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

\[(s, e) \approx^p (s', e')\]  \hspace{1cm} (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 $\phi$ yield an illuminant invariant code $\kappa(s, e)$. We can denote a protanopic version of this model with:

\[(s, e) \approx^p (s', e') \Leftrightarrow \kappa_i(s, e) = \kappa_i(s', e'), \quad i = 2, 3.\]  \hspace{1cm} (4.2)

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

\[\kappa_i(s, e) = \frac{1}{\sum_j^N \phi_i(e \cdot s_j)} \cdot \phi_i(e \cdot s), \quad i = 2, 3.\]  \hspace{1cm} (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

\[\kappa_i(s, e) = \frac{1}{\phi_i(e \cdot (t * w))} \cdot \phi_i(e \cdot s) = \frac{1}{t \cdot \phi_i(e)} \cdot \phi_i(e \cdot s), \quad i = 2, 3.\]  \hspace{1cm} (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:

$$
\phi^P(e \cdot s) = \phi^P(e' \cdot s) \quad \text{and} \quad \phi_i(e \cdot s) \neq \phi_i(e' \cdot s).
$$

(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

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

(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

$$
\phi^P(e \cdot s) = \phi^P(e' \cdot s) \quad \text{and} \quad \phi^P(e) = \phi^P(e')
$$

(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

$$
\phi^P(e) = \phi^P(e') \Rightarrow \phi^P(e \cdot s) = \phi^P(e' \cdot s)
$$

(4.8)

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

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:

$$\phi_i(e \cdot s) = \phi_i(e) \cdot c_i^s, \quad i = 2, 3,$$

(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):

$$\phi_i(e \cdot s) = \phi_i(e) \cdot \phi_i(s), \quad i = 1, 2, 3.$$

(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

$$\phi^P(e) \neq \phi^P(e'), \quad \text{and}$$

$$\phi^P(e \cdot s) \neq \phi^P(e' \cdot s).$$

(4.11)

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

\[
\phi_1(e) = \phi_1(e') \quad \text{and} \quad \phi_1(e \cdot s) = \phi_1(e' \cdot s)
\]

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; Chichilnsky & Wandell, 1995)—simple von-Kries receptor scaling is not sufficient to predict the color appearance of the center (Chichilnsky & 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 \( \phi_{s,e} \). Let us then define a relation of metamerism ‘\( \approx \)’ where

\[
(s,e) \approx (s', e')
\]

(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 \( \kappa \) (denoted as vectors) are identical:

\[
(s,e) \approx (s', e') \iff \kappa_{s,e} = \kappa_{s',e'}.
\]

(4.14)
The models differ in the specification of these constancy codes. However, in each of these models the constancy codes \( \kappa \) are given by a linear transformation of receptor codes \( \phi \) which can be written as:

\[
\kappa_{s,e} = M_e \phi_{s,e},
\]

(4.15)
where \( M_e \) is a \( 3 \times 3 \) 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 \( M_e \) can be expressed as diagonal matrices \( D_e \). 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 \( \psi_{s,e} \). 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 \( \psi \) as a linear transformation of constancy codes \( \kappa \). This transformation is given by a \( 3 \times 3 \) matrix \( T \) (see Equation 2.29):

\[
\psi_{s,e} = T(M_e \phi_{s,e}).
\]

(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:

\[
A(a,e) \iff \psi_1(a,e) = \psi_2(a,e) = 0.
\]

(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:

\[
A(a, e) \Leftrightarrow T(M_e \phi_{a,e}) = (0, 0, k)'.
\]  

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, \( D_e^+ \) for increments and \( D_e^- \) 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 \( \Delta_{e,r} \), we can write the affine ID model as follows:

\[
A(a, e) \Leftrightarrow TD_e^+ \Delta_{e,r} = (0, 0, k^+)'; \quad \text{if } \Delta_i(e, r) \geq 0;
\]

\[
A(a, e) \Leftrightarrow TD_e^- \Delta_{e,r} = (0, 0, k^-)'; \quad \text{if } \Delta_i(e, r) < 0; \quad \text{for } i = 1, 2, 3.\tag{4.19}
\]

As noted above, the expression \( A(a, e) \) refers to the observation that surface \( a \) appears achromatic under illuminant \( e \). The matrix \( T \) 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 \( D_e^+ \) and \( D_e^- \) and six additive constants of corresponding linear regressions.

The diagonal model can be derived from Equation (4.18) by simply replacing matrix \( M_e \) with a corresponding diagonal matrix \( D_e \):

\[
A(a, e) \Leftrightarrow TD_e \phi_{a,e} = (0, 0, k)'.
\]  

\(^6\)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 \( D_e \).

A special case of model (4.20) is the spatial mean model where the diagonal elements \( d_i(e) \) of \( D_e \) are specified as the reciprocals of corresponding mean cone excitations across the whole scene. We will denote this matrix with \( D^*_e \) 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:

\[
d^*_i(e) = \frac{1}{\sum_j^N \phi_i(s_j, e)}, \quad i = 1, 2, 3, \tag{4.21}
\]

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:

\[
A(a, e) \iff TD_e^* \phi_{a,e} = (0, 0, k)', \tag{4.22}
\]

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

\[
\psi_{a,e} = \psi_{a',e'} \tag{4.23}
\]

holds for these stimuli which is equivalent to

\[
T(M_e \phi_{a,e}) = T(M_{e'} \phi_{a',e'}). \tag{4.24}
\]

For the moment, we choose unspecified \( 3 \times 3 \) transformation matrices \( M_e, M_{e'} \) which will be replaced later by matrices specific to the respective model. The matrix \( T \) 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 \( T^{-1} \) and see that in this case constancy codes as defined in Equation (4.15) are identical:

\[
M_e \phi_{a,e} = M_{e'} \phi_{a',e'}. \tag{4.25}
\]

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

\[
(a, e) \approx (a', e') \iff M_e \phi_{a,e} = M_{e'} \phi_{a',e'}. \tag{4.26}
\]

which is equivalent to

\[
(a, e) \approx (a', e') \iff \psi_{a,e} = \psi_{a',e'}. \tag{4.27}
\]
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 $e'$:

$$(a, e) \approx (a', e') \Leftrightarrow \phi_{a,e} = M_e^{-1}M_e\phi_{a',e'}.$$  \hspace{1cm} (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 Chapter 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

$$\phi_{a,e} = D_e^{-1}D_{D65}\phi_{a',D65}. \hspace{1cm} (4.29)$$

The matrix $D_e^{-1}D_{D65}$ 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.\footnote{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.}

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

$$y = Xb + \varepsilon. \hspace{1cm} (4.30)$$

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

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

\hspace{1cm} (4.31)
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 \( \varepsilon \) 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 \( b_i^+ = \frac{d_i^+(D65)}{d_i^+(e)} \). The additive constant \( c_i^+ \) can be written as

\[
c_i^+ = \phi_i(r, e) - b_i^+ \phi_i(r', D65),
\]

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

\[
\begin{pmatrix}
\phi_i(a_1, e) \\
\vdots \\
\phi_i(a_N, e)
\end{pmatrix} =
\begin{pmatrix}
\phi_i(a_1', D65) \\
\vdots \\
\phi_i(a_N', D65)
\end{pmatrix} \cdot b_i + \varepsilon.
\]

(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

\[
b_i^* = \frac{\sum_j^K \phi_i(s_j, e)}{\sum_j^K \phi_i(s_j, D65)},
\]

(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

\[
H_0 : \quad b_i^{ID} = (b_i^D, 0, b_i^D, 0)',
\]

(4.35)

where \( b_i^{ID} \) 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

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

(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):

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

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 $UV$ around $P$ until the pulsing disappeared. At this point the subject perceived a homogeneous disk. In this case, $P$ and all points on the line $U'V'$ 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 $UV$ 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).

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
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 $D_65$ 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 $CIE-XYZ$-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:

$$X_{s,e} = \sum_{\lambda=380}^{800} s(\lambda) \cdot e(\lambda) \cdot \bar{x}(\lambda),$$ (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 $D_65$ 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 $1^\circ 46' \times 1^\circ 46'$. The background extended over the whole screen and subtended 40.5cm × 30cm ($15^\circ \times 11^\circ$ of visual angle).
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/m². Using illuminations with luminance of 100 cd/m² 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 $D_{65}$. 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-
characteristics 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.
CHAPTER 4. EXPERIMENTS I

4.2.3 Apparatus

The stimuli were presented to the subjects on an EIZO FlexScan F980 CRT screen with resolution $1.024 \times 768$ 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 $D_{65}$ 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).

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. 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 \( m_{D65} \) represent \( u'v' \)-coordinates of the mean achromatic setting of the subject under neutral daylight \( D65 \) and \( m_{data} \) corresponding mean achromatic settings under test illumination. We will denote settings that indicate perfect color constancy with \( m_{perf} \). This index can be written as:

\[
CI = 1 - \frac{|m_{data} - m_{perf}|}{|m_{D65} - m_{perf}|},
\]

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
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.
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 $D_{40}$ 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 $s^I_c$ and $s^D_c$ of the two separate regression lines were determined:

\[ q_c = \frac{s^I_c}{s^D_c}. \] 

(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 $D_{65}$. 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 $\alpha$-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 \( x \) indicates mean cone coordinates across the scene and the open circle M-receptor codes of the estimated reference stimulus.
CHAPTER 4. EXPERIMENTS I

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 adjacent patches. The resulting stimulus can be imagined as a $5 \times 5$ 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 $b^m_i$ 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 $b_{mi}^m$ 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.
CHAPTER 4. EXPERIMENTS I

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 $s_{D40}$ and $s_{D40}$ 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
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
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 Bäuml (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...
Observations

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 $\mu$ and $\sigma$. We can draw a sample from the population of indices with parameters $\mu$ and $\sigma$ 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 $\sigma$ 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 $\sigma$, 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 $s'_I$ and $s'_D$ 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 $-0.61$ to $-0.94$. 
CHAPTER 4. EXPERIMENTS I

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 $\sigma$ representing the mean standard deviation of the normal subjects participating in Experiment 3 (see text for details).
CHAPTER 4. EXPERIMENTS I

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 x indicates mean cone coordinates across the scene and the open circle S-receptor codes of the estimated reference stimulus.
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-
CHAPTER 4. EXPERIMENTS I

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.

Performance 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 $b_{m}^{i}$ 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-
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).