CHAPTER 10

COLOUR CONSTANCY: DEVELOPING EMPIRICAL TESTS OF COMPUTATIONAL MODELS

DAVID H. BRAINARD, JAMES M. KRAFT, AND PHILIPPE LONGÈRE

Preface

I first become interested in studying vision when, as an undergraduate, I read the first chapter of David Marr's book Vision *(Marr 1982). In that chapter, he articulates the view that vision can be understood as a system that extracts an explicit representation of the world from the retinal image, and that our understanding of human vision is usefully informed by consideration of machine vision algorithms that accomplish the same task. My studies, to that point, had focused on physics and computer science, and this was my first exposure to the notion that psychological questions (e.g. How does vision work?) could be connected to physics (e.g. image formation) and computer science (e.g. image processing). I found the idea sufficiently exciting that I pursued study in psychology.*

Subsequently, I ended up studying colour constancy because I viewed it as a relatively simple model problem that embodies the general processing task faced by vision: how can the visual system create a useful representation of surface properties (e.g. colour appearance) from a retinal image that confounds the physical properties of surfaces with those of the illuminant? An attractive feature of colour constancy is that there has been substantial progress both in our understanding of human performance and also in our understanding of how to achieve constancy in computer vision systems. In principle, though, these two lines can stand separately—one need not model human performance by drawing on the computational work, and computational solutions to colour constancy have application in digital image processing whether or not they connect to human performance. Indeed, in much of the literature the promise of connections between computation and performance has not been explicitly pursued. The idea that an understanding of the computational requirements of colour constancy can inform our study of human performance has, however, remained tantalizing.

In my own work, I have pursued both quantitative measurements of human constancy and have considered the computer vision problem presented by colour constancy. My hope remains that the two lines of research can indeed be brought together in a satisfactory fashion. In the present chapter, we review the current state of this enterprise, with particular emphasis on how psychophysical experiments can be structured so that the results speak directly to whether a particular computational theory is a good model of human colour vision.

David H. Brainard

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 colour, 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 colour signal reflected to the eye is shown schematically in Fig. 10.1. 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 colour 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 Fig. 10.2 illustrate the large effect that a typical change in natural illumination (see Wyszecki and Stiles 1982) can have on the colour signal. This effect might lead us to expect that the colour 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 colour indicates

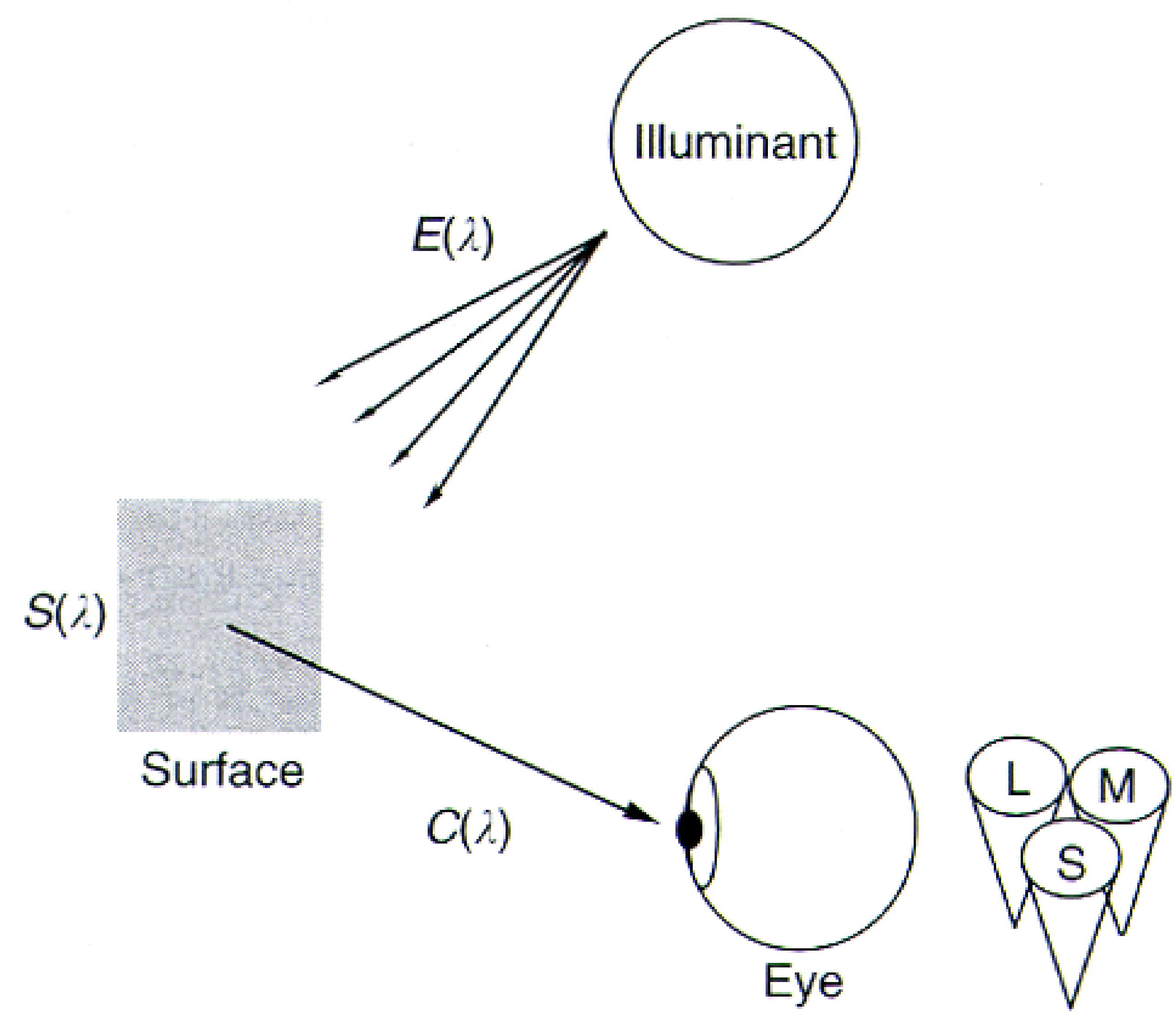

Figure 10.1 Effect of changing the illuminant on light reflected to the eye. The light incident on a surface is characterized by its spectral power distribution $E(\lambda)$. A small surface element reflects a fraction of the incident illuminant to the eye. The surface reflectance function $S(\lambda)$ specifies this fraction as a function of wavelength. The spectrum of light reaching the eye is called the colour 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. Note that this is a simplified imaging model. In general, the function $S(\lambda)$ depends on the geometry of the observer, illuminant, and object.

otherwise. Somehow our visual system stabilizes the colour appearance of objects against changes in illumination, a perceptual effect that is referred to as colour constancy.

Because the illumination is the most salient object-extrinsic factor that affects the colour signal, it is natural that emphasis has been placed on understanding how changing the illumination affects object colour appearance. In a typical colour constancy experiment, the independent variable is the illumination and the dependent variable is a measure of colour appearance (Helson 1938; Helson and Jeffers 1940; Helson and Michels 1948; Hunt 1950; Burnham *et al.* 1957; McCann *et al.* 1976; Arend and Reeves 1986; Valberg and Lange-Malecki 1990; Arend *et al.* 1991; Brainard and Wandell 1992; Lucassen and Walraven 1993, 1996; Bauml 1994, 1995; Brainard *et al.* 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 colour constancy.

Recall that the top two patches of Fig. 10.2 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

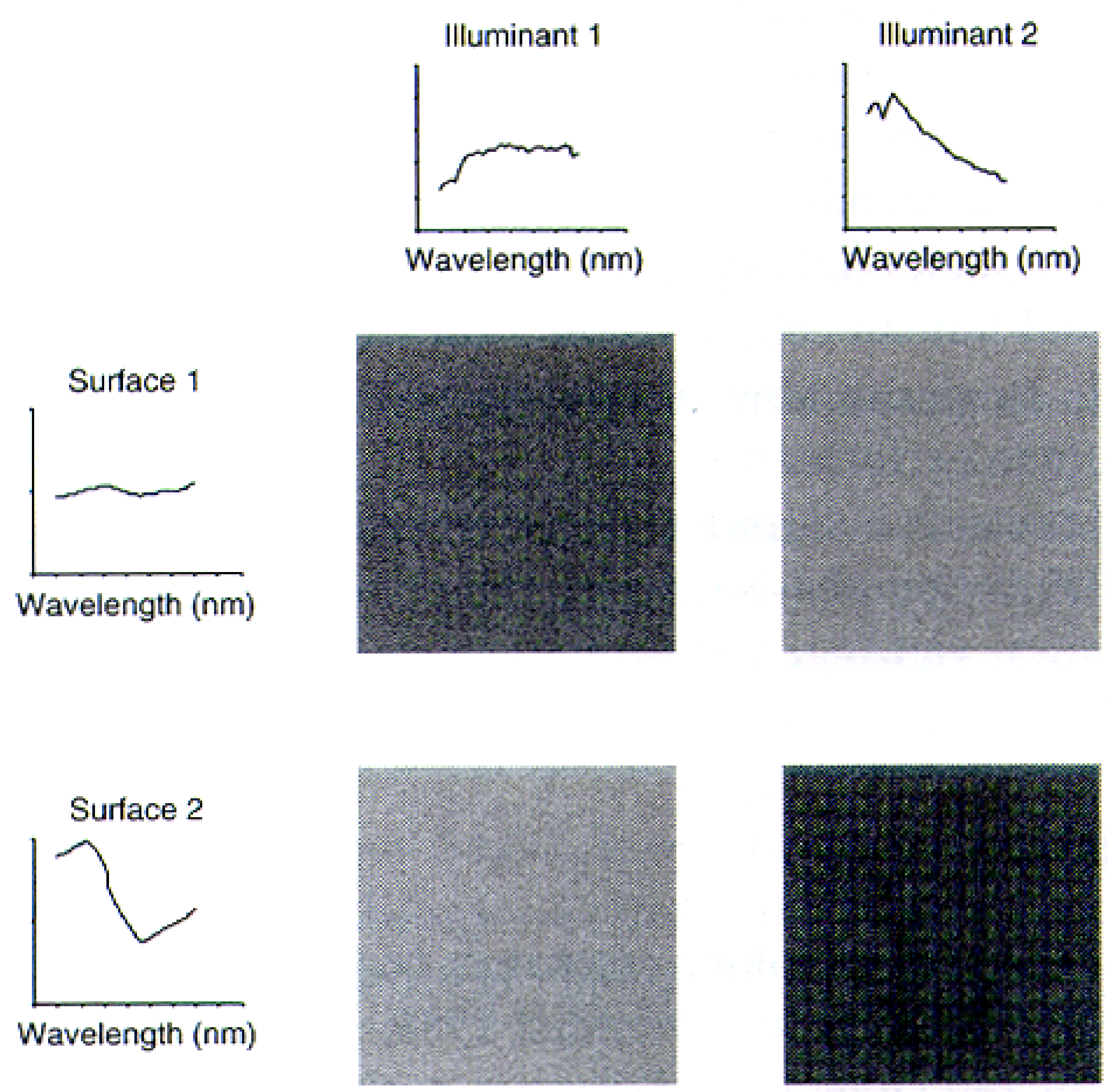

Figure 10.2 Renderings of two surfaces under two illuminants. The top row shows the same surface rendered under two different illuminants. Each rendering was obtained using an illuminant spectral power distribution and surface reflectance function to compute the spectrum of the colour signal. From this the Smith–Pokorny estimates (Smith and Pokorny 1975; DeMarco *et al.* 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 colour signal. These, in turn, were used, together with typical red, green, and blue phosphor emission spectra and monitor gamma curves, to compute 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 colour signal being rendered. The RGB coordinates were used to produce the figure by methods outside of the authors' control. The spectral plots show the surface reflectance functions and illuminant spectral power distributions used for this example. (See also colour Plate 28 in the centre of this book.)

faced by a visual system that attempts to achieve colour constancy. This is the ambiguity created because of the interaction between illuminant and surface reflectance, an ambiguity illustrated if we consider Fig. 10.2 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 colour 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 colour 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 Fig. 10.2.

Given that colour constancy is not possible in general, it makes little sense to provide a simple answer to the question of how colour constant human vision is. It is more sensible 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 colour constancy. The goal of computational theories is to define particular ensembles of scenes in which some degree of colour constancy is possible, and to express algorithms that achieve constancy for these ensembles. Computational theories of colour 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 programme designed to understand human colour 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 colour 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 colour constancy (e.g. Buchsbaum 1980; D'Zmura and Lennie 1986; Lee 1986; Maloney and Wandell 1986; Trussell and Vrhel 1991; D'Zmura and Iverson 1993; Funt and Drew 1993; D'Zmura *et al.* 1995; Brainard and Freeman 1997; Finlayson *et al.* 1997) share the same basic two-step framework. In the first step, the image is analysed 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 Buchsbaum'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 jth 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 nth entry of $\mathbf{s}_j$ is $S_j(\lambda_n)$. Given this representation, the spectral power distribution reflected to the eye from the jth surface is

$$\mathbf{c}_j = \mathrm{diag}(\mathbf{e})\ \mathbf{s}_j = \mathrm{diag}(\mathbf{s}_j)\ \mathbf{e}, \tag{10.1}$$

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 $\mathrm{L}(\lambda)$ be the spectral sensitivity of the L cones, $\mathrm{M}(\lambda)$ the spectral sensitivity of the M cones, and $\mathrm{S}(\lambda)$ the spectral sensitivity of the S cones. Form the 3 by N_λ matrix $\mathbf{R}$, where the nth entry of the first row of $\mathbf{R}$ is $\mathrm{L}(\lambda_n)$, the nth entry of the second row is $\mathrm{M}(\lambda_n)$, and the nth 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

$$\mathbf{r}_j = \mathbf{R}\,\mathbf{c}_j \tag{10.2}$$

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 *colour statistics* of the image.

It is straightforward to show that when we restrict attention to the Mondrian World, colour constancy remains an under-determined computational problem. It is possible to choose two illuminants and two collections of surfaces that produce identical colour 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 10.1 and 10.2 to obtain

$$\mathbf{r}_j = \mathbf{R}\,\mathbf{c}_j = \mathbf{R}\,\mathrm{diag}(\mathbf{e})\,\mathbf{B}_s\mathbf{w}_{sj} = \mathbf{R}\,\mathrm{diag}(\mathbf{s}_j)\,\mathbf{B}_e\mathbf{w}_e. \tag{10.3}$$

There is considerable evidence that small-dimensional linear models provide a reasonable description of many illuminants and surfaces (e.g. Cohen 1964; Judd *et al.* 1964; Maloney 1986; Parkkinen *et al.* 1989; Jaaskelainen *et al.* 1990; Romero *et al.* 1997; see Maloney 1992).

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

To understand how colour 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

$$\bar{\mathbf{r}} = \mathbf{R}\,\mathrm{diag}(\bar{\mathbf{s}})\mathbf{B}_e\mathbf{w}_e. \tag{10.4}$$

This follows because the spatial averaging operation commutes with the linear process of image formation described by Equation 10.3. If the spatial average of the surface reflectance is known, then Equation 10.4 may be inverted to solve for the illuminant:

$$\hat{\mathbf{e}} = \mathbf{B}_e\,\mathbf{M}_s^{-1}\,\bar{\mathbf{r}} \tag{10.5}$$

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 illuminants (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

$$\mathbf{s}_j = \mathbf{B}_s\,\mathbf{M}_e^{-1}\,\mathbf{r}_j \tag{10.6}$$

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 surfaces (3) is also matched to the number of human cone types (L, M, and S).

Equation 10.5 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 10.5 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 (1) the scene conforms to the Mondrian World assumptions; (2) the linear models $\mathbf{B}_e$ and $\mathbf{B}_s$ describe the illuminant and surfaces that comprise the scene; and (3) the actual spatial average of surfaces matches the assumed $\bar{\mathrm{s}}$.

Note that in Equation 10.5 the illuminant estimate depends on the scene only through the spatial average of the receptor responses, $\bar{\mathrm{r}}$. 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 and Wandell 1986; Forsyth 1990; Trussell and Vrhel 1991; D'Zmura *et al.* 1995; Brainard and Freeman 1997; Finlayson *et al.* 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 colour statistics only through their covariance matrix, while that returned by Forsyth's (1990) algorithm depends only on the convex hull of the colour statistics. As we will see below, understanding which properties of the colour 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 (D'Zmura and Lennie 1986; Hurlbert 1986; Lee 1986; Tominaga and Wandell 1989; D'Zmura and Iverson 1993; Funt and Drew 1993; see Hurlbert 1998; Maloney 1999). The algorithms associated with these theories generally estimate the illuminant using both information contained in the colour 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 judgements are not of the direct spectral functions but rather assess, in one way or another, the colour 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 $\boldsymbol{\sigma}$ is a vector whose entries describe the perceptual experience of colour. To connect an algorithm such as Buchsbaum's to human performance, we can suppose that $\boldsymbol{\sigma}$ is related to estimated surface reflectance $\hat{\mathrm{s}}$ by some unknown but fixed function f(), so that $\boldsymbol{\sigma} = \mathrm{f}(\hat{\mathrm{s}})$. Although the form of 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 colour 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 colour matching (e.g. Burnham *et al.* 1957; Stiles 1967; Arend and Reeves 1986; Brainard and Wandell 1992; Brainard *et al.* 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 10.5 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 eqn 10.6). 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 colour constancy, with the goal of connecting the experiments to the ideas discussed above.

Colour constancy in the nearly natural image

The effect of the illuminant

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

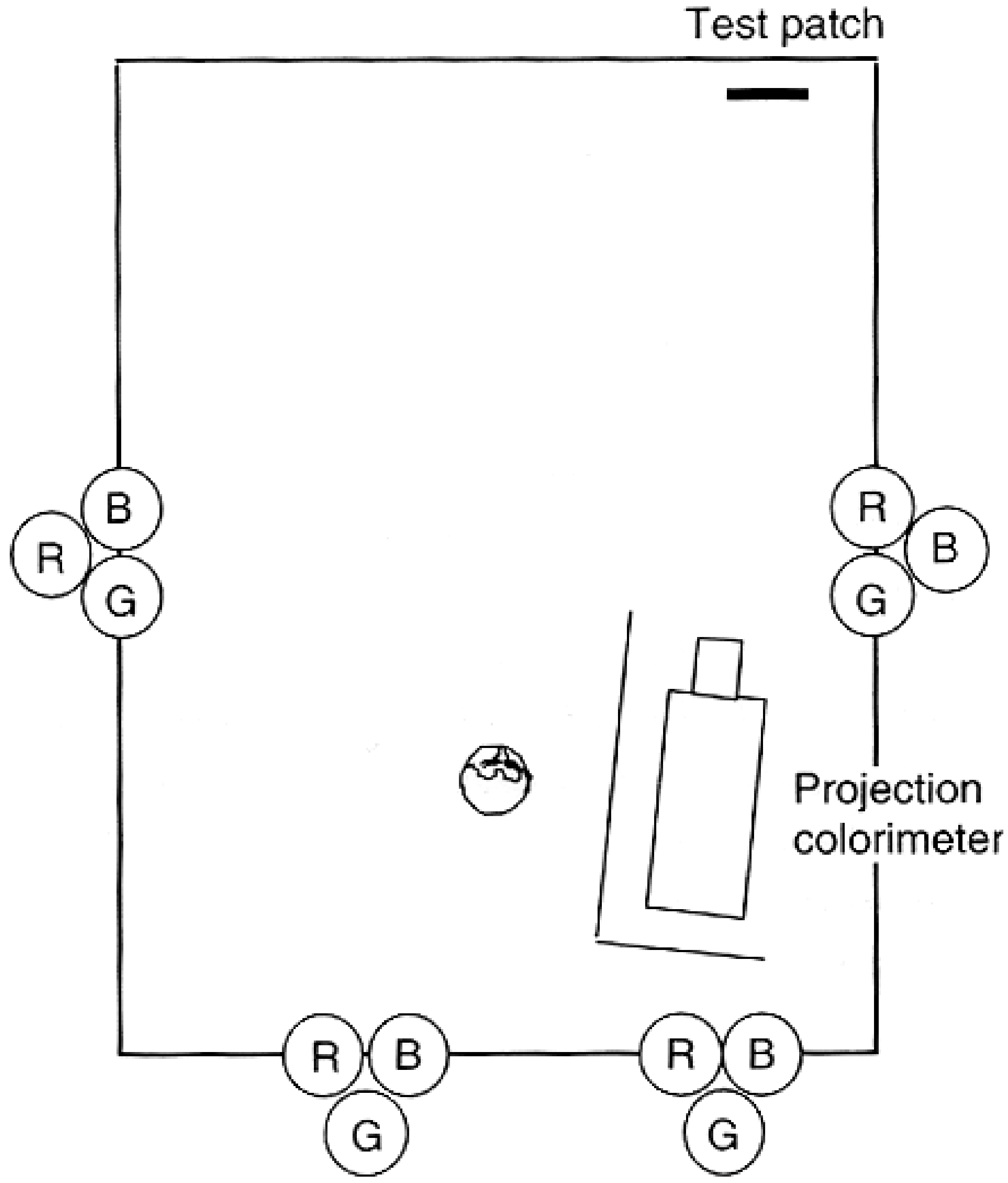

Figure 10.3 Room apparatus. Schematic of the experimental room. The dimensions of the experimental room were approximately 4 m × 3 m. Four triads of computer-controlled lights provided the ambient illumination. A projection colorimeter allowed adjustment of the colour appearance of a test patch located on the far wall of the room. (Adopted from Figure 1 of Brainard 1998.)

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 colour 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 Adelson 1999; Gilchrist *et al.* 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 colour 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 Fig. 10.3 and described in detail elsewhere (Speigle and Brainard 1996; Brainard *et al.* 1997; Brainard 1998). The ambient illumination of the room is produced by three 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 coloured shadows. By varying the intensities of the three sets of lamps, we can vary the spectral power distribution of the ambient illumination.

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 colour appearance of the test surface. Essentially, we have taken the stimulus configuration exploited by Gelb (Gelb 1950; see also Katz 1935; Koffka, 1935) and brought it under computer control (see also Uchikawa *et al.* 1989; Valberg and Lange-Malecki 1990; Kuriki and Uchikawa 1996, 1998).

As noted above, asymmetric colour matching provides a convenient and natural experimental method for linking computational theory and human performance. This procedure is particularly well suited to studying colour constancy when there is a spatial change in the illumination (simultaneous colour 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 1997). It is also possible to use asymmetric matching to study colour constancy for the situation of interest here, uniformly illuminated scenes where the illuminant varies from one time to another (successive colour constancy; Brainard and Wandell 1991, 1992; 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, Chapter 11 this volume). 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 colour 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 10.4 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]

[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).

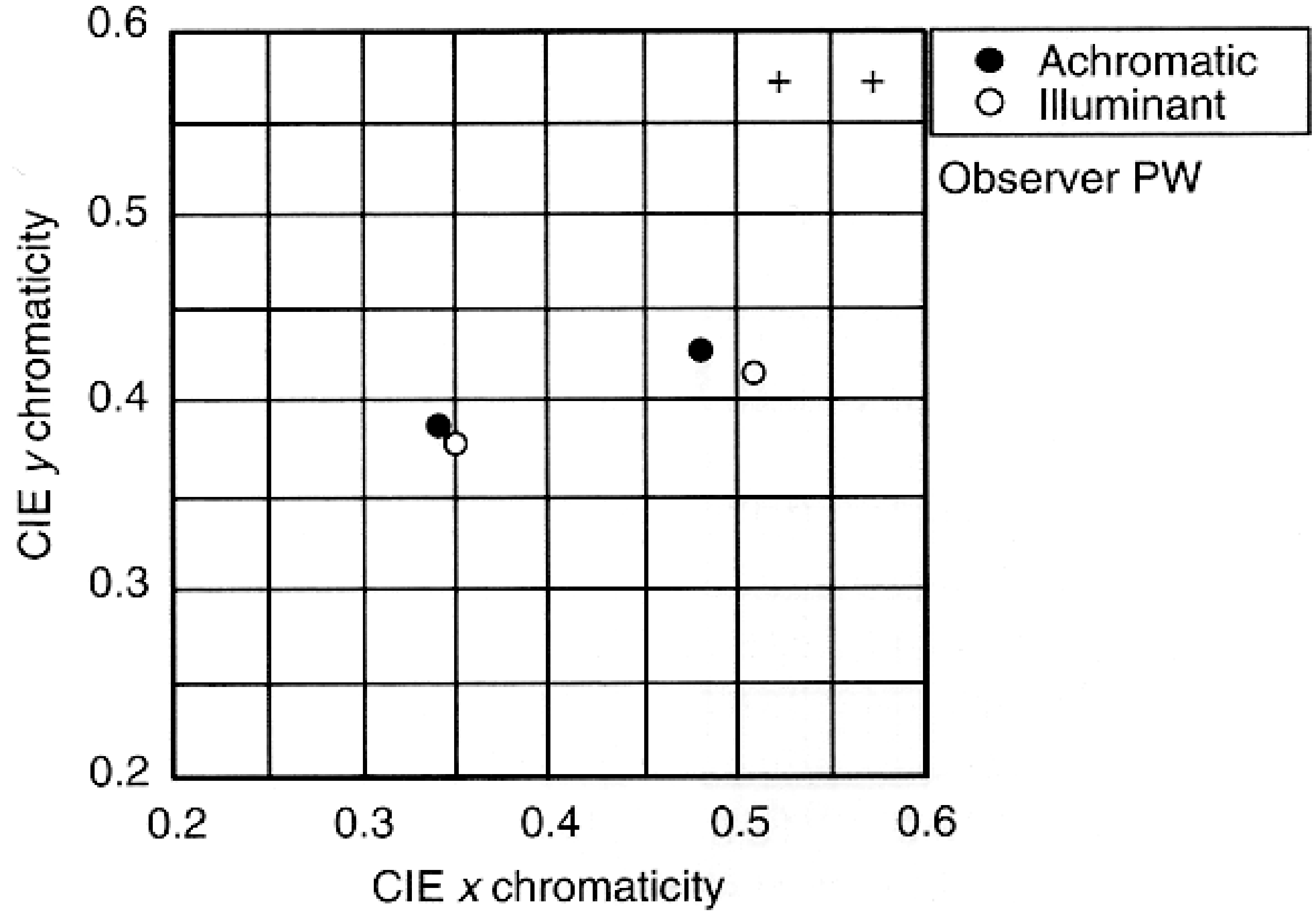

Figure 10.4 Basic achromatic results. The figure shows the CIE 1931 chromaticities of the achromatic loci (solid circles) measured under two experimental illuminants (chromaticity shown by open circles) for one observer. The between-session standard error of the mean is smaller than the plotted points. 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.)

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 colour 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 10.5 plots the chromaticity of the light reflected from a Munsell N 9.5/surface under two illuminants. This surface appears achromatic when seen under the standard viewing conditions for which the Munsell system is defined, and for a colour-constant visual system it will continue to appear achromatic under other viewing conditions. Thus for a colour-constant visual system, the chromaticity of the achromatic locus should coincide with the chromaticity of the light reflected from this surface. We conclude that colour constancy is indicated when the chromaticity of the achromatic loci lies near that of the illuminants (see Fig. 10.5). This pattern is roughly what is seen in the data shown in Fig. 10.4.

It is possible to go from the data shown in Fig. 10.4 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 Fig. 10.4 is 0.80, and the mean value across a wide range of conditions (different objects in the room,

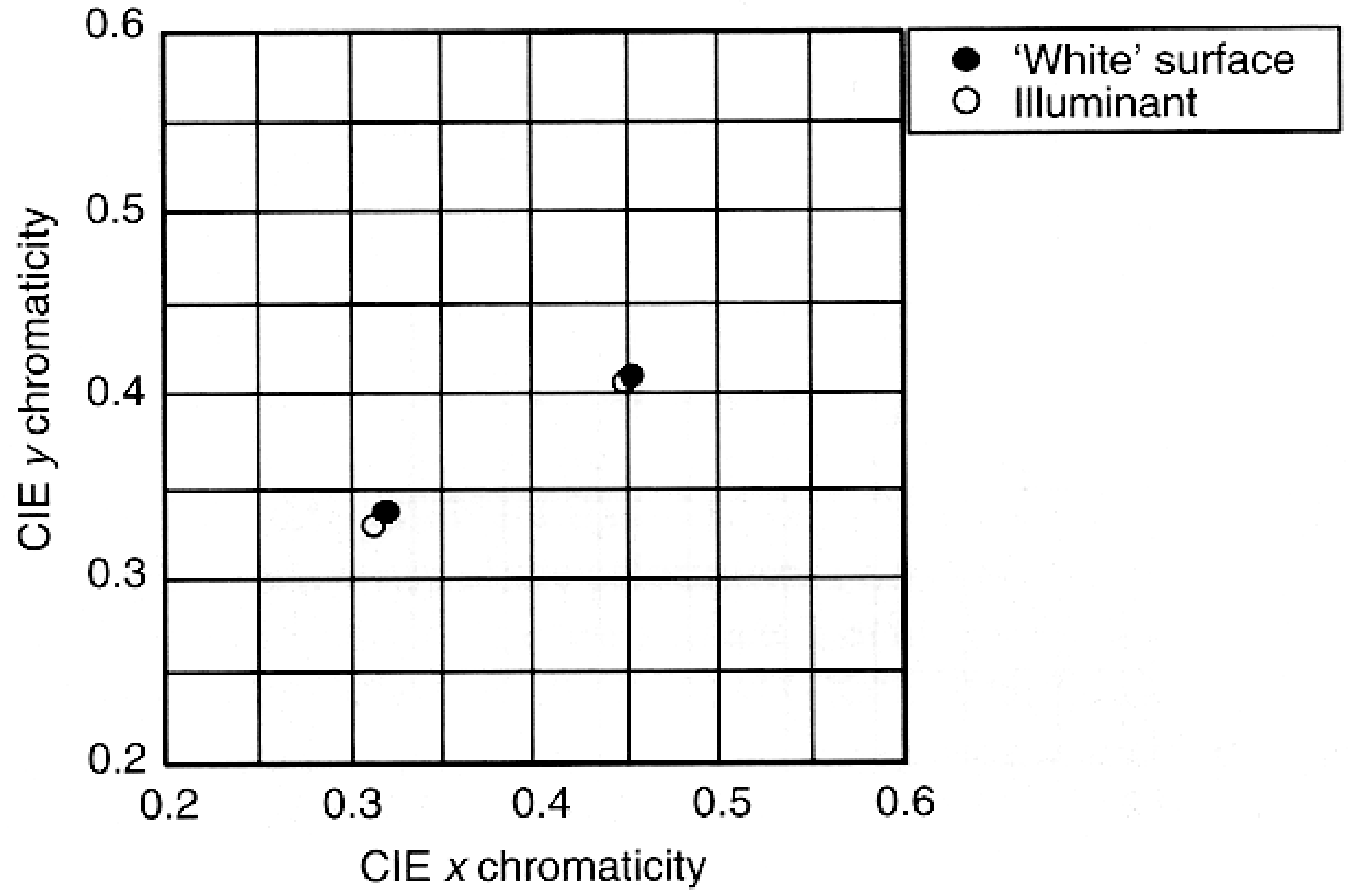

Figure 10.5 Data expected for a colour-constant visual system. The figure plots the chromaticity of the light reflected from a Munsell N 9.5/surface (solid circles) under two illuminants. The chromaticities of the illuminants are indicated by the open circles.

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; Fairchild and Lennie 1992; Brainard *et al.* 1993.) The relatively high constancy index shown by observers in our experiments is consistent with everyday experience: object colours 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 colour 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 colour constancy, eliminates from the stimulus ensemble the illuminant–surface ambiguity, described in the introduction, which 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 1 m × 1 m) chamber in which the spectrum of the illuminant and the spectral reflectance of all visible surfaces could be controlled independently. Figure 10.6 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 ($\bar{s}$) in the scene is quite different in the

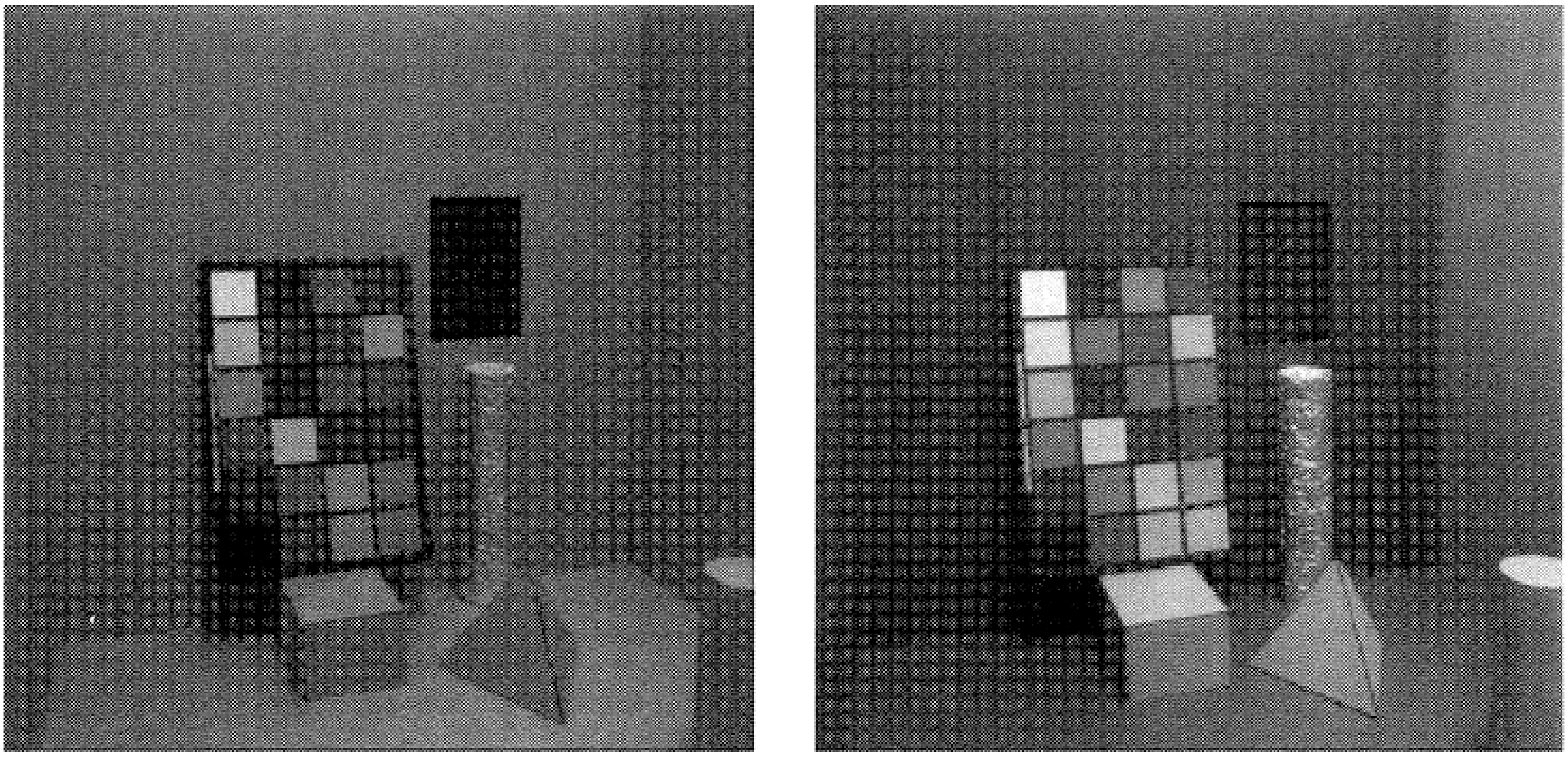

Figure 10.6 Pictures of the experimental chamber when the spectral average has been equated. This plate shows pictures of the experimental chamber used by Kraft and Brainard (1999). Across the two images, both the illuminant and the surfaces in the scene have been changed. The two changes have a reciprocal 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 narrow-band (approximately 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. (Adopted from Figure 1 of Kraft and Brainard 1999.) (See colour Plate 29 in the centre of this book.)

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 10.7 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 colour appearance. This,

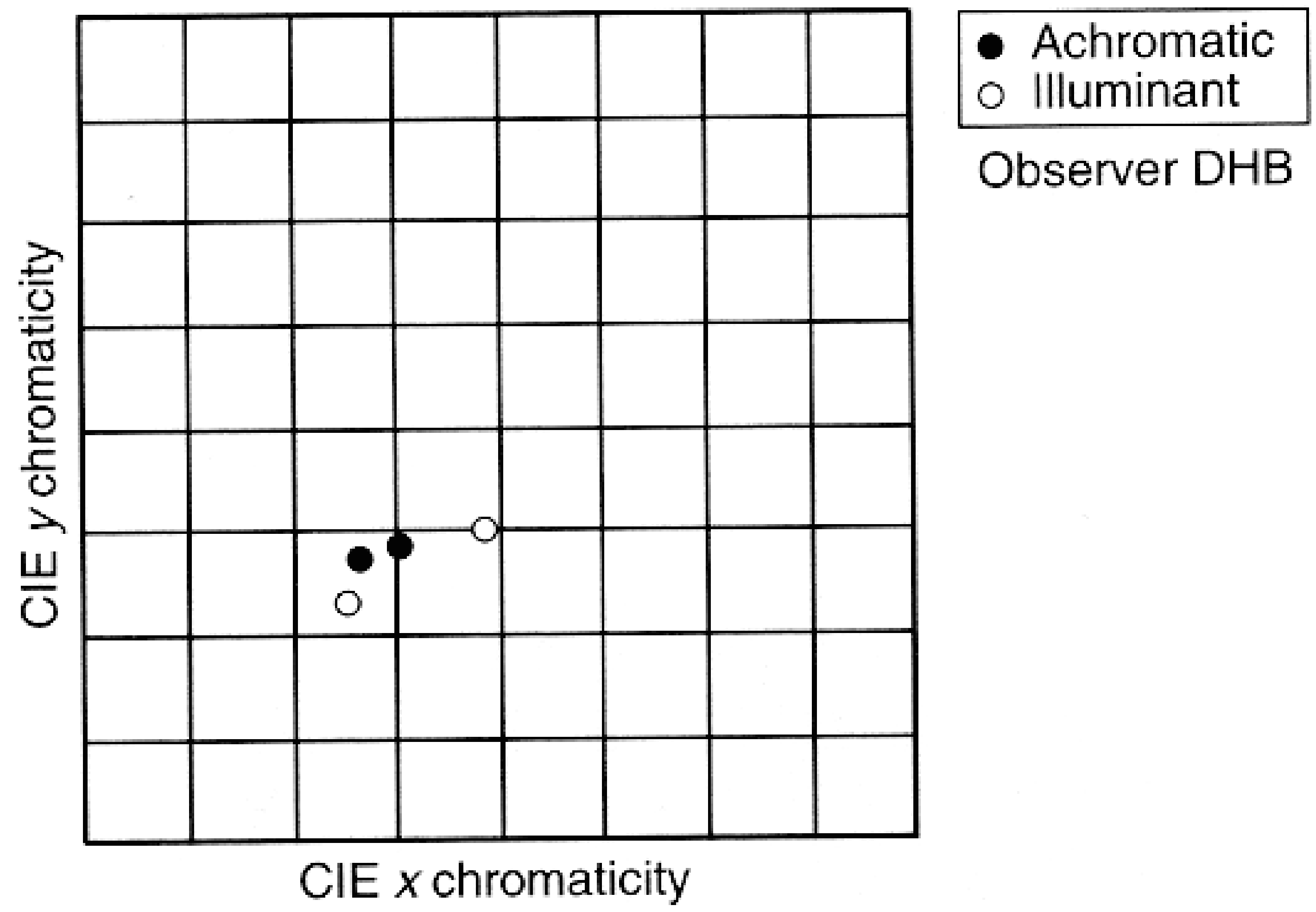

Figure 10.7 Achromatic settings with spatial average equated. The format of the figure is the same as for Fig. 10.4. Here the achromatic settings were made in the context of two images where the illuminant differed (open circles) but the spatial average of the image was held constant. The between-session standard error of the mean is smaller than the plotted points. (Data are replotted from Figure 2 of Kraft and Brainard 1999.)

in turn, says that Buchsbaum's algorithm cannot completely describe human performance. Perhaps it is worth noting that this result does not rule out the possibility that the algorithm would describe performance if the stimuli conformed strictly to the assumptions of the Mondrian World.

The constancy index for the data shown in Fig. 10.7 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 colour constancy and human performance. In doing so, we have implicitly endorsed what Maloney refers to as the *illumination estimation hypothesis* (Maloney and Yang, Chapter 11 this volume). 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 colour (see also Speigle and Brainard 1996; Brainard *et al.* 1997; Mausfeld 1998; Gilchrist *et al.* 1999). The work reviewed here does not directly test the illuminant estimation hypothesis, since observers do not make any judgements of perceived illumination. Recent work (Rutherford 2000) suggests that the illuminant estimation hypothesis is at best an approximation (see also Beck 1959, 1961; Oyama 1968; Kozaki and Noguchi 1976; Noguchi and Kozaki 1985; Logvinenko and Menshikova 1994). Even if human surface colour appearance does not depend on an explicit illuminant estimate, we need not refrain from using computational theory to develop and test models of which image statistics influence the perception of surface colour. Indeed, the models we have elaborated are designed to make predictions about asymmetric surface colour 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 colour 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 colour information.[2] The experiments do not, however, rule out a role 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 colour 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, Chapter 11 this volume) 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 (Gilchrist and Jacobsen 1984; McCann 1994; 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. More experiments are needed where the essential ambiguity between surfaces and illuminants is restored to the experimental situation.

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

[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 L-, M-, and S-cone responses (again Brainard and Wandell 1986; Land 1986). Strictly 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 (Singer and D'Zmura 1994; Jenness and Shevell 1995; Brown and MacLeod 1997).

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 colour 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 colour 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 colour appearance even when the colour statistics of the image are held fixed (Gilchrist 1977, 1980; Knill and Kersten 1991; Bloj *et al.* 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, Chapter 11 this volume) and mutual illumination (Funt *et al.* 1991; Funt and Drew 1993; see Bloj *et al.* 1999). Whether this is the case, or whether for our scenes the colour 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 (Adelson 1999; Gilchrist *et al.* 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 colour 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

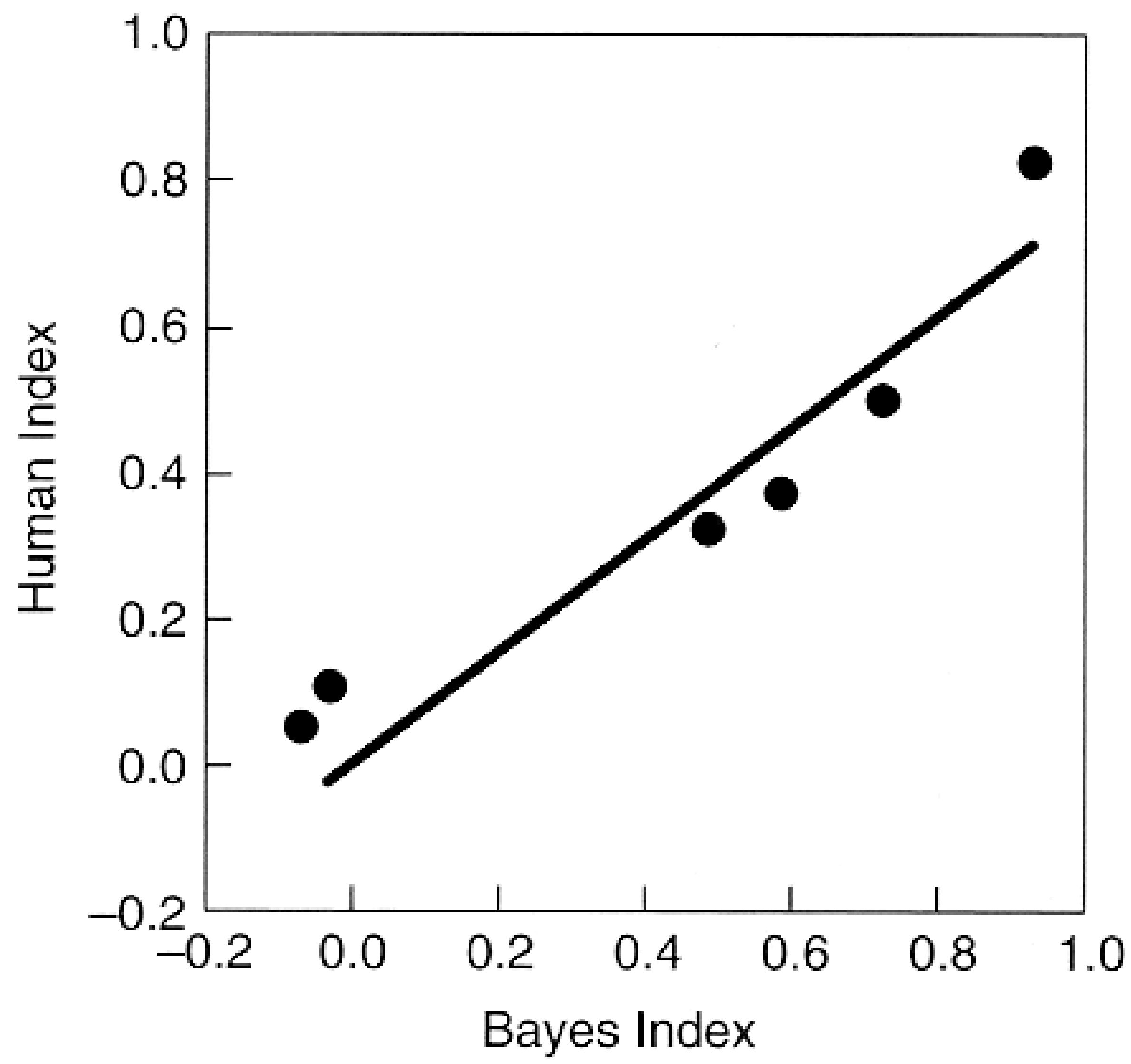

Figure 10.8 Comparison of human performance with Bayesian algorithm. The figure plots the constancy indices obtained for human observers against constancy indices obtained for the Bayesian algorithm of Brainard and Freeman (1997). The algorithm was run using points selected at random from calibrated LMS images of the stimulus. The image acquisition procedure is described in the caption for Fig. 10.6. 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 analysing 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.

such applications, one predicts the performance of an *ideal observer* that uses all of the information in the stimulus optimally 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 colour 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 its illuminant estimates in the same way that we treat the achromatic loci measured for human observers. Figure 10.8 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

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

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

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 colour 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 colour 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 colour 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.

Acknowledgements

Supported by NIE EY 10016. We thank E. Adelson, P. Delahunt, W. Freeman, and L. Maloney for useful discussions.

References

Adelson, E. H. (1999). Lightness perception and lightness. In *The New Cognitive Neurosciences* (ed. M. Gazzaniga) (2nd edn), pp. 339–351. MIT Press, Cambridge, MA.

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

Arend, L. E., Reeves, A., Schirillo, J., and **Goldstein, R.** (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.

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

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

Brainard, D. H. (1995). Colorimetry. In *Handbook of optics: Vol. 1. Fundamentals, techniques, and design,* (ed. M. Bass), pp. 26.1–26.54. McGraw-Hill, New York.

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. and **Freeman, W. T.** (1997). Bayesian color constancy. *Journal of the Optical Society of America A* 14, 1393–1411.

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

Brainard, D. H. and **Wandell, B. A.** (1991). A bilinear model of the illuminant's effect on color appearance. In *Computational models of visual processing* (ed. M. S. Landy and J. A. Movshon). MIT Press, Cambridge, MA.

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

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

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

Brown, R. O. and **MacLeod, D. I. A.** (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., Evans, R. M., and **Newhall, S. M.** (1957). Prediction of color appearance with different adaptation illuminations. *Journal of the Optical Society of America* 47, 35–42.

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

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

D'Zmura, M. and **Iverson, G.** (1993). Color constancy. I. Basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces. *Journal of the Optical Society of America A* 10, 2148–2165.

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

D'Zmura, M., Iverson, G., and **Singer, B.** (1995). Probabilistic color constancy. In *Geometric representations of perceptual phenomena: Papers in honor of Tarow Indow's 70th birthday* (ed. R. D. Luce, M. D'Zmura, D. Hoffman, G. Iverson, and A. K. Romney), pp. 187–202. Lawrence Erlbaum Associates, Mahwah, NJ.

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

DeMarco, P., Pokorny, J., and **Smith, V. C.** (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 **Lennie, P.** (1992). Chromatic adaptation to natural and incandescent illuminants. *Vision Research* 32, 2077–2085.

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

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

Funt, B. V. and **Drew, M. S.** (1993). Color space analysis of mutual illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **15**, 1319–1326.

Funt, B. V., **Drew, M. S.**, and **Ho, J.** (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. In *Source book of gestalt psychology* (ed. W. D. Ellis), pp. 196–209. The Humanities Press, New York.

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

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

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.

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

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

Green, D. M. and **Swets, J. A.** (1966). *Signal detection theory and psychophysics.* John Wiley and Sons, New York.

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

Helson, H. and **Michels, W. C.** (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. In *Perceptual constancy: Why things look as they do* (ed. V. Walsh and J. Kulikowski), pp. 283–322. Cambridge University Press, Cambridge.

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

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

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

Judd, D. B., **MacAdam, D. L.**, and **Wyszecki, G. W.** (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.* Kegan, Paul, Trench, Truber, London.

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

Koffka, K. (1935). *Principles of gestalt psychology*. Harcourt, Brace, New York.

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

Kraft, J. M. and **Brainard, D. H.** (1999). Mechanisms of color constancy under nearly natural viewing. *Proceedings of the National Academy of Sciences USA* **96**, 307–312.

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

Kuriki, I. and **Uchikawa, K.** (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 **McCann, J. J.** (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 **Menshikova, G.** (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 **Brainard, D. H.** (2001). Simulation of digital camera images from hyperspectral input. In *Vision models and applications to image and video processing* (ed. C. van den Branden Lambrecht), pp. 123–150. Kluwer, Dordrecht.

Lucassen, M. P. and **Walraven, J.** (1993). Quantifying color constancy: evidence for nonlinear processing of cone-specific contrast. *Vision Research* **33**, 739–757.

Lucassen, M. P. and **Walraven, J.** (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. In *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (ed. D. E. Meyer and S. E. Kornblum), MIT Press, Cambridge, MA.

Maloney, L. T. (1999). Physics-based approaches to modeling surface color perception. In *Color vision: From genes to perception* (ed. K. T. Gegenfurtner and L. T. Sharpe), pp. 387–416. Cambridge University Press, Cambridge.

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

Marr, D. (1982). *Vision*. W. H. Freeman, San Francisco.

Mausfeld, R. (1998). Color perception: from Grassman codes to a dual code for object and illumination colors. *Color vision—perspectives from different disciplines* (ed. W. G. K. Backhaus, R. Kliegl, and J. S. Werner), pp. 219–250. Walter de Gruyter, Berlin.

Mausfeld, R. and **Niederee, R.** (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, pp. 397–401.

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

Noguchi, K. and **Kozaki, A.** (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., Hallikainen, J., and **Jaaskelainen, T.** (1989). Characteristic spectra of Munsell colors. *Journal of the Optical Society of America* **6**, 318–322.

Romero, J., Garcia-Beltran, A., and **Hernandez-Andres, J.** (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. Ph.D. thesis, Department of Psychology. University of California, Santa Barbara.

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

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

Speigle, J. M. and **Brainard, D. H.** (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 **Brainard, D. H.** (1999). Predicting color from gray: The relationship between achromatic adjustment and asymmetric matching. *Journal of the Optical Society of America A* **16**, 2370–2376.

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

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

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

Trussell, H. J. and **Vrhel, M. J.** (1991). Estimation of illumination for color correction. *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing Society*, 2513–2516. Toronto, Canada May 14–17, 1991.

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

Valberg, A. and **Lange Malecki, B.** (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 **Walraven, J.** (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 **Shevell, S. K.** (1992). Color perception within a chromatic context—changes in red green equilbria caused by noncontiguous light. *Vision Research* **32**, 1623–1634.

Wyszecki, G. and **Stiles, W. S.** (1982). *Color science—concepts and methods, quantitative data and formulae.* John Wiley & Sons, New York.

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

Commentaries on Brainard, Kraft, and Longère

Surface colour perception and its environments

Laurence T. Maloney

The chapter by Brainard and colleagues begins with a brief summary of the computational literature on colour constancy and ends with a partial summary of the elegant and important experimental work on colour perception undertaken by Brainard and his colleagues over the past decade. The intent of the chapter is not simply to juxtapose model and experiment, but to emphasize the importance of the interplay between experimentation and theory in colour science, and it fulfils its intention very well.

In comparison with many other subdivisions of visual perception, colour science is in particular need of theory. There are (at least) two reasons why this is so. First of all, when we study how well human observers judge properties of the environment, including shape, or depth, and separation (length), we usually know, or can determine, the correct answer to any question that we pose to the observer. Put a bit more formally, we have agreed-upon measurement procedures that allow us to determine which of two lengths is longer, or which of two objects is further away (e.g. a ruler). We do not expect to disagree with other observers concerning judgements of length (Asch 1956) and we readily resolve differences between what we see and what we measure in favour of the measurements, explaining the discrepancy as due to a visual illusion (Coren and Girgus 1978). In such studies, theory still plays a role, and an important one, but its role is primarily to explain the patterns of deviation from what our measurement procedures tell us is 'ground truth'.

For colour perception, we typically don't know what counts as the right answer. We don't have measuring devices to tell us the (true) colour of an object; some researchers (e.g. Brown, Chapter 8 this volume) even reject the possibility that we could ever identify measurable, physical properties of objects that correspond to the subjective experience of colour. Consequently, the study of colour is typically framed in terms of invariances or constancies: the experimenter doesn't know what colour a homogeneous object 'should' be, but has the intuition that whatever it might be, it should remain the same under changes of illumination in the scene. If we were challenged to justify the claim that colour is invariant under a scene transformation, we could not do so. We would reply to the same challenge in the case of length by simply verifying that measured length remained invariant under the specified transformation.

It is interesting, then, that many of the computational theories of colour constancy that Brainard and colleagues mention begin with explicit models of light–surface interaction that include parametric descriptions of surface spectral reflectance functions. The ensuing colour constancy algorithm is a recipe for estimating these parametric surface descriptors, wholly or in part. These descriptors are, of course, measurable properties of the hypothetical surfaces postulated within the framework of each model, and the analogue of colour perception for these models is the explicit estimation of properties of surface in the environment. While such *estimation theories* noisily compete to describe human colour judgements, they quietly agree that colour is the subjective correlate of unspecified physical properties of surfaces (for a review, see Maloney 1999).

The first contribution of theory to the study of colour perception, then, is development of explicit models of what might count as the physical properties corresponding to colour. Implicit in the structure of such a theory is a claim that there is no fundamental difference between colour, on the one hand, and length or shape, on the other. We are simply less familiar with the rules governing colour in our world. If we eventually conclude that no estimation theory is an adequate description of human colour perception, then we will likely gain insight into the radical difference between perceptual attributes, such as length, that have agreed upon measurement procedure, and perceptual attributes, such as colour, that do not.

The study of colour vision is in need of theory for a second reason. The geometric structure of the environment around us is extremely well described as a Euclidean geometry, inside the laboratory and out. Three numbers characterize a location, and relations such as collinearity, orthogonality, parallelism, and so on are so well mirrored by ordinary geometry that we can pass from computational description to physical measurement and back with confidence. In contrast, we know of no accurate parametric descriptions of lights and surfaces in the natural environment that require as few as three parameters to characterize a light, and three parameters to characterize a surface. That is not to say that there are not models of lights and of surfaces that provide excellent approximations (Maloney 1999), but that a critical observer will likely be able to detect the difference between a natural environment and an approximation based on whatever three-parameter models of lights and surfaces. Further, the computational theories we have so far require that descriptions of light and surface use no more than three free parameters (Maloney 1999: MacLeod and Golz, Chapter 7 this volume). Consequently, the models are designed to operate in abstractions or idealizations of the natural environment that a human observer can discriminate from the natural environment, at least under some circumstances. I term these abstractions *environments* (Maloney, Chapter 9 this volume).

If we wish to test such a theory as a model of human performance, we can usefully divide our task into two. First, we examine the human observer's performance when placed in the idealized environment assumed by the theory. Typically, the prediction is that the observer will have perfect colour constancy across the range allowed by the environment. The model of MacLeod and Golz (Chapter 7 this volume) is an exception, in that it predicts failures of lightness constancy even within its environment. Then we can examine the behaviour of the model outside of its environment and compare the performance of the observer to the same altered environment.

The key in both cases is to simulate accurately an environment composed of idealizations of lights and surfaces specified mathematically. This sort of experiment is described in Maloney and Yang (Chapter 11 this volume). I argue that this approach is the correct way to test the kinds of computational models that have been developed in the past two decades.

In conclusion, then, our lack of understanding of the physics of light–surface interaction in the environment requires a tighter link between theory and experiment than in other areas of perception.

While Brainard and colleagues would likely agree with this general conclusion, they have taken a different tack in dealing with the uncertainty surrounding the proper environment for the study of human colour perception. They have developed a series of carefully controlled three-dimensional environments that they refer to as 'nearly natural'. These approximations, constructed of known surface materials and illuminants, allow them to measure human colour constancy performance under something like natural viewing conditions. Since the idealized environments accompanying computational models of colour constancy are invariably intended as approximations of the natural environment, the nearly natural environment cannot be too far away from their normal 'operating range'. As Brainard and colleagues note, the very high degree of colour constancy they find under some experimental conditions, and the very low colour constancy they find under others, is a strong indication that they have built an environment appropriate for the study of surface colour perception.

And yet I would argue that the future belongs to accurate simulations of arbitrary 'unnatural' environments, environments chosen to match the assumptions of a particular theory under test. The technology needed to do this sort of experiment is complex: it includes high-intensity binocular display devices with a wide spectral gamut, as well as accurate computer graphics, rendering software for simulating light–surface interactions in complex, three-dimensional scenes. Once this sort of equipment is readily available, it should be possible to explore the match between human colour performance and computational theories systematically.

References

Asch, S. E. (1956). Studies of independence and conformity. A minority of one against a uniform majority. *Psychological Monographs* **70** (9).

Coren, S. and Girgus, J. S. (1978). *Seeing is deceiving: The psychology of visual illusions.* Lawrence Erlbaum, Hillsdale, NJ.

Maloney, L. T. (1999). Physics-based models of surface color perception. In *Color vision: From genes to perception* (ed. K. R. Gegenfurtner and L. T. Sharpe), pp. 387–418. Cambridge University Press, Cambridge.

Commentaries on Brainard, Kraft, and Longère

Comparing the behaviour of machine vision algorithms and human observers

Vebjørn Ekroll and Jürgen Golz

> The indefiniteness in the concept of colour lies, above all, in the indefiniteness of the concept of the sameness of colours, i.e. of the method of comparing colours.
>
> Wittgenstein (1977)

The inner coherence of the previously published work of Brainard and his research associates[1] makes it tempting to speak of a concentrated research effort which seems to be less well defined by basic assumptions than by its dedication to tracking them down, being explicit about them, and subjecting them to empirical test. Their present contribution is no exception in this regard. In this case Brainard and colleagues isolate an assumption which is implicit to a large body of research on colour constancy, formalize it, and discuss its implications.

Although this is not always recognized, interdisciplinary cross-talk between psychologists and psychophysicists, on the one hand, and computer vision scientists and physicists, on the other, is not only impeded by the fact that we speak different academic languages, but also by the fact that these disciplines are not separate due to mere historical chance. They are separate mainly because they have different realms of phenomena as their subject, and any vaguely promising attempt to bridge the gap between them is worthy of our attention.

The *match-linking proposition hypothesis* proposed by Brainard and colleagues represents a conceptual clarification which eases the comparison of the predictions of colour constancy algorithms and human performance. Interesting implications of this linking proposition are deduced in an elegant and conclusive manner, and it is clear that the linking proposition makes a large number of interesting theoretical questions accessible to empirical investigation. The much commendable explicitness of the analysis by Brainard and colleagues also makes it a very interesting target for critical discussion.

Due to the important implications of the *match-linking proposition hypothesis*, it should be considered carefully whether the assumptions upon which it is based can be regarded as correct in any given experimental situation.

The basis for the linking proposition of Brainard *et al.* (see p. 314) is the assumption that there exists a function f() which relates estimated surface reflectances $\hat{s}$—algorithm output—to perceived colours, σ. This ensures that if the estimated surface reflectances $\hat{s}(A)$ and $\hat{s}(B)$ corresponding to two surfaces A and B are equal, their perceived colours $\sigma(A)$ and $\sigma(B)$ must also be equal. (The assumption that the function f() is one-to-one ascertains that the converse is also true.) It is assumed implicitly that the perceived colours, σ, which are the basis for human matching behaviour, are the appropriate counterparts to estimated surface reflectances $\hat{s}$. This implicit assumption seems to be a very plausible one, but it is by no means ascertained that it is correct. And if it is not, application of the *match-linking proposition hypothesis* may lead to erroneous or misleading conclusions, as will become clear in the following example.

A priori, it is not clear whether a colour match between two patches made by a human observer was made on the basis of perceived surface colour or some other variable. An obvious alternative is unasserted colour, a term introduced by Arend (1994) and defined as the chromatic counterpart of brightness, that is, an aspect of perceived colour which is presumed to be more elementary and prior to any parsing into surface and illumination performed by the visual system. In order to ascertain that subjects actually base their matches on perceived surface colour, and not on any other aspects which are not intended in investigations of colour constancy, they may be appropriately instructed. This is known to have a substantial effect on the matches made (Arend 1994). However, even when clear

[1] Among the more recent work: Brainard *et al.* (1997); Brainard (1998); Kraft and Brainard (1999); Speigle and Brainard (1999); Delahunt and Brainard (2000); Kraft *et al.* (2002); Rutherford and Brainard (2002).

instructions are given, asymmetric colour matches are difficult to make. The subjective difficulties associated with making asymmetric colour matches are well known (Katz 1911; Gelb 1929; Whittle 1994*a*,*b*), although it is less well understood why they appear. An immediately plausible explanation would be that these difficulties are due to the subjective uncertainties which may be associated with the comparably impoverished and artificial stimuli which are typically employed in such experiments (see p. 315).

As Brainard *et al.* note, there are many good reasons for studying more natural images than the typical cathode ray tube (CRT) displays. The subjective difficulty of making asymmetric matches gives a further good reason; if the subjective matching problems are due to the artificiality of the stimulus, they ought to disappear when more realistic stimuli are used, since they would be more likely to trigger a unique perceptual parsing into illumination and surface reflectance components. Interestingly, this is not the case, as is clear from the comments made by Brainard *et al.* (1997) in their study using a nearly naturalistic stimulus set-up. They state: 'The observers were able to set reliably what they regarded as the best match. At this match point, however, the test and the match surfaces looked different, and the observers felt as if further adjustments of the match surface should produce a better correspondence. Yet turning any of the knobs or combinations of knobs only increased the perceptual difference. We verified that the observers' adjustments near the best match were not limited by the gamut of our apparatus.' They suggest the following explanation for this phenomenon: 'One intriguing possibility is that our color experience at a location is described by more than three variables. This is possible if the influence of the illuminant (or, more generally, of the viewing context) has the effect of changing the perceptual representation of color in a way that cannot be compensated for simply by varying the tristimulus coordinates at a single location. Such an effect might be expected if the visual system uses color to code both surface and illuminant identity.'

This interpretation is supported by a recent analysis made by Niederée (1998), in which it is deduced from standard assumptions that a perceptually complete colour code for stimuli as simple as infield-surround configurations must be at least four-dimensional. If the visual system 'uses color to code both surface and illuminant identity' the function f( ) relating algorithm output to perceived colour would probably be more appropriately assumed to depend on both the estimated surface reflectance $\hat{s}$ and the estimated illumination $\hat{\imath}$, thus challenging the rationale for the *match-linking proposition hypothesis.* In this case, equal estimated surface reflectances obviously do not imply equal perceived colours.

The subjective matching problem and the presumably related fact that the perceived colour cannot be adequately represented by a three-dimensional colour code, even in simple stimulus configurations (Katz 1911; Evans 1949, 1964, 1974; Niederée 1998; see also Ekroll *et al.* 2002*a*,*b*, for a related phenomenon), should be taken seriously, not only as interesting phenomena in their own right, but also because our failure to understand them deprives us of a fuller understanding of results from the very promising research programme suggested and pursued by Brainard *et al.*

A preliminary strategy that could contribute to our understanding of these phenomena would be to investigate under which experimental conditions the subjective matching problems are particularly prominent, and under which conditions they are less prominent or even absent, as suggested by Bäuml (1999). It is, for instance, interesting to note that they have been reported to be absent in experiments with haploscopically superimposed displays (Whittle 1994*a*, *b*). However, it also seems imperative to develop ideas about which functional role the higher dimensionality of perceived colour might play. Some ideas about this can be found in Mausfeld (Chapter 13 this volume) and Niederée (1998).

Although we have focused on a problematic aspect of the match-linking hypothesis proposed by Brainard and colleagues, there is every reason to pay close attention to their research. The explicitness and empirical rigour of the approach is, in our opinion, bound to enhance our understanding of colour perception one way or another. An aspect of Brainard and colleagues' research which we have not addressed, but are particularly enthusiastic about, is the study of more naturalistic scenes under laboratory conditions. This line of research is likely to further our understanding of colour perception for several reasons, some of them mentioned by the authors themselves. A simple point demonstrating the value of this research is that, as far as results from artificial displays and naturalistic scenes differ, this discrepancy may draw our attention to factors and cues influencing colour perception which have been overlooked by present theory (Kraft *et al.* 2002; Logvinenko *et al.* 2002). And, as already noted, the experimental study of naturalistic displays will ultimately show whether problems and phenomena often attributed to the artificiality of the experimental stimulus may be discarded as such, or rather reflect problems of our theoretical concepts. Of course, naturalness of stimuli does not imply that the tasks subjects are asked to perform (e.g. asymmetric matching), and which seem natural on the basis of our—potentially misleading—theoretical preconceptions about what the visual system does or aims to do (e.g. estimate reflectances), are in fact natural with respect to the internal structure of the visual system (cf. Mausfeld, Chapter 13 this volume).

References

Arend, L. (1994). Surface colors, illumination, and surface geometry: Intrinsic-image models of human color perception. In *Lightness, brightness and transparency* (ed. A. L. Gilchrist), pp. 159–213. Lawrence Erlbaum Associates, Hillsdale, NJ.

Bäuml, K. H. (1999). Simultaneous color constancy: How surface color perception varies with the illuminant. *Vision Research* **39**(8), 1531–1550.

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

Brainard, D. H., Brunt, W. A., and **Speigle, J. M.** (1997). Color constancy in the nearly natural image. I. Asymmetric matches. *Journal of the Optical Society of America A. Optics and Image Science* **14**(9), 2091–2110.

Delahunt, P. B. and **Brainard, D. H.** (2000). Control of chromatic adaptation: signals from separate cone classes interact. *Vision Research* **40**(21), 2885–2903.

Ekroll, V., Faul, F., Niederée, R., and **Richter, E.** (2002*a*). *Das natürliche Zenrum der Chromatizitätsebene ist nicht immer achromatisch: Ein neuer methodischer Zugang zur Untersuchung relationer Farbkodierung.* Poster presented at the 5th Tübinger Wahrnehmungs-Konferenz. Available online at http://www.psychologie.uni-kiel.de/~vekroll/research.htm.

Ekroll, V., Faul, F., Niederée, R., and **Richter, E.** (2002*b*). The natural centre of chromaticity space is not always achromatic: A new look on colour induction. *Proceedings of the National Academy of Sciences USA* **99**(20), 13352–6.

Evans, R. M. (1949). On some aspects of white, gray and black. *Journal of the Optical Society of America* **39**(9), 774–779.

Evans, R. M. (1964). Variables of perceived colour. *Journal of the Optical Society of America* **54**(12), 1467–1474.

Evans, R. M. (1974). *The perception of color.* Wiley, New York.

Gelb, A. (1929). Die 'Farbenkonstanz' der Sehdinge. In *Handbuche der normalen und pathologischen Physiologie* (ed. A. Bethe, G. v. Bergman, G. Embden, and A. Ellinger), pp. 594–687. Springer, Berlin.

Katz, D. (1911). *Die Erscheinungsweisen der Farben und ihre Beeinflussung durch die individuelle Erfahrung.* Barth, Leipzig. (English translation of later, revised, edition available as Katz, D. (1935). *The world of color.* Kegan, Paul, Trench, Truber, London.)

Kraft, J. M. and **Brainard, D. H.** (1999). Mechanisms of color constancy under nearly natural veiwing. *Proceedings of the National Academy of Sciences USA* **96**(1), 307–312.

Kraft, J. M., Maloney, S. I., and **Brainard, D. H.** (2002). Surface-illuminant ambiguity and color constancy: effects of scene complexity and depth cues. *Perception* **31**, 247–263.

Logvinenko, A. S., Kane, J., and **Ross, D. A.** (2002). Is lightness induction a pictoral illusion? *Perception* **31**, 73–82.

Niederée, R. (1998). *Die Erscheinungsweisen der Farben und ihre stetigen Übergangsformen: Theoretische und Experimentelle Untersuchungen zur relationalen Farbkodierung und zur Dimensionalität vollständiger perzeptueller Farbcodes.* Postdoctoral thesis, Philosophische Fakultät der Christian-Albrechts-Universität, Kiel.

Rutherford, M. D. and **Brainard, D. H.** (2002). Lightness constancy: A direct test of the illumination estimation hypothesis. *Psychological Science* **13**, 142–149.

Speigle, J. M. & Brainard, D. H. (1999). Predicting color from gray: The relationship between achromatic adjustment and asymmetric matching. *Journal of the Optical Society of America A. Optics and Image Science Vis* **10**, 2370–2376.

Whittle, P. (1994*a*). Contrast brightness and ordinary seeing. In *Lightness, brightness and transparency* (ed. A. L. Gilchrist), pp. 111–158. Lawrence Erlbaum Associates, Hillsdale, NJ.

Whittle, P. (1994*b*). The psychophysics of contrast brightness. In *Lightness, brightness and transparency* (ed. A. L. Gilchrist), pp. 35–110. Lawrence Erlbaum Associates, Hillsdale, NJ.

Wittgenstein, L. (1977). *Remarks on colour.* Basil Blackwell, Oxford.

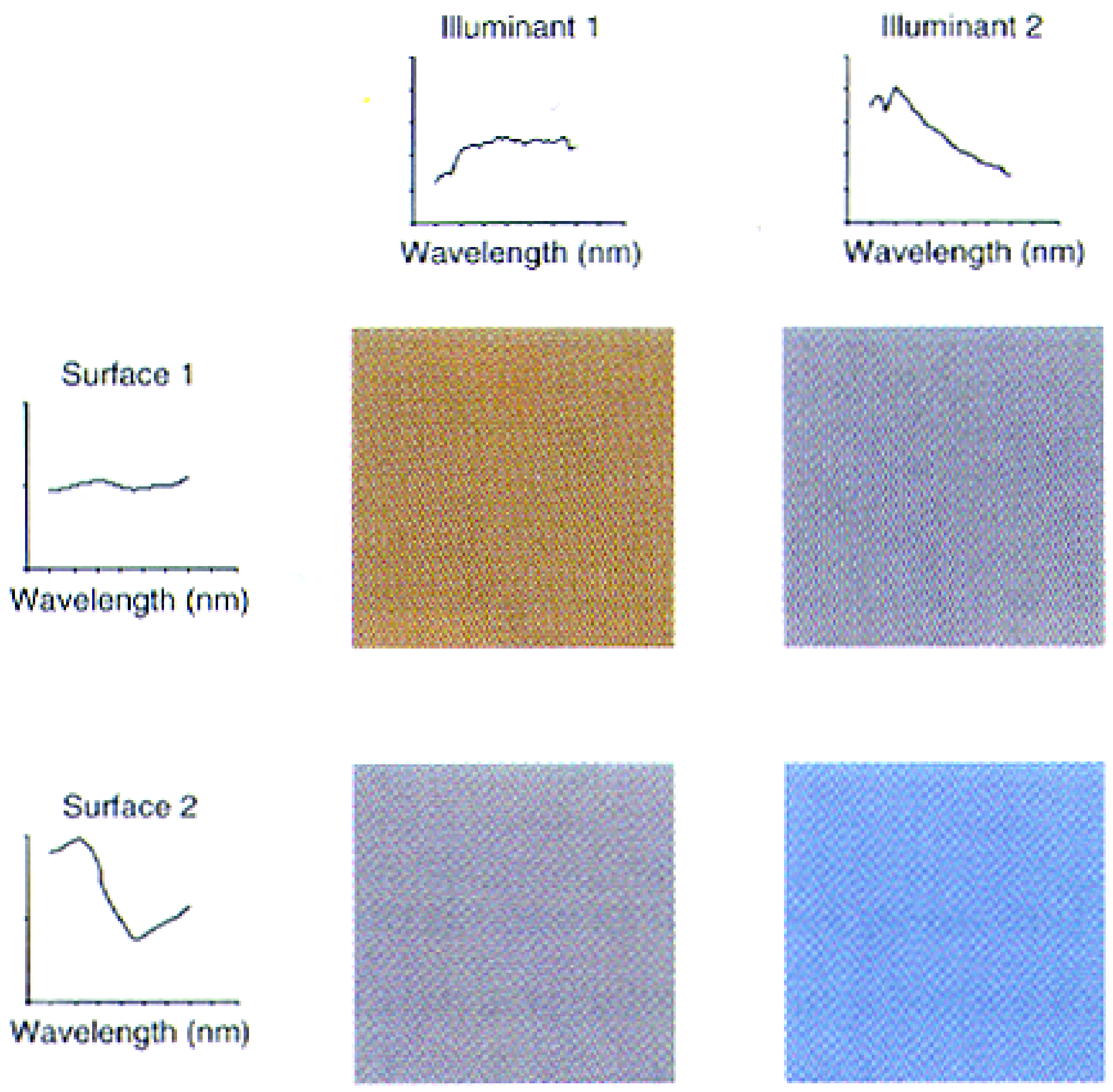

Plate 28 Renderings of two surfaces under two illuminants. The top row shows the same surface rendered under two different illuminants. Each rendering was obtained using an illuminant spectral power distribution and surface reflectance function to compute the spectrum of the colour signal. From this the Smith–Pokorny estimates (Smith and Pokorny 1975; DeMarco *et al.* 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 colour signal. These, in turn, were used, together with typical red, green, and blue phosphor emission spectra and monitor gamma curves, to compute 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 colour signal being rendered. The RGB coordinates were used to produce the figure by methods outside of the authors' control. The spectral plots show the surface reflectance functions and illuminant spectral power distributions used for this example (See Fig. 10.2.)

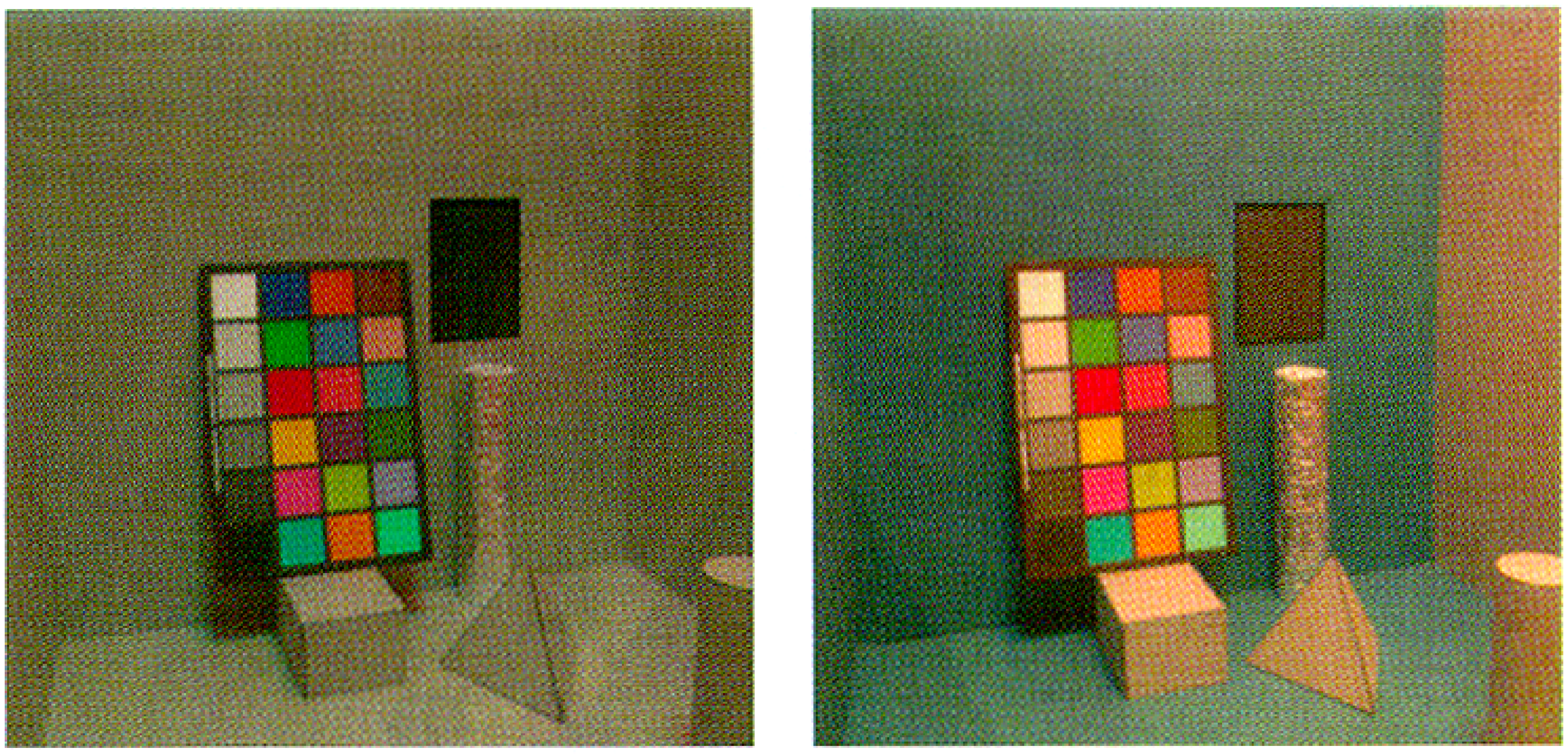

Plate 29 Pictures of the experimental chamber when the spectral average has been equated. This plate shows pictures of the experimental chamber used by Kraft and Brainard (1999). Across the two images, both the illuminant and the surfaces in the scene have been changed. The two changes have a reciprocal 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 narrow-band (approximately 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. (Adopted from Figure 1 of Kraft and Brainard 1999.) (See Fig. 10.6.)