Color Constancy: Developing Empirical Tests of Computational Models

David H. Brainard, James M. Kraft*, & Philippe Longère

Department of Psychology

University of California Santa Barbara Santa Barbara, CA 93106

To appear in: R. Mausfeld & D. Heyer (eds.), Colour Perception: From Light to Object, Oxford University Press.

Draft of July, 2000.

*Present address: Center for Neural Science, New York University, 4 Washington Place, New York, NY 10003

INTRODUCTION

Object recognition is difficult because there is no simple relation between an object's properties and the retinal image. Where the object is located, how it is oriented, and how it is illuminated also affect the image. Moreover, the relation is under-determined: multiple physical configurations can give rise to the same retinal image.

In the case of object color, the spectral power distribution of the light reflected from an object depends not only on the object's intrinsic surface reflectance but also factors extrinsic to the object, such as the illumination. The relation between intrinsic reflectance, extrinsic illumination, and the color signal reflected to the eye is shown schematically in Figure 1. The light incident on a surface is characterized by its spectral power distribution $E(\lambda)$. A small surface element reflects a fraction of the incident illuminant to the eye. The surface reflectance function $S(\lambda)$ specifies this fraction as a function of wavelength. The spectrum of the light reaching the eye is called the color signal and is given by $C(\lambda) = E(\lambda)S(\lambda)$. Information about $C(\lambda)$ is encoded by three classes of cone photoreceptors, the L-, M-, and S-cones.

The top two patches rendered in Plate 1 illustrate the large effect that a typical change in natural illumination (see Wyszecki and Stiles, 1982) can have on the color signal. This effect might lead us to expect that the color appearance of objects should vary radically, depending as much on the current conditions of illumination as on the object’s surface reflectance. Yet the very fact that we can sensibly refer to objects as having a color indicates otherwise. Somehow our visual system stabilizes the color appearance of objects against changes in illumination, a perceptual effect that is referred to as color constancy.

Insert Figure 1 about here.

Insert Plate 1 about here

Because the illumination is the most salient object-extrinsic factor that affects the color signal, it is natural that emphasis has been placed on understanding how changing the illumination affects object color appearance. In a typical color constancy experiment, the independent variable is the illumination and the dependent variable is a measure of color appearance (Helson, 1938; Helson and Jeffers, 1940; Helson and Michels, 1948; Burnham, Evans and Newhall, 1957; Hunt, 1950; McCann, McKee and Taylor, 1976; Arend and Reeves, 1986; Arend, Reeves, Schirillo and Goldstein, 1991; Valberg and Lange-Malecki, 1990; Brainard and Wandell, 1992; Bauml, 1994; Bauml, 1995; Lucassen and Walraven, 1993; Lucassen and Walraven, 1996; Brainard, Brunt and Speigle, 1997; Brainard, 1998). These various experiments employ different stimulus configurations and psychophysical tasks, but taken as a whole they support the view that human vision exhibits a reasonable degree of color constancy.

Recall that the top two patches of Plate 1 illustrate the limiting case where a single surface reflectance is seen under multiple illuminations. Although this case illustrates the effect of the illuminant, it fails to capture an essential feature of the computational problem faced by a visual system that attempts to achieve color constancy. This is the ambiguity created because of the interaction between illuminant and surface reflectance, an ambiguity illustrated if we consider Plate 1 in its entirety. The rendered patches in the second row of the plate show the effect of the same illuminant change on the information encoded about an additional surface. Note that when seen under the first illuminant, this second surface presents the same spectral signature as does the first surface under the second illuminant.

When we consider both illuminant and surface variation, the essential ambiguity underlying color constancy emerges: how can the visual system determine which object is present in the world if the information reaching the eye is identical for two different object-illuminant configurations? Clearly color constancy is not possible in general, since the visual system cannot distinguish the two simple scenes rendered in the top right and bottom left patches of Plate 1.

Given that color constancy is not possible in general, it makes little sense to provide a simple answer to the question of how color constant human vision is. More sensible is to investigate constancy for some specified ensemble of scenes (Maloney, 1999). Of particular interest are ensembles that are representative of scenes we encounter in daily viewing.

In this chapter, our aim is to link two lines of research. The first is theoretical work on the computational problem of color constancy. The goal of computational theories is to define particular ensembles of scenes in which some degree of color constancy is possible, and to express algorithms that achieve constancy for these ensembles. Computational theories of color constancy stand independent of their relevance to human vision. None-the-less, we have found that the computational work provides useful guidance for a research program designed to understand human color vision. Our treatment of the computational work is intended primarily to clarify how computational models can be elaborated to make predictions about human performance.

The second line of research is empirical measurements of human color constancy made in our laboratory. Here the emphasis is on studies of performance for stimulus conditions closely related to natural viewing and on measurements that connect to computational theory.

COMPUTATIONAL THEORY

Most computational theories of color constancy (e.g. Buchsbaum, 1980; Maloney and Wandell, 1986; Lee, 1986; D'Zmura and Lennie, 1986; Trussell and Vrhel, 1991; Funt and Drew, 1993; D'Zmura and Iverson, 1993; D'Zmura, Iverson and Singer, 1995; Brainard and Freeman, 1997; Finlayson, Hubel and Hordley, 1997) share the same basic two step framework. In the first step, the image is analyzed to yield an estimate of illuminant properties. In the second step, this estimate is used to process the light reflected to the eye from each surface. The second step produces a description of surface properties that is approximately independent of the actual illuminant. Within this two step framework, individual theories are distinguished by the ensemble of scenes to which they are meant to apply and by how they accomplish each step.

To illustrate how computational work can provide a basis for developing statements about human performance, it is useful to consider one theory in some detail. For this illustrative purpose we have chosen Buchsbuam’s classic (1980) theory, expressed with respect to the human visual system.

As emphasized above, any computational theory must define a restricted ensemble of scenes to which it applies. In the case of Buchsbaum’s theory a single scene in the ensemble consists of a collection of flat matte surfaces arranged in a single plane and illuminated diffusely by spatially uniform illumination. Light from each surface in the scene is reflected to the eye. The eye contains three classes of cone photoreceptors (L-, M-, and S-cones) that encode the spectral properties of the light reflected from each surface to the eye. Thus the image may be specified by the quantal absorption rates of the L-, M-, and S-cones at each image location. This simplified ensemble of visual scenes is sometimes referred to as the *Mondrian world* because of the resemblance of its individual scenes to paintings by the Dutch artist Piet Mondrian (Land and McCann, 1971; see also Maloney, 1999).

For any scene from the Mondrian world, we can describe the spectral power distribution of the illuminant by a function of wavelength $E(\lambda)$ and the spectral reflectance of each surface by a function $S_j(\lambda)$. The light reflected from the j^{th} surface to the eye then has spectral power distribution $C_j(\lambda) = E(\lambda)S_j(\lambda)$. It is convenient to discretize these spectral quantities and express them as vectors (e.g. Wandell, 1987; Brainard, 1995). Thus we can use the vector $\mathbf{e}$ to describe $E(\lambda)$, where $\mathbf{e}$ is an N_λ–dimensional column vector. The entries of $\mathbf{e}$ represent the power of the illuminant at N_λ sample wavelengths λ_n spaced evenly across the visible spectrum. Similarly we can represent the surface reflectance functions by the N_λ–dimensional column vector $\mathbf{s}_j$, where the n^{th} entry of $\mathbf{s}_j$ is $S_j(\lambda_n)$. Given this representation the spectral power distribution reflected to the eye from the j^{th} surface is

$$\mathbf{c}_j = \mathrm{diag}(\mathbf{e})\ \mathbf{s}_j = \mathrm{diag}(\mathbf{s}_j)\ \mathbf{e}, \qquad (1)$$

where the function diag() creates a diagonal matrix with the entries of its argument on the diagonal.

The information about the spectrum of light encoded by a single class of cones is the rate at which photons are absorbed by the photopigment contained within the cone. This rate may be computed from the cone's spectral sensitivity. Let $L(\lambda)$ be the spectral sensitivity of the L-cones, $M(\lambda)$ the spectral sensitivity of the M-cones, and $S(\lambda)$ the spectral sensitivity of the S-cones. Form the 3 by N_λ matrix $\mathbf{R}$, where the n^{th} entry of the first row of $\mathbf{R}$ is $L(\lambda_n)$, the n^{th} entry of the second row is $M(\lambda_n)$, and the n^{th} entry of the third row is $S(\lambda_n)$. We can then compute the quantal absorption rates of the three classes of cones in response to a spectral power distribution $\mathbf{c}_j$ through the equation

$$\mathbf{r}_j = \mathbf{R}\ \mathbf{c}_j \qquad (2)$$

where $\mathbf{r}_j$ is a three-dimensional column vector whose entries are the quantal absorption rates for the L-, M-, and S-cones respectively.

A feature of the Mondrian world is that the minimal spatial structure of the images does not carry information about the illuminant. Thus we can summarize the information available from the image about the illuminant by the list of quantal absorption rates $\{\mathbf{r}_j\}$. In addition, the ordering of the elements in the list is not important. We refer to the list $\{\mathbf{r}_j\}$ as the *color statistics* of the image.

It is straightforward to show that when we restrict attention to the Mondrian world, color constancy remains an underdetermined computational problem. It is possible to choose two illuminants and two collections of surfaces that produce identical color statistics. Thus Buchsbaum added additional constraints to the ensemble of scenes to which his theory applies. The first constraint concerned the spectral form of individual illuminants and surfaces. Rather than allowing arbitrary choices of $\mathbf{e}$ and the $\mathbf{s}_j$, Buchsbaum assumed that both illuminants and surfaces were constrained to lie within three-dimensional linear models. For illuminants, this assumption is that the illuminant $\mathbf{e}$ can be written as $\mathbf{e} = \mathbf{B}_e \mathbf{w}_e$ where $\mathbf{B}_e$ is an N_λ by 3 dimensional matrix and $\mathbf{w}_e$ is a three dimensional column vector. The columns of the matrix $\mathbf{B}_e$ are referred to as the basis vectors for the model, while the entries of the vector $\mathbf{w}_e$ are referred to as the model weights for the particular illuminant $\mathbf{e}$. For surfaces, the linear model assumption is similar. In this case we write $\mathbf{s}_j = \mathbf{B}_s \mathbf{w}_{sj}$ with where $\mathbf{B}_s$ is an N_λ by 3 dimensional matrix and $\mathbf{w}_{sj}$ is a three dimensional column vector. We can combine the linear model constraints with Equations 1 and 2 to obtain

$$\mathbf{r}_j = \mathbf{R}\,\mathbf{c}_j = \mathbf{R}\,\mathrm{diag}(\mathbf{e})\,\mathbf{B}_s\,\mathbf{w}_{sj} = \mathbf{R}\,\mathrm{diag}(\mathbf{s}_j)\,\mathbf{B}_e\,\mathbf{w}_e\,. \quad (3)$$

There is considerable evidence that small-dimensional linear models provide a reasonable description of many illuminants and surfaces (e.g. Judd, MacAdam and Wyszecki, 1964; Cohen, 1964; Maloney, 1986; Parkkinen, Hallikainen and Jaaskelainen, 1989 Jaaskelainen, Parkkinen and Toyooka, 1990; Romero, Garcia-Beltran and Hernandez-Andres, 1997; see Maloney, 1992)

A second constraint on the scenes was that the spatial average of the surfaces in any particular scene was constant across scenes. This is often referred to as the *gray world assumption*.

To understand how color constancy is possible in a Mondrian world with scenes constrained as described above, let $\bar{\mathbf{s}}$ be the spatial average of the $\mathbf{s}_j$ and $\bar{\mathbf{r}}$ be the spatial average of the corresponding $\mathbf{r}_j$. Then we can write

$$\bar{\mathbf{r}} = \mathbf{R}\,\mathrm{diag}(\bar{\mathbf{s}})\,\mathbf{B}_e\,\mathbf{w}_e\ . \tag{4}$$

This follows because the spatial averaging operation commutes with the linear process of image formation described by Equation 3. If the spatial average of the surface reflectance is known, then Equation 4 may be inverted to solve for the illuminant:

$$\hat{\mathbf{e}} = \mathbf{B}_e\ \mathbf{M}_s^{-1}\ \bar{\mathbf{r}} \tag{5}$$

where $\mathbf{M}_s$ is the three-by-three matrix given by $[\mathbf{R}\,\mathrm{diag}(\bar{\mathbf{s}})\,\mathbf{B}_e]$. The matrix $\mathbf{M}_s$ is invertible because the dimension of the linear model for surfaces (3) is matched to the number of human cone types (L, M, and S).

Given the estimate of the illuminant $\hat{\mathbf{e}}$ computation of the individual $\mathbf{s}_j$ is obtained through

$$\mathbf{s}_j = \mathbf{B}_s\ \mathbf{M}_e^{-1}\ \mathbf{r}_j \tag{6}$$

where $\mathbf{M}_e = [\mathbf{R}\,\mathrm{diag}(\hat{\mathbf{e}})\,\mathbf{B}_s]$. The matrix $\mathbf{M}_e$ is invertible because the dimension of the linear model for illuminants (3) is also matched to the number of human cone types (L, M, and S).

Equation 5 is the key to Buchsbaum's algorithm. By assuming that the spatial average of surface reflectances in the scene $\bar{\mathbf{s}}$ is known, it is possible to form the matrix $\mathbf{M}_s$ and apply Equation 5 to estimate the illuminant. Although Buchsbaum's theory is designed for the Mondrian world with linear model constraints, the estimation procedure may be applied to any set of image data. The estimate will be accurate to the extent that a) the scene conforms to the Mondrian world assumptions, b) the linear models $\mathbf{B}_e$ and $\mathbf{B}_s$ describe the illuminant and surfaces that comprise the scene, and c) the actual spatial average of surfaces matches the assumed $\bar{\mathbf{s}}$.

Note that in Equation 5 the illuminant estimate depends on the scene only through the spatial average of the receptor responses $\bar{\mathbf{r}}$. In this sense, the spatial average summarizes the scene with respect to the illuminant estimate obtained by Buchsbaum's algorithm. Several other theories (e.g. Maloney, 1986; Forsyth, 1990; Trussell and Vrhel, 1991; D'Zmura, Iverson and Singer, 1995; Brainard and Freeman, 1997; Finlayson, Hubel and Hordley, 1997) are also designed for the Mondrian world. As with Buchsbaum's theory, the algorithms associated with these theories work in two steps, first estimating the illuminant and then using the illuminant estimate to obtain surface reflectance estimates. These theories differ from Buchsbaum's primarily in what information is used to make the illuminant estimate. For example, the illuminant estimate returned by Maloney and Wandell's (1986) algorithm depends on the color statistics only through its covariance matrix, while that returned by Forsyth's (1990) algorithm depends only on the convex hull of the color statistics. As we will see below, understanding which properties of the color statistics affect an algorithm's estimate makes possible empirical tests of the algorithm's usefulness as a model of human performance.

Although we will not consider them further in this chapter, it is worth noting that there is a growing literature on theories that operate for richer scenes than those within the Mondrian world (Lee, 1986; D'Zmura and Lennie, 1986; Hurlbert, 1986; Tominaga and Wandell, 1989; Funt and Drew, 1993; D'Zmura and Iverson, 1993; see Hurlbert, 1998; Maloney, 1999). The algorithms associated with these theories generally estimate the illuminant using both information contained in the color statistics and information contained in the spatial structure of the image.

Linking computation and performance

How can we employ Buchsbaum’s (1980) theory (or any computational algorithm) as a model of human performance? It is not entirely obvious how to proceed. For example, the algorithm produces estimates of the illuminant spectral power distributions and surface reflectance functions, whereas human observers make psychophysical judgements. Such judgments are not of the direct spectral functions but rather assess, in one way or another, the color appearance of illuminants and surfaces in the scene. Thus the algorithm output and human judgements are not commensurate. To develop an algorithm into a model requires additional linking theory.

Suppose that σ is a vector whose entries describe the perceptual experience of color. To connect an algorithm such as Buchsbaum’s to human performance, we can suppose that σ is related to estimated surface reflectance $\hat{\mathbf{s}}$ by some unknown but fixed function $\mathbf{f}()$, so that $\sigma = \mathbf{f}(\hat{\mathbf{s}})$. Although the form of $\mathbf{f}()$ is unknown, we will assume that it does not depend on context and that it is one-to-one. This simple linking assumption does not allow us to predict color names from algorithm output. But it does allow the following general prediction to be made about the relation between human performance and algorithm output: two surfaces seen in the context of different images should appear the same if and only if the algorithm estimates the same surface reflectance for each surface. We will refer to this idea as the *match-prediction linking hypothesis*.

If we accept the match-prediction linking hypothesis, we can make predictions about human performance. Using a psychophysical procedure we establish pairs of stimuli that, when seen in the context of different images, appear the same. A typical procedure would be asymmetric color matching (e.g. Stiles, 1967; Burnham, Evans and Newhall, 1957; Arend and Reeves, 1986; Brainard and Wandell, 1992; Brainard, Brunt and Speigle, 1997). Given pairs of stimuli that match across contexts, we ask whether the surface reflectances estimated by an algorithm for these stimuli also match. To the extent that they do, the algorithm provides a good description of human performance.

The difficulty with taking this approach is that an algorithm's specific estimates depend on a number of parameter choices. For example, in Buchsbaum's algorithm the choice of linear models $\mathbf{B}_e$ and $\mathbf{B}_s$ will affect the estimated surface reflectances. These would either have to be set through parameter search or clever guess. Although this is not necessarily prohibitive, it seems desirable to investigate more directly whether the core principles of a computational theory can be used to understand human performance.

For Buchsbaum's algorithm, Equation 5 shows that the illuminant estimate it returns depends on the image only through the spatial average $\bar{\mathbf{r}}$; if we have two different images with the same spatial average ($\bar{\mathbf{r}}$), the algorithm will return the same illuminant estimate. In addition, the surface reflectance function estimated at a location depends on the image only through the light reflected from the surface at that location ($\mathbf{r}_j$) and the illuminant estimate (see Equation 6). Thus if two images have the same spatial average and we embed a surface that reflects the same light to the eye in each image, Buchsbaum's algorithm is a candidate model for human performance only if the two surfaces appear the same. This prediction holds independent of the choice of linear models $\mathbf{B}_e$ and $\mathbf{B}_s$.

In the next section we consider experiments that measure human color constancy, with the goal of connecting the experiments to the ideas discussed above.

COLOR CONSTANCY IN THE NEARLY NATURAL IMAGE

The effect of the illuminant

To allow precise stimulus specification and control, many experiments that attempt to quantify color constancy employ rather simple stimuli. One configuration that has been used extensively in recent years is a computer simulation of a scene consisting of a flat matte surfaces seen under diffuse illumination (e.g. Arend and Reeves, 1986; Arend, 1993; Troost and de Weert, 1991; Brainard and Wandell, 1992; Bauml, 1994; Bauml, 1995; Lucassen and Walraven, 1996). These stimuli are essentially instantiations of scenes from the Mondrian world. Recent experiments on color appearance also employ closely related stimuli (e.g. Wesner and Shevell, 1992; Jenness and Shevell, 1995; Singer and D'Zmura, 1994; Delahunt and Brainard, 2000).

When Mondrian world scenes are simulated on monitors, however, they appear somewhat artificial. This is probably not due to problems of the simulation but rather to the fact that the scenes that match the Mondrian world assumptions are rare in nature and the visual system may not treat them in the same way as it does natural images. Indeed, one can argue that seemingly simple scenes are very difficult for the visual system to parse. We might expect that before using color statistics to estimate the illuminant, the visual system attempts to determine which regions are objects and which are light sources, which image variations represent illumination boundaries, and which represent variations in reflected light due to geometric factors (see Gilchrist et al., 1999; Adelson, 1999). If this is the case, the processes that normally make these determinations may produce unstable or conflicting results when presented with impoverished stimuli. As a result, performance measured for simple stimuli could be much more difficult to understand than performance for stimuli which provide a rich set of cues.

These considerations motivated us to study color constancy using stimuli consisting of actual illuminated surfaces, configured in three-dimensions. By doing so, we hoped to study constancy as it operates in natural viewing. In the work reported here, however, we focus on results obtained using scenes that are (approximately) uniformly illuminated. This simplifies the comparison of human and algorithmic performance, since it is not necessary to consider processes that segment the image into distinctly illuminated regions.

The apparatus used in the first set of experiments is an entire room, shown schematically in Figure 2 and described in detail elsewhere (Speigle and Brainard, 1996; Brainard, Brunt and Speigle, 1997; Brainard, 1998). The ambient illumination of the room is produced by 3 sets of computer controlled stage lamps arranged in four triads. One set has red filters, one has green filters, and one has blue filters. The light from each triad passes through a diffuser to minimize colored shadows. By varying the intensities of the three sets of lamps, we can vary the spectral power distribution of the ambient illumination.

Insert Figure 2 about here.

A test surface on the far wall of the room is located so that it can be illuminated by a projection colorimeter. The illumination from the colorimeter consists of a mixture of red, green, and blue primaries. This illumination is focused and aligned so that it is spatially coincident with the test surface: it is not explicitly visible to the observer. The overall light reflected to the observer from the test surface thus consists of two components. The first is the normal reflection of the ambient illumination, while the second is generated by the colorimeter. Varying the intensity of the colorimeter primaries has the perceptual effect of changing the color appearance of the test surface. Essentially, we have taken the stimulus configuration exploited by Gelb (Gelb, 1950; see also Koffka, 1935; Katz, 1935) and brought it under computer control (see also Valberg and Lange-Malecki, 1990; Uchikawa, Uchikawa and Boynton, 1989; Kuriki and Uchikawa, 1996; Kuriki and Uchikawa, 1998).

As noted above, asymmetric color matching provides a convenient and natural experimental method for linking computational theory and human performance. This procedure is particularly well-suited to studying color constancy when there is a spatial change in the illumination (simultaneous color constancy) so that the matches can be made between two surfaces that are viewed at the same time (e.g. Arend and Reeves, 1986; Brainard, Brunt and Speigle, 1997). It is also possible to use asymmetric matching to study color constancy for the situation of interest here, uniformly illuminated scenes where the illuminant varies from one time to another (successive color constancy; Brainard and Wandell, 1992; Brainard and Wandell, 1991; Bauml, 1995; Jin and Shevell, 1996). In this case, however, the matches typically involve a memory component and are more difficult for observers.

A simpler experimental task is to measure the *achromatic locus* by having observers adjust the chromaticity of a surface (or image region) until it appears achromatic (Helson and Michels, 1948; Werner and Walraven, 1982; Fairchild and Lennie, 1992; Arend, 1993; Bauml, 1994; Chichilnisky and Wandell, 1996; Maloney and Yang, 2001). This task is performed easily and reliably by even the most naive of observers. A direct comparison of asymmetric matching and achromatic adjustment in a simultaneous color constancy experiment indicates that the two tasks tap the performance of the same visual mechanisms (Speigle and Brainard, 1999).

We measured how the achromatic locus depends on changes in illumination. Figure 3 shows typical results. Each of the open circles shows the chromaticity of an experimental illuminant. Each of the corresponding closed circles shows the chromaticity of the achromatic locus, measured for one observer, under the corresponding illuminant. The achromatic loci were determined by averaging loci determined in separate sessions. The x and y standard errors of measurement for each locus are smaller than the plotted points.[1]

Insert Figure 3 about here.

The achromatic loci plotted are the chromaticities of the light reflected to the eye that appeared achromatic (i.e. the chromaticities of the proximal stimulus). To interpret the data in terms of color constancy, consider the chromaticity of the light reflected from a surface that appears white under typical daylight. Such a surface has a reflectance spectrum that is nearly constant across wavelength and thus the light reflected from it always has a chromaticity close to that of the illuminant. Figure 4 plots the chromaticity of the light reflected from a Munsell N 9.5/ surface under two illuminants. This surface appears achromatic when seen

[1] We verified that for our conditions the chromaticity of observers' achromatic adjustments does not depend on luminance (Brainard, 1998). This invariance does not hold in general (Helson and Michels, 1948; Werner and Walraven, 1982; Chichilnisky and Wandell, 1996; see also Mausfeld and Niederee, 1993; Mausfeld, 1998; Delahunt and Brainard, 2000) but is obeyed for decrements seen against uniform surrounds (Chichilnisky and Wandell, 1996).

under the standard viewing conditions for which the Munsell system is defined, and for a color constant visual system it will continue to appear achromatic under other viewing conditions. Thus for a color constant visual system, the chromaticity of the achromatic locus should coincide with the chromaticity of the light reflected from this surface. We conclude that color constancy is indicated when the chromaticity of the achromatic loci lies near that of the illuminants (see Figure 4). This pattern is roughly what is seen in the data shown in Figure 3.

Insert Figure 4 about here.

It is possible to go from the data shown in Figure 3 to a constancy index. The calculations are described in detail elsewhere (Brainard, 1998). The index takes on a value of 0 for the case when the achromatic loci are unaffected by the illuminant (no constancy) and 1 when the achromatic loci track the illuminant perfectly (complete constancy). For intermediate cases, the index may be thought of as describing the extent to which the achromatic loci track the illuminant change. The value of the index for the data shown in Figure 3 is 0.80, and the mean value across a wide range of conditions (different objects in the room, different illuminant changes) was 0.82 (Brainard, 1998). Interestingly, this is more constancy than is typically seen in studies conducted with monitor displays. (Comparable indices are generally in the range 0.50-0.60, see Brainard and Wandell, 1991; Brainard, Wandell and Chichilnisky, 1993; Fairchild and Lennie, 1992). The relatively high constancy index shown by observers in our experiments is consistent with everyday experience: object colors do not change much with changes in illuminant. We believe that laboratory experiments employing the sort of nearly natural stimuli described above assess constancy as it operates in the real world.

Testing computational models

The experiment described above quantifies color constancy across changes of illumination. It does not, however, tell us much about how the visual system achieves the measured constancy. In the experiment, the surfaces that make up the scene remain constant as the illuminant is varied. Such a design, almost ubiquitous in studies of color constancy, eliminates from the stimulus ensemble the illuminant-surface ambiguity that makes constancy a difficult computational task. Indeed, most computational theories can predict good constancy under circumstances where the same collection of surfaces is viewed under an unknown illuminant. To test these theories it is necessary to conduct experiments where both the surfaces in the scene and the illuminants are varied.

To do so, we (Kraft and Brainard, 1999) had observers look into a small (approximately 3 ft. on a side) chamber in which the spectrum of the illuminant and the spectral reflectance of all visible surfaces could be independently controlled. Plate 2 shows images of the chamber in two different configurations. Between the two, some of the objects in the chamber were changed, so that the mean surface reflectance ($\bar{\mathbf{s}}$) in the scene is quite different in the two cases. In addition, the illumination in the two chambers is also different. The combined effect of the surface and illuminant manipulations is to make the spatial mean of the two images ($\bar{\mathbf{r}}$) identical. As with the experiments in the full room, the appearance of a test patch in the chamber could be adjusted through the use of the projection colorimeter. The observers' task was again to adjust the chromaticity of the test patch until it appeared achromatic.

The prediction of Buchsbaum’s algorithm for our experimental situation is straightforward. Given that the spatial average of the two images is the same, the match-prediction hypothesis says that when two test patches seen in the respective images match in appearance, the light reflected to the eye should be the same. Achromatic adjustments do not establish complete perceptual matches. But it is plausible that each point on the achromatic locus measured in one image matches some point on the achromatic locus measured in the other image. Given that we find that the chromaticity of light that appears achromatic is independent of test luminance (see footnote 1), we arrive at the prediction that the achromatic locus should have the same chromaticity when measured in the two images.

Figure 5 plots the achromatic loci measured for one observer in this experiment. The achromatic loci are significantly different from each other, as they were for three other observers (keep in mind that the standard errors for the achromatic loci are smaller than the plotted points; see Kraft and Brainard, 1999). From this fact, we can conclude directly that the spatial average of the image is not the only statistic governing color appearance. This in turn says that Buchsbaum's algorithm cannot completely describe human performance.

Insert Figure 5 about here.

The constancy index for the data shown in Figure 5 is 0.29. The mean index for four observers in the same experiment was 0.39. These indices are considerably lower than the value of 0.82 found for the experiments conducted in the full room. The reduction is not due to the fact that observers were looking into a chamber rather than sitting in an entire room: control experiments with the chamber, where only the illuminant was varied, yielded constancy indices of about 0.83.

DISCUSSION

In this chapter we have emphasized the link between computational theories of color constancy and human performance. In doing so, we have implicitly endorsed what Maloney refers to as the *illumination estimation hypothesis* (Maloney and Yang, 2001). This is the idea, central to the motivation here, that the visual system estimates the illuminant and that the estimate is used to govern the perception of surface color (see also Mausfeld, 1998; Gilchrist et al., 1999; Speigle and Brainard, 1996; Brainard, Brunt and Speigle, 1997). The work reviewed here does not directly test the illuminant estimation hypothesis, since observers do not make any judgments of perceived illumination. Recent work (Rutherford, 2000) suggests that the illuminant estimation hypothesis is at best an approximation (see also Beck, 1959; Beck, 1961; Oyama, 1968; Kozaki and Noguchi, 1976; Noguchi and Kozaki, 1985; Logvinenko and Menshikova, 1994). Even if human surface color appearance does not depend on an explicit illuminant estimate, we need not refrain from using computational theory to develop and test models of what image statistics influence the perception of surface color. Indeed, the models we have elaborated are designed to make predictions about asymmetric surface color matches (or closely related measures of appearance). In this sense, they are agnostic about whether the visual system computes an estimate of illuminant or whether such an estimate plays a governing role in surface color perception.

Our experimental logic can be used to show that a particular theory does not provide a complete description of human performance. In the case of Buchsbaum's algorithm, we learn that something other than the spatial average of the cone responses in the image contributes to how the visual system processes color information.[2] The experiments do not, however, rule out a role

---

[2] We should note that theories that postulate that the spatial average is the statistic that sets the visual system's effective estimate of the illuminant vary in terms of exactly how the average is computed. In our experiment, we matched the spatial average taken over image pixels, equally weighted. One can consider variants that weight distinct image regions identically (e.g. Gershon and Jepson, 1989), that take a spatially weighted average for each local image region (e.g. Land, 1986; see Brainard and Wandell, 1986), and that use the geometric rather than the arithmatic average of the LMS responses (again Land, 1986; Brainard and Wandell, 1986). Strictly

for the spatial average. Indeed, the fact that the constancy index is greatly reduced when the spatial average is held constant suggests that this statistic may play an important role in color perception. A more definitive statement is not possible based on our experiments, since by silencing the spatial average we also affected other image statistics. Yang and Maloney (Yang, 1999; Maloney and Yang, 2001) have recently taken an empirical approach complementary to ours, where they make small perturbations to one image statistic while holding others constant. Experiments of this sort can be used to establish that particular statistics are used by the visual system.

A crucial feature of our experimental design is that we manipulate both the illuminant and surfaces in the scene. Without doing so, we could not match the spatial average in the image while at the same time changing the illuminant. This is a point of wide applicability. Most computational theories derive their estimate of the illuminant from specific scene statistics. To test whether a particular theory provides a complete description of human performance, we can proceed by silencing the statistics used by that algorithm. To do so in a non-trivial manner, it is necessary to vary both the surfaces in the scene and the illuminant. To date, only a few other experimentalists have explored conditions where both the surfaces and illuminants varied (McCann, 1994; Gilchrist and Jacobsen, 1984; Kuriki and Uchikawa, 1998; see also Gilchrist, 1988). It is our opinion that further experiments where only the illuminant is varied are unlikely to advance our knowledge of constancy much beyond its current state. What is needed is more experiments where the essential ambiguity between surfaces and illuminants is restored to the experimental situation.

---

speaking, additional experiments would be needed to rule out all of these variants for the class of rich stimulus configurations we used. There are, however, a growing number of results for simpler laboratory images that make it difficult to adhere to any of these variants (Jenness and Shevell, 1995; Singer and D'Zmura, 1994; Brown and MacLeod, 1997).

In addition to conducting the experiment described above, where the spatial average of the image was held constant across a change of illuminant, we have measured achromatic loci in a variety of other images where surfaces in the scene were varied across an illuminant change. We will not review the particulars of these manipulations here; most are described in Kraft and Brainard (1999). Across the conditions we studied, constancy indices (mean across observers) varied considerably, ranging from 0.06 to 0.83. The lowest indices corresponded to spatially simple scenes where the surfaces were changed to reduce information about the illuminant change. The highest indices were obtained when the surfaces in the scene were held constant across an illuminant change. The variation of constancy index with experimental conditions emphasizes the fact that how well the visual system adjusts to a change of illuminant depends on the stimulus ensemble: when little information is available about the illuminant change the visual system is not very color constant.

We find it encouraging that we have found stimulus manipulations that cause the constancy indices to vary widely. This indicates that we have brought into the laboratory a set of factors that operate in rich images and that have a substantial impact on human performance. Identifying these factors more precisely and bringing them under parametric control should allow more systematic investigation of how color appearance is governed in complex natural scenes.

Although our stimuli consisted of real illuminated three-dimensional objects, we did not manipulate the spatial structure of the scenes. The spatial structure (either actual or perceived) of a scene can affect color appearance even when the color statistics of the image are held fixed (Gilchrist, 1977; Gilchrist, 1980; Knill and Kersten, 1991; Bloj, Kersten and Hurlbert, 1999). Such effects are not captured by the experiments and models described here. It is possible that for our stimulus configurations, the visual system takes advantages of cues such as specular highlights (Lee, 1986; D'Zmura and Lennie, 1986; Tominaga and Wandell, 1989; see Yang, 1999; Maloney and Yang, 2001) and mutual illumination (Funt, Drew and Ho, 1991; Funt and Drew, 1993; see Bloj, Kersten and Hurlbert, 1999). Whether this is the case or whether for our scenes the color statistics alone provide most of the information used by the visual system is an interesting and open question.

Another simplified aspect of our scenes is that the illumination was close to spatially uniform. Thus the task of segmenting the image according to different illuminants has a particularly simple solution for our images. How such segmentation operates in images with multiple illuminants (simultaneous constancy) remains a central unsolved problem that is not addressed by our work. Recent theories (Gilchrist et al., 1999; Adelson, 1999) have identified a number of heuristics that might guide the segmentation process. These theories also suggest that once the image has been segmented into separate regions, visual processing within regions is guided by the color statistics or some summary of them. Our work focuses on exactly how the image statistics are used within uniformly illuminated regions. Within the context of these recent theories, our work is complementary to explorations of how the segmentation processes operate.

It may be possible to quantify the relation between human performance and the information about the illuminant change that is actually available in a pair of images. Up to this point, we have considered computational theories as potential models for human performance. But computational models can also be used to provide a benchmark against which to compare human performance. This sort of analysis has been very successful in understanding data obtained from experiments that measure performance on objective psychophysical tasks such as detection and discrimination (e.g. Green and Swets, 1966; Geisler, 1989). In such applications, one predicts the performance of an *ideal observer* that optimally uses all of the information in the stimulus to perform some task. An ideal observer benchmark provides a principled method for evaluating how efficiently a real observer performs a particular task and thus to identify sites of information loss in visual processing.

Brainard and Freeman (1997) used Bayesian decision theory to develop an ideal observer for color constancy in the Mondrian world. Their work assumes that in any scene, the surface and illuminant spectra are drawn at random from a population whose distribution is known. When the prior assumptions are met, the algorithm returns an estimate of the illuminant that is optimal in the sense that it minimizes the expected illuminant estimation error.[3]

The Brainard and Freeman algorithm may be applied to each image for which Kraft and Brainard (1999) measured achromatic loci. We can compute a constancy index for the algorithm by treating the chromaticity of is illuminant estimates in the same way that we treat the achromatic loci measured for human observers. Figure 6 shows the constancy indices obtained for human observers plotted against the constancy indices obtained for the Bayesian algorithm. What is apparent in the plot is that there is a strong correlation between the human and Bayesian indices. If we take the performance of the Bayesian algorithm as a measure of how much information is available for an ideal observer to estimate the illuminant, we see that the variation in human performance across the conditions is well-explained by information differences between the various conditions. The slope of the regression line between the human and Bayes indices is 0.77. This could be taken as a measure of the

[3] See Brainard and Freeman (1997) for a detailed description of exactly what error is minimized.

degree of human constancy, relative to ideal performance, across the whole set of image manipulations.

Insert Figure 6 about here.

We do not wish to claim that the Brainard and Freeman (1997) algorithm provides a good model of human performance, even for stimulus configurations where the color statistics alone drive the visual system's estimate of the illuminant. A strong test of the particular algorithm requires that we apply the same logic as we developed earlier in the chapter -- find two images for which the algorithm predicts the same illuminant estimate and then measure color appearance for these two images. Doing so will require development of more sophisticated stimulus control techniques than we currently have at our disposal. The algorithm does, however, measure the information available from the color statistics about the illumination change across a pair of images. It is therefore intriguing that the algorithm is able to make accurate predictions of how human performance varies across a wide range of experimental conditions.

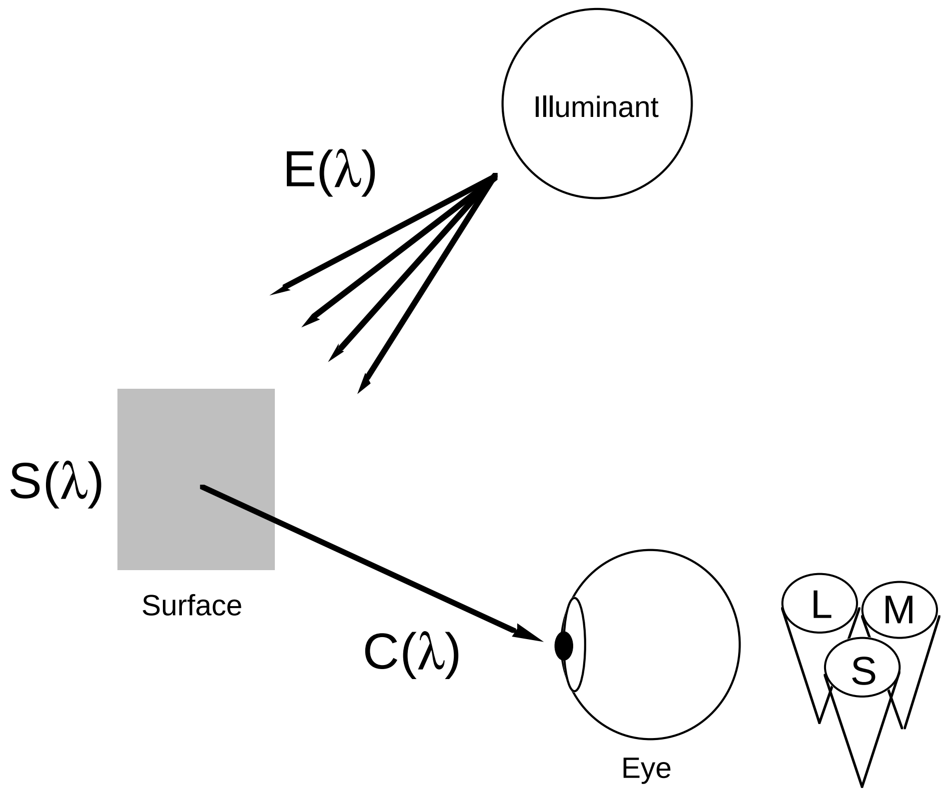

Figure 1. Effect of changing the illuminant on light reflected to the eye. The light incident on a surface is characterized by its spectral power distribution E(λ). A small surface element reflects a fraction of the incident illuminant to the eye. The surface reflectance function S(λ) specifies this fraction as a function of wavelength. The spectrum of the light reaching the eye is called the color signal and is given by C(λ) = E(λ)S(λ). Information about C(λ) is encoded by three classes of cone photoreceptors, the L-, M-, and S-cones. Note that this is a simplified imaging model. In general, the function S(λ) depends on the geometry of the observer, illuminant, and object.

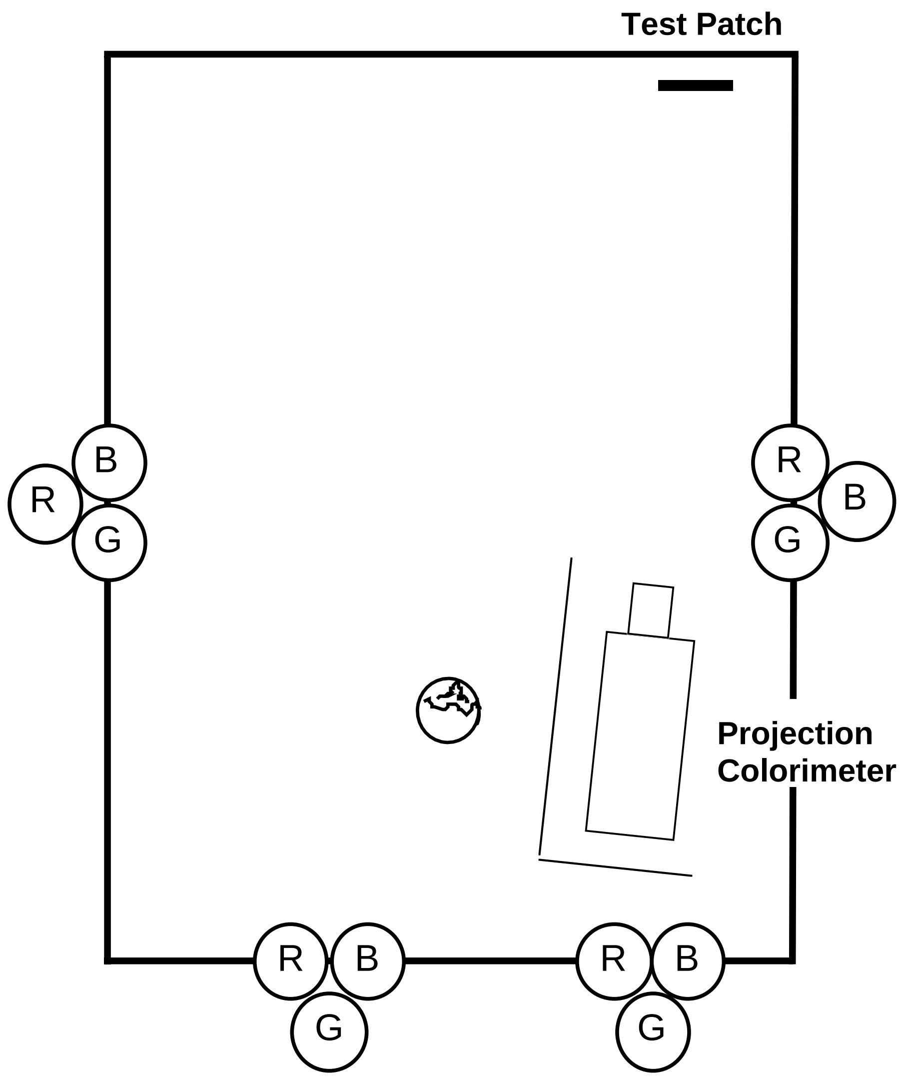

Figure 2. Room apparatus. Schematic of the experimental room. The experimental room was approximately 12’ by 9’ in dimension. Four triads of computer controlled lights provided the ambient illumination. A projection colorimeter allowed adjustment of the color appearance of a test patch located on the far wall of the room. Adopted from Figure 1 of Brainard (1998).

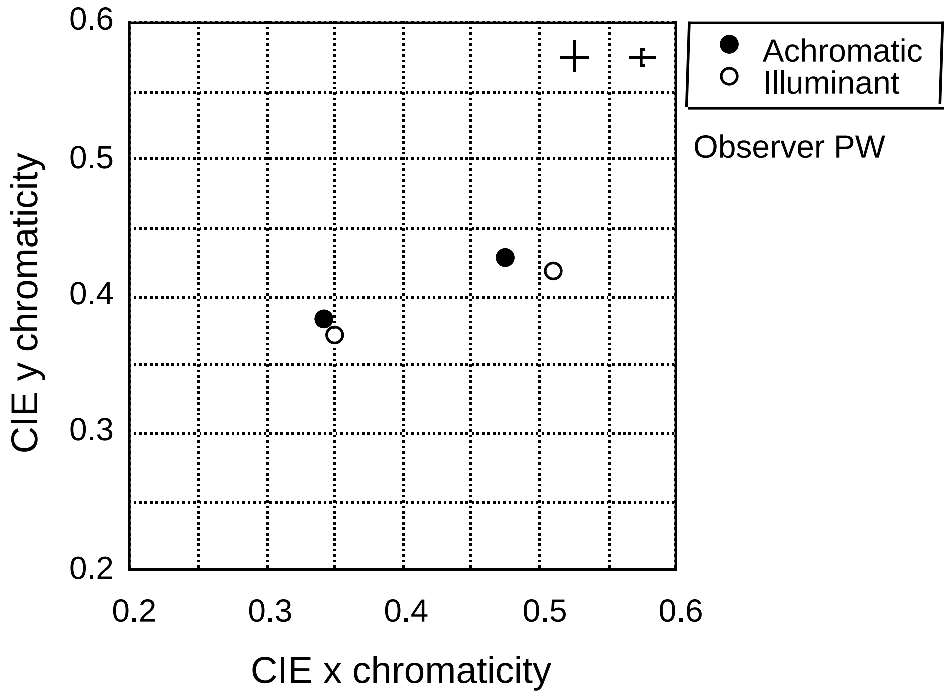

Figure 3. Basic achromatic results. The figure shows the CIE 1931 chromaticities of the achromatic loci (solid circles) measured under two experimental illuminants (chromaticity shown by open circles) for one observer. The between session standard error of the mean is smaller than the plotted points. The maximum within session standard deviation of the individual achromatic settings is indicated by the crosses at the upper left of the figure. Adopted from Figure 3 of Brainard (1998).

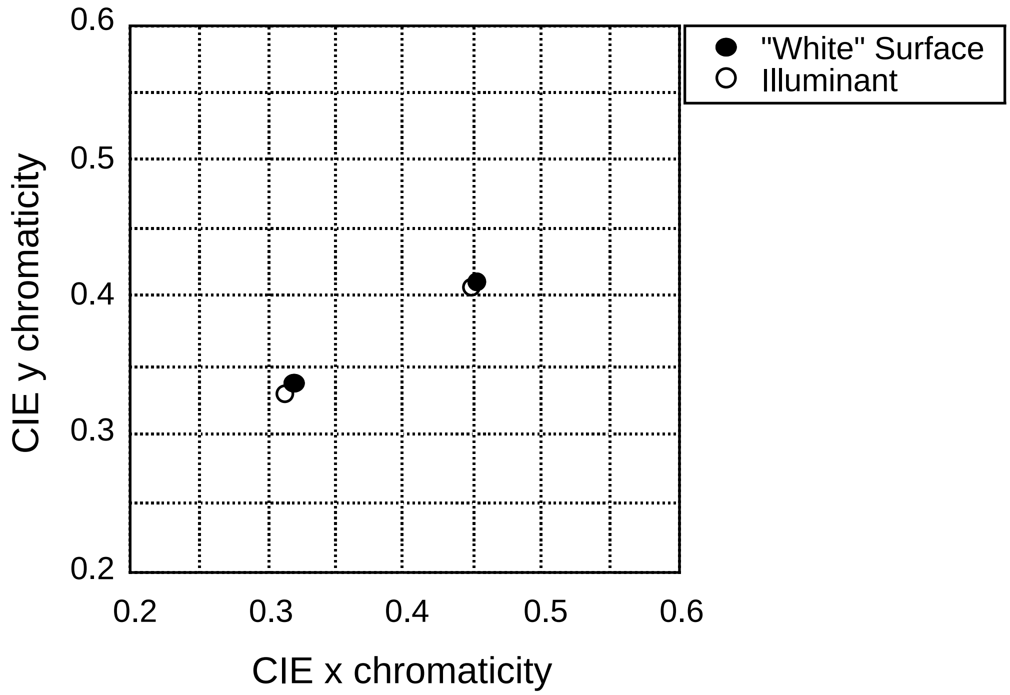

Figure 4. Data expected for a color constant visual system. The figure plots the chromaticity of the light reflected from a Munsell N 9.5/ surface (solid circles) under two illuminants. The chromaticities of the illuminants are indicated by the open circles.

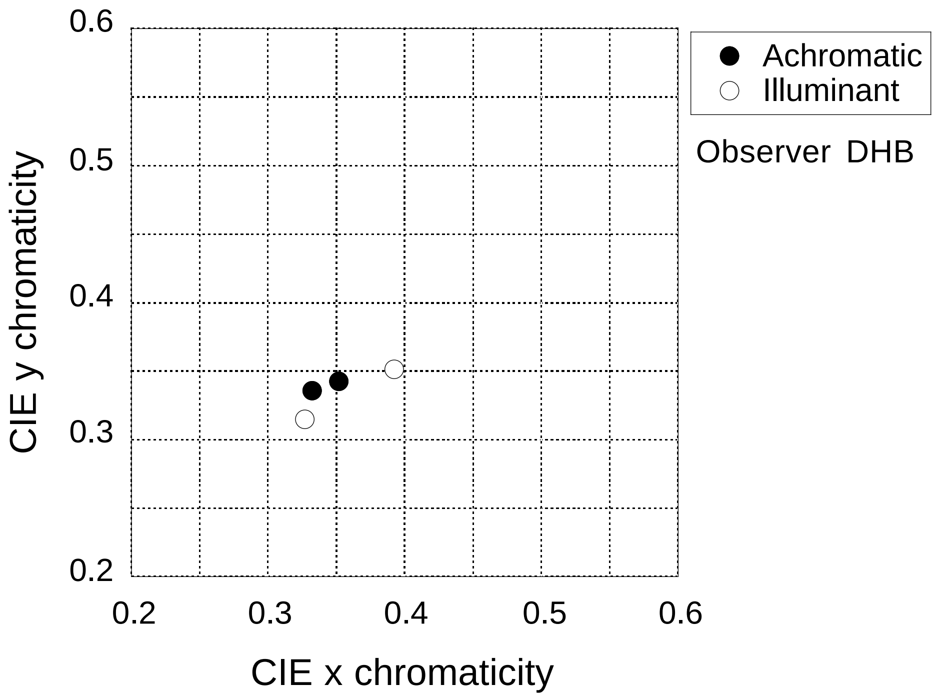

Figure 5. Achromatic settings with spatial average equated. The format of the figure is the same as for Figure 3. Here the achromatic settings were made in the context of two images where the illuminant differed (open circles) but the spatial average of the image was held constant. The between session standard error of the mean is smaller than the plotted points. Data are replotted from Figure 2 of Kraft and Brainard (1999).

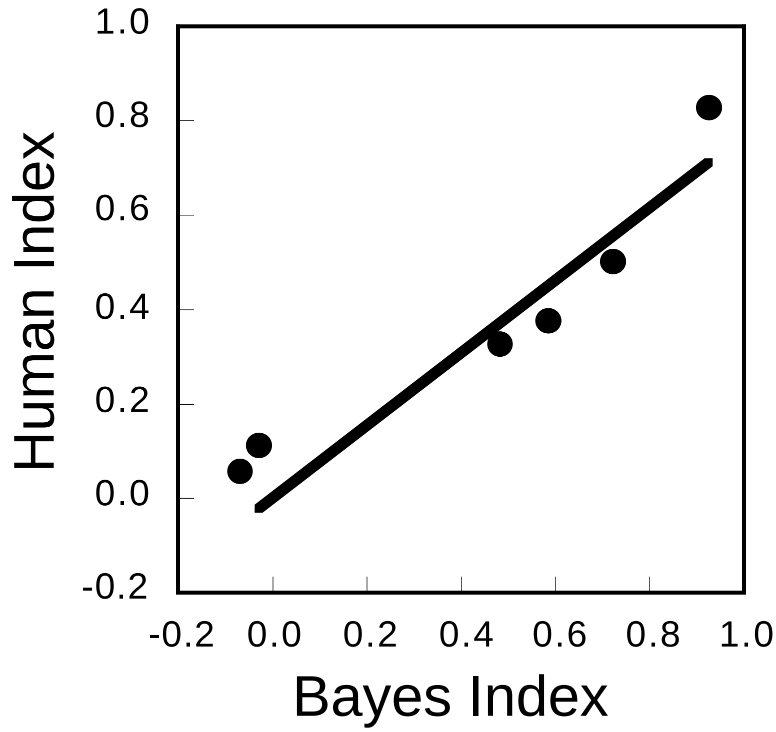

Figure 6. **Comparison of human performance with Bayesian algorithm.** The figure plots the constancy indices obtained for human observers against constancy indices obtained for the Bayesian algorithm of Brainard and Freeman (1997). The algorithm was run using points selected at random from calibrated LMS images of the stimulus. The image acquisition procedure is described in the caption for Plate 2 of this chapter. The prior distribution for illuminants was constructed to match the range of illuminants that our apparatus could produce. The prior distribution of surfaces was obtained by analyzing a measurements of Munsell papers, as described in Brainard and Freeman (1997). The small negative constancy indices obtained in some cases occur because the illuminant estimate shifts slightly in a direction opposite to the actual illuminant change.

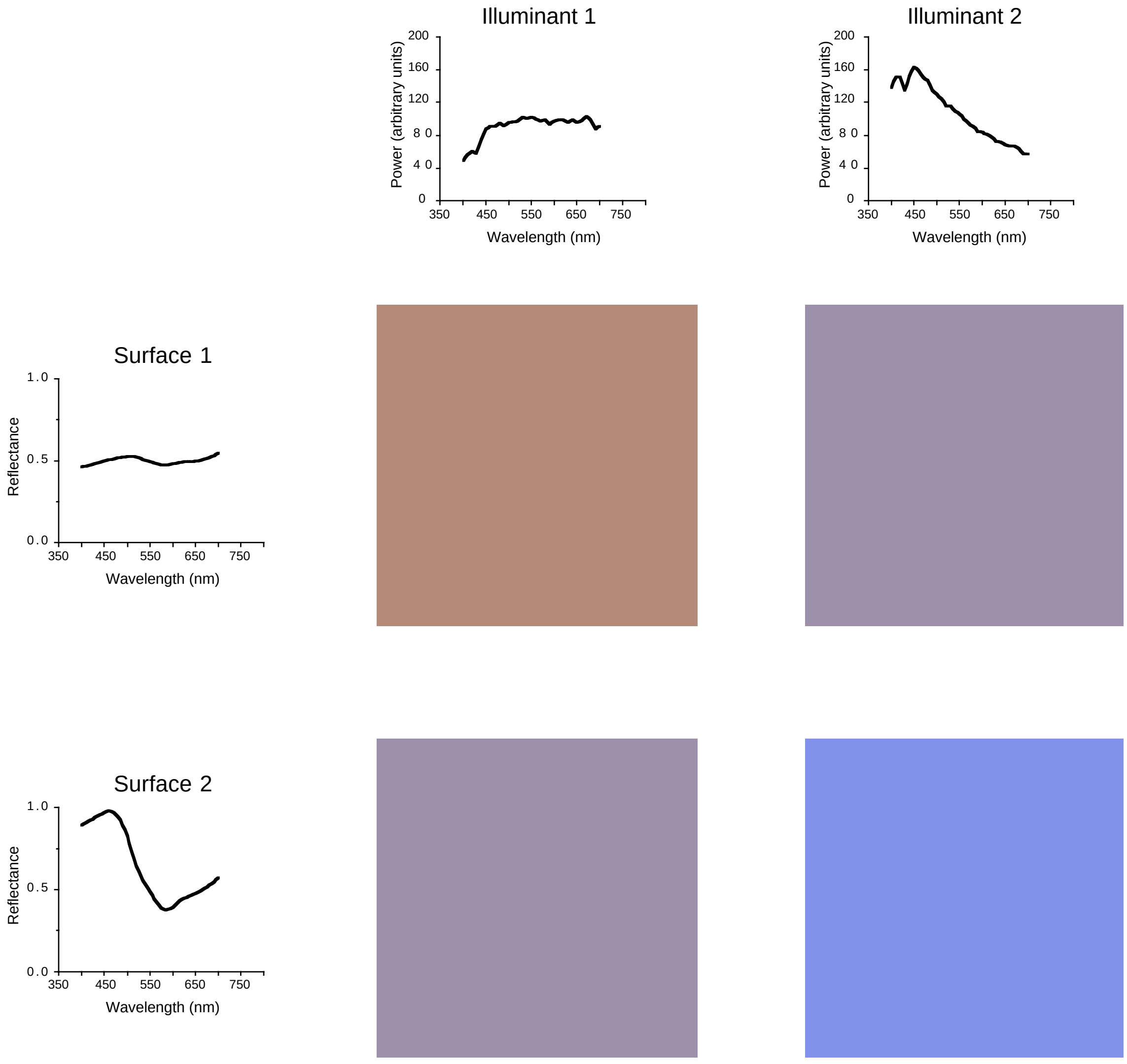

Plate 1. Renderings of two surfaces under two illuminants. The top row shows the same surface rendered under two different illuminants. Each rendering was obtained by using an illuminant spectral power distribution and surface reflectance function to compute the spectrum of the color signal. From this, the Smith-Pokorny estimates (Smith and Pokorny, 1975; DeMarco, Pokorny and Smith, 1992) of the L-, M-, and S-cone spectral sensitivities were used to obtain the quantal absorption rates of each cone class in response to the color signal. These in turn were used together with typical red, green, and blue phosphor emission spectra and monitor gamma curves to compute (cont.)

(Plate 1 cont.) RGB coordinates for the rendering. The RGB coordinates were chosen using standard methods (e.g. Brainard, 1995) so that the light they cause to be emitted from the monitor has the same effect on the cones as the color signal being rendered. The RGB coordinates were used to produce the plate by methods outside of the authors' control. The spectral plots show the surface reflectance functions and illuminant spectral power distributions used for this example.

Plate 2. Pictures of experimental chamber when spatial average has been equated. This plate shows pictures of the experimental chamber used by Kraft and Brainard (1999). Across the two images, both the illuminant and the surfaces in the scene have been changed. The two changes have a reciprocal effect, so that the spatial average of the L-, M-, and S-cone quantal absorption rates is the same in both images. The images shown are rendered versions of hyperspectral images taken of the stimuli. The hyperspectral imaging system (Longère and Brainard, 2001) provided 31 narrowband (approx. 10 nm bandwidth at 10 nm spacing between 400 and 700 nm) images of the scene. The hyperspectral images were also used to determine the spatial average of the cone quantal absorption rates. Plate adopted from Figure 1 of Kraft and Brainard (1999).

ACKNOWLEDGEMENTS

Supported by NIE EY10016. We thank E. Adelson, P. Delahunt, W. Freeman, and L. Maloney for useful discussions.

REFERENCES

Adelson, E. H. (1999). “Lightness perception and lightness illusions”. The New Cognitive Neurosciences, 2nd ed. M. Gazzaniga (Ed.). Cambridge, MA, MIT Press: 339-351.

Arend, L. E. (1993). “How much does illuminant color affect unattributed colors?” Journal of the Optical Society of America A **10**: 2134-2147.

Arend, L. E. and A. Reeves (1986). “Simultaneous color constancy.” Journal of the Optical Society of America A **3**: 1743-1751.

Arend, L. E., A. Reeves, J. Schirillo and R. Goldstein (1991). “Simultaneous color constancy: papers with diverse Munsell values.” Journal of the Optical Society of America A **8**: 661-672.

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

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

Beck, J. (1959). “Stimulus Correlates for the judged illumination of a surface.” Journal of Experimental Psychology **58**: 267-274.

Beck, J. (1961). “Judgments of surface illumination and lightness.” Journal of Experimental Psychology **61**: 368-373.

Bloj, M., D. Kersten and A. C. Hurlbert (1999). “Perception of three-dimensional shape influences colour perception through mutual illumination.” Nature **402**: 877-879.

Brainard, D. H. (1995). “Colorimetry”. Handbook of Optics: Volume 1. Fundamentals, Techniques, and Design. M. Bass (Ed.). New York, McGraw-Hill**:** 26.1-26.54.

Brainard, D. H. (1998). “Color constancy in the nearly natural image. 2. achromatic loci.” Journal of the Optical Society of America A **15**: 307-325.

Brainard, D. H., W. A. Brunt and J. M. Speigle (1997). “Color constancy in the nearly natural image. 1. asymmetric matches.” Journal of the Optical Society of America A **14**: 2091-2110.

Brainard, D. H. and W. T. Freeman (1997). “Bayesian color constancy.” Journal of the Optical Society of America A **14**: 1393-1411.

Brainard, D. H. and B. A. Wandell (1986). “Analysis of the retinex theory of color vision.” Journal of the Optical Society of America A **3**: 1651-1661.

Brainard, D. H. and B. A. Wandell (1991). “A bilinear model of the illuminant's effect on color appearance”. Computational Models of Visual Processing. M. S. Landy and J. A. Movshon (Ed.^Eds). Cambridge, MA, MIT Press.

Brainard, D. H. and B. A. Wandell (1992). “Asymmetric color-matching: how color appearance depends on the illuminant.” Journal of the Optical Society of America A **9**: 1433-1448.

Brainard, D. H., B. A. Wandell and E.-J. Chichilnisky (1993). “Color constancy: from physics to appearance.” Current Directions in Psychological Science **2**: 165-170.

Brown, R. O. and D. I. A. MacLeod (1997). “Color appearance depends on the variance of surround colors.” Current Biology **7**: 844-849.

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

Burnham, R. W., R. M. Evans and S. M. Newhall (1957). “Prediction of color appearance with different adaptation illuminations.” Journal of the Optical Society of America **47**: 35-42.

Chichilnisky, E. J. and B. A. Wandell (1996). “Seeing gray through the on and off pathways.” Visual Neuroscience **13**: 591-596.

Cohen, J. (1964). “Dependency of the spectral reflectance curves of the Munsell color chips.” Psychon. Sci **1**: 369-370.

D'Zmura, M. and G. Iverson (1993). “Color constancy. I. Basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces.” J. Opt. Soc. Am. A **10**: 2148-2165.

D'Zmura, M., G. Iverson and B. Singer (1995). “Probabilistic color constancy”. Geometric Representations of Perceptual Phenomena: Papers in Honor of Tarow Indow's 70th Birthday. R. D. Luce, M. D'Zmura, D. Hoffman, G. Iverson and A. K. Romney (Eds). Mahwah, NJ, Lawrence Erlbaum Associates**:** 187-202.

D'Zmura, M. and P. Lennie (1986). “Mechanisms of color constancy.” Journal of the Optical Society of America A **3**: 1662-1672.

Delahunt, P. B. and D. H. Brainard (2000). “Control of chromatic adaptation: signals from separate cone classes interact.” Vision Research **in press**.

DeMarco, P., J. Pokorny and V. C. Smith (1992). “Full-spectrum cone sensitivity functions for X-chromosome-linked anomalous trichromats.” Journal of the Optical Society A **9**: 1465-1476.

Fairchild, M. D. and P. Lennie (1992). “Chromatic adaptation to natural and incandescent illuminants.” Vision Research **32**: 2077-2085.

Finlayson, G. D., P. H. Hubel and S. Hordley (1997). “Color by correlation.” Proceedings of the IS&T/SID Fifth Color Imaging Conference: Color Science, Systems, and Applications, Scottsdale, AZ, 6-11.

Forsyth, D. A. (1990). “A novel algorithm for color constancy.” International Journal of Computer Vision **5**: 5-36.

Funt, B. V. and M. S. Drew (1993). “Color space analysis of mutual illumination.” IEEE Transactions On Pattern Analysis And Machine Intelligence **15**: 1319-1326.

Funt, B. V., M. S. Drew and J. Ho (1991). “Color constancy from mutual reflection.” International Journal of Computer Vision **6**: 5-24.

Geisler, W. S. (1989). “Sequential ideal-observer analysis of visual discriminations.” Psychological Review **96**: 267-314.

Gelb, A. (1950). “Colour constancy”. Source Book of Gestalt Psychology. W. D. Ellis (Ed.). New York, The Humanities Press: 196-209.

Gershon, R. and A. D. Jepson (1989). “The computation of color constant descriptors in chromatic images.” Color Research and Application **14**: 325-334.

Gilchrist, A. and A. Jacobsen (1984). “Perception of lightness and illumination in a world of one reflectance.” Perception **13**: 5-19.

Gilchrist, A., C. Kossyfidis, F. Bonato, T. Agostini, J. Cataliotti, et al. (1999). “An anchoring theory of lightness perception.” Psychological Review **106**: 795-834.

Gilchrist, A. L. (1977). “Perceived lightness depends on perceived spatial arrangement.” Science **195**: 185.

Gilchrist, A. L. (1980). “When does perceived lightness depend on perceived spatial arrangement?” Perception and Psychophysics **28**: 527-538.

Gilchrist, A. L. (1988). “Lightness contrast and failures of constancy: a common explanation.” Perception and Psychophysics **43**: 415-424.

Green, D. M. and J. A. Swets (1966). Signal Detection Theory and Psychophysics. New York, John Wiley and Sons.

Helson, H. (1938). “Fundamental problems in color vision. I. The principle governing changes in hue, saturation and lightness of non-selective samples in chromatic illumination.” Journal of Experimental Psychology **23**: 439-476.

Helson, H. and V. B. Jeffers (1940). “Fundamental problems in color vision. II. Hue, lightness, and saturation of selective samples in chromatic illumination.” Journal of Experimental Psychology **26**: 1-27.

Helson, H. and W. C. Michels (1948). “The effect of chromatic adaptation on achromaticity.” Journal of the Optical Society of America **38**: 1025-1032.

Hunt, R. W. G. (1950). “The effects of daylight and tungsten light-adaptation on color perception.” Journal of the Optical Society of America **40**: 336-371.

Hurlbert, A. (1986). “Formal connections between lightness algorithms.” Journal of the Optical Society of America **3**: 1684-1694.

Hurlbert, A. C. (1998). “Computational models of color constancy”. Perceptual Constancy: Why Things Look As They Do. V. Walsh and J. Kulikowski (Ed.^Eds). Cambridge, Cambridge University Press: 283-322.

Jaaskelainen, T., J. Parkkinen and S. Toyooka (1990). “A vector-subspace model for color representation.” Journal of the Optical Society of America A **7**: 725-730.

Jenness, J. W. and S. K. Shevell (1995). “Color appearance with sparse chromatic context.” Vision Research **35**: 797-805.

Jin, E. W. and S. K. Shevell (1996). “Color memory and color constancy.” Journal of the Optical Society of America A **13**: 1981-1991.

Judd, D. B., D. L. MacAdam and G. W. Wyszecki (1964). “Spectral distribution of typical daylight as a function of correlated color temperature.” Journal of the Optical Society of America **54**: 1031-1040.

Katz, D. (1935). The World of Colour. London, Kegan, Paul, Trench, Truber & Co., Ltd.

Knill, D. C. and D. Kersten (1991). “Apparent surface curvature affects lightness perception.” Nature **351**: 228--230.

Koffka, K. (1935). Principles of Gestalt Psychology. New York, Harcourt, Brace & Company.

Kozaki, A. and K. Noguchi (1976). “The relationship between perceived surface-lightness and perceived illumination.” Psychological Research **39**: 1-16.

Kraft, J. M. and D. H. Brainard (1999). “Mechanisms of color constancy under nearly natural viewing.” Proc. Nat. Acad. Sci. USA **96**: 307-312.

Kuriki, I. and K. Uchikawa (1996). “Limitations of surface-color and apparent-color constancy.” Journal of the Optical Society of America A **13**: 1622-1636.

Kuriki, I. and K. Uchikawa (1998). “Adaptive shift of visual sensitivity balance under ambient illuminant change.” Journal of the Optical Society of America A **15**: 2263-2274.

Land, E. H. (1986). “Recent advances in retinex theory.” Vision Research **26**: 7-21.

Land, E. H. and J. J. McCann (1971). “Lightness and retinex theory.” Journal of the Optical Society of America **61**: 1-11.

Lee, H. (1986). “Method for computing the scene-illuminant chromaticity from specular highlights.” Journal of the Optical Society of America A **3**: 1694-1699.

Logvinenko, A. and G. Menshikova (1994). “Trade-off between achromatic colour and perceived illumination as revealed by the use of pseudoscopic inversion of apparent depth.” Perception **23**: 1007-1023.

Longère, P. and D. H. Brainard (2001). “Simulation of digital camera images from hyperspectral input”. Vision Models and Applications to Image and Video Processing. C. van den Branden Lambrecht (Ed.), Kluwer. In press.

Lucassen, M. P. and J. Walraven (1993). “Quantifying color constancy: evidence for nonlinear processing of cone-specific contrast.” Vision Research: 739-757.

Lucassen, M. P. and J. Walraven (1996). “Color constancy under natural and artificial illumination.” Vision Research **36**: 2699-2711.

Maloney, L. T. (1986). “Evaluation of linear models of surface spectral reflectance with small numbers of parameters.” Journal of the Optical Society of America A **3**: 1673-1683.

Maloney, L. T. (1992). “Color constancy and color perception: the linear models framework”. Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience. D. E. Meyer and S. E. Kornblum (Eds). Cambridge, MA, MIT Press.

Maloney, L. T. (1999). “Physics-based approaches to modeling surface color perception”. Color Vision: From Genes to Perception. K. T. Gegenfurtner and L. T. Sharpe (Eds). Cambridge, Cambridge University Press**:** 387-416.

Maloney, L. T. and B. A. Wandell (1986). “Color constancy: a method for recovering surface spectral reflectances.” Journal of the Optical Society of America A **3**: 29-33.

Maloney, L. T. and J. N. Yang (2001). “The illuminant estimation hypothesis and surface color perception”. Colour Perception: From Light to Object. R. Mausfeld and D. Heyer (Eds). Oxford, Oxford University Press. This volume.

Mausfeld, R. (1998). “Color perception: from Grassman codes to a dual code for object and illumination colors”. Color Vision - Perspectives from Different Disciplines. W. G. K. Backhaus, R. Kliegl and J. S. Werner (Eds). Berlin, Walter de Gruyter & Co.**:** 219-250.

Mausfeld, R. and R. Niederee (1993). “An inquiry into relational concepts of colour, based on incremental principles of colour coding for minimal relational stimuli.” Perception **22**: 427-462.

McCann, J. J. (1994). “Psychophysical experiments in search of adaptation and the gray world.” Proceedings of IS&T's 47th Annual Conference, Rochester, NY, IS&T, 397-401.

McCann, J. J., S. P. McKee and T. H. Taylor (1976). “Quantitative studies in retinex theory: A comparison between theoretical predictions and observer responses to the 'Color Mondrian' experiments.” Vision Research **16**: 445-458.

Noguchi, K. and A. Kozaki (1985). “Perceptual scission of surface lightness and illumination: An examination of the Gelb effect.” Psychological Research **47**: 19-25.

Oyama, T. (1968). “Stimulus determinants of brightness constancy and the perception of illumination.” Japanese Psychological Research **10**: 146-155.

Parkkinen, J. P. S., J. Hallikainen and T. Jaaskelainen (1989). “Characteristic spectra of Munsell colors.” Journal of the Optical Society of America **6**: 318-322.

Romero, J., A. Garcia-Beltran and J. Hernandez-Andres (1997). “Linear bases for representation of natural and artificial illuminants.” Journal of the Optical Society of America A **14**: 1007-1014.

Rutherford, M. D. (2000). “The role of illumination perception in color constancy”. Unpublished Ph.D. thesis, Department of Psychology. UC Santa Barbara.

Singer, B. and M. D'Zmura (1994). “Color contrast induction.” Vision Research **34**: 3111-3126.

Smith, V. and J. Pokorny (1975). “Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm.” Vision Research **15**: 161-171.

Speigle, J. M. and D. H. Brainard (1996). “Luminosity thresholds: effects of test chromaticity and ambient illumination.” Journal of the Optical Society of America A **13**: 436-451.

Speigle, J. M. and D. H. Brainard (1999). “Predicting color from gray: the relationship between achromatic adjustment and asymmetric matching.” J. Opt. Soc. Am. A **16**: 2370-2376.

Stiles, W. S. (1967). “Mechanism concepts in colour theory.” Journal of the Colour Group **11**: 106-123.

Tominaga, S. and B. A. Wandell (1989). “The standard surface reflectance model and illuminant estimation.” Journal of the Optical Society of America **6**: 576-584.

Troost, J. M. and C. M. de Weert (1991). “Naming versus matching in color constancy.” Perception & Psychophysics **50**: 591-602.

Trussell, H. J. and M. J. Vrhel (1991). “Estimation of illumination for color correction.” Proceedings of IEEE ICASSP, 2513-2516.

Uchikawa, K., H. Uchikawa and R. M. Boynton (1989). “Partial color constancy of isolated surface colors examined by a color-naming method.” Perception **18**: 83-91.

Valberg, A. and B. Lange-Malecki (1990). “"Color constancy" in mondrian patterns: a partial cancellation of physical chromaticity shifts by simultaneous contrast.” Vision Research **30**: 371-380.

Wandell, B. A. (1987). “The synthesis and analysis of color images.” IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-9**: 2-13.

Werner, J. S. and J. Walraven (1982). “Effect of chromatic adaptation on the achromatic locus: the role of contrast, luminance and background color.” Vision Research **22**: 929-944.

Wesner, M. F. and S. K. Shevell (1992). “Color perception within a chromatic context - changes in red green equilbria caused by noncontiguous light.” Vision Research **32**: 1623-1634.

Wyszecki, G. and W. S. Stiles (1982). Color Science - Concepts and Methods, Quantitative Data and Formulae. New York, John Wiley & Sons.

Yang, J. N. (1999). “Illuminant estimation in surface color perception”. Department of Psychology. Unpublished Ph.D. thesis, New York University.