slow adaptational processes affect color appearance.

Chapter 6

General Discussion

6.1 Summary

Before discussing the empirical findings let us summarize the main results of the two experimental studies in brief:

- Color deficient protanopic observers show slightly better adjustment to illuminants than predicted by the reduction hypothesis of dichromacy.
- The strength of increment-decrement asymmetries increases with decreasing correlated color temperature of daylights.
- Color constancy performance of observers is in general better for decremental than for incremental stimuli.
- Observers' adjustment to daylights is not better than to illuminations off the daylight locus.
- Color constancy performance of observers measured with hue scaling is comparable to performance obtained with established quantitative measures.
- The contribution of chromatic adaptation to color constancy is greater than expected.

6.2 Discussion of Central Findings

6.2.1 The Reduction Hypothesis

The results do not strictly argue against the reduction hypothesis of dichromacy. However, the ambiguity of the present findings is in line with recent results from a number of studies (Montag, 1994; Wachtler, Dohrmann & Hertel, 2004; Bonnardel, 2006). These results indicate that dichromacy is by far a much more

complex framework than suggested by the simple reduction model. Explanations that may account for the ambiguous findings are variations of optical density across the retina and across observers (Crognale, Teller, Yamaguchi, Motulsky & Deeb, 1999). Learned strategies may also account for variation within classes of dichromatic observers (Bonnardel, 2006).

Genetic analysis of color deficient observers revealed that it is difficult to draw a clear dividing line between anomalous trichromats and dichromats (Neitz, Neitz, He & Shevell, 1999). The authors found observers with two distinct photopigments who still experienced trichromatic vision probably due to differences in the pigments optical density. In this sense, investigating the linkage between genotypical characteristics and phenotypical performance might be a promising approach toward a more adequate model of dichromacy (Shevell, He, Kainz, Neitz & Neitz, 1998; Sharpe, Stockman, Jägle, Knau, Klausen, Reitner & Nathans, 1999).

Genetic singularities and the related neurophysiological substrate are probably not the only source that account for aspects of dichromatic vision. Jameson and Hurvich (1978) suggested that acquired cognitive strategies might be responsible for the amazing color naming abilities of dichromatic observers (Montag, 1994). Recently, Bonnardel (2006) examined this assumption more systematically. She compared results from a color naming and two color ordering tasks and found that performance of dichromats was comparable to normal observers only in the color naming task. Bonnardel concluded that 'this result is attributed to a normative language system that dichromatic observers developed by learning subtle visual cues to compensate for their impoverished color system.' (Bonnardel, 2006, p. 637). In addition, it is a well established fact that developmental aspects play an important role in human color vision (Sugita, 2004; Lotto, 2004).

Taken together, there is strong evidence that the reduction hypothesis provides an oversimplified description of dichromatic vision. More advanced models should consider genotypical and corresponding phenotypical characteristics of dichromatic observers and cognitive developmental factors.

6.2.2 Increment-Decrement Separation

The separation of incoming incremental and decremental light signals is a substantial feature of the visual system across a number of different viewing conditions. The present work provides further evidence for the assumption that incremental and decremental stimuli trigger two different classes of codes, illuminant and object code (Bäuml, 1997; Mausfeld, 1998). Findings from psychophysical and neurophysiological studies suggest that the S-cone system and related pathways play a crucial role in this segregation process (Chichilnisky & Wandell, 1996; McLellan & Eskew, 2000).

At which point in the visual path might the separation of S-cone incremental and decremental stimuli be located? Results from a recent study by McLellan and Eskew (2000) suggest that one candidate could be S-ON and S-OFF ganglion

cells (see also Schiller, Sandell & Maunsell; 1986). These cells are fed by S-cone signals and different relative amounts of L- and M-cone input.

Further support for this assumption comes from a psychophysical study by Chichilnisky and Wandell (1996). They found that the assumption of cone-independent rescaling held for decremental but not for incremental stimuli. The authors argue that this result reflects different response patterns of S-ON and S-OFF ganglion cells which are related with increments and decrements respectively. In this sense, postreceptoral adaptation in ON-cells is supposed to be mediated by more than one class of cones whereas this is not the case in OFF-cells. However, due to the experimental design the issue of cone-independent rescaling was not investigated in the present study.

In addition, results from the present study suggest that asymmetric processing of increments and decrements seems to be optimally tuned to the locus of natural daylights. Observed asymmetries were strongest under yellowish illumination. One may speculate if the strong distinction between incremental and decremental signals under yellowish illumination yield any advantage for the visual system in natural environments. For example, the visual system may interpret incremental stimuli rather as self-luminous objects, e.g. punctate light sources, and decrements as reflecting surfaces (Mausfeld, 1998). Support for this assumption comes from the finding that in general, observers' adjustment to the illuminant is better for decrements than for increments (see also Chichilnisky & Wandell, 1996; Bäuml, 2001).

In natural environments where our visual system has evolved punctate light sources, such as the sun or light from a fire, appear mostly yellowish to the observer. Apparently, the mechanism of increment-decrement separation is optimally adjusted to this situation. It enables us to perceive surfaces approximately invariant in color and to detect the illuminant source and its corresponding color at the same time. In this sense, the incomplete adjustment to yellowish light sources may provide us with an additional, less biased cue to the color of the illuminant.

The notion that the visual system uses the spatial mean of cone excitations to derive invariant designators of surfaces across illuminant changes has been discussed extensively in the color constancy literature (Land & McCann, 1971; Buchsbaum, 1980; Brainard & Wandell, 1986; Kraft & Brainard, 1999; Gegenfurtner, 2003). However, several observations argue against the assumption that the spatial mean is the exclusive mechanism underlying color constancy (Brainard & Wandell, 1986; Brown & MacLeod, 1997; Kraft & Brainard, 1999; see also Section 3.2.2). Results from the present work suggest that taking into account the spatial mean provides a useful heuristic when adjusting to the illuminant for decremental but not for incremental stimuli. It is assumed that the integration of cone signals across the whole scene is located at higher levels in the visual path (Gegenfurtner, 2003). Hence, a constancy mechanism based on the spatial mean is presumably acting after signals from incremental and decremental

stimuli had already been separated. One may speculate if such a hypothetical spatial mean mechanism is then mainly acting on signals in the OFF-pathway. This assumption is to be tested in future research.

6.2.3 The Daylight Hypothesis

Both experimental studies did not provide evidence for the hypothesis that color constancy is better under daylight conditions than under non-daylight illumination. This result is line with findings from a number of recent studies (Brainard, 1998; Delahunt & Brainard, 2004b; Hansen, Walter & Gegenfurtner, 2007). Apparently, our visual system does not make use of the regularities of natural daylights.

6.2.4 Measures of Color Constancy

Findings from a number of studies indicate that human color constancy is far from perfect (Arend & Reeves, 1986; Kuriki & Uchikawa, 1996; Kraft & Brainard, 1999). This is not surprising if we recall the main purpose of the color constancy mechanism: the recognition of one and the same surface under different (natural) illumination conditions.[1] Color constancy is mostly obtained with quantitative measures such as asymmetric matching and achromatic setting but it is questionable if these methods provide an adequate access to human color constancy (Foster, 2003; Logvinenko & Maloney, 2006). As the assignment of color to an object rather operates on color categories than on chromaticity coordinates it would be desirable to determine the degree of color constancy using categorical methods such as color naming or hue scaling.

There are at least two problems related with color naming: First, the number of color categories is very limited if one follows for example the approach of Berlin and Kay (1969) where eleven basic color categories are assumed. Second, it is difficult to compare color naming results with traditional quantitative measures. These limitations are also relevant for the study by Troost and deWeert (1991). The authors reported good categorical color constancy of observers but they were not able to quantify the degree of color constancy.

Very recently, Hansen et al. (2007) tried to overcome the limitation of the small number of color categories by obtaining color naming data for more than 400 different surfaces. By using a close-meshed array of isoluminant stimuli they were able to trace category boundaries across different illuminations and to quantify the degree of color constancy. Hansen et al. found that color constancy of observers was nearly perfect (99%) when they were provided with spatial and temporal cues to the illuminant.

[1] Color Plate D.2 (Appendix D) gives an impressive demonstration of the strength of our color constancy mechanism.

Hue scaling represents a compromise between quantitative measures and color naming. The advantage of this method is that it does not suffer from the limitations related with color naming. The data presented here indicates that hue scaling is an appropriate technique to investigate color constancy which leads to results comparable with quantitative measures. As it is based on opponent color categories, it may provide a more adequate access to human color constancy in the spirit of Foster (2003).

6.2.5 Chromatic Adaptation

Chromatic adaptation is often understood as sensory process which is presumably located at an early stage in the visual path (Bäuml, 1997). The chromatic adaptation mechanism enables our visual system to optimally tune to ambient daylight when slow changes of illumination occur across time. The related phenomenon is referred to as successive color constancy (Bäuml, 1997; Brainard, 1998). In this sense, chromatic adaptation provides a fundamental mechanism which maintains successive color constancy.

The contribution of the chromatic adaptation mechanism to color constancy is often underestimated in the color constancy literature. Elaborated models of color constancy, such as Lands Retinex model (Land & McCann, 1971) or the Ratio model by Foster and coworkers (Foster & Nascimento, 1994; Foster, Amano & Nascimento, 2001b) focus only on spatial characteristics of the scene. In these models temporal aspects of color constancy are largely ignored.

The results of the experiments presented in Chapter 5 demonstrate the strength of the chromatic adaptation mechanism in isolation. Moreover they indicate that our visual system is able to integrate information from isolated stimuli over time. A similar effect in scenes where single test surfaces were presented in sequence against a fixed background has been reported by Smithson and Zaidi (2004).

Recently, the effect of chromatic adaptation on color constancy has been investigated by Hansen et al. (2007). They estimated that the contribution of chromatic adaptation to color constancy is about 15% when temporal and spatial cues are given to the observer. The isolated effect of chromatic adaptation that I measured was even about 50% (see Figure 5.9).

6.3 Outlook

In the present work different aspects of human color constancy were investigated. In the remaining section open questions that are related with these aspects and directions for future research will be outlined.

There is strong evidence that the reduction hypothesis of dichromacy does not hold in general. Hence it seems necessary to test alternative explanations, namely differences in optical density of photopigments and the contribution of

cognitive factors. One approach may be to investigate carefully if differences in optic density of pigment have an effect on color matching performance of dichromats. Another goal of future research should be to separate genetic based aspects from acquired cognitive factors of dichromatic color vision.

A number of studies provided evidence for the assumption that separation of incremental and decremental signals is a fundamental principle of vision which may enable us to distinguish illumination and surfaces (Walraven, 1977; Mausfeld & Niederée, 1993; Bäuml, 1997). The results presented here suggest that the spatial mean may serve as a useful heuristic to maintain stable perception of surface colors. Therefore, the relationship between the separate processing of incremental and decremental stimuli and the use of the spatial mean as cue to the illuminant should be investigated more systematically.

The results presented in Chapter 4 indicate that differences in the processing of incremental and decremental stimuli increase with decreasing correlated color temperature of daylights. The visual system could in principle use this information as cue to the illumination in the scene. This question is to be tested in elaborated experiments.

In recent years, the notion that there is no single mechanism which accounts exclusively for the phenomenon of color constancy has gained acceptance (Kraft & Brainard, 1999). This idea is also reflected by the results of the experiments presented here. A promising approach which considers the possibility of a dynamic interplay of several mechanisms is the cue combination framework by Maloney (Yang & Maloney, 2001; Maloney, 2002; Boyaci, Doerschner & Maloney, 2006). This approach can serve as a starting point to systematically test the validity of several cues, interactions between different cues and conditions under which they are used.

Bibliography

Abramov, I. & Gordon, J. (1994). Color appearance: on seeing red-or yellow, or green, or blue. *AnnuRevPsychol*, *45*, 451–485.

Abramov, I., Gordon, J. & Chan, H. (1990). Using hue scaling to specify color appearances and to derive color differences. *Perceiving, Measuring, and Using Color. Proceedings of SPIE.*, *1250*, 40–51.

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

Arend, L. E., Reeves, A., Schirillo, J. & Goldstein, R. (1991). Simultaneous color constancy: paper with diverse munsell values. *Journal of the Optical Society of America A*, *8*, 661–672.

Augenstein, E. J. & Pugh, E. N. J. (1977). The dynamics of the pi 1 colour mechanism: further evidence for two sites of adaptation. *Journal of Physiology*, *272*, 247–281.

Bäuml, K. H. (1994). Color appearance: effects of illuminant changes under different surface collections. *Journal of the Optical Society of America A*, *11*, 531–542.

Bäuml, K. H. (1995). Illuminant changes under different surface collections: examining some principles of color appearance. *Journal of the Optical Society of America A*, *12*, 261–271.

Bäuml, K. H. (1997). *Sensory and perceptual mechanisms of color constancy.* Unveröffentlichte Habilitationsschrift. Universität Regensburg.

Bäuml, K. H. (1999a). Color constancy: the role of image surfaces in illuminant adjustment. *Journal of the Optical Society of America A*, *16*, 1521–1530.

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

Bäuml, K. H. (2001). Increments and decrements in color constancy. *Journal of the Optical Society of America A*, *18*, 2419–2429.

Berlin, B. & Kay, P. (1969). *Basic color terms: Their universality and evolution.* Berkeley: University of California Press.

Bonnardel, V. (2006). Color naming and categorization in inherited color vision deficiencies. *Visual Neuroscience*, *23*, 637–643.

Boyaci, H., Doerschner, K. & Maloney, L. T. (2006). Cues to an equivalent lighting model. *Journal of Vision*, *6(2)*, 106–118.

Boyaci, H., Maloney, L. T. & Hersh, S. (2003). The effect of perceived surface orientation on perceived surface albedo in three-dimensional scenes. *Journal of Vision*, *3*, 541–553.

Boynton, R. M. & Gordon, J. (1965). Bezold-brücke hue shift measured by color-naming technique. *Journal of the Optical Society of America*, *55*, 78–86.

Boynton, R. M. & Olson, C. X. (1987). Locating basic colors in the osa space. *Color Res. Applic.*, *12*, 94–105.

Brainard, D. H. (1989). Calibration of a computer controlled color monitor. *Color Res. Applic.*, *14*, 23–34.

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

Brainard, D. H. (2004). Color constancy. In L. Chalupa & J. Werner (Eds.), *The Visual Neurosciences* (pp. 948–961). MIT Press.

Brainard, D. H., Brunt, W. A. & Speigle, J. M. (1997). Color constancy in the nearly natural image. I. asymmetric matches. *Journal of the Optical Society of America A*, *14*, 2091–2110.

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.

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*, 1433–1447.

Brill, M. H. (1978). A device performing illuminant-invariant assessment of chromatic relations. *Journal of Theoretical Biology*, *71*, 473.

Brown, R. O. & MacLeod, D. I. A. (1997). Color appearance depends on the variance of surround colors. *Current Biology*, *7*, 844–849.

Brunswik, E. (1929). Zur Entwicklung der Albedowahrnehmung. *Zeitschrift für Psychologie*, *109*, 40–115.

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

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

Chichilnisky, E.-J. & Wandell, B. A. (1995). Photoreceptor sensitivity changes explain color appearance shifts induced by large uniform backgrounds in dichoptic matching. *Vision Research*, *35*, 239–254.

Chichilnisky, E.-J. & Wandell, B. A. (1996). Seeing gray through the ON and OFF pathways. *Visual Neuroscience*, *13*, 591–596.

Cohen, J. (1964). Dependency of the spectral reflectance curves of the munsell color chips. *Psychonomic Science*, *1*, 369–370.

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

Crognale, M. A., Teller, D. Y., Yamaquchi, T., Motulsky, A. G. & Deeb, S. S. (1999). Analysis of red/green color discrimination in subjects with a single x-linked photopigment gene. *Vision Research*, *39*, 707–719.

Dannemiller, J. L. (1993). Rank orderings of photoreceptor photon catches from natural objects are nearly illuminant-invariant. *Vision Research*, *33*, 131–140.

De Valois, R. L., Abramov, I. & Jacobs, G. H. (1966). Analysis of response patterns of lgn cells. *Journal of the Optical Society of America*, *56*, 966–977.

De Valois, R. L., De Valois, K. K., Switkes, E. & Mahon, L. (1997). Hue scaling of isoluminant and cone-specific lights. *Vision Research*, *37*, 885–897.

Delahunt, P. B. & Brainard, D. H. (2000). Control of chromatic adaptation: signals from seperate cone classes interact. *Vision Research*, *40*, 2885–2903.

Delahunt, P. B. & Brainard, D. H. (2004a). Color constancy under changes in reflected illumination. *Journal of Vision*, *4*, 764–778.

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

Derrington, A. M., Krauskopf, J. & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology*, *357*, 241–265.

D'Zmura, M. & Lennie, P. (1986). Mechanisms of color constancy. *Journal of the Optical Society of America A*, *3*, 1662–1672.

Eisner, A. & MacLeod, D. I. (1980). Blue-sensitive cones do not contribute to luminance. *Journal of the Optical Society of America*, *70*, 121–123.

Ekroll, V., Faul, F. & Niederée, R. (2004). The peculiar nature of simultaneous colour contrast in uniform surrounds. *Vision Research*, *44*, 1765–1786.

Endler, J. A. (1993). The color of light in forests and its implications. *Ecological Monographs*, *63*, 1–27.

Ernst, M. O. & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433.

Fairchild, M. D. & Reniff, L. (1995). Time course of chromatic adaptation for color-appearance judgements. *Journal of the Optical Society of America A*, *12*, 824–833.

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

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

Foster, D. H., Amano, K. & Nascimento, S. M. C. (2001b). How temporal cues can aid color constancy. *Color Res. Applic.*, *26*, S180–S185.

Foster, D. H. & Linnell, K. J. (1995). Evidence for relational colour constancy in red-green colour-deficient human observers. *Journal of Physiology*, *485.P*, 23P–24P.

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

French, J. W. (1921). The interocular distance. *Transactions of the Optical Society.*, *23*, 44–55.

Gegenfurtner, K. R. (2003). Cortical mechanisms of colour vision. *NatRevNeurosci*, *4*, 563–572.

Gegenfurtner, K. R., Wichmann, F. A. & Sharpe, L. T. (1998). The contribution of color to visual memory in X-chromosome-linked dichromats. *Vision Research*, *38*, 1041–1045.

Gelb, A. (1929). Die Farbenkonstanz der Sehdinge. In A. Bethe, G. von Bergman, G. Embden & A. Ellinger (Eds.), *Handbuch der normalen und pathologischen Physiologie* (pp. 594–687). Berlin: Springer.

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

Grassmann, H. (1853). Zur Theorie der Farbenmischung. *Poggendorffs Annalen der Physik und Chemie*, *89*, 69–84.

Halir, R. & Flusser, J. (1998). Numerically stable direct least squares fitting of ellipses. In V. Skala (Ed.), *Proceedings of the International Conference in Central Europe on Computer Graphics and Visualisation* (pp. 125–132). WSCG98: Plzen.

Hansen, T. & Gegenfurtner, K. R. (2006). Color scaling of discs and natural objects at different luminance levels. *Visual Neuroscience*, *23*, 603–610.

Hansen, T., Olkkonen, M., Walter, S. & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience*, *9(11)*, 1367–1368.

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

Hecht, S., Schlaer, S. & Pirenne, M. H. (1942). Energy, quanta, and vision. *Journal of General Physiology*, *25*, 819 – 840.

Helmholtz, H. L. F. v. (1867). *Handbuch der physiologischen Optik*. Leipzig: Leopold Voss.

Hering, E. (1920). *Grundzüge der Lehre vom Lichtsinn*. Berlin: Springer.

Heyer, D. (1997). *Psychophysikalische Modelle der primären Farbcodierung. Experimentelle und theoretische Studien zur Beziehung von Kontrastkodierung und opponenten Mechanismen.* Habilitationsschrift, Philosophische Fakultät der Christian-Albrechts-Universität Kiel.

Hurlbert, A. & Wolf, K. (2004). Color contrast: a contributory mechanism to color constancy. *Perception and Psychophysics*, *144*, 147–160.

Hurvich, L. M. (1981). *Color vision*. Sunderland: Sinauer.

Hurvich, L. M. & Jameson, D. (1955). Some quantitative aspects of an opponent-colors theory. II. brightness, saturation, and hue in normal and dichromatic vision. *Journal of the Optical Society of America*, *45*, 602–616.

Ishihara, S. (1997). *Ishihara's Test for Colour Deficiency*. Tokyo: Kanehara and Co.

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.

Jameson, D. & Hurvich, L. M. (1978). Dichromatic color language: "reds" and "greens" don't look alike but their colors do. *Sens Processes*, *2(2)*, 146–155.

Jameson, D. J. & Hurvich, L. M. (1989). Essay concerning color constancy. *Annual Review of Psychology*, *40*, 1–22.

Johnson, R. A. & Wichern, D. W. (2002). *Applied Multivariate Statistical Analysis*. Upper Saddle River: Prentice Hall.

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.

Kaiser, P. K. & Boynton, R. M. (1996). *Human color vision* (2 Ed.). Washington DC: Optical Society of America.

Katz, D. (1911). *Die Erscheinungsweisen der Farben und ihre Beeinflussung durch die individuelle Erfahrung.* Leipzig: Barth.

Katz, D. (1935). *The World of Colour*. London: Kegan Paul.

Kay, P. & Regier, T. (2003). Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences of the USA*, *100*, 9085–9089.

Knoblauch, K., Sirovich, L. & Wooten, B. R. (1985). Linearity of hue cancellation in sex-linked dichromacy. *Journal of the Optical Society of America A*, *2*, 136–146.

König, A. & Dieterici, C. (1893). Die Grundempfindungen in normalen und anomalen Farbensystemen und ihre Intensitätsverteilungen im Spektrum. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *4*, 241–247.

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

Krantz, D. H. (1975). Color measurement and color theory: I. representation theorem for grassmann structures. *Journal of Mathematical Psychology*, *12*, 283–303.

Krauskopf, J., Williams, D. R. & Heeley, D. W. (1982). Cardinal directions of color space. *Vision Research*, *22*, 1123–1131.

Kries, J. v. (1905). Die Gesichtsempfindungen. In W. Nagel (Ed.), *Handbuch der Physiologie des Menschen*. Braunschweig: Vieweg.

Kuriki, I. & Uchikawa, K. (1996). Limitations of surface color and apparent color constancy. *Journal of the Optical Society of America A*, *13*, 1622–1636.

Lakowski, R. & Oliver, K. (1979). Discrimination steps of dichromats. Technical report, Visual Laboratory, Psychology Department, U.B.C., Vancouver, Canada.

Lakowski, R., Young, K. & Kozak, J. (1982). The SPP test-a quantitative analysis. *Docum Ophthal Proc Ser*, *33*, 191–197.

Land, E. H. (1977). The retinex theory of color vision. *Scientific American*, *218*, 108–128.

Land, E. H. (1986). Recent advances in retinex theory. *Vision Research*, *26*, 7–21.

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

Landy, M. S., Maloney, L. T., Johnston, E. B. & Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research*, *35*, 389–412.

Larimer, J., Krantz, D. & Cicerone, C. (1974). Opponent process additivity - I. red/green equilibria. *Vision Research*, *14*, 1127–1140.

Larimer, J., Krantz, D. H. & Cicerone, C. M. (1975). Opponent process additivity. II. yellow/blue equilibria and nonlinear models. *Vision Research*, *15(6)*, 146–155.

Larson, G. W. & Shakespeare, R. (1996). *Rendering with radiance; The art and science of lighting and visualization.* San Francisco: Morgan Kaufmann Publishers, Inc.

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.

Lee, H.-C., Breneman, E. J. & Schulte, C. P. (1990). Modeling light reflection for computer color vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*, 402–409.

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

Logvinenko, A. D. & Maloney, L. T. (2006). The proximity structure of achromatic surface colors and the impossibility of asymmetric lightness matching. *Perception and Psychophysics*, *68*, 76–83.

Lotto, R. B. (2004). Visual development: experience puts the colour in life. *Current Biology*, *14*, 619–621.

Lotto, R. B. & Purves, D. (2004). Perceiving color. *Rev Prog Coloration*, *34*, 12–25.

MacAdam, D. L. (1970). *Sources of Color Vision.* Cambridge, Massachusetts: MIT Press.

Maloney, L. T. (1984). *Computational Approaches to Color Constancy.* PhD thesis, Stanford University, Stanford, CA.

Maloney, L. T. (1986). Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *Journal of the Optical Society of America A*, *3*, 1673–1683.

Maloney, L. T. (1999). Physics-based approaches to modeling surface color perception. In K. R. Gegenfurtner & L. T. Sharpe (Eds.), *Color Vision: From Genes to Perception* (pp. 387–422). Cambridge, UK: Cambridge University Press.

Maloney, L. T. (2002). Illuminant estimation as cue combination. *Journal of Vision*, *2*, 493–504.

Maloney, L. T. & Landy, M. S. (1989). A statistical framework for robust fusion of depth information. In W. Pearlman (Ed.), *Visual Communications and Image Processing* (pp. 1154–1163). IV Proceedings of the SPIE.

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*, 29–32.

Mausfeld, R. (1998). Color perception: From grassmann codes to a dual code for object and illumination colors. In W. Backhaus, R. Kliegl & J. S. Werner (Eds.), *Color vision Perspectives from Different Disciplines* (pp. 219–250). Berlin: De Gruyter.

Mausfeld, R. & Andres, J. (2002). Second-order statistics of colour codes modulate transformations that effectuate varying degrees of scene invariance and illumination invariance. *Perception*, *31*, 209–224.

Mausfeld, R. & Niederée, R. (1993). Inquiries into relational concepts of colour based on an incremental principle of colour coding for minimal relational stimuli. *Perception*, *22*, 427–462.

Maxwell, S. E. & Delaney, H. D. (2003). *Designing experiments and analyzing data: A model comparison perspective* (2 Ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

McCann, J. J., McKee, S. P. & Taylor, T. H. (1976). Quantitative studies in retinex theory: A comparision between theoretical predictions and observer responses to the 'Color mondrian' experiments. *Vision Research*, *16*, 445–458.

McLellan, J. S. & Eskew, R. T. (2000). ON and OFF S-cone pathways have different long-wave cone inputs. *Vision Research*, *40*, 2449–2465.

McMahon, M. J. & MacLeod, D. I. A. (1998). Dichromatic color vision at high light levels: red/green discrimination using the blue-sensitive mechanism. *Vision Research*, *38*, 973–983.

Montag, E. D. (1994). Surface color naming in dichromats. *Vision Research*, *34*, 2137–2151.

Montag, E. D. & Boynton, R. M. (1987). Rod influence in dichromatic surface color perception. *Vision Research*, *27*, 2153–2162.

Morgan, M. J., Adam, A. & Mollon, J. D. (1992). Dichromats detect colour-camouflaged objects that are not detected by trichromats. *Proceedings of the Royal Society of London B*, *248*, 291–295.

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

Nascimento, S. M. C. & Foster, D. H. (1997). Detecting natural changes of cone-exitation ratios in simple and complex coloured images. *Proceedings of the Royal Society of London B*, *264*, 1395–1402.

Nathans, J., Piantanida, T. P., Eddy, R. L., Shows, T. B. & Hogness, D. S. (1986). Molecular genetics of inherited variation in human color vision. *Science*, *232*, 203–210.

Neitz, J., Neitz, M., He, J. C. & Shevell, S. K. (1999). Trichromatic color vision with only two spectrally distinct photopigments. *Nat Neurosci*, *2(10)*, 884–888.

Neitz, M. & Neitz, J. (2000). Molecular genetics of color vision and color vision defects. *Arch Ophthalmol*, *118(5)*, 691–700.

Neitz, M., Neitz, J. & Jacobs, G. H. (1995). Genetic basis of photopigment variations in human dichromats. *Vision Research*, *35*, 2095–2103.

Paramei, G. V., Bimler, D. L. & Cavonius, C. R. (1998). Effect of luminance on color perception of protanopes. *Vision Research*, *38*, 3397–3401.

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.

Pokorny, J., Smith, V. C., Verriest, G. & Pinckers, A. J. L. G. (1979). *Congenital and acquired color vision defects.* New York: Grune and Stratton.

Reeves, A. (1982). Exchange thresholds for green tests. *Vision Research*, *22*, 961–966.

Richter, E. (2002). *Modelle der Kontrastcodierung.* PhD thesis, Martin-Luther-Universität Halle-Wittenberg.

Rinner, O. & Gegenfurtner, K. R. (2000). Time course of chromatic adaptation for color appearance and discrimination. *Vision Research*, *40*, 1813–1826.

Rinner, O. & Gegenfurtner, K. R. (2002). Cone contribution to colour constancy. *Perception*, *31*, 733–746.

Rüttiger, L., Braun, D. I., Gegenfurthner, K. R., Petersen, D., Schönle, P. & Sharpe, L. T. (1999). Selective colour constancy deficits after circumscribed unilateral brain lesions. *Journal of Neuroscience*, *19*, 3094–3106.

Rüttiger, R., Mayser, H., Serey, L. & Sharpe, L. T. (2001). The color constancy of the red-green color blind by. *Color Research and Application*, *26*, 209–213.

Scheibner, H. M. O. & Boynton, R. M. (1968). Residual red-green discrimination in dichromats. *Journal of the Optical Society of America A*, *58*, 1151–1158.

Scheifler, R. W. & Gettys, J. (1996). *X Window system: Core library and standards.* Boston: Digital Press.

Schiller, P. H., Sandell, J. H. & Maunsell, J. H. (1986). Functions of the on and off channels of the visual system. *Nature*, *322*, 824–825.

Schnapf, J. L., Kraft, T. W. & Baylor, D. A. (1987). Spectral sensitivity of human cone photoreceptors. *Nature*, *325*, 439–441.

Schrödinger, E. (1920). Grundlinien einer Theorie der Farbenmetrik im Tagessehen. Erste und Zweite mitteilung. *Annalen der Physik*, *63*, 397–456.

Shafer, S. A. (1985). Using color to separate reflectance components. *Color Research and Application*, *10*, 210–218.

Sharpe, L. T., Stockman, A., Jägle, H., Knau, H., Klausen, G., Reitner, A. & Nathans, J. (1998). Red, green, and red-green hybrid pigments in the human retina: correlations between deduced protein sequences and psychophysically measured spectral sensitivities. *Journal of Neuroscience*, *18*, 10053–10069.

Shepard, R. (1994). Perceptual-cognitive universals as reflections of the world. *Psychonomic Bulletin & Review*, *1*, 2–28.

Shevell, S. K. (2001). The time course of chromatic adaptation. *Color Res. Applic.*, *26*, S170–S173.

Shevell, S. K., He, J. C., Kainz, P., Neitz, J. & Neitz, M. (1998). Relating color discrimination to photopigment genes in deutan observers. *Vision Research*, *38*, 3371–3376.

Simunovic, M. P., Regan, B. C. & Mollon, J. D. (2001). Is color vision deficiency an advantage under scotopic conditions? *Investigative Ophthalmology and Visual Science*, *42*, 3357–3364.

Smith, V. C. & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research*, *15*, 161–171.

Smith, V. C. & Pokorny, J. (1977). Large-field trichromacy in protanopes and deuteranopes. *Journal of the Optical Society of America*, *2(2)*, 213–220.

Smithson, H. & Zaidi, Q. (2004). Colour constancy in context: Roles for local adaptation and levels of reference. *Journal of Vision*, *4*, 693–710.

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

Snyder, J. L., Doerschner, K. & Maloney, L. T. (2005). Illumination estimation in three-dimensional scenes with and without specular cues. *Journal of Vision*, *5(10)*, 863–877.

Spectral Database. University of Joensuu Color Group. http://spectral.joensuu.fi/.

Speigle, J. M. & Brainard, D. H. (1999). Predicting color from gray: the relationship between achromatic adjustment and asymmetric matching. *Journal of the Optical Society of America A*, *16*, 2370–2376.

Sugita, Y. (2004). Experience in early infancy is indispensable for color perception. *Current Biology*, *14*, 1267–1271.

Svaetichin, G. (1956). Spectral response curves from single cones. *Acta Physiol Scand Suppl*, *39(134)*, 17–46.

Troost, J. M. & de Weert, C. M. (1991). Naming versus matching in color constancy. *Perception and Psychophysics*, *50(6)*, 591–602.

Valberg, A. (2001). Unique hues: an old problem for a new generation. *Vision Research*, *41*, 1645–1657.

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.

Verhulst, S. & Maes, F. W. (1998). Scotopic vision in colour-blinds. *Vision Research*, *38*, 3387–3390.

Wachtler, T., Dohrmann, U. & Hertel, R. (2004). Modeling color percepts of dichromats. *Vision Research*, *44*, 2843–2855.

Walraven, J. (1976). Discounting the background - the missing link in the explanation of chromatic induction. *Vision Research*, *16*, 289–295.

Walraven, J. (1977). Colour signals from incremental and decremental light stimuli. *Vision Research*, *17*, 71–76.

Walraven, J., Benzschawel, T. L. & Rogowitz, B. E. (1989). Color-constancy interpretation of chromatic induction. *Die Farbe*, *34*, 269–273.

Walraven, J., Enroth-Cugell, C., Hood, D. C., MacLeod, D. I. A. & Schnapf, J. L. (1990). The control of visual sensitivity. In L. Spillmann & J. S. Werner (Eds.), *Visual perception:the neurophysiological foundations*. San Diego: Academic.

Webster, M. A., Miyahara, E., Malkoc, G. & Raker, V. E. (2000). Variations in normal color vision. ii. unique hues. *Journal of the Optical Society of America A*, *17*, 1545–1555.

Werner, A., Sharpe, L. T. & Zrenner, E. (2000). Asymmetries in the time-course of chromatic adaptation and the significance of contrast. *Vision Research*, *40*, 1101–1113.

Werner, J. S. & Wooten, B. R. (1979). Opponent chromatic mechanisms: relation to photopigments and hue naming. *Journal of the Optical Society of America*, *69(3)*, 422–434.

Worthey, J. A. (1985). Limitations of color constancy. *Journal of the Optical Society of America A*, *2*, 1014–1026.

Wyszecki, G. & Stiles, W. S. (1982). *Color science-concepts and methods, quantitative data and formulae* (2 Ed.). New York: Wiley.

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

Appendix A

Approximating Daylight Spectra

Blackbody Radiator

Spectral power distributions of thermal radiators are functions of the wavelength λ and depend on the characteristics of the radiator and on its temperature. In physics the concept of the *blackbody radiator* has proven useful. Such a radiator can be imagined as an idealized black object for which the spectral power distribution of emitted light depends only on one parameter, the temperature T of the blackbody radiator. This temperature T is termed *color temperature* of the blackbody radiator and it is taken on the absolute scale in Kelvin ($K =^{\circ} C + 273.15$). Corresponding chromaticities of the blackbody radiator at different color temperatures can be arranged on a curve in xy-chromaticity diagram which is called the *Planckian locus* (Figure A.1).

Correlated Color Temperature

The chromaticities of many light sources, including natural daylights, do not match exactly with any of the chromaticities along the Planckian locus. In this case light sources are often characterized by their *correlated color temperature* T_c. The concept of correlated color temperature is defined as the corresponding color temperature of the blackbody radiator which matches most closely with the given light source under the same brightness and specified viewing conditions. The correlated color temperature of a given light source with chromaticity close to the Planckian locus can be determined using *isotemperature lines*. These lines can be understood as straight lines in xy-chromaticity diagram crossing the Planckian locus (Figure A.1). Each isotemperature line represents chromaticities of light sources with the same correlated color temperature.

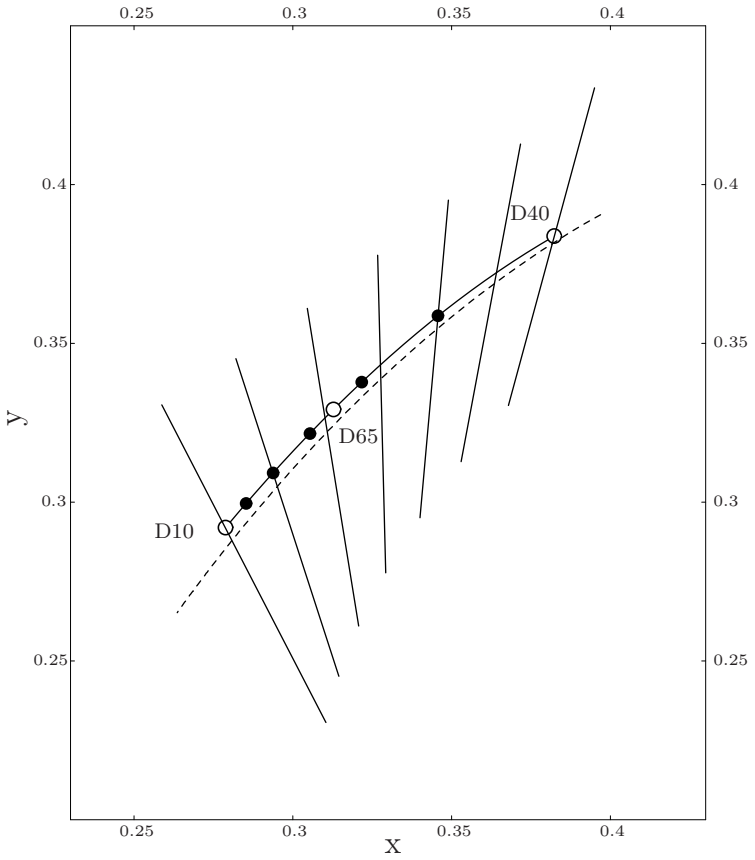

Figure A.1: Daylight locus (solid curve with open and filled circles) and Planckian locus (dashed curve) in xy-chromaticity diagram. The daylights D40 (4000 K), D65 (6500 K) and D10 (10000 K) that were used throughout this work are shown as open circles. The closed circles represent daylights of 5000 K, 6000 K, 7000 K, 8000 K and 9000 K correlated color temperature. The solid lines crossing both loci show selected isotemperature lines.

Daylight Reconstruction

Natural daylights consist of two components, the direct light from sun and sunlight scattered by the atmosphere. Many measurements of daylight spectra at different times of the day and at different locations have been reported in the literature. Detailed analyses of these measurements have shown that the spectral power distributions of natural daylights follow certain regularities (Judd, MacAdam & Wyszecki, 1964). For example, these distributions are fairly smooth and show similar patterns of variation. Moreover, chromaticities of natural daylights can be arranged on a curve close to the Planckian locus in xy-chromaticity diagram (Figure A.1). Hence, daylights are often characterized by corresponding correlated color temperatures. Roughly speaking, daylights which represent a

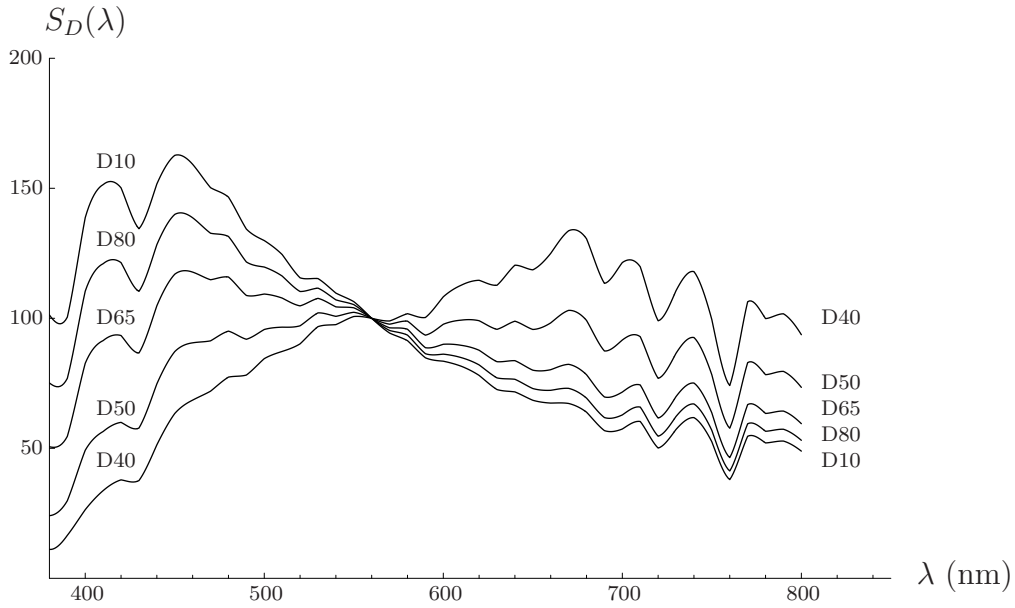

Figure A.2: Reconstructed relative spectral power distributions of daylights D40, D50, D65, D80 and D10 using the CIE Method of calculating D-illuminants.

mixture of direct sunlight and light from the sky have correlated color temperatures in the range of 5000–7000 K. Daylight from the sky with the sun occluded results in correlated color temperatures above 7000 K. Direct sunlight without the sky component or daylights at dusk or dawn have color temperatures below 5000 K.

The found regularities of natural daylights led to the idea to approximate spectral power distributions of daylights using only a small number of parameters. Judd et al. performed a Principal Components Analysis (PCA) on a large set of measured daylight spectra.[1] They were able to reconstruct spectral power distributions of daylights extremely well. Reconstituted daylight spectra were expressed as linear combinations of only three basis functions. These basis functions are the first three eigenvectors determined by the PCA. The method of reconstructing daylight spectra suggested by Judd et al. has been adopted by the CIE and is often referred to as the *CIE Method of calculating D-illuminants*. As this method was used throughout the present work for approximating daylight spectra I will now describe this method in some more detail. Using this method we only need the correlated color temperature T_c of the given daylight and the eigenvectors $S_0(\lambda)$, $S_1(\lambda)$ and $S_2(\lambda)$ in order to derive an estimate of the spectral power distribution. From T_c we can calculate chromaticity coordinates (x_D, y_D) of the given daylight. For correlated color temperatures of 4000–7000 K, x_D is

[1]See Johnson and Wichern (2002) for an introduction to the method of PCA.

given by:

$$x_D = -4.6070\frac{10^9}{T_c^3} + 2.9678\frac{10^6}{T_c^2} + 0.09911\frac{10^3}{T_c} + 0.244063. \tag{A.1}$$

If the correlated color temperature is in the range of 7000–25000 K, x_D is:

$$x_D = -2.0064\frac{10^9}{T_c^3} + 1.9018\frac{10^6}{T_c^2} + 0.24748\frac{10^3}{T_c} + 0.237040. \tag{A.2}$$

From x_D we get y_D with:

$$y_D = -3x_D^2 + 2.87x_D - 0.275. \tag{A.3}$$

Natural daylights lie approximately on a curve close to the Planckian locus which is defined by (A.3). This curve is often referred to as the *daylight locus*. The spectral power distribution $S_D(\lambda)$ of the given daylight can be approximated using the eigenvectors $S_0(\lambda)$, $S_1(\lambda)$ and $S_2(\lambda)$:[2]

$$S_D(\lambda) = S_0(\lambda) + M_1 S_1(\lambda) + M_2 S_0(\lambda), \tag{A.4}$$

where M_1 and M_2 are scalars with

$$M_1 = \frac{-1.3515 - 1.7703x_D + 5.9114y_D}{0.0241 + 0.2562x_D - 0.7341y_D} \tag{A.5}$$

and

$$M_2 = \frac{0.03 - 31.4424x_D + 30.0717y_D}{0.0241 + 0.2562x_D - 0.7341y_D}. \tag{A.6}$$

Examples of reconstructed daylight spectra using the CIE Method of calculating D-illuminants are shown in Figure A.2. An excellent source for further reading is the book by Wyszecki and Stiles (1982).

[2]In previous chapters I denoted approximated daylight spectra with $\hat{E}(\lambda)$. I chose this notation in the text because it is common in the color constancy literature to denote illuminants with E and surfaces with S. The notation $S_D(\lambda)$ refers to terminology used by the CIE.

Appendix B

Stimuli and Simulated Illuminants

Table B.1: xy- and $u'v'$-chromaticities of the illuminants used in the experiments described in Chapter 4.

illuminant	x	y	u'	v'
D40	0.382	0.384	0.223	0.505
D65	0.313	0.329	0.198	0.468
D10	0.279	0.292	0.188	0.442
D45	0.362	0.371	0.215	0.496
D50	0.346	0.359	0.209	0.488
D55	0.333	0.348	0.205	0.481
D60	0.322	0.338	0.201	0.474
D70	0.305	0.322	0.195	0.463
D80	0.294	0.309	0.192	0.454
D90	0.285	0.300	0.189	0.448
green	0.268	0.347	0.162	0.471
red	0.354	0.313	0.234	0.466
green (confusion)	0.265	0.339	0.162	0.466
red (confusion)	0.359	0.323	0.233	0.472

Table B.2: Munsell notation, chromaticity coordinates and luminances (under $D65$) of the stimuli used in the experiments described in Chapter 4.

Munsell	x	y	L
5R 6/1	0.325	0.329	22.94
5R 5/4	0.373	0.332	15.15
5YR 4/1	0.336	0.340	9.62
7.5YR 5/4	0.393	0.377	15.49
10YR 6/2	0.348	0.359	23.87
5Y 4/2	0.353	0.375	9.53
10Y 6/1	0.327	0.350	23.25
7.5GY 5/2	0.316	0.365	15.63
10GY 4/1	0.310	0.343	9.59
2.5G 5/4	0.293	0.375	15.20
5G 6/2	0.298	0.347	23.91
10G 4/2	0.290	0.342	9.73
5BG 6/1	0.300	0.330	23.87
5BG 5/4	0.261	0.333	15.43
5B 4/1	0.297	0.323	9.45
7.5B 5/4	0.254	0.295	15.27
10B 6/2	0.289	0.314	23.35
5PB 4/2	0.284	0.296	9.54
10PB 6/1	0.306	0.316	23.67
10PB 5/4	0.286	0.278	15.00
10P 4/1	0.316	0.316	9.47
2.5RP 5/4	0.332	0.299	15.81
5RP 6/2	0.327	0.319	22.61
10RP 4/2	0.342	0.321	9.90

Table B.3: xy- and $u'v'$-chromaticities of the illuminants used in the experiments described in Chapter 5.

illuminant	x	y	u'	v'
D40	0.382	0.384	0.223	0.505
D65	0.313	0.329	0.198	0.468
D10	0.279	0.292	0.188	0.442
green	0.268	0.347	0.231	0.454
red	0.354	0.313	0.163	0.483

Table B.4: Munsell notation, chromaticity coordinates and luminances (under $D65$) of the stimuli used in the experiments described in Chapter 5.

Munsell	x	y	L	u'	v'
5R7/4	0.357	0.333	18.620	0.228	0.477
10R7/4	0.366	0.346	19.150	0.228	0.485
5YR7/4	0.373	0.362	18.820	0.226	0.494
10YR7/4	0.375	0.383	18.830	0.219	0.504
5Y7/4	0.369	0.397	18.370	0.210	0.509
10Y7/4	0.360	0.405	18.740	0.201	0.511
5GY7/4	0.341	0.400	17.880	0.192	0.506
10GY7/4	0.313	0.387	18.310	0.178	0.496
5G7/4	0.290	0.365	17.970	0.170	0.483
10G7/4	0.281	0.353	18.360	0.168	0.476
5BG7/4	0.271	0.338	17.960	0.166	0.467
10BG7/4	0.265	0.323	18.120	0.167	0.458
5B7/4	0.262	0.307	18.240	0.170	0.449
10B7/4	0.272	0.301	18.360	0.179	0.446
5P7/4	0.304	0.296	18.560	0.205	0.448
5RP7/4	0.337	0.314	18.260	0.222	0.464

Appendix C

Supplementary Data

Table C.1: Estimated confusion points (xy) and normalized missing fundamentals (LMS; Smith & Pokorny, 1975) of the dichromatic observers who took part in the experiments described in Chapter 4.

Observer	Confusion point		Missing fundamental			Class of
	x	y	L	M	S	dichromat
LE	0.755	0.254	1	-0.005	-0.0003	Protanope
MB	0.798	0.243	0.9985	-0.0552	-0.0013	Protanope
SS	0.664	0.254	0.9979	0.0654	0.0027	Protanope

Table C.2: Estimated S-cone parameters from Experiment 1 (Chapter 4). The parameters are S-cone rescaling coefficients of the three models ID model (ID; parameters are listed for decrements and increments), diagonal model (DIA) and spatial mean model (MEAN). They are given for each subject and illuminant condition. Subjects KH, SA and TK are normal observers and subjects SS, MB and LE are protanopes.

condition	sub.	ID (D)	ID (I)	DIA	MEAN
Green	KH	1.02	1.02	1.02	1.04
	SA	1.02	1.02	1.02	1.04
	TK	1.02	1.01	1.02	1.04
	SS	1.04	0.98	1	1.04
	MB	1.04	0.99	1.01	1.04
	LE	1.02	0.97	1.01	1.04
Red	KH	0.98	0.99	0.99	1
	SA	0.98	0.99	0.98	1
	TK	0.98	0.99	0.98	1
	SS	1	0.94	0.97	1
	MB	0.99	0.98	0.98	1
	LE	0.97	1	0.99	1
D40	KH	0.57	0.84	0.71	0.57
	SA	0.67	0.79	0.7	0.57
	TK	0.56	0.72	0.62	0.57
	SS	0.43	0.72	0.62	0.57
	MB	0.56	1.07	0.75	0.57
	LE	0.61	0.97	0.77	0.57
D65	KH	0.96	1.14	1.05	1
	SA	1.01	1.05	1.03	1
	TK	1.03	1.01	0.97	1
	SS	0.84	1.15	1.02	1
	MB	0.94	1.31	1.09	1
	LE	1.24	1.38	1.24	1
D10	KH	1.34	1.44	1.39	1.35
	SA	1.4	1.42	1.39	1.35
	TK	1.39	1.29	1.29	1.35
	SS	1.13	1.33	1.29	1.35
	MB	1.14	1.43	1.33	1.35
	LE	1.24	1.55	1.47	1.35

Table C.3: Color constancy indices of subjects for all illumination conditions in Experiments 1 and 2 (Chapter 4). Subjects KH, SA and TK are normal observers and subjects SS, MB and LE are protanopes.

Illumination	KH	TK	SA	LE	MB	SS
Blue (D10)	0.92	0.82	0.88	0.66	0.60	0.83
Yellow (D40)	0.74	0.89	0.75	0.61	0.70	0.88
Red (Exp. 1)	0.69	0.86	0.82	0.60	0.79	0.60
Red (Exp. 2)	–	–	0.65	0.28	0.46	–0.03
Green (Exp. 1)	0.93	0.75	0.94	0.32	0.61	0.17
Green (Exp. 2)	–	–	0.91	–0.13	0.73	0.13

Table C.4: Estimated S-cone parameters from Experiment 3 (Chapter 4). The parameters are S-cone rescaling coefficients of the three models ID model (ID; parameters are listed for decrements and increments), diagonal model (DIA) and spatial mean model (MEAN). They are given for each subject and illuminant condition. Subjects KH, LA, KUH, RL and SA are normal observers and subjects SS, MB and LE are protanopes.

		ID (D)	ID (I)	DIA	MEAN			ID (D)	ID (I)	DIA	MEAN
	KH	0.601	0.89	0.697	0.569		KH	0.963	1.213	1.1	1.
	LA	0.506	0.673	0.625	0.569		LA	0.968	1.045	0.994	1.
	KUH	0.591	0.717	0.62	0.569		KUH	0.961	1.001	0.947	1.
D40	RL	0.589	0.67	0.648	0.569	D65	RL	0.925	1.044	1.018	1.
	SA	0.674	0.795	0.703	0.569		SA	1.007	1.047	1.03	1.
	SS	0.43	0.717	0.619	0.569		SS	0.843	1.15	1.016	1.
	MB	0.56	1.072	0.755	0.569		MB	0.938	1.308	1.091	1.
	LE	0.614	0.968	0.769	0.569		LE	1.237	1.379	1.242	1.
	KH	0.688	0.878	0.775	0.665		KH	1.096	1.193	1.145	1.062
	LA	0.596	0.739	0.716	0.665		LA	1.029	0.907	1.03	1.062
	KUH	0.724	0.818	0.737	0.665		KUH	1.07	1.069	1.028	1.062
D45	RL	0.653	0.833	0.734	0.665	D70	RL	1.01	1.182	1.101	1.062
	SA	0.766	0.901	0.793	0.665		SA	1.133	1.251	1.156	1.062
	SS	0.562	0.801	0.71	0.665		SS	0.883	1.198	1.095	1.062
	MB	0.627	1.137	0.823	0.665		MB	0.997	1.353	1.115	1.062
	LE	0.71	1.13	0.872	0.665		LE	1.095	1.307	1.216	1.062
	KH	0.738	1.011	0.877	0.761		KH	1.146	1.357	1.242	1.179
	LA	0.706	0.781	0.808	0.761		LA	1.132	1.09	1.136	1.179
	KUH	0.747	0.82	0.778	0.761		KUH	1.127	1.003	1.073	1.179
D50	RL	0.73	0.888	0.799	0.761	D80	RL	1.13	1.298	1.204	1.179
	SA	0.843	0.967	0.869	0.761		SA	1.256	1.304	1.259	1.179
	SS	0.62	0.888	0.796	0.761		SS	0.998	1.251	1.209	1.179
	MB	0.706	1.168	0.902	0.761		MB	1.075	1.442	1.247	1.179
	LE	0.886	1.255	0.937	0.761		LE	1.196	1.604	1.351	1.179
	KH	0.834	1.104	0.94	0.849		KH	1.148	1.338	1.315	1.266
	LA	0.788	0.823	0.835	0.849		LA	1.169	0.979	1.168	1.266
	KUH	0.784	1.014	0.867	0.849		KUH	1.145	1.23	1.185	1.266
D55	RL	0.844	1.047	0.906	0.849	D90	RL	1.184	1.303	1.219	1.266
	SA	0.953	1.	0.942	0.849		SA	1.37	1.32	1.347	1.266
	SS	0.713	0.949	0.876	0.849		SS	1.119	1.335	1.273	1.266
	MB	0.828	1.234	0.968	0.849		MB	1.154	1.603	1.337	1.266
	LE	0.99	1.116	1.028	0.849		LE	1.116	1.437	1.39	1.266
	KH	0.842	1.117	1.023	0.929		KH	1.256	1.594	1.405	1.348
	LA	0.845	1.107	0.937	0.929		LA	1.28	1.225	1.302	1.348
	KUH	0.826	1.045	0.931	0.929		KUH	1.213	1.227	1.152	1.348
D60	RL	0.933	1.01	0.953	0.929	D10	RL	1.275	1.271	1.271	1.348
	SA	0.953	1.114	0.999	0.929		SA	1.405	1.42	1.387	1.348
	SS	0.759	1.026	0.963	0.929		SS	1.127	1.335	1.293	1.348
	MB	0.886	1.259	1.032	0.929		MB	1.136	1.436	1.327	1.348
	LE	0.838	1.245	1.118	0.929		LE	1.245	1.554	1.467	1.348

Table C.5: Color constancy indices of subjects for all illumination conditions in Experiments 1–4 (Chapter 5). Indices that refer to Experiments 1–3 are derived from mean hue scalings. Indices in Experiment 4 are calculated from subject's achromatic settings. The last column shows the indices of the mean data of all subjects.

illumination	exp.	AHR	AST	LGS	XH	ALL
Blue (D10)	1	0.74	0.71	0.67	0.91	0.69
	2	0.72	0.50	0.59	0.73	0.75
	3	0.17	-0.18	0.34	0.68	0.08
	4	0.4	-0.01	0.57	0.75	0.51
Yellow (D40)	1	0.70	0.63	0.62	0.76	0.64
	2	0.29	0.56	0.23	0.18	0.38
	3	-0.04	0.13	0.06	-0.23	-0.05
	4	0.2	0.21	-0.69	0.68	0.04
Red	1	0.65	0.58	0.58	0.85	0.75
	2	0.16	0.46	0.37	0.63	0.46
	3	0.06	0.21	0.22	0.05	0.03
	4	-0.03	0.12	0.08	0.75	0.3
Green	1	0.75	0.65	0.73	0.73	0.68
	2	0.60	0.42	0.72	0.44	0.57
	3	0.07	-0.04	0.49	-0.01	-0.11
	4	0.21	0.03	0.22	0.73	0.24

Appendix D

Color Plates

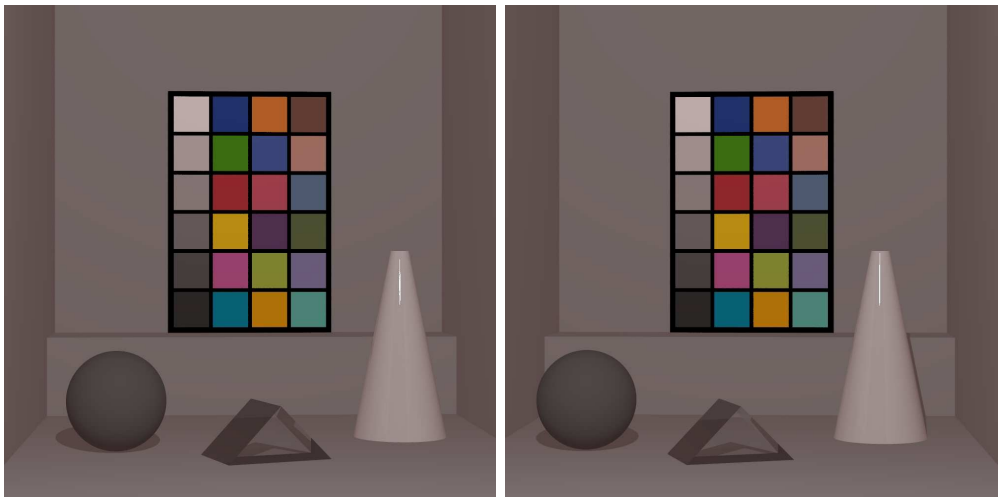

Plate D.1: A virtual scene under simulated neutral illumination. When the left and the right image of the simulated scene are fused by the observer a three dimensional impression results. The images were rendered with Radiance (Larson & Shakespeare, 1996). The scene resembles the stimuli used by Delahunt and Brainard (2004b).

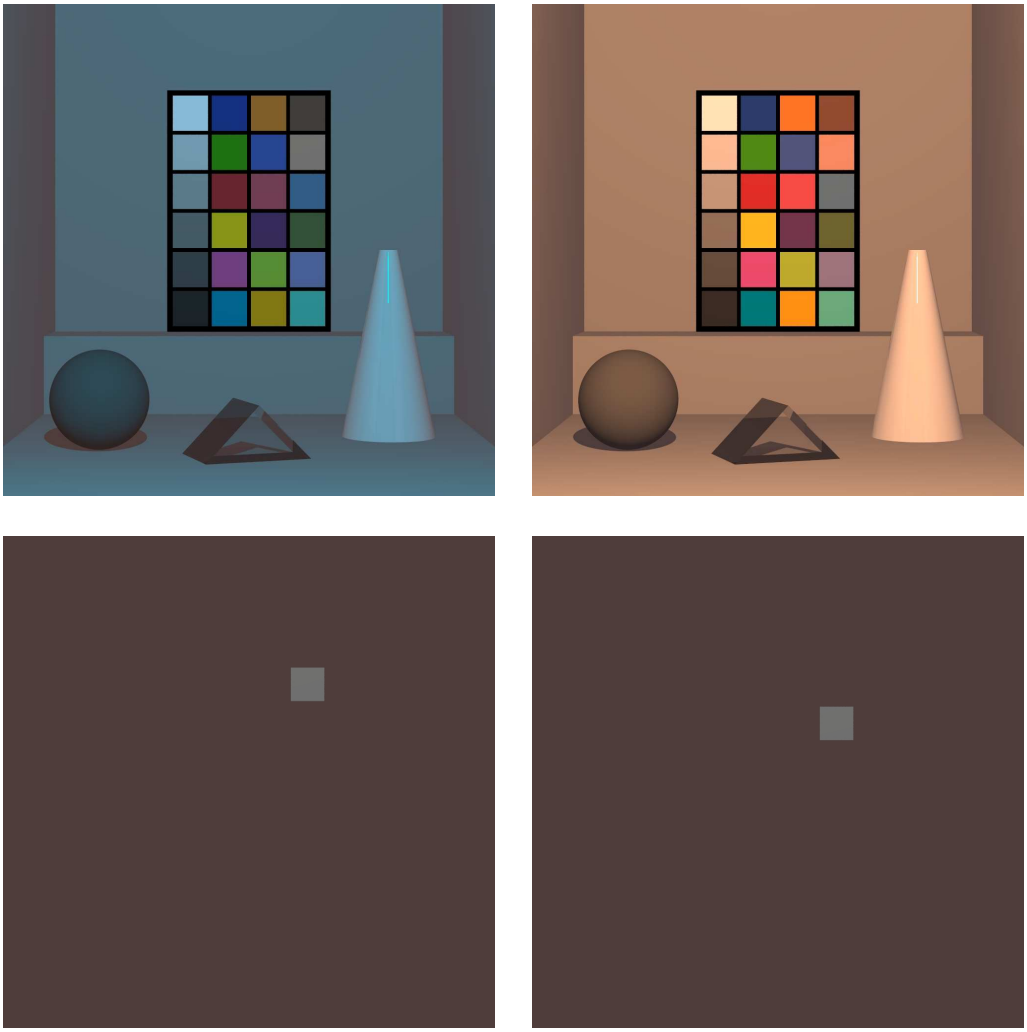

Plate D.2: A virtual scene under two different chromatic illuminants (top row). The same scene is shown in Color Plate D.1 under simulated neutral illumination. A simulation of a GretagMacbeth ColorChecker is leaning on the back wall of the chamber. The patch in the second row (from top) and fourth column (from left) serves as a reference for 'light skin'. The patch below (third row, fourth column) is labeled 'blue sky'. The illuminants are chosen so that the color codes of the physical light signals from the 'light skin'-patch under bluish illumination and from the 'blue sky'-patch under yellowish illumination are identical. This fact is illustrated by the simulations in the bottom row. The scene is masked except for the 'light skin'-patch under bluish illumination (left) and the 'blue sky'-patch under yellowish illumination (right).

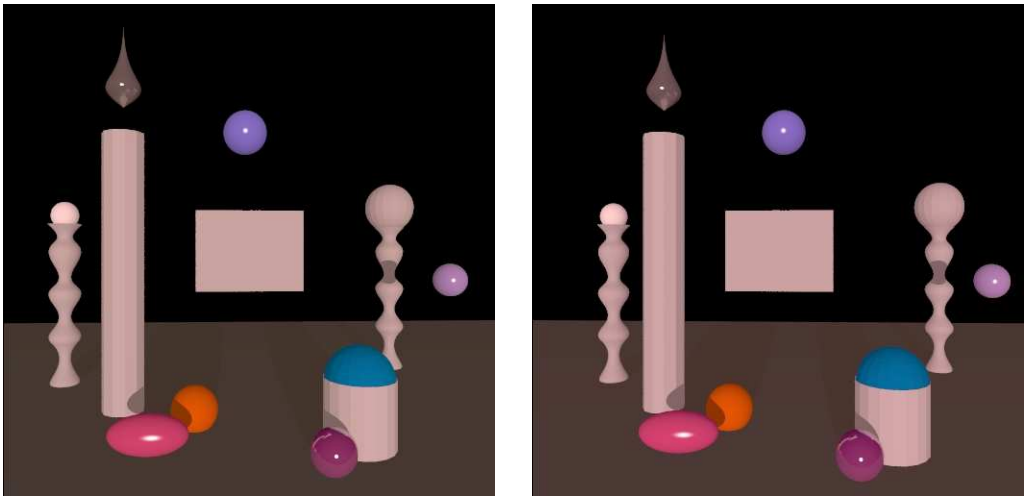

Plate D.3: A typical stimulus which was used in the experiments described in Chapter 5. The left and the right image of the simulated scene under neutral illumination are shown. Scenes were presented to the subject stereoscopically. The task of the subject was either to set hue scalings or to make achromatic settings for the test patch in the center of the scene. The images were rendered with Radiance (Larson & Shakespeare, 1996).

Appendix E

Zusammenfassung

Einführung

Das Phänomen der Farbkonstanz bezieht sich auf die Beobachtung, dass uns die Farbe einer Oberfläche unter wechselnder Beleuchtung annähernd gleich erscheint. Ein solcher Mechanismus trägt maßgeblich zur Orientierung in unserer Umwelt bei. Zur Verdeutlichung des Phänomens der Farbkonstanz können wir uns ein gelbes Objekt vorstellen, das zunächst bei Tageslicht und anschließend unter künstlicher Beleuchtung betrachtet wird. In beiden Fällen erscheint uns das Objekt in etwa in demselben Gelb, obwohl sich die von diesem Objekt abgestrahlten Lichtspektren sehr stark voneinander unterscheiden können. Wie erstaunlich die Leistung ist, die unser visuelles System dabei vollbringt, wird in Farbtafel D.2 demonstriert.

In der vorliegenden Arbeit wurden in zwei experimentellen Studien zahlreiche Aspekte der menschlichen Farbkonstanz untersucht. Über diese Studien wird in den Kapiteln 4 und 5 ausführlich berichtet. Zuvor werden in zwei theoretischen Kapiteln Einführungen in wichtige Konzepte der Farbwahrnehmungsforschung gegeben. In Kapitel 2 werden die Grundlagen der Primärkodierung von Lichtreizen und daraus resultierende Farbmetriken behandelt. Eine umfangreiche Einführung in das Phänomen der Farbkonstanz wird in Kapitel 3 gegeben.

Im Zentrum der ersten Studie, die in Kapitel 4 berichtet wird, steht die Untersuchung der Farbkonstanz einer Gruppe farbfehlsichtiger Beobachter, den *Dichromaten*. Dabei wird überprüft, ob Farbkonstanzleistungen von Dichromaten in Einklang mit solchen Vorhersagen stehen, die auf der Grundlage der vorherrschenden Konzeption von der Dichromasie gemacht werden. Des weiteren wird in der ersten Studie der Frage nachgegangen, ob komplexe inkrementelle und dekrementelle Reize durch das visuelle System in unterschiedlicher Weise verarbeitet werden und ob Verarbeitungsunterschiede in einem systematischen Zusammenhang mit der Chromatizität von Tageslichtern stehen.

In der zweiten Studie, die in Kapitel 5 dargestellt ist, wird eine alterna-

tive Methode zur Bestimmung der Farbkonstanzleistungen von Versuchspersonen, die Farbskalierung (*hue scaling*) eingeführt. Diese Methode bietet gegenüber herkömmlichen quantitativen Methoden den Vorteil, einen stärkeren Bezug zum eigentlichen Phänomen der menschlichen Farbkonstanz aufzuweisen. Ein weiterer Teilaspekt der zweiten Serie von Experimenten ist die Untersuchung des Einflusses von chromatischer Adaptation auf Farbkonstanzurteile von Versuchspersonen.

Beide Studien unterscheiden sich hinsichtlich des verwendeten Reizmaterials und der damit verbundenen Modellvorstellungen von Umgebungsbedingungen. Der ersten Studie liegt ein sehr einfaches Modell der Umgebung zugrunde, das *Flat World* genannt werden soll. In der zweiten Untersuchung wird ein komplexeres dreidimensionales Modell von Umgebungsbedingungen verwendet. Die zugehörige Klasse von dreidimensionalen Modellen der Umgebung wird als *Shape World* bezeichnet.

Farbkonstanz normaler und dichromatischer Beobachter

Fragestellung

Dichromaten zeichnen sich dadurch aus, dass ihnen gegenüber normalsichtigen Trichromaten, die drei Farbrezeptortypen besitzen, ein Rezeptortyp fehlt. Die vorherrschende Annahme besagt, dass sich der dichromatische Farbraum als eine Projektion des dreidimensionalen Farbraumes der Trichromaten in einen zweidimensionalen Unterraum auffassen lässt. Das bedeutet, dass dieser Vorstellung zufolge die Dichromasie gleichsam eine reduzierte Form der Trichromasie ist. Diese Annahme soll *Reduktionshypothese* genannt werden.

Neuere Befunde zeigen jedoch, dass Dichromaten unter bestimmten Bedingungen zu erstaunlichen Diskriminationsleistungen in der Lage sind, die mit dieser Auffassung nicht in Einklang stehen (Montag, 1994; Wachtler, Dohrmann & Hertel, 2004; Bonnardel, 2006). Diese Befunde scheinen insgesamt darauf hinzudeuten, dass die klassische Reduktionshypothese nur eingeschränkte Gültigkeit besitzen könnte oder gänzlich verworfen werden muss. In der in Kapitel 4 berichteten Untersuchung soll die Reduktionshypothese einem weiteren Test unterzogen werden. Es wird überprüft, ob Farbkonstanzurteile von Protanopen, einer Klasse von Dichromaten, denen der langwellige L-Rezeptor fehlt, mit den Vorhersagen dieser Modellvorstellung übereinstimmen. Farbkonstanzleistungen von Dichromaten sind bisher sehr selten untersucht worden und bisherige Befunde scheinen eher gegen die Reduktionshypothese zu sprechen (Rüttiger, Mayser, Sérey & Sharpe, 2001).

Neben der Untersuchung von Farbkonstanzleistungen dichromatischer Beobachter soll in dieser experimentellen Studie auch der Frage nachgegangen werden, ob sich die Verarbeitung von inkrementellen und dekrementellen Reizen systema-

tisch unterscheidet. Zahlreiche Befunde deuten darauf hin, dass Inkremente und Dekremente sowohl in einfachen Infeld-Umfeld Konfigurationen (Mausfeld & Niedereé, 1993; Heyer, 1997) als auch in komplexen Szenen (Bäuml, 2001) auf unterschiedliche Weise kodiert werden. Eine mögliche Erklärung für diese Beobachtung bietet die Annahme, das visuelle System würde Inkremente als selbstleuchtende Objekte auffassen, während Dekremente vom visuellen System als reflektierende Oberflächen interpretiert würden (Bäuml, 1997; Mausfeld, 1998). In der vorliegenden Untersuchung soll überprüft werden, ob das Ausmaß der unterschiedlichen Verarbeitung von Inkrementen und Dekrementen systematisch mit der korrelierten Farbtemperatur von Tageslichtern zusammenhängt, wie eine eigene Reanalyse der Daten von Bäuml (2001) vermuten lässt.[1]

Methoden

In den Experimenten wurden Simulationen von zweidimensionalen matten Mondrianreizen verwendet, die unter verschiedenfarbigen simulierten Beleuchtungen auf einem Computermonitor präsentiert wurden (vgl. Abb. 4.4). Aufgabe der Versuchsperson war es dabei stets, die zentrale Testfläche so einzustellen, dass sie weder bläulich noch gelblich und weder rötlich noch grünlich erschien. Einstellungen dieser Art werden auch als *Urgraueinstellungen* bezeichnet.

Die drei Experimente unterschieden sich hinsichtlich der simulierten Beleuchtungen: Im ersten Experiment wurden drei Tageslichter (D10, D65, D40) und zwei künstliche Lichter verwendet, die normalsichtigen Beobachtern rötlich beziehungweise grünlich erschienen. Die beiden Beleuchtungen in Experiment 2 hatten die Eigenschaft, dass ihre Chromatizitäten auf einer protanopen Konfusionslinie angeordnet waren. Diese Lichter erschienen normalsichtigen Betrachtern rötlich bzw. grünlich, waren für Protanope jedoch ununterscheidbar. Im dritten Experiment wurden insgesamt zehn simulierte Tageslichter verwendet, deren korrelierte Farbtemperaturen aus dem Intervall 4000–10000 K stammten (vgl. Abb. 4.6).

An den Experimenten nahmen insgesamt sechs normalsichtige und drei protanope Versuchspersonen teil.

Ergebnisse

Die dichromatischen Versuchspersonen zeigten im ersten Experiment erstaunlich gute Farbkonstanzleistungen. Unter den simulierten Beleuchtungen in Experiment 2 nahmen sie häufig Einstellungen vor, die nicht mit den Vorhersagen der Reduktionshypothese übereinstimmten. In diesem Experiment brach die Farbkonstanz von dichromatischen Versuchspersonen lediglich in zwei von sechs Fällen – wie von der Reduktionshypothese vorhergesagt – zusammen.

[1] Für eine kurze Einführung in das Konzept der *korrelierten Farbtemperatur* von Tageslichtern sei auf Anhang A verwiesen.

In den Urgraueinstellungen der Versuchspersonen unter simulierten Tageslichtern zeigten sich deutliche Unterschiede zwischen inkrementellen und dekrementellen Testreizen. Es fand sich ein systematischer Zusammenhang zwischen der Stärke der Inkrement–Dekrement Asymmetrie und der korrelierten Farbtemperatur der Beleuchtung. Vereinfacht gesprochen traten mit zunehmender Gelblichkeit eines Tageslichtes stärkere Unterschiede zwischen Inkrementen und Dekrementen auf. Generell waren die Farbkonstanzleistungen der Versuchspersonen besser für dekrementelle als für inkrementelle Reize.

Ein Test dreier Farbkonstanzmodelle zeigte, dass ein Modell, welches Inkrement–Dekrement Asymmetrien berücksichtigt, die Urgraueinstellungen der Versuchspersonen sehr viel besser vorherzusagen vermochte, als ein einfaches von-Kries-artiges Diagonalmodell und eine Variante des Landschen Retinex–Models. Letzteres Modell, das vorschlägt, beleuchtungsunabhängige Codes durch die Gewichtung von lokalen Rezeptorcodes an den mittleren Rezeptorerregungen in der Szene zu gewinnen, lieferte allerdings erstaunlich gute Vorhersagen wenn dekrementelle Reize betrachtet wurden.

Farbskalierung und Farbkonstanz

Fragestellung

Der Farbkonstanzmechanismus unseres visuellen Systems erlaubt es uns, einem Objekt unter wechselnden Beleuchtungsbedingungen nahezu identische Farbeindrücke zuzuordnen. Farbkonstanzurteile von Versuchspersonen werden häufig mit Hilfe quantitativer Methoden, wie asymmetrischen Abgleichen oder Urgraueinstellungen erhoben. Es ist jedoch aus verschiedenen Gründen fraglich, ob diese Methoden eine adäquate Bestimmung der menschlichen Farbkonstanz erlauben (Foster, 2003; Logvinenko & Maloney, 2006; vgl. auch Abschnitt 3.1.3). Eine direkte Methode der Erhebung von Farbkonstanzurteilen ist die Farbbenennung. Nachteile eines solchen Vorgehens bestehen darin, dass die Anzahl der verwendeten Farbkategorien sehr gering ist und dass sich die so gewonnenen Daten nur schwer mit quantitativen Einstellungen vergleichen lassen.

In der in Kapitel 5 dargestellten Untersuchung wird Farbskalierung (*hue scaling*) als alternative Methode zur Erhebung von Farbkonstanzurteilen eingeführt. Diese Methode berücksichtigt phenomenologische Aspekte der Farbkonstanz stärker als die üblichen quantitativen Methoden und erlaubt im Vergleich zu Farbbenennungen sehr viel feinere Urteile. Die mit Hilfe von Farbskalierungen gewonnenen Farbkonstanzurteile der Versuchspersonen sollen mit zugehörigen Urgraueinstellungen verglichen werden.

Zeitliche Aspekte werden in Modellen der Farbkonstanz häufig nur unzureichend berücksichtigt (Land & McCann, 1971; Foster & Nascimento, 1994). Ein wichtiger zeitlicher Mechanismus, der Farbkonstanz unterstützt, ist chromatische

Adaptation. In der vorliegenden Untersuchung wird die isolierte Wirkung von chromatischer Adaptation bei Kontrolle der räumlichen Konfiguration demonstriert.

Ein Einwand, der häufig gegen die Verwendung von Urgraueinstellungen angeführt wird, ist die Feststellung, dass in diesem Fall streng genommen nur Aussagen über achromatisch erscheinende Reize gemacht werden können. Ergebnisse einer vergleichenden Untersuchung von Speigle und Brainard (1999) deuten daraufhin, dass asymmetrische Abgleiche mit Hilfe von Urgraueinstellungen vorhergesagt werden können. In der vorliegenden Untersuchung soll überprüft werden, ob sich Farbskalierungen von Versuchspersonen unter einer festen Beleuchtung aus dem geschätzten zugehörigen Urgraupunkt unter dieser Beleuchtung vorhersagen lassen.

Eine weitere Besonderheit dieser Untersuchung ist die Verwendung von simulierten dreidimensionalen Szenen als Reizmaterial. Eine solche Konfiguration verspricht eine realistischere Repräsentation von natürlichen Beleuchtungsbedingungen herzustellen.

Methoden

Als Reizmaterial wurden den Versuchspersonen Simulationen von dreidimensionalen Szenen präsentiert, die aus mehreren Objekten und einer Testfläche im Zentrum der Szene bestanden (vgl. Abb. D.3). In den Experimenten 2 und 3 wurden die Testflächen isoliert vor schwarzem Hintergrund präsentiert. Die Beleuchtung der Szenen bestand stets aus Simulationen je einer neutralen diffusen und einer chromatischen punktuellen Lichtquelle. Fünf verschiedene punktuelle Beleuchtungen wurden verwendet, die dem Betrachter weiß, blau, gelb, rot und grün erschienen (vgl. Abb. 5.2). Insgesamt wurden 16 verschiedene Testflächen verwendet, die sich in Luminanz und Sättigung nicht unterschieden (vgl. Abb. 5.1). Die Reize wurden auf zwei CRT-Monitoren stereoskopisch dargeboten.

Aufgabe der Versuchsperson war es in den Experimenten 1–3, jeweils auf einer Skala von null (nicht) bis sechs (stark gesättigt) einzuschätzen, wie blau, gelb, rot und/oder grün ihr die Testfläche erschien. Im vierten Experiment sollten die Versuchspersonen jeweils Urgraueinstellungen für die Testfläche vornehmen. Fünf normalsichtige Versuchspersonen nahmen an den Experimenten teil.

Ergebnisse

Die Versuchspersonen nahmen im ersten Experiment, in dem vollständige Szenen präsentiert wurden, für eine feste Testfläche ähnliche Einstellungen unter verschiedenen Beleuchtungsbedingungen vor. Das Ausmaß der mit Hilfe von Farbskalierungen gemessenen Farbkonstanz stimmte in etwa mit Ergebnissen früherer Studien überein, in denen quantitative Methoden verwendet wurden (Kraft &

Brainard, 1999). Die Urgraueinstellungen der Versuchspersonen, die in Experiment 4 erhoben wurden, deuten hingegen eher auf eine geringe Farbkonstanzleistung hin. Die Ergebnisse aus Experiment 1 zeigen ferner, dass sich die Skalierungsdaten recht gut durch korrespondierende Urgraupunkte im Sinne von Speigle und Brainard (1999) vorhersagen lassen.

Die Einschätzungen der Versuchspersonen in Experiment 2, in dem die Testflächen einer Beleuchtungsbedingung geblockt und isoliert dargeboten wurden, hingen nicht wie erwartet von den Farbkoordinaten der Testfläche, sondern von der Sequenz zuvor präsentierter Flächen ab. Im Gegensatz dazu wurden die Einstellungen in Experiment 3, in dem die Testflächen isoliert und vollständig randomisiert dargeboten wurden, ausschließlich durch die Farbkoordinaten der präsentierten Testfläche determiniert.

Diskussion

Die Ergebnisse der ersten Untersuchung scheinen eher gegen die Gültigkeit einer strengen Reduktionshypothese zu sprechen. Dieser Befund steht in Einklang mit Ergebnissen aus zahlreichen Studien zu verschiedenen Aspekten dichromatischer Wahrnehmung, die ebenfalls Zweifel an der Gültigkeit dieser Annahme aufkommen lassen (Crognale, Teller, Yamaguchi, Motulsky & Deeb, 1999; Wachtler, Dohrmann & Hertel 2004; Bonnardel, 2006). Darüber hinaus haben genetische Analysen gezeigt, dass einer jeden phänotypischen Form der Dichromasie eine Reihe verschiedener Genotypen zugrunde liegen können (Neitz, Neitz, He & Shevell, 1999). Zusammengenommen deuten all diese Befunde daraufhin, dass die klassische Reduktionshypothese ein sehr vereinfachendes Modell der Dichromasie darstellt, das der Komplexität dieser Klasse von Beobachtern nicht gerecht wird.

Die Ergebnisse zur unterschiedlichen Verarbeitung von inkrementellen und dekrementellen Reizen stehen in Einklang mit ähnlichen Befundmustern, die für Infeld-Umfeld-Konfigurationen (Mausfeld Niederée, 1993; Heyer, 1997) und komplexe Mondrianreize (Bäuml, 2001) berichtet wurden. Sie legen nahe, dass es sich bei der getrennten Kodierung von Inkrementen und Dekrementen um einen fundamentalen Mechanismus unseres visuellen Systems handelt. Der systematische Zusammenhang zwischen der Stärke des Verarbeitungsunterschiedes und der Chromatizität von Tageslichtern deutet in Übereinstimmung mit früheren Befunden (Chichilnisky & Wandell, 1996) daraufhin, dass dieser Mechanismus eng mit der Verarbeitung von Signalen des kurzwelligen S-Rezeptorkanals verknüpft zu sein scheint. Ergebnisse einer Studie von McLellan und Eskew (2000) lassen vermuten, dass die Trennung von inkrementellen und dekrementellen Signalen schon frühzeitig in den S–ON und S–OFF Ganglienzellen erfolgt.

Eine weitere Implikation der Trennung von Inkrementen und Dekrementen ergibt sich für Modelle der Farbkonstanz. Grundsätzlich sollte diese Unterscheidung in Farbkonstanzmodellen berücksichtigt werden. Die Ergebnisse zeigen aber

auch, dass die Gewichtung von lokalen Rezeptorcodes an den mittleren Rezeptorerregungen in der Szene eine nützliche Heuristik zur Gewinnung beleuchtungsunabhängiger Codes sein kann, wenn dekrementelle Reize betrachtet werden.

Die Ergebnisse der zweiten Studie legen nahe, dass Farbskalierung eine nützliche Methode zur Untersuchung einer stärker phenomenologisch gefassten Konzeption der Farbkonstanz sein kann. Diese Technik stellt gleichsam ein Bindeglied zwischen kategorialen und quantitativen Methoden dar. Ein Vorteil der Farbskalierung ist die gute Vergleichbarkeit mit quantitativen Daten.

Ein zentraler Befund der zweiten Studie ist der starke Einfluss von chromatischer Adaptation auf Farbkonstanzurteile. Die Ergebnisse zeigen, dass unser visuelles System in der Lage ist, Informationen von sequentiell präsentierten isolierten Lichtreizen zu integrieren. Die Bedeutsamkeit von chromatischer Adaptation für die Farbkonstanz wird auch durch weitere aktuelle Befunde unterstrichen (Smithson & Zaidi, 2004; Hansen, Walter & Gegenfurtner, 2007). Chromatische Adaptation ist ein Prozess, der verhältnismäßig langsam abläuft. Er scheint es uns daher zu erlauben, uns optimal an langsame graduelle Änderungen des Tageslichts anzupassen.

Eine Frage, der in beiden Untersuchungen nachgegangen wurde, bezieht sich auf die Annahme, dass das visuelle System gewisse Regelmäßigkeiten von natürlichen Tageslichtern bei der Anpassung an die Beleuchtungsbedingungen nutzen könnte. Ergebnisse aus beiden Studien deuten daraufhin, dass das Ausmaß der erreichten Farbkonstanz für natürliche Tageslichter nicht höher ist als für rötlich und grünlich erscheinende künstliche Lichter. Ähnliche Befunde wurden auch von Delahunt und Brainard (2004b), sowie von Hansen et al. (2007) berichtet.

Zusammenfassend deuten die Ergebnisse aus beiden Untersuchungen daraufhin, dass es keinen singulären Mechanismus zu geben scheint, der dem Phänomen der Farbkonstanz zugrunde liegt. Vielmehr könnte das visuelle System eine Reihe von verschiedenen Informationen aus einer Szene auswerten und miteinander kombinieren. Diese Vorstellung der Kombination mehrerer cues und Mechanismen steht in Einklang mit zahlreichen empirischen Befunden (Kraft & Brainard, 1999; Yang & Maloney, 2001). Sie stellt einen vielversprechenden Ansatz in der Farbkonstanzforschung dar, der sich für die Wahrnehmung von Tiefe bereits als sehr fruchtbar erwiesen hat (Landy, Maloney, Johnson & Young, 1995).