Vision Research 51 (2011) 674–700

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Review

Color constancy

David H. Foster

Department of Electrical and Electronic Engineering, University of Manchester, Sackville Street, Manchester, M13 9PL England, UK

A R T I C L E I N F O

Article history:
Received 8 March 2010
Received in revised form 24 August 2010
Available online 16 September 2010

Keywords:
Color constancy
Color appearance
Chromatic adaptation
Von Kries coefficient
Surface color
Asymmetric color matching
Color naming
Achromatic adjustment
Relational color constancy
Spatial ratios of cone excitations
Color-constancy indices
Illuminant estimation
Natural scene statistics
Spectral basis functions

A B S T R A C T

A quarter of a century ago, the first systematic behavioral experiments were performed to clarify the nature of color constancy—the effect whereby the perceived color of a surface remains constant despite changes in the spectrum of the illumination. At about the same time, new models of color constancy appeared, along with physiological data on cortical mechanisms and photographic colorimetric measurements of natural scenes. Since then, as this review shows, there have been many advances. The theoretical requirements for constancy have been better delineated and the range of experimental techniques has been greatly expanded; novel invariant properties of images and a variety of neural mechanisms have been identified; and increasing recognition has been given to the relevance of natural surfaces and scenes as laboratory stimuli. Even so, there remain many theoretical and experimental challenges, not least to develop an account of color constancy that goes beyond deterministic and relatively simple laboratory stimuli and instead deals with the intrinsically variable nature of surfaces and illuminations present in the natural world.

1. Introduction

In its modern formulation, color constancy is usually taken as the effect whereby the perceived or apparent color of a surface remains constant despite changes in the intensity and spectral composition of the illumination.[1] The resulting changes in the spectrum of the light reflected from a scene are readily apparent over the course of a day (Romero, Hernández-Andrés, Nieves, & García, 2003), with the gamut of colors at sunset almost doubling under the mixture of direct and indirect illuminations (Hubel, 2000). Fig. 1 shows an example of the effect of an extreme change in daylight spectrum: a pelargonium illuminated by reddish direct sunlight and by bluish light from the north sky, along with the corresponding reflected spectra from the surface of a petal.[2] Isolated lights with these two spectra appear very different, but, in context, the petal surface reflecting these lights appears to be the same.

Color constancy has had a long history of analysis, with contributions from, among others, Monge (1789), Young (1807), von Helmholtz (1867), Hering (1920), and von Kries (1902, 1905), and later Helson and Jeffers (1940), Judd (1940), and Land and McCann (1971). Over much of this period, two opposing theoretical views of the phenomenon of color constancy held: one that it was the result of unconscious inference (Judd, 1940; von Helmholtz, 1867) and the other that it was the result of sensory adaptation (Helson, 1943; Hering, 1920). The experimental data that were available did not easily discriminate between these positions. As Mausfeld (2003) noted, as late as the 1970s, standard textbooks on color science were almost silent on the phenomenon: Boynton (1979) had a short section (pp. 183, 185) describing chromatic adaptation and essentially the coefficient rule of von Kries (1902, 1905), considered here in Section 5.2; Wyszecki and Stiles (1982) were similarly reserved, judging that for the more general problem of surface-color perception in complicated scenes, "the science of color has not advanced far enough to deal with this problem quantitatively" (p. 173).

In the 1980s, however, major developments took place with the first systematic behavioral experiments, by Arend and Reeves (1986), aimed at clarifying the nature of observers' color-constancy judgments, and the appearance of new models of color constancy, physiological data on cortical mechanisms, and photographic colorimetric measurements of natural scenes. In these and related works, the main questions of concern were as follows.

E-mail address: d.h.foster@manchester.ac.uk

[1] In a slightly different formulation, Beck (1972) stated that the perceived color of a surface *tends* to remain constant, but the modifier is more commonly omitted and color constancy itself described in graded terms.

[2] A flower from the same family revealed John Dalton's own particular loss in color constancy (Dalton, 1794; Hunt, Dulai, Bowmaker, & Mollon, 1995).

0042-6989/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2010.09.006

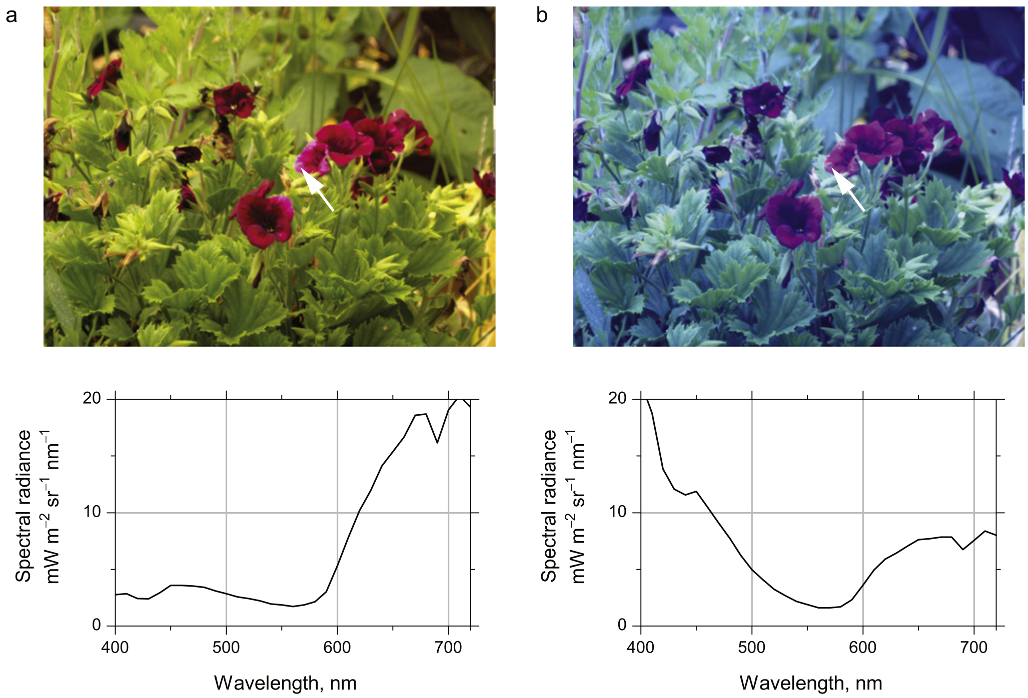

Fig. 1. Images of a pelargonium under sunlight and skylight with respective correlated color temperatures (a) 4000 K and (b) 25,000 K and the corresponding radiance spectra reflected from the arrowed region of a petal (simulated from author's unpublished hyperspectral data).

(1) *How is color constancy physically possible?* Following theoretical analyses by Brill (1978), and Buchsbaum (1980) using low-dimensional linear representations of illumination and reflectance spectra, a succession of linear models were developed and formalized by Maloney and Wandell (1986), D'Zmura and Lennie (1986), Hurlbert (1986), and others. These models made explicit the kinds of physical information required for color constancy, including the minimum number of degrees of freedom necessary to recover surface and illumination spectra uniquely. In practice, application of these models proved to be more difficult than expected.

(2) *What do observers judge?* The color-matching experiment by Arend and Reeves (1986) revealed the importance of using appropriate criteria to distinguish explicitly between adaptational and inferential aspects of subjects' judgments.[3] Although later researchers introduced criteria that probed other aspects of color constancy, this experiment was critical in quantifying the dual nature of subjects' judgments about stimulus appearance.

(3) *What experimental methods are suitable?* The approaches by Arend and Reeves (1986) and afterwards by Troost and de Weert (1991) adapted traditional psychophysical techniques of asymmetric color matching and color naming to the measurement of color constancy, and they introduced useful indices of subjects' performance. Other approaches followed, but the increasing variety of methods, adaptational conditions, and decision criteria led to uncertainty in assessing the varying levels of observer performance. To help comparisons, a table is presented here, in Section 4.1, in which constancy indices from a range of experiments are listed.

(4) *What physical scene properties are relevant?* Many theoretical approaches to color constancy made implicit assumptions about the properties of surface spectral reflectances and illuminants, in particular, that the effects of an illumination change could be offset by adapting cone photoreceptor responses to reflected light according to the coefficient rule of von Kries (1902, 1905). The rule was understood to be an approximation with real scenes, but little was known of its limits.

(5) *What neural mechanisms support color constancy?* Early experiments pointed to a special role of cortical area V4. Influential single-cell experiments by Zeki (1980, 1983) showed that monkey V4 cells responded to the surface color of a stimulus irrespective of its local spectral composition, and behavioral experiments by Wild, Butler, Carden, and Kulikowski (1985) showed that color identification under different illuminations was impaired when V4 was lesioned. Later experiments produced a more nuanced understanding of cortical and other mechanisms contributing to human color constancy.

(6) *Are natural scenes and surfaces special?* The properties of natural scenes are very different from those of traditional laboratory stimuli. Burton and Moorhead (1987), using photographic colorimetry, supplied the first detailed colorimetric and spatial analysis of the structure of natural scenes. Subsequent studies using different imaging techniques furnished more comprehensive data for modeling color constancy and for testing and interpreting observer performance. Even so, a careful study by Arend (2001) identified fundamental problems with defining color constancy in complex natural visual environments.

The aim of this review is to consider more closely how each of these questions has been addressed over the last quarter of a century; what new issues have arisen; what broadly constitutes the current state of knowledge; and what new challenges have emerged. The concern throughout is with human color constancy. References to constancy in animals and in machine vision are included only where they seem directly relevant. The review does not deal with color constancy in color-vision deficiency, whether inherited or acquired, or with lightness constancy, which has its

[3] These terms, adaptational and inferential, are interpreted here in relation to two kinds of observer judgments (Section 3.1), but see e.g. Mausfeld (2003).

own specialist literature, or, likewise, with color induction and color contrast, except where they also seem directly relevant. Previous reviews of color constancy have been provided by Smithson (2005) and Shevell and Kingdom (2008), and, with a more theoretical emphasis, by Hurlbert (1998), and Maloney (1999). Some of the historical literature has been summarized by Beck (1972) and Jameson and Hurvich (1989). Introductions to the work of Monge (1789) have been made available by Kuehni (1997) and Mollon (2006).

2. How is color constancy physically possible?

Although rarely articulated, it is implicit in the analysis of color constancy that the source of illumination is not known directly to the observer. This presumption is not as unnatural as it might seem, for even if the light source or sources are visible (e.g. sun, clear or cloudy sky, incandescent or narrow-band lamp), direct inspection with a trichromatic eye allows only a three-dimensional representation of the spectrum to be inferred.[4] Moreover, in the natural environment, the source itself may not be well defined in that the illumination at a particular point in a scene is usually a complex mixture of direct and indirect irradiation distributed over a range of incident angles, in turn modified by local occlusion and mutual reflection, all of which may vary with time and position. Accordingly, two rationales have been advanced for color constancy. First, it affords a stable percept of surface identity, independent of variations in the illumination spectrum, enabling the observer to interact veridically with the world (von Helmholtz, 1867; Zeki, 1993). Second, it makes possible an estimate of the illumination spectrum, including the phase of daylight, and, by inference, of the time of day and weather (Jameson & Hurvich, 1989; Reeves, 1992).[5]

The challenge for color constancy is that neither the spectral reflectance of a surface nor the spectral irradiance of the incident illumination can be readily estimated directly from the pattern of spectral radiance reflected from the surface into the eye. Mathematically, the recovery of reflectance from the image spectrum is an ill-posed problem, and, in general, does not have a unique solution: if at wavelength λ the spectral reflectance at a point is $r(\lambda)$ and the spectrum of the illumination is $e(\lambda)$, then $r(\lambda)$ cannot be recovered from the product, i.e. the reflected spectrum $c(\lambda) = r(\lambda)e(\lambda)$, without knowing $e(\lambda)$. Consequently, as Maloney and Wandell (1986) pointed out, without restrictions on spectral reflectances and illuminants, color constancy is impossible.

There have been several theoretical approaches to the problem of human color constancy, the main ones concentrating on so-called lightness algorithms, on directly estimating the illumination spectrum, on applying low-dimensional linear models, and on appealing to Bayes' rule. Despite differences in origin, they have some features in common.

2.1. Lightness algorithms

The first explicit algorithms designed to recover an approximation of surface spectral reflectance were described by Land in his Retinex models (Land, 1983, 1986; Land & McCann, 1971). These models were conceived as a description of human surface-color perception and were supported by demonstrations with illuminated Mondrian patterns[6] of colored papers (illustrated in Section 4.2). The principle of the models was that the spectral reflectance of any area could be approximated by the ratio of the light reflected from that area to the light reflected from one or more other areas along a path or paths evaluated within each of several spectral channels or bands. Bands were defined typically by the spectral sensitivities of the long-, medium-, and short-wavelength-sensitive cones. Although difficult to reconcile with human constancy performance (e.g. Brainard & Wandell, 1986; Shapley, 1986; Valberg & Lange-Malecki, 1990), Retinex algorithms became popular in color-management systems, with their parameters, including the number of integration paths, thresholds, and iterations, optimized for a range of applications (Funt, Ciurea, & McCann, 2004; Provenzi, De Carli, Rizzi, & Marini, 2005). Land's Retinex models were a prototype for color-constancy algorithms referred to generically as lightness algorithms (Hurlbert, 1986).

Retinex models contained two important assumptions. One has already been alluded to: that processing in each spectral channel was effectively independent of any other channel. This assumption is contingent on the overlap of cone spectral sensitivities and is closely related to the coefficient rule of von Kries (1902, 1905). The other assumption provided the method of normalizing the calculation of the triplets of ratios with respect to the illumination spectrum. In this way, a neutral surface in a Mondrian pattern would produce equal values in all three spectral bands. The particular method of normalization was essentially a procedure for estimating the illumination spectrum.

Both the independent-channels assumption and the illuminant-estimation assumption are statistical assertions about the sampling properties of surface spectral reflectances and illuminants in relation to the spectral sensitivities of the cones. Both have a significance that goes beyond lightness models. Some of the methods used to estimate the illuminant are considered in the next section and the independent-channels assumption and von Kries' coefficient rule are considered in Section 5.2.

2.2. Estimators of the illuminant

In the absence of any information about the illuminant, including the family of radiant spectra from which it is drawn, a common device has been to assume simply that the spatial average of scene reflectances is spectrally neutral, so that the space-average chromaticity of the reflected light provides an estimate of the illuminant chromaticity (Evans, 1946/1951). This "gray-world" assumption was part of both Buchsbaum's (1980) model of color constancy and one of Land's (1983, 1986) Retinex models, although the interpretation of what constitutes "gray" has since varied with the application and the population of reflectances. As D'Zmura, Iverson & Singer (1995) pointed out, it is sufficient that the space-average spectral reflectance is known, not that it is gray.

In applications, the gray-world assumption is easily violated, but if the surfaces of a scene truly form an unbiased sample from the population of such surfaces, then space-average chromaticity can give a reliable estimator of illuminant chromaticity (Barnard, Cardei, & Funt, 2002; Hurlbert, 1986).

An alternative approach to estimating the illuminant is to assume that the surface in the scene with highest luminance (or brightness) reflects maximally and uniformly over the spectrum (Brill & West, 1981; Land & McCann, 1971), sometimes known as

[4] Specifications of natural surface and illuminant spectra generally need more than three degrees of freedom (Sections 2.3 and 7.2–7.3), and, by definition, metameric lights, whether from sources or reflected from surfaces, cannot be distinguished by direct inspection. But there are arguments that observers can, in some viewing conditions, infer more degrees of freedom than the number of cone classes available (Broackes, 2010).

[5] In machine vision, because failures in color constancy may be dominated by illuminant estimation, the recovery of illuminant information is often given priority over other aspects of the problem (e.g. Hordley, 2006; Gijsenij, Gevers, & van de Weijer, 2010).

[6] These patterns, and similar checkerboard versions, although conventionally referred to as Mondrian patterns, differ from the majority of Piet Mondriaan's "Neo-Plasticist" paintings, which had black gridlines and a constrained color gamut. Mondrian patterns have the theoretical and experimental advantage of consisting entirely of patches of uniform spectral reflectance separated by step edges: reflectance variations are therefore not confounded with illumination variations.

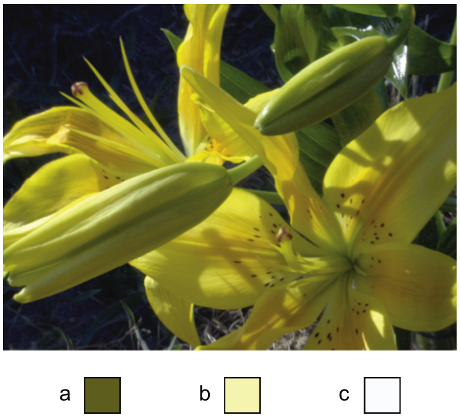

Fig. 2. Image of scene under a daylight with correlated color temperature 6500 K and samples of (a) space-average image color, (b) a bright non-specular region, and (c) a specular highlight due to moisture (image from Foster, Amano, Nascimento & Foster, 2006).

the "bright-is-white" or in other applications as the "scale-by-max", "max-RGB", or "white-patch" assumption (Barnard et al., 2002).[7] As with a gray-world estimate, the highest-luminance estimate from matte surfaces may be chromatically biased, but an estimate from a specular highlight is more likely to be reliable (D'Zmura and Lennie, 1986; Lee, 1986) or a combination of reflections from multiple surfaces (Tominaga, 1991). Fig. 2 shows an image of a scene where both the space-average chromaticity and the chromaticity of a bright non-specular reflection produced biased estimates of the illumination, but where a specular highlight is reliable, signaling the very slightly bluish cast of the illumination. Yet specular reflections need not always be spectrally neutral, even with non-metallic surfaces (Angelopoulou & Poger, 2003).

In general, estimates of illumination chromaticity from space-average chromaticity and from the brightest patch covary across scenes. For example, with illumination chromaticity inferred from a moving spotlight on a variegated scene, space-average chromaticity was found to be as good a model of observer estimates as a model weighting the brightest patches (Khang & Zaidi, 2004). By contrast, when space-average chromaticity and the chromaticity of the brightest patch were independently manipulated in Mondrian patterns and the illuminant cues provided were pitted against each other, space-average chromaticity was found to dominate observers' estimates of illumination except in patterns with the fewest patches (Linnell & Foster, 2002), a result that was confirmed in a much larger study of surface-color matching with Mondrian patterns (Amano & Foster, 2004).

From a theoretical stand, estimators based on space-average chromaticity and from the brightest patch in a scene fall at the extremes of a continuum, and intermediate versions may be defined by varying the way that signals from individual surfaces are summed.[8] But cues to the illuminant may be better combined within a more comprehensive cue-combination framework (Maloney, 2002).

Higher-order statistical properties offer other potential cues to the illuminant. One such property proposed by Golz and MacLeod (2002) was a correlation between the color of the image and its luminance. Such a property would allow a reddish scene under a neutral illuminant to be distinguished from a neutral scene under a reddish illuminant. The visual extent over which this correlation is used by observers seems to be local (Golz, 2008; Granzier, Brenner, Cornelissen, & Smeets, 2005). More importantly, the correlation may be a property of particular kinds of scenes where there is a preponderance of foliage. Ciurea and Funt (2004) showed that for images simulated from a more uniformly sampled set of hyperspectral images, the predicted correlation was weak, and for a very large database of digital camera images, the luminance-redness correlation failed completely. There are, however, other regularities of scenes (Hordley, 2006) whose relevance to human vision has yet to be tested. The extent to which observers might take advantage of illuminant estimators is considered in Section 5.1.

2.3. Low-dimensional linear models

Rather than making ad hoc assumptions about the properties of ensembles of surface spectral reflectances in a scene, as described in Section 2.2, some theoretical approaches to color constancy have taken a more principled line. Developed within a well-defined linear framework, the emphasis was on how individual surface spectral reflectances and illuminant spectra could be described analytically. The key idea was that if reflectance and illuminant spectra can be expressed as a weighted sum of a few spectral basis functions (Brill, 1978, 1979), and these basis functions are known to the observer (Dannemiller, 1991; Maloney & Wandell, 1986), then spectral reflectances can be recovered exactly, without, for example, assumptions about the mean such as the gray-world assumption (Section 2.2, cf. Buchsbaum, 1980).

The constraint on the numbers of basis functions, i.e. the dimensionality of the representations, is central to such analyses. Maloney and Wandell (1986) showed that if the numbers of reflectance and illuminant basis functions are n and m, respectively, the number of surfaces in the scene is s, and the number of photoreceptor classes is k, then providing that $k > n$ and $s > m$, both the illuminant and reflectance at each location can be recovered exactly. The practical difficulty is that there are normally only three classes of cones and that at least three and generally more basis functions are needed to adequately represent surface spectral reflectances (Section 7.2), so the condition $k > n$ is rarely satisfied with real scenes. Significantly, the advantages of more cone classes, i.e. $k > 3$, may not be great as anticipated (Mosny & Funt, 2007).

One way to overcome the problem of dimensionality is to take multiple views of a scene. D'Zmura and Iverson (1993a, 1993b) showed that if the same scene is illuminated by different illuminants, then higher-dimensional descriptors can be obtained for both surfaces and illuminants. Certain conditions have to be satisfied, which were enumerated in detail by D'Zmura and Iverson (1994). As D'Zmura (1992) nicely observed, a change in illumination, which creates the problem of color constancy, also supplies its solution.

In a different approach to the problem of dimensionality, Funt, Drew, and Ho (1991) showed that if there are regions where light reflected from one surface illuminates another, i.e. where mutual reflection is present, then the condition for recovery becomes $k \geqslant (2n + m)/3$, which is satisfied with three basis functions for surface spectral reflectances and three for illuminants (see e.g. Bloj, Kersten, & Hurlbert, 1999). Funt and Ho (1989) showed that the chromatic aberration of the eye could also be used to address the dimensionality problem.

Low-dimensional linear models have helped identify what is possible in color constancy and what constraints need to be satisfied, although they provide little insight into the perceptual correlates of these models. They also prompt other questions, including how many spectral basis functions are needed to properly represent surface spectral reflectances and illuminant spectra in real scenes. Consideration of these questions is postponed until Sections 7.2 and 7.3.

[7] It suffices to have something in the scene reflecting maximally over short-, medium-, and long-wavelength portions of the spectrum, but not necessarily all at one point; thus, a white surface is not essential for this approach.

[8] Given a Minkowski norm which has index p and which is defined on the set of surface signals, the norm with $p = 1$ corresponds to the gray-world assumption and with $p = \infty$ to the bright-is-white assumption (Finlayson & Trezzi, 2004).

2.4. Bayesian models

Bayes' rule gives a formal way of incorporating the known statistical structure of scenes into estimates of surface color, and it effectively generalizes the gray-world assumption (D'Zmura et al., 1995). It requires two functions: a prior distribution, i.e. the probability $p(r, e)$ of surface spectral reflectance r and illuminant e, usually expressed in a parametric form $p(\theta)$, where θ represents the weight in a linear model of reflectances and illuminants, and the likelihood $f(c|\theta)$, which is proportional to the probability of the observed reflected spectrum c given θ. Bayes' rule is used to calculate the posterior distribution $p(\theta|c)$, i.e. the probability of the reflectance and illuminant given the observed reflected spectrum. Provided that a solution can be calculated, it leads to very effective use of data (Forsyth, Haddon, & Ioffe, 2001). Bayesian models have been incorporated into a more comprehensive framework for analyzing and modeling color constancy known as color by correlation (Finlayson, Hordley, & Hubel, 2001).

Disappointingly, the application of Bayesian methods to constancy judgments in the natural world has had limited success. There is a basic problem in accurately specifying the prior distribution $p(r, e)$, and strong assumptions may need to be made. For example, Brainard and Freeman (1997) and Brainard et al. (2006) assumed that the set of natural spectral reflectances could be represented as randomly weighted combinations of basis functions extracted by principal component analysis (PCA) from a set of Munsell surface reflectances (Munsell Color Company, Baltimore, MD). Three basis functions were used and the weights were modeled by a truncated multivariate normal distribution. They made an analogous assumption about illuminants drawn from the daylight locus.

The assumptions of the relatively low dimensionality of Munsell reflectances (Section 7.2) and the normality of the distributions of weights were not critical, unlike the assumptions about the parameters of the distributions. As with the gray-world assumption, a Bayesian model using a prior for reflectances based on a particular mean (and variance), here that of the Munsell reflectances, may fail drastically when the spectral properties of the application no longer match those assumptions, for example, when the spectra are biased away from the assumed mean (e.g. Endler, 1993; Webster, Mizokami, & Webster, 2007). Even when there is less uncertainty about the prior, as with the distribution of daylights, observers' judgments may be difficult to predict from the expected distributions (Delahunt & Brainard, 2004b).

Some of the problems with specifying priors may be avoided by calculating color signals that represent relative rather than absolute quantities (Fine, MacLeod, & Boynton, 2003), an approach which is analogous to that in relational analyses of surface-color judgments (Sections 3.2 and 5.3).

A different way of relating the statistical structure of natural scenes to color appearance has been advanced by Long and Purves (2003). They proposed that color constancy (and many other perceptual phenomena) associated with any particular aspect of the visual stimulus is predicted by the relative frequency of the occurrence of that stimulus in relation to all the other instances that have been experienced in the same context (Howe, Lotto, & Purves, 2006). For example, a stimulus consisting of a yellowish patch in a reddish background appears greener because that is the typical chromaticity of such a patch in backgrounds with similar spatial complexities in natural scenes (Long & Purves, 2003). As with Bayesian methods, however, there remains the problem of determining the appropriate relative frequency of chromaticities and spatial complexities in natural scenes.

3. What do observers judge?

An intrinsic difficulty with measuring color constancy is that there is often more than one sense in which the perceived or

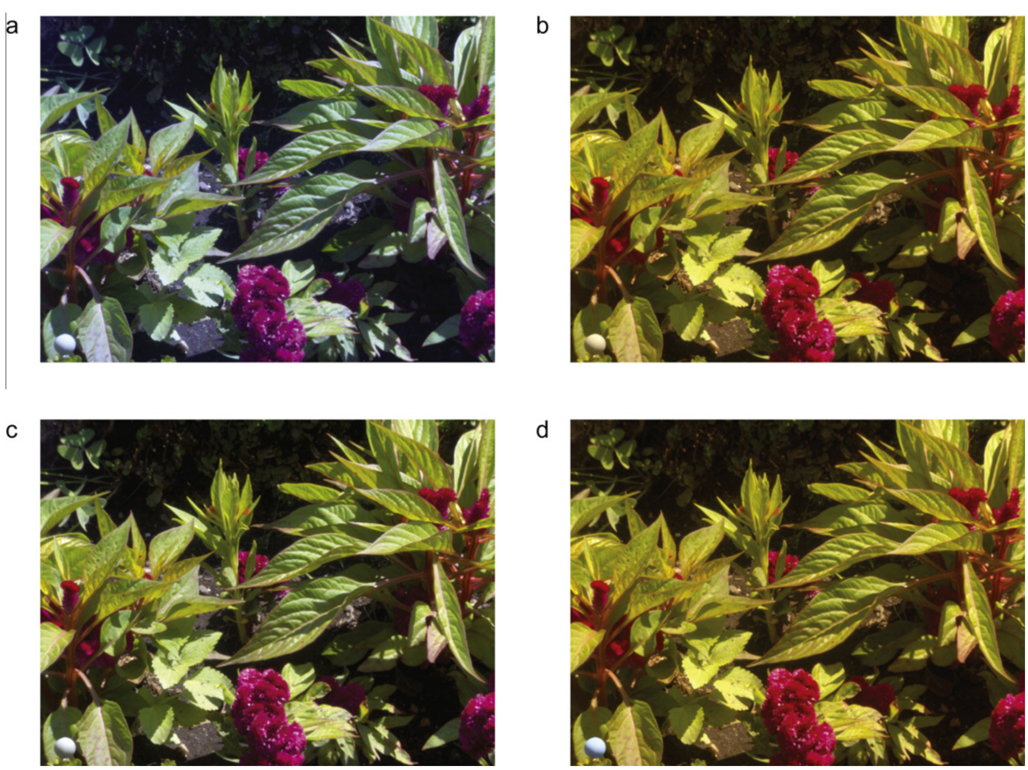

Fig. 3. Images of a natural scene under different daylights. The scene illuminant is (a) skylight, with correlated color temperature 17,000 K, (b) sunlight, 4000 K, (c) a mixture of sunlight and skylight, 6500 K, and (d) sunlight, 4000 K. In images a–c, the sphere (bottom left corner) is covered with the same gray paint, but in image d it is covered with blue-gray paint to give the same reflected light as in image a. In a–c the color relations are largely preserved; in d they are not. Images from hyperspectral data in Foster, Amano, Nascimento et al. (2006).

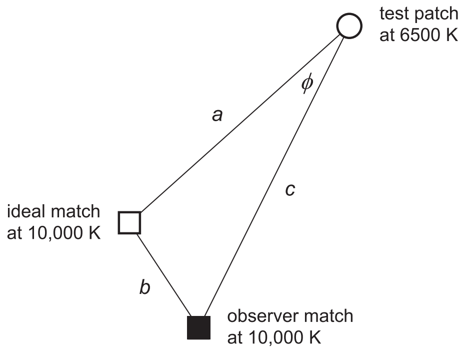

Fig. 4. Definition of color-constancy index and Brunswik ratios. The coordinates of a test surface under daylight with correlated color temperature 6500 K (open circle), the ideal match under daylight with correlated color temperature 10,000 K (open square), and the corresponding observer match (solid square) are plotted in the CIE (x, y) chromaticity diagram. The constancy index is defined by $\mathrm{CI} = 1 - b/a$; the Brunswik ratio by $\mathrm{BR} = c/a$; and its projection by $\mathrm{BR}_\phi = c \cos \phi / a$. Data from Arend and Reeves (1986, Fig. 4, middle right).

apparent color of a surface may be judged as being constant. Observers may make decisions based on different internal criteria; they may draw on different invariant properties of images; and their ability to extract those invariants may depend on other image properties that covary with illuminant changes.

3.1. Hue, saturation, and brightness constancy vs. surface-color constancy

Chromatic adaptation to the prevailing illumination is reasonably complete only if the spectrum of the illuminant does not differ too much from daylight (Judd, 1940). But as already indicated in Section 1, with the normal variation in daylight, reflected lights may differ in appearance even after full adaptation, as Arend (1993) confirmed with Mondrian patterns uniformly illuminated by daylights. With natural scenes, the variation of illumination is usually more complicated, with some regions in shadow or illuminated by reflected light from the sky and other regions in direct sunlight. With the eye in continual movement over a scene, some time-averaged adaptation necessarily takes place, depending on the general spatial and chromatic structure of the scene; yet differences in illumination remain visible. The same surface partly in shadow and partly in direct sunlight may be seen clearly to be under different illuminations and still have the same surface color.

This apparent paradox in natural viewing has been long recognized (Evans, 1974; Katz, 1935; Lichtenberg, 1793), but was not addressed routinely in experimental practice. Before Arend and Reeves' (1986) work there had been measurements of observers' responses to the color of the reflected light—hue, saturation, and brightness—and to the color of reflecting surfaces, but adaptational and inferential processes were not easily distinguished. For example, in a study by McCann, McKee, and Taylor (1976) of color matching under different illuminants, a Mondrian pattern of Munsell papers and the Munsell chips used to match a test paper within the pattern were viewed monocularly with different eyes. As Arend and Reeves (1986) pointed out, the adaptational states of the two eyes were different; and the surrounds of the test and match papers were also different, producing a potential confound with the differences in illumination; and, crucially, the task given to subjects was unspecific as to whether they should match for hue, saturation, and brightness or for surface color.

The asymmetric color-matching experiment by Arend and Reeves (1986), also using Mondrian patterns, contained several novel features, detailed in Section 4.2, but the most important element was that in separate tasks subjects were given two specifically differentiated criteria for matching patches across the pairs of patterns under different illuminants: with one criterion, the patches were to have the same hue and saturation (a "hue-saturation match" or in other circumstances a "hue-saturation-brightness match"); with the other criterion, the patches were to look as if they were "cut from the same piece of paper" (a "paper match" or "surface match"). To make these matches, subjects controlled the stimuli in two dimensions (the third dimension corresponding to brightness or lightness variation was omitted to separate color constancy from lightness constancy). Subjects were able to make these color judgments reliably (see also Arend, Reeves, Schirillo, & Goldstein, 1991, Fig. 5), much as they can with judgments of size and of shape.[9] The different levels of performance with the two criteria (Arend & Reeves, 1986; Arend et al., 1991), i.e. constancy indices roughly two-times higher with paper matches than with hue-saturation matches, were replicated by other authors (e.g. Bäuml, 1999; Cornelissen & Brenner, 1995; Troost & de Weert, 1991) and by comparison with other kinds of stimulus judgments (Reeves, Amano, & Foster, 2008).

Arend and Reeves (1986) concluded (p. 1749) that observers' hue-saturation judgments were determined mainly by sensory or adaptational mechanisms and paper matches mainly by perceptual-computational or inferential mechanisms. The different kinds of processing required by the two criteria have been objectively demonstrated in subjects' eye movements recorded in a simultaneous asymmetric color-matching experiment by Cornelissen and Brenner (1995). It was found that when making paper matches across Mondrian patterns, subjects usually spent more time looking at the surround of the matching patch than at the patch itself, whereas the opposite was true when they were making hue-saturation matches. A complementary finding was reported by Golz (2010) in an experiment on achromatic adjustment where subjects had to adjust a test patch in a variegated surround so that it appeared gray (Section 4.4). The accuracy of subjects' achromatic settings evaluated by their closeness to the mean surround chromaticity was better when they were instructed to explore the surround than when they fixated the test patch.

To distinguish between changes in sensitivity and changes in response criterion, van Es, Vladusich, and Cornelissen (2007) presented observers with a colored checkerboard pattern undergoing an illuminant change and asked observers in one condition whether the central color patch in the pattern kept the same hue, saturation and brightness and in another condition, with the same stimuli, whether there was an overall illuminant change across the entire pattern. Randomization of the surround affected subjects' criteria, but not their discrimination performance d' from signal-detection theory (Macmillan & Creelman, 2005) nor their constancy indices.

Reports of the ease of applying hue-saturation and paper-match criteria have varied. Troost and de Weert (1991) stated that their subjects found making the equivalent of paper matches much more difficult than making hue-saturation matches, despite much higher performance levels being obtained with the former than with the latter. Brainard, Brunt, and Speigle (1997) reported difficulty in distinguishing the two tasks, yet Bäuml (1999) stated that his subjects found paper matches very natural, seeming "to perceive that a certain color was the 'right' one, in order for the matching field to represent the same surface as shown in the test field" (p. 1548).

The fact that subjects can make dual judgments suggests that using an undifferentiated criterion for color matching may lead to greater response uncertainty. Indeed, the variability found in some constancy judgments was claimed by Katz (1935, Section 27)

[9] For some discussion of dual color codes in this context, see e.g. Mausfeld (2003).

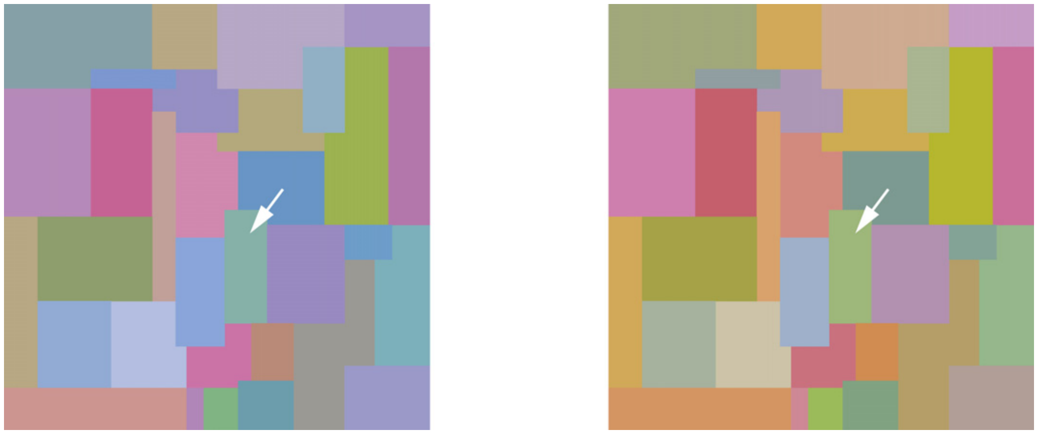

Fig. 5. Mondrian patterns used by Arend and Reeves (1986) in simultaneous asymmetric color matching. The patterns consisted of Munsell matte colored papers of Munsell Value 5 simulated under daylight and sunlight with correlated color temperatures 6500 K on the left and 4000 K on the right. Patch luminances were varied by ±10%. The variable "match" patch arrowed in the right pattern was matched against the corresponding test patch arrowed in the left pattern (arrows absent in the original). Recreated from Fig. 1 of Arend and Reeves (1986, pp. 1744–1745).

and Judd (1940, Section III, 2(c)) to be due to indeterminacy either generally in observer attitude or specifically in assigning chromatic effects to an appropriate physical origin. With a properly differentiated criterion for matching, therefore, fewer extreme values should occur. A direct comparison of asymmetric color matching with undifferentiated and differentiated criteria appears not to have been reported. There is evidence that with a paper-match criterion subjects' responses are at least close to being normally distributed, with little evidence of outliers. In an experiment on simultaneous asymmetric color matching with a paper-match criterion (Foster, Amano, & Nascimento, 2001), the distribution of constancy indices from 20 subjects (Section 4.1) was found to have a standard deviation of 0.14 about a mean of 0.66. No subject scored less than 2 s.d. below the mean.[10]

Implicit in Arend and Reeves' (1986) experimental procedure was the assumption that their two kinds of judgments were based on two 3-dimensional spaces: one concerned with hue, saturation, and brightness, the other with surface color per se. Brainard et al., (1997) have also proposed that in asymmetric color matching with an undifferentiated color-match criterion, more than three dimensions are involved in subjects' judgments. The question of the number of dimensions underlying color judgments with surfaces under variegated illumination was addressed directly by Tokunaga and Logvinenko (2010) in a multidimensional scaling experiment. Subjects were asked to judge the dissimilarity of surfaces in a scene with multiple illuminants. Their responses were best modeled with three dimensions associated with surfaces and another three with the illuminants, but with just one illuminant, responses could be modeled with the usual three dimensions.

3.2. Relational color constancy

As shown later (Sections 4.2 and 4.5), many experiments aimed at measuring color constancy have actually measured a different phenomenon, namely, relational color constancy. This refers to the constancy of the perceived relations between the colors of surfaces under illuminant changes, rather than of the perceived colors themselves[11] (Foster & Nascimento, 1994; Nascimento & Foster, 1997). For example, in Fig. 3, the scene is illuminated by different daylights, with correlated color temperatures (a) 17,000 K, (b) 4000 K, (c) 6500 K, and (d) 4000 K. The color of the light reflected from the sphere in the bottom left corner in a–c is clearly different. Nevertheless, given the limits of the color reproduction of these images on the printed page, it can be seen that the sphere has the same or similar surface color in each image by comparing it with the nearby foliage and by looking over each image as a whole. By contrast, in d, although the color of the light reflected from the sphere is the same as in a, it can be seen that the sphere has a different surface color, now more bluish, again by comparing it with nearby foliage or over the image as a whole. In a–c, the perceived relations between the colors are largely preserved, and in d, they are not.

Relational color constancy has been given an operational meaning, independent of its subjective content, namely, the ability of an observer to correctly attribute changes in the color appearance of a scene either to changes in the spectral composition of the illuminant or to changes in the reflecting properties of that scene, i.e. its materials (Craven & Foster, 1992; Foster, Craven, & Sale, 1992). A similar issue has been emphasized by Zaidi (1998). The formal equivalence of perceptual and operational interpretations of relational color constancy was set out by Foster and Nascimento (1994, Appendix 1), and its experimental application is described here in Section 4.5.

The phenomenology of illuminant and material changes has been found to be particularly compelling when the changes occur as a temporal sequence without an intervening delay. Thus, when subjects were presented with successive Mondrian patterns related by illuminant or material changes (Craven & Foster, 1992), they reported that the "changes of illuminant tended to be perceived as a coloured wash over the display, whereas changes of material led to a distinctively uneven appearance" (p. 1364).

As Fig. 3 illustrates, reliable discriminations can also be made between simultaneously presented images related by an illuminant change, as in a and b, and by an additional material change, as in a and d. This ability persists with presentations lasting < 200 ms and led to the suggestion that the ability to judge whether color relations were preserved or violated was the result of fast, relatively low-level, spatially parallel visual processing (Foster et al., 1992). This notion was supported by subsequent measurements in which during successive illuminant changes, material changes in one or more surfaces in an array of other surfaces were shown to be readily detected almost independently of the numbers of surfaces (Foster, Nascimento, et al., 2001). One

[10] The kurtosis of the sample, a measure of the potential of the distribution for outliers, was in fact less than that for a normal distribution.

[11] The perceived relations between the colors of surfaces should not be confused with what defines related colors, such as brown and olive, which require the presence of other colors, achromatic or chromatic, to be perceived. The former refers to pairs (or larger groupings) of arbitrary colors; the latter to particular individual colors.

potential role for this mechanism may be to provide the visual system with information about a rapidly changing world in advance of the generation of a more elaborate and stable surface-color representation.

Relational color constancy has a natural physical substrate: the ratios of cone excitations generated in response to light reflected from pairs of surfaces or groups of surfaces, which may or may not be adjacent. Such ratios, which can also be determined across post-receptoral combinations and spatial averages of cone signals, have the remarkable property of being nearly invariant under changes in illuminant and may explain performance in several color-constancy tasks (Section 5.3).

Van Es et al. (2007) have proposed that relational color constancy and color constancy based on judgments of hue, saturation, and brightness involve mechanisms with distinct spatial dependencies. Some support for this proposal comes from data on the effect of the surround field on, variously, asymmetric color matching, achromatic adjustment, and color naming (Amano & Foster, 2004; Brenner & Cornelissen, 1991, 2002) and from data on eye movements in asymmetric color matching (Cornelissen & Brenner, 1995) and achromatic adjustment (Golz, 2010).

3.3. Positional and atmospheric color constancy

Changing the spectrum of reflected light from an object by changing the spectrum of the illumination constitutes the most fundamental test of color constancy. But there are other, less direct ways in which the illumination on an object may change with a change in viewing conditions. One of the most natural ways arises with a change in position or context. For surface color to be perceived as constant, the context of the surface needs to be taken into account; otherwise, the spectral reflecting properties of the surface and the spectral properties of the illumination cannot be disconfounded. Yet color constancy also requires the context to be discounted in some sense; otherwise, perceived surface color would be an accident of position (Wachtler, Albright, & Sejnowski, 2001). This apparent contradiction was examined in an experiment (Amano & Foster, 2004) with Mondrian patterns of simulated Munsell surfaces whose average spatial and spectral properties could be accurately controlled. It was found that subjects could make simultaneous asymmetric color matches with a paper-match criterion (as in Arend & Reeves, 1986) across simultaneous changes in test-patch position and illuminant almost as well as across changes in illuminant alone. Performance was no poorer when the surfaces surrounding the test patch were randomly permuted (Amano & Foster, 2004; cf. van Es et al., 2007); Provided that changes in context do not entail a change in composition and that they are not systematic, it seems that color constancy is preserved, at least in Mondrian patterns. This invariance to position is not necessarily inconsistent with classical color-contrast or chromatic-induction effects where changes to the surrounds are made in a systematic way (Section 5.4).

Another natural change to viewing conditions that affects the spectrum of the reflected light is a change in viewing medium. In the natural world, fog, mist, and smoke can all modify spectral transmission, by an amount that depends on the composition of the suspended particles and their density, the ambient illumination, and the distance of the reflecting surface from the observer. Despite the ubiquity of this experience, little is known about the degree of color constancy under these conditions, except for one study with Mondrian patterns by Hagedorn and D'Zmura (2000). Subjects made asymmetric color matches of the patterns with and without a colored fog and their performance was represented by an affine combination of simulated reflected and scattered light. All the subjects compensated for the loss in contrast due to the colored fog, but to differing extents. An automatic compensation for contrast loss may explain the observation that the color of a test patch appears more colorful against a low-contrast, neutral background than against a high-contrast, multicolored background of the same space-average color (Brown & MacLeod, 1997).

4. What experimental methods are suitable?

Four main kinds of psychophysical methods have been used to measure color constancy: asymmetric color matching, color naming, achromatic adjustment, and discriminating illuminant from reflectance changes. Each of these methods involves design factors that can influence observed performance in different ways. Most applications described in the following used the same kind of experimental apparatus, namely, a computer-controlled RGB color monitor, although some used physical materials and lights. Before considering the advantages and disadvantages of each of these methods, it is useful to summarize the more common ways in which the level of color constancy has been quantified. Some of the logical content of this section is based on that in Foster (2003).

4.1. Indices for color constancy

The basis for quantifying the degree to which color constancy succeeds—or fails—is the difference between an observer's match or setting and its ideal value in some appropriate color space. This space is typically either a two-dimensional chromaticity space such as the CIE 1931 (x, y) or CIE 1976 (u', v') chromaticity diagram or a three-dimensional space such as CIELAB or CIELUV color space (CIE, 2004). Differences between particular measures center on how the difference between observed and ideal settings is expressed in the chosen space.

A simple Euclidean distance may be used to quantify the difference, but instead a constancy index CI (Arend et al., 1991) or a Brunswik ratio BR (Troost & de Weert, 1991) has often been preferred to scale the difference to yield a dimensionless quantity. Fig. 4, adapted from data in Arend and Reeves (1986, Fig. 4, middle right), shows in CIE (x, y) space the coordinates of a test patch under 6500 K (open circle), the ideal match under 10,000 K (open square), and the observer match under 10,000 K (solid square), along with the Euclidean distances a, b, and c between them and the angle ϕ between ideal and observer matches. The constancy index CI is defined as $1 - b/a$; the Brunswik ratio BR as c/a; and its projection BR_ϕ on the ideal-match line as $c \cos \phi/a$. In principle, perfect constancy corresponds to an index or ratio of unity and the complete absence of constancy, with no account taken of the illuminant, corresponds to an index or ratio of zero.

All three measures, BR, BR_ϕ, and CI, coincide when the coordinates of the observer match fall on the line segment joining the coordinates of the test surface and ideal match, i.e. when $\phi = 0$ and $BR < 1$. But only with Arend et al.'s index does the error get smaller as the index gets larger, so that in the limit $CI = 1$ implies $b = 0$, i.e. a perfect match. With the Brunswik ratio, the error b is not a unique function of BR, so that $b > 0$ and $BR \geqslant 1$ can both be true. The result is that both BR and BR_ϕ tend to underestimate the error. The differences need not be trivial. For the example in Fig. 4, the match is clearly imperfect, and $CI = 0.59$, whereas $BR = 1.14$ and $BR_\phi = 1.06$.

The inflationary property of ratio measures does not disappear when means are taken over groups of observations. Thus, when BR, BR_ϕ, and CI were applied to data from 20 subjects making simultaneous asymmetric surface-color matches (Foster, Amano, et al., 2001), the mean values of BR and BR_ϕ were found to be higher than the mean value of CI by 27% and 22%, respectively.

The fact that CI does not underestimate the error gives it an advantage over BR and BR_ϕ. Of course, no single measure is perfect.

A problem common to BR, BR_{ϕ}, and CI is that their magnitudes tend to infinity as the physical difference a tends to zero. Dividing by a is intended to scale the expected error by the size of the change in illuminants, but values of b can be fairly stable across extensive variations in a (see de Almeida, Fiadeiro, & Nascimento, 2004, Fig. 3, where a different notation is used). In practice, however, a is usually large and held constant over different conditions.

Because these three measures are mainly of the magnitude of the error, they lose information about its direction (Ling & Hurlbert, 2008; Troost & de Weert, 1991). In the machine-vision literature, it is more common to record an unscaled measure of the difference between ideal and empirical matches or one converted to an angular measure with respect to the origin of the color space of interest (Barnard et al., 2002; Hordley & Finlayson, 2006). Perceptually weighted measures have been described by Vazquez-Corral, Parrága, Vanrell, and Baldrich (2009) and by Gijsenij, Gevers, and Lucassen (2009). Within the context of a given scene, it is also possible to measure the effect of errors in an information-theoretic sense (Foster, Marín-Franch, Amano, & Nascimento, 2009).

Some authors report constancy indices or Brunswik ratios by making fits to multiple response categories, for example, by estimating boundaries between regions of color space classified according to unique hues or basic color categories. This procedure can remove significant variance in the data and lead to higher estimates of performance, but may limit comparisons with indices from single response categories.

Table 1 shows constancy indices and Brunswik ratios from a sample of experimental studies grouped by method. Data from some relevant studies were omitted because of the difficulty in extracting indices. Care should be exercised in making simple numerical comparisons of constancy values across different methods, especially given differing observer adaptational states and decision criteria, and the unexplained variation within some methods.

4.2. Asymmetric color matching

The method of asymmetric color matching described by Wyszecki and Stiles (1982) involves stimuli being compared under different viewing conditions, here different illuminants. Stimuli may be viewed simultaneously or successively or in an alternating sequence, binocularly or dichoptically. A critical factor for color constancy is whether the adaptational state covaries with the change in illuminant (Section 3.1). In the method of asymmetric color matching used by Arend and Reeves (1986), the Mondrian patterns consisted of matte, colored Munsell surfaces simulated under different illuminants and presented side by side on an RGB monitor, as illustrated in Fig. 5. The subject, who viewed the stimuli binocularly, adjusted a variable match patch in the right pattern to match the corresponding fixed test patch in the left pattern.[12] The match shown is a perfect paper match. Novel features of the design, in addition to the experimental task (Section 3.1), were the simultaneous presentation of the patterns to minimize confounding adaptational effects and the use of identical spectral reflectances for corresponding fixed surfaces in the test and match patterns so that only the illuminant would affect performance. The mean constancy index obtained by Arend and Reeves (1986) with a paper-match criterion was 0.52 averaged over three subjects, but in a closely similar replication of this experiment by Bäuml (1999) the mean index was 0.79 averaged over eight subjects. With von-Kries-approximated illuminant changes (Section 4.8), Troost and de Weert (1991) obtained 0.81 averaged over 14 subjects. Values from some other studies are listed in Table 1.

A potential nonadaptational confound in asymmetric color matching can come from the introduction into the test scene of a surface that duplicates the test surface. As Maloney (1999) cautioned, by setting the match patch to have the same hue, saturation, and brightness as the duplicate patch in the match scene, a subject can demonstrate perfect constancy when in fact they have none. Duplicate surfaces were actually eliminated by Arend and Reeves (1986), and, in randomly composed scenes such as Mondrian patterns (Fig. 5), a more stringent constraint may be imposed requiring a minimum chromatic difference between the test and other patches in the scene (Foster, Amano, et al., 2001).

Given a particular stimulus geometry, slightly higher constancy indices may be obtained in asymmetric color matching by presenting the two differently illuminated Mondrian patterns successively, one rapidly after the other in the same position, rather than simultaneously side by side. The improvement is about 14% on average (Foster, Amano, et al., 2001; compare Troost & de Weert, 1991). It seems probable that successive presentation, either as a "one shot" or as an alternating sequence, without an intervening delay, can generate a useful cue to changes in reflectance (Section 5.3).

A novel form of sequential asymmetric color matching was introduced by Barbur, de Cunha, Williams, and Plant (2002) in which the two patterns, each presented for 800 ms, were continuously alternated. The subject adjusted the match field so that alternating test and match fields appeared invariant during the alternating surround illuminant. This dynamic matching technique, applied with a hue, saturation, and brightness criterion, yielded a mean Brunswik ratio of 0.50, which is relatively high given the criterion and short illuminant durations. The results were interpreted as being due to "instantaneous" constancy mechanisms (Barbur, de Cunha, Williams, & Plant, 2004; Barbur et al., 2002).

These successive or sequential methods of asymmetric color matching with short illuminant durations should be distinguished from more traditional forms of successive asymmetric color matching (Brainard & Wandell, 1992; Bäuml, 1995) in which a subject is given greater chance to adapt to the differently illuminated patterns or scenes.[13] For example, in a successive asymmetric color-matching experiment by Murray, Daugirdiene, Vaitkevicius, Kulikowski, and Stanikunas (2006), subjects were given up to 60 s to adapt to a background field of various sizes. With a 120-degree adapting field, a mean constancy index of 0.91 was obtained with judgments based on a hue, chroma, and value. In principle, successive asymmetric color matching depends on memory but dichoptic simultaneous asymmetric color matching does not, and it allows complete or almost complete adaptation of the eye to each illuminant (Chichilnisky & Wandell, 1995). Levels of constancy can be very high. Using this dichoptic technique, a Brunswik ratio of 0.89 was obtained by Bramwell and Hurlbert (1996) with an undifferentiated forced-choice color-matching criterion. Somewhat lower values were obtained by Kuriki and Uchikawa (1996) but, significantly, paper matches and the equivalent of hue-saturation-brightness matches produced almost identical indices (Table 1).

In general, asymmetric color matching, whether simultaneous or successive, binocular or dichoptic, offers precision and flexibility, with little constraint on the reflectance or geometry of the test stimulus. But it has a limitation: it can establish only an equivalence of stimuli (Foster, 2003). To see this, consider Fig. 5 again.

[12] There is an inconsistent nomenclature for these surfaces: a fixed "test" and variable "match" (or "reference" or "comparison") are common, and used here, but in some reports "test" is replaced by "standard" (or "reference") and "match" by "test".

[13] The functional difference between simultaneous and successive presentation is not always well defined: viewing two simultaneous patterns with alternating gaze produces almost the same retinal stimulation as viewing successive patterns with gaze fixed.

Table 1
Levels of color constancy from some experimental studies. Average constancy indices CI and Brunswik ratios BR are tabulated against experimental method, stimulus configuration, illuminants, judgment by subject, experimental apparatus, illuminant change, cues other than those defined by the experimental method, constraints on stimulus variation, the number of subjects *N*, and source of data. Values of CI and BR are averages over subjects and conditions, with some entries estimated from published figures. Not all entries were calculated exactly as in Section 4.1, but they retain the properties that CI values cannot exceed unity and BR values can. See text for further details.

Experimental method	Stimulus configuration	Illuminants	Judgment	Experimental apparatus	Illuminant change	Other cues	Constraints	*N*	CI	BR	Source
Simultaneous asymmetric matching											
	2D multielement	Daylights	Same paper	Monitor	Spectral product		Chromaticity	3	0.52		Arend and Reeves (1986, Fig. 4)
	2D multielement	Daylights	Same hue, saturation	Monitor	Spectral product		Chromaticity	3	0.20		Arend and Reeves (1986, Fig. 4)
	2D patch in surround	Daylights, others	Same hue, saturation	Monitor	Spectral product		Chromaticity	4	0.30		Tiplitz Blackwell and Buchsbaum (1988b, Tables 1,3,6,8,10)
	2D patch in far surround	Daylights, others	Same hue, saturation	Monitor	Spectral product		Chromaticity	4	0.15		Tiplitz Blackwell and Buchsbaum (1988b, Tables 1,3,6,8,10)
	2D multielement	Daylights	Same paper	Monitor	Spectral product			4	0.44		Arend et al. (1991, Fig. 7)
	2D multielement	Daylights	Same hue, saturation, brightness	Monitor	Spectral product			4	0.18		Arend et al. (1991, Fig. 7)
	2D patch in surround	Daylights	Same paper	Monitor	Spectral product	Gray surround		3	0.35		Arend et al. (1991, Fig. 9)
	2D patch in surround	Daylights	Same hue, saturation, brightness	Monitor	Spectral product	Gray surround		3	0.11		Arend et al. (1991, Fig. 9)
	2D multielement	Colored lights	Same paper[a]	Monitor	von Kries shift	Gray background	Chromaticity	14	0.81	0.82	Troost and de Weert (1991, Table 2)
	2D multielement	Colored lights	Same hue, saturation[a]	Monitor	von Kries shift	Gray background	Chromaticity	14	0.46	0.46	Troost and de Weert (1991, Table 2)
	2D multielement	Daylights	Same paper	Monitor[b]	von Kries shift		Chromaticity	5	0.37		Cornelissen and Brenner (1995, Fig. 4B)
	2D multielement	Daylights	Same hue, saturation	Monitor[b]	von Kries shift		Chromaticity	5	0.18		Cornelissen and Brenner (1995, Fig. 4B)
	3D room	Colored lights	Undifferentiated color match	Illuminated surfaces	Physical reflection	Gray room		5	0.61		Brainard et al. (1997, Table 3)
	2D multielement	Daylights	Same paper	Monitor	Approx. spectral product			8	0.79		Bäuml (1999, pp. 1537, 1541)
	2D multielement	Daylights	Same hue, saturation, brightness	Monitor	Approx. spectral product			6	0.23		Bäuml (1999, pp. 1537, 1541)
	2D multielement	Daylights	Same paper	Monitor	Spectral product			20	0.60		Foster, Amano et al. (2001, Table 2)
	2D multielement	Daylights	Same paper	Monitor	Spectral product		Chromaticity	20	0.66		Foster, Amano et al. (2001, Table 2)
	2D multielement	Daylights	Same paper	Monitor	Spectral product			11	0.74		Amano and Foster (2004, Table 1)
	2D multielement, transposed test	Daylights	Same paper	Monitor	Spectral product			11	0.70		Amano and Foster (2004, Table 1)
	2D multielement, permuted	Daylights	Same paper	Monitor	Spectral product			11	0.75		Amano and Foster (2004, Table 2)
	3D tableau	Daylight metamers	Same paper	Illuminated surfaces	Physical reflection	Gray background		4	0.86		de Almeida et al. (2004, Fig. 5)
	3D tableau	Daylight, colored lights	Undifferentiated color match	Stereo monitor	Spectral product	Gray background		4	0.23		Delahunt and Brainard (2004a, Fig. 11)
	2D multielement	Daylights	Same paper	Monitor	Spectral product			6	0.73		Amano et al. (2005, Table 1)
	2D pair[c]	Daylights	Same paper	Monitor	Spectral product			6	0.72		Amano et al. (2005, Table 1)
	3D tableau	Colored lights	Physical match	Illuminated surfaces	Physical reflection			7		0.84	Granzier et al. (2009a, Fig. 7)
Dichoptic simultaneous asymmetric matching											
	2D multielement	Daylights, other	Undifferentiated color match	Monitor	Spectral product		Chromaticity	3		0.89[d]	Bramwell and Hurlbert (1996, Table 2)
Successive asymmetric matching											
	2D multielement	Colored lights	Same paper[a]	Monitor	von Kries shift	Gray background	Chromaticity	8	0.59	1.38	Troost and de Weert (1991, Table 3)
	2D multielement	Colored lights	Same hue, saturation[a]	Monitor	von Kries shift	Gray background	Chromaticity	8	0.41	0.42	Troost and de Weert (1991, Table 3)
	2D multielement	Daylights	Same paper	Monitor	Spectral product			20	0.69		Foster, Amano et al. (2001, Table 2)

(continued on next page)

Table 1 (*continued*)

Experimental method	Stimulus configuration	Illuminants	Judgment	Experimental apparatus	Illuminant change	Other cues	Constraints	*N*	CI	BR	Source
	2D multielement	Daylights	Same paper	Monitor	Spectral product		Chromaticity	20	0.75		Foster, Amano et al. (2001, Table 2)
	2D multielement	Daylight, tungsten	Same hue, saturation, brightness	Monitor	Spectral product		Chromaticity	3		0.50	Barbur et al. (2004, Fig. 2b)
	2D patch in surround	Daylight, tungsten	Same hue, saturation, brightness	Monitor	Spectral product	Gray surround	Chromaticity	3		0.44	Barbur et al. (2004, Fig. 2b)
Successive asymmetric matching with adaptation											
	2D multielement	Daylights	Same paper	Illuminated surfaces	Physical reflection	Gray background		4	0.69		Kuriki and Uchikawa (1996, Fig. 6)
	2D multielement	Daylights	Same hue, saturation, brightness[a]	Illuminated surfaces	Physical reflection	Gray background		4	0.56		Kuriki and Uchikawa (1996, Fig. 6)
	2D patch in background	Colored lights	Undifferentiated color match	Illuminated surfaces	Physical reflection	Gray background		6	0.46–0.63[e]		Kulikowski and Vaitkevicius (1997, Fig. 3)
Dichoptic successive asymmetric matching with adaptation											
	2D multielement	Daylights	Same paper	Illuminated surfaces	Physical reflection	Gray background		4	0.77		Kuriki and Uchikawa (1996, Fig. 8)
	2D multielement	Daylights	Same hue, saturation, brightness[a]	Illuminated surfaces	Physical reflection	Gray background		4	0.72		Kuriki and Uchikawa (1996, Fig. 8)
	2D multielement	Daylights	Same hue, saturation, brightness	Monitor	Spectral product			3	0.69		Lucassen and Walraven (1996, Table 3)
	2D multielement	Metamers[f]	Same hue, saturation, brightness	Monitor	Spectral product			3	0.47		Lucassen and Walraven (1996, Table 3)
Successive asymmetric memory matching with adaptation											
	2D multielement	Daylights, other	Physical match	Illuminated surfaces	Physical reflection	Gray background		2	0.64[g]		Uchikawa et al. (1998, Fig. 5)
	2D multielement	Daylights, others	Same paper[a]	Monitor	Approx. spectral product			3	<0.44		Nieves, García-Beltrán, and Romero (2000, p. 55)
	2D patch in 120-deg background	Daylights, tungsten	Undifferentiated color match	Monitor	Spectral product	Gray background		4	0.91		Murray et al. (2006, Fig. 7B)
	2D patch in background	Daylight metamers, others	Physical match	Illuminated surfaces	Physical reflection	White background	Samples	7	0.92[a,h]		Ling and Hurlbert (2008, Table 6)
	3D tableau	Daylights, others	Same paper	Illuminated surfaces	Physical reflection	White background	Samples	28		0.79[i]	Hedrich et al. (2009, p. 12)
	2D multielement	Daylights, others	Same paper	Illuminated surfaces	Physical reflection		Samples	28		0.58[i]	Hedrich et al. (2009, p. 12)
Color naming and related methods											
	2D patch in background	Colored lights	From 12 colors	Monitor	von Kries shift	Gray background		30	0.65	0.65	Troost and de Weert (1991, Table 4)
	2D multielement	Daylights	From unique hues and gray	Monitor	Spectral product		Chromaticity	2	0.66		Arend (1993, Fig. 8)
	2D patch in surround	Daylights	From unique hues and gray	Monitor	Spectral product	Gray surround	Chromaticity	2	0.70		Arend (1993, Fig. 8)
	2D patch in void	Daylights	From unique hues and gray	Monitor	Spectral product		Chromaticity	2	0.63		Arend (1993, Fig. 8)
	2D multielement	Daylights	From 4 colors	Monitor	Spectral product	Global illuminant		3		0.83[j]	Smithson and Zaidi (2004, Fig. 5)
	2D multielement	Daylights	From 4 colors	Monitor	Spectral product	Test illuminant		2		0.86[j]	Smithson and Zaidi (2004, Fig. 12)

Table 1 (*continued*)

Experimental method	Stimulus configuration	Illuminants	Judgment	Experimental apparatus	Illuminant change	Other cues	Constraints	*N*	CI	BR	Source
	3D tableau	Daylights, colored lights	From 4 colors, 7 numbers	Stereo monitor	Spectral product			5	0.70		Schultz et al. (2006, Table 3)
	2D patch in background	Neutral and colored lights	From 8 colors	Monitor	Chromatic shift	Gray background		11		0.99[j]	Hansen et al. (2007, Fig. 5b)
	2D patch in far surround	Neutral and colored lights	From 8 colors	Monitor	Chromatic shift	Gray surround		11		0.49[k]	Hansen et al. (2007, Fig. 5b)
	2D patch in background	Neutral and colored lights	From 8 colors	Monitor	Spectral product	Gray background		4		0.75[j]	Olkkonen et al. (2009, Fig. 10)
	2D patch in background	Neutral and colored lights	Gray	Monitor	Spectral product	Gray background		4		0.94[l]	Olkkonen et al. (2009, Fig. 10)
	2D fruit images	Colored lights	Typical color	Monitor	Chromatic shift	Gray background	Chromaticity	10		0.76	Olkkonen et al. (2008, Fig. 7A)
Achromatic adjustment											
	3D room	Colored lights	Achromatic	Illuminated surfaces	Physical reflection	Gray room	Chromaticity	5	0.85		Brainard (1998, Table 2)
	3D tableau	Colored lights	Achromatic	Illuminated surfaces	Physical reflection	Gray chamber	Chromaticity	4	0.83[m]		Kraft and Brainard (1999, p. 310)
	3D tableau	Daylight, tungsten	Achromatic	Stereo monitor	Spectral product	Specularity	Chromaticity	3		0.45	Yang and Shevell (2002, Fig. 8)
	3D tableau, no stereo	Daylight, tungsten	Achromatic	Stereo monitor	Spectral product	Specularity	Chromaticity	3		0.32	Yang and Shevell (2002, Fig. 8)
	3D tableau	Daylight, tungsten	Achromatic	Stereo monitor	Spectral product		Chromaticity	3		0.28	Yang and Shevell (2002, Fig. 8)
	3D tableau, complex	Colored lights	Achromatic	Illuminated surfaces	Physical reflection	Gray wall	Chromaticity	10		0.86	Kraft et al. (2002, p. 255)
	3D tableau, simple	Colored lights	Achromatic	Illuminated surfaces	Physical reflection	Gray wall	Chromaticity	10		0.87	Kraft et al. (2002, p. 255)
	3D tableau	Daylight, colored lights	Achromatic	Stereo monitor	Spectral product	Gray background	Chromaticity	7	0.73		Delahunt and Brainard (2004b, Fig. 7)
	2D multielement	Neutral and colored lights	Achromatic	Stereo monitor	Chromatic shift			2		0.69	Werner (2006, Fig. 2)
	2D multielement, test in depth	Neutral and colored lights	Achromatic	Stereo monitor	Chromatic shift			2		0.61	Werner (2006, Fig. 2)
	2D multielement, moving test	Neutral and colored lights	Achromatic	Monitor	Chromatic shift			3		0.82	Werner (2007, Fig. 7)
	2D multielement, static test	Neutral and colored lights	Achromatic	Monitor	Chromatic shift			3		0.63	Werner (2007, Fig. 7)
Successive illuminant-reflectance-change judgment											
	3D object, uniform surround	Daylight metamers	Same material	Illuminated surfaces	Physical reflection		Chromaticity	4	0.80		Nascimento, de Almeida, et al. (2005, Fig. 2b)
	3D object, complex surround	Daylight metamers	Same material	Illuminated surfaces	Physical reflection		Chromaticity	4	0.79[n]		Nascimento, de Almeida, et al. (2005, Fig. 2b)
	2D natural scenes	Daylights	Same material	Monitor	Spectral product	Sky, sphere	Chromaticity	12	0.71		Amano et al. (2006, Fig. 3, 4)
	2D natural scenes	Daylights	Same material	Monitor	Spectral product		Chromaticity	12	0.69–0.97[o]		Foster, Amano, and Nascimento (2006, Fig. 3)
	2D multielement	Daylights	Illuminant change	Monitor	von Kries shift			5	0.74[p]		van Es et al. (2007, p. 151)
	2D multielement	Daylights	Same hue, saturation, brightness	Monitor	von Kries shift			5	0.23[p]		van Es et al. (2007, p. 151)
	2D multielement	Daylights	Material appearance	Monitor	Spectral product		Daylight locus	8	0.75		Reeves et al. (2008, Table 2)
	2D multielement	Daylights	Hue-saturation appearance	Monitor	Spectral product		Daylight locus	8	0.35		Reeves et al. (2008, Table 2)

(*continued on next page*)

Table 1 (*continued*)

Experimental method	Stimulus configuration	Illuminants	Judgment	Experimental apparatus	Illuminant change	Other cues	Constraints	*N*	CI	BR	Source
	2D multielement	Daylights	Same material	Monitor	Spectral product		Daylight locus	8	0.77		Reeves et al. (2008, Table 2)
	3D tableau	Daylight metamers	Same material	Illuminated surfaces	Physical reflection		Chromaticity	4	0.85		de Almeida et al. (2010, Fig. 5)
	2D projection[q]	Daylight metamers	Same material	Illuminated surfaces	Physical reflection		Chromaticity	4	0.83		de Almeida et al. (2010, Fig. 5)

[a] Actual definition differs slightly from that used by Arend and Reeves (1986).
[b] Eye movements also recorded.
[c] Same size patches as in multielement array.
[d] Bias-corrected ratio 0.91.
[e] Range over 40 Munsell hues.
[f] Daylights vs. two monochromatic lights.
[g] Excluding data with no illuminant shift.
[h] Eliminates memory shift.
[i] Relative to memory shift; also 40% subjects excluded.
[j] Fitted color-classification boundaries.
[k] Fitted neutral convergence point.
[l] Gray-category centroid.
[m] Control condition.
[n] Average of 2 scene types.
[o] Range over 21 scenes.
[p] Ratio of responses.
[q] 2D collage, without shading.

The bluish-green test patch in the left pattern has a perfect paper match with the greenish match patch in the right pattern (recall that the patterns were produced by simply changing the spectrum of illumination on the same set of Munsell surfaces). Suppose a colored filter is placed exactly over the test patch in the left pattern. Because the filter is localized and coextensive with the patch, the change in reflected spectrum of the pattern is perceived as a change in the test spectral reflectance. Consequently, the paper match with the match patch in the right pattern no longer holds. If, now, the same colored filter is applied to the rest of the left pattern, the paper match is restored.[14] Because the filter is then global, the change in reflected spectrum of the pattern is perceived as a change in illuminant, not as a localized change in reflectance. Therefore, by a manipulation of the surround, two different spectral reflectances in the left pattern (the original bluish-green test patch and the same patch with the coextensive filter superimposed) can be matched by the same spectral reflectance in the right pattern. Since the same argument may be applied to any other test patch, and—within limits—to any other colored filter, it follows that a paper match does not determine the test spectral reflectance uniquely, even up to metamerism. What it does determine is the chromatic relationship between the test patch and the other patches in the left pattern. This limitation on the scope of asymmetric color matching is revealed most clearly with very simple patterns consisting of just two surfaces (Section 5.1). In short, asymmetric color matching is a measure of relational color constancy (Section 3.2) rather than of color constancy.

Arend and Reeves (1986) were aware of the importance of within-scene comparisons. They informed subjects that other surfaces in the Mondrian patterns had closely related colors and "that the relations among such groups might be useful" (p. 1745). A similar approach was used by Arend et al. (1991), Cornelissen and Brenner (1995), and Bäuml (1999) but not by Brainard et al. (1997) and Foster, Amano et al. (2001). Informing subjects about these cues seems to have little effect on the recorded degree of color constancy (Table 1).

4.3. Color naming and related methods

Unlike asymmetric color matching, color naming provides a direct method of measuring color constancy, since it concentrates on identification rather than equivalence. It is also a very natural measure, as Jameson (1983) has convincingly argued. Names may be drawn from a fixed, small repertoire, e.g. the 11 monolexemic basic color terms (Berlin & Kay, 1969; Boynton & Olson, 1987), or applied without constraint.

There is, in principle, a problem of chromatic resolution in that the number of discernible surface colors is more than two million, either by theory (Pointer & Attridge, 1998) or from the computational analysis of natural scenes from hyperspectral data (Linhares, Pinto, & Nascimento, 2008), more than can be named accurately or consistently. Even when the lightness dimension is ignored, there remain about 26,000 discernible surface colors (Linhares et al., 2008). Determining the degree of constancy must therefore be established by some measure that describes how the distribution of the set of names changes with change in illuminant, one of the simplest measures being a location measure such as the centroid. The choice of this measure and the choice of the set of names both require care. Too large a set of names can lead to uncertainty in individual definitions, and too small a set, e.g. red, green, blue, and yellow, may place too much weight on the particular distributional measure, especially with a limited gamut of surface colors (Speigle & Brainard, 1996). As explained later, if distributional

[14] The restoration need not be exact, as made clear in Section 5.3.

measures are replaced by boundary tracking, some of these problems with small gamuts are avoided. Still, constancy estimates based on particular sets of colors such as red, green, blue, and yellow may be misleading, since these surface colors may be more stable perceptually than others and therefore lead to higher constancy indices (Kulikowski & Vaitkevicius, 1997).

For the reasons of naturality just mentioned, color naming was used by Troost and de Weert (1991) as a contrast with their experiment on asymmetric color matching. The stimulus was a colored disc presented on a background simulated under different illuminants, to which the subject was allowed to become adapted. Subjects assigned color names from a set of 12 modified from the categories empirically determined by Boynton and Olson (1987). The location in the CIE 1964 (u, v) chromaticity diagram of each of the color names was calculated by taking the average u, v values of all color stimuli that were called red, green, purple, and so on. The effect of the change in illuminant was quantified by the shift in this average. Despite differences in stimulus configuration, adaptation, and task, the mean Brunswik ratio of 0.65 fell between the means for simultaneous and successive asymmetric color matching with a paper-match criterion (Table 1).

The precision of color naming can be increased but only by losing some of its advantages. One way to improve it is by adjoining a numerical scale, although the method is then no longer strictly categorical. Thus, in an experiment by Speigle and Brainard (1996) with real surfaces and illuminants, subjects named a flat test surface using 11 basic color categories and gave a rating from 0 to 9. Rather than being described by a stimulus centroid, judgments were expressed as a multi-dimensional vector with differences quantified by a city-block metric.[15] Performance was compared with asymmetric color matching according to an undifferentiated color-match criterion and with achromatic adjustment, and was reported to be similar. In another experiment by Schultz, Doerschner, and Maloney (2006) with three-dimensional scenes containing objects of various shapes and materials simulated on stereo RGB monitors, subjects rated how red, green, blue, and yellow a test patch appeared. From these numbers, judgments were expressed as two-dimensional vectors, which in turn yielded a mean color-constancy index of 0.70.

A different way of improving the precision of color naming is to determine the location of color boundaries. The method has most frequently been applied using the four unique hues, i.e. unique red and green, and unique blue and yellow, where each is defined so that it contains none of the other pair (unique red, for example, is neither bluish nor yellowish). Unique gray is the special case that is neither reddish nor greenish nor bluish nor yellowish. For a given continuum in some appropriate color space, these unique colors can define a precise location measure (Valberg, 1971). The stability of these locations was measured by Arend (1993) with a test patch presented in a Mondrian pattern of Munsell surfaces, in a neutral surround, and with an empty surround (i.e. black), each simulated under different daylights. Over all conditions, the unique hues set by subjects, when adapted, produced a mean constancy index of 0.66. The mean index for the uniform and Mondrian surrounds was 0.68, slightly higher than for the test patch alone, which yielded 0.63. The last result is important in showing how categorical judgments may be formed independent of an immediate spatial chromatic reference.

This approach was extended by Smithson and Zaidi (2004) to demarcate whole regions of color space with samples being classified as either red or green in one set of trials and either yellow or blue in another set. Test patches were presented on variegated backgrounds of natural and manufactured surfaces simulated under different daylights. In one condition the illuminant varied consistently across the test and background and in another condition it varied inconsistently, with one illuminant for the test and a different one for the background, so that on a single trial there was no information about the test illuminant (cf. Arend, 1993). A form of Brunswik ratio was calculated from the achromatic point estimated from the intersection of the two classification boundaries. In the consistent condition, the mean Brunswik ratio was 0.83 and in the inconsistent condition it was 0.65. This result shows that some level of spatially highly localized adaptation can be maintained across successive test presentations (Smithson & Zaidi, 2004).

Boundary location can be deployed with larger repertoires of color names. As with unique hues, it requires making fits to multiple response categories to estimate the degree of color constancy. This technique was used in an experiment by Hansen, Walter, and Gegenfurtner (2007) in which subjects categorized colored patches as belonging to one of eight categories (the equivalent of red, orange, yellow, green, turquoise, blue, purple, and gray) in the presence of a large neutral background that had the chromaticity of the illuminant. The illuminant was one of four colored lights and one effectively white and illuminant changes were approximated by uniform chromatic shifts of the patches (Section 4.8). Subjects were adapted to the stimuli. A mean Brunswik ratio was estimated from the neutral convergence point of the fitted color boundaries: as the background was reduced to a peripheral stimulus, the ratio fell from 0.99 to 0.49.

An experiment similar to that by Hansen et al. (2007) was performed by Olkkonen, Hansen, and Gegenfurtner (2009) but with simulations of real surfaces drawn from the Munsell set and illuminants from fluorescent illuminant spectra. Again, there was a large neutral background that had the chromaticity of the illuminant, and subjects were allowed to adapt. The mean Brunswik ratio from the chromatic category boundaries was 0.75 with the full background. A higher ratio was obtained from the location of the gray category centroid, but it is unclear whether the categorization of gray was helped by the presence of the gray surround, an issue that is considered in the next section.

With an undifferentiated color-naming criterion, levels of color constancy recorded in the foregoing experiments presumably reflected primarily adaptational effects. Curiously, color naming seems not to have been used in the differentiated way introduced by Arend and Reeves (1986) in simultaneous asymmetric color matching, that is, with one criterion, observers describing the perceived surface color and with the other criterion the hue, saturation, and brightness of the stimulus. A revealing test would then be to record naming behavior across simultaneously presented surfaces under different illuminants (Section 3.1 and Tokunaga & Logvinenko, 2010).

4.4. Achromatic adjustment

The method of achromatic adjustment is normally applied in an undifferentiated way. A subject typically sets a test stimulus so that it appears "achromatic", i.e. somewhere on the continuum from gray to white (Fairchild & Lennie, 1992; Werner & Walraven, 1982). Measurements are easy to make, and easy to analyze (Arend, 1993), but there are interpretational difficulties in that, depending on the criterion used by the subject, the achromatic setting provides only an estimate of the illumination spectrum at that point or region in the scene (Foster, 2003), i.e. the subject's local white point (Webster & Leonard, 2008). Like any local measurement, it may or may not be influenced by manipulations of the scene elsewhere; for example, as with asymmetric color matching of Mondrian patterns, an achromatic setting should be indepen-

[15] The city-block metric is a Minkowski metric with index $p = 1$; i.e. the sum of the absolute coordinate differences.

dent of resampling changes in the background (Section 3.3). As noted elsewhere (Foster, 2003; Schultz et al., 2006), it is uncertain whether achromatic settings can be used to estimate the stability of perceived surface color away from the neutral point (see e.g. Delahunt & Brainard, 2004a). The bias in the estimate has, however, been used successfully to explore the effects of scene structure such as local chromatic context (Brainard, 1998; Kraft & Brainard, 1999) and to test the role of memory in surface-color judgments (Hansen, Olkkonen, Walter, & Gegenfurtner, 2006).

As with color naming, achromatic adjustment could be used in the differentiated way introduced by Arend and Reeves (1986); that is, with one criterion, observers making a stimulus look as if it were made of gray or white paper and with the other criterion making it appear devoid of hue (again see Section 3.1 and Tokunaga & Logvinenko, 2010). With a differentiated criterion, achromatic adjustments might correlate with observers' eye movements (Golz, 2010), much as asymmetric color matching does (Cornelissen & Brenner, 1995).

A potential methodological confound with achromatic adjustment, analogous to the one identified by Maloney (1999) with asymmetric color matching (Section 4.2), can come from the introduction into the scene of a white or gray surface that the subject assumes or is led to believe is spectrally neutral. The subject, rather than making an independent achromatic setting of the test stimulus, can instead match it against this surface, uninfluenced by whether the test stimulus appears achromatic. The assumed neutral surface may be a single patch (e.g. part of a familiar color palette) or the whole room in which the experiment was undertaken. There have been reports of subjects' being asked to ignore such cues (e.g. Delahunt & Brainard, 2004b; Kraft, Maloney, & Brainard, 2002), yet apparently rarely. Subjects' settings may also depend on whether they are given explicit instructions about eye movements (Golz, 2010) (Section 3.1).

4.5. Discriminating illuminant changes from reflectance changes

If surface color is considered as a proxy for surface spectral reflectance, a more objectively oriented method of measuring color constancy is to ask subjects to distinguish between changes in illumination spectrum and in surface reflectance (or material) in a scene (Craven & Foster, 1992). In this sense, color constancy is interpreted operationally, with reference not to subjectively defined qualia but to the objective properties of the world, namely, the stability of surface spectral reflectance under illuminant changes (Section 3.2). To take an illustration given by Craven and Foster (1992, p. 1360), turning on an incandescent lamp in a room lit partly by daylight may lend a yellowish cast to the surfaces of the objects it illuminates, but we do not infer that the reflecting properties of the illuminated objects have changed. In an experimental test of this approach using Mondrian patterns of Munsell surfaces simulated under different daylights, it was found that subjects were able to make these discriminations quickly, accurately, and effortlessly (Craven & Foster, 1992; Foster, Nascimento, et al., 2001). As with asymmetric color matching, the task can be performed with images of scenes presented simultaneously, side by side, or sequentially, with a variable interval. But also like asymmetric color matching, it can establish only an equivalence of stimuli (Section 4.2).

A direct comparison of observers' performance in this objective task with their performance in more subjective rating measurements was performed by Reeves et al. (2008) using sequentially presented Mondrian patterns of Munsell surfaces simulated under different daylights. Subjects judged whether a change of color originated from a change in material and, separately, they rated the stimuli for sameness of material appearance and sameness of hue and saturation. Binary judgments of origin were very closely correlated with material-appearance ratings (Pearson's $\rho = 0.93$) and produced similar constancy indices of, on average, 0.76. With hue-saturation ratings, indices fell, on average, to 0.34. Fittingly, when the different indices were plotted against each other, judgments of origin were found to be linearly separable from hue-saturation ratings (Reeves et al., 2008, Fig. 4).

This objective discrimination task was also used to compare the degree of color constancy across rural and urban natural scenes with the aid of a test probe, a matte gray sphere, physically embedded in each scene (Foster, Amano, & Nascimento, 2006). Hyperspectral images of 21 scenes were rendered under two successive daylights on an RGB monitor. Subjects reported whether there was a change in reflectance of the test probe. Mean constancy indices were found to range from 0.60 to 0.97 depending on the scene and illuminant change. The highest index of 0.97 was obtained with the scene shown in Fig. 2 (the image was cropped to exclude the specular highlight). The main explanatory factor for this dependence is discussed in Section 5.3.

The same task has also been used to demonstrate color constancy with brief simultaneous presentations of Mondrian patterns of less than 200 ms duration, and, with one subject, just 1 ms duration (Foster et al., 1992); to quantify the color constancy of red-green dichromats and anomalous trichromats (Baraas, Foster, Amano, & Nascimento, 2010) and of tritanopes (Foster, Amano, & Nascimento, 2003); and to compare color constancy in real three-dimensional tableaux with their two-dimensional planar projections (de Almeida, Fiadeiro, & Nascimento, 2010), discussed in Section 4.6.

Even so, there are difficulties in extending judgments about material constancy under changes in illuminant to judgments about material constancy when illuminant changes are combined with changes in viewpoint. The problem was demonstrated by Zaidi (2001) with textured surfaces which changed their perceived surface texture with a change in orientation. This problem is considered further in Section 7.4. In some object-identification tasks, subjects may actually adopt a suboptimal strategy (Zaidi & Bostic, 2008).

Nevertheless, with only illuminant changes on a scene, observers seem to be able to separate their visual experiences from what those experiences are of. To continue the earlier illustration from Craven and Foster (1992), despite a yellowish cast to objects illuminated by an incandescent lamp, we can still tell what the surface color is and recognize its constancy under the altered lighting.

4.6. Differences between real and simulated scenes

There has been a persistent albeit reasonable expectation, sometimes implicit, that color constancy should be better with natural, three-dimensional stimuli than with flat, coplanar, geometric scenes, usually generated on an RGB monitor and exemplified by Mondrian patterns (Boyaci, Doerschner, Snyder, & Maloney, 2006; Brainard et al., 1997; Hedrich, Bloj, & Ruppertsberg, 2009; Schultz et al., 2006; Smithson, 2005). Natural scenes offer more cues to surface structure allowing spectral reflectance and the illumination spectrum to be more easily disconfounded; the illumination itself may be more readily identified; and, unlike simulations on a monitor, natural scenes contain a clear physical referent for the notion of a paper match.

Yet, when quantified by constancy indices and Brunswik ratios, summarized in Table 1, there appears little systematic difference in performance that can be attributed to different classes of stimuli and methods of presentation, providing that variations in adaptational state are allowed for. Exact comparisons across studies are difficult, especially because not all used the same decision criterion and some used constancy indices and others Brunswik ratios.

Nonetheless, some general conclusions can be drawn, as follows. All reported indices and ratios represent means.

First, there is little difference in the effects of presentations with simulated and real stimulus materials. Simultaneous asymmetric color matching with e.g. Mondrian patterns simulated on an RGB monitor produced constancy indices of 0.79 (Troost & de Weert, 1991) and 0.81 (Bäuml, 1999) and with e.g. a tableau of real solid objects an index of 0.86 (de Almeida et al., 2004) and a Brunswik ratio of 0.84 (Granzier, Brenner, & Smeets, 2009a), all based on a paper-match criterion.

Second, there is little effect of stimulus complexity, as defined by the number of surfaces in the scene (cf. Section 7.4). Simultaneous asymmetric color matching with Mondrian patterns of two and 49 simulated Munsell surfaces produced indices of 0.72 and 0.73, respectively (Amano, Foster, & Nascimento, 2005), based on a paper-match criterion.[16] Achromatic adjustment of a test surface in a tableau of real surfaces and solids produced indices of 0.82–0.87 over two and three dimensions and differing complexities (Kraft et al., 2002). And discriminating sequential illuminant-material changes with a real three-dimensional test object in an empty uniform field and surrounded by many objects of differing shapes produced indices of 0.80 and 0.83, respectively (Nascimento, de Almeida, Fiadeiro, & Foster, 2005).

Third, there is little difference in the effects of geometric and natural images on an RGB monitor. Judging successive illuminant-reflectance changes with colored checkerboards and natural scenes rendered from hyperspectral data yielded constancy indices of 0.74 (van Es et al., 2007) and 0.81 (Foster, Amano, & Nascimento, 2006), with the latter a mean over a wide range of values (Table 1).

Fourth, and last, there is little difference in the effects of three-dimensional and two-dimensional scenes. In an experiment by de Almeida et al. (2010), subjects viewed by means of a novel optical system a tableau of real three-dimensional objects or its two-dimensional planar projection without depth cues and shading. The spectrum of the illumination on a test object or surface changed either consistently or inconsistently with the scene illuminant, and subjects reported whether the object underwent a change in material. The constancy indices for three- and two-dimensional stimuli were 0.85 and 0.83, respectively. This result contradicts the outcome of an experiment by Hedrich et al. (2009), in which subjects memorized the color of either a solid object in a tableau or a surface in a flat array under one illumination spectrum and afterwards selected the closest colored patch from an array under a different illumination spectrum. When subjects' color memory bias was allowed for (compare Section 4.7), Brunswik ratios were found to be 0.79 and 0.58, respectively. Comparison of these ratios is problematic, however, as the three- and two-dimensional learning environments were not exactly the same and 40% of subjects tested could not perform the memory task.

Stereo per se was found by Yang and Shevell (2002) to give improved achromatic settings with rendered three-dimensional objects, but Brunswik ratios were relatively low, both with and without stereo, at 0.45 and 0.32, respectively.

Why, then, given an appropriate decision criterion, does the structure of the stimulus seem to have so little effect? One possibility is that with real surfaces in real scenes it may be difficult to attend to properties other than those related to surface color, whether instructions are undifferentiated, namely "to make color matches", or specific, namely to make paper matches in the sense of Arend and Reeves (1986). With simulated surfaces on an RGB monitor, it may be easier to attend to each of the two kinds of properties as required (Reeves et al., 2008). This indeterminacy (Section 3.1) may account for the similar degrees of constancy recorded in real scenes with an undifferentiated criterion and in simulated scenes with a paper-match criterion, and for the different degrees of constancy recorded in real and simulated scenes with an undifferentiated criterion (Table 1).

4.7. Effects of familiarity and memory

Most experiments on color constancy have been designed so that performance is not confounded by the familiarity or the semantic content of the object or surface being judged. But real objects in natural scenes are usually familiar and familiarity might be expected to modify perceived surface color, an idea that may be traced back to Hering (1920). Extending earlier experiments, Siple and Springer (1983) asked subjects to select typical colors of individual fruits and vegetables presented variously as photographic images with texture and as silhouettes and collapsed to a disk. The typical-hue judgment was accurate with respect to a reference matching condition, but typical chroma was higher than for matched chroma, for all three types of presentation (Siple & Springer, 1983, Fig. 4), although the authors eventually concluded that a preference for increased saturation occurs only for objects, and not for color patches (pp. 367–368). Memory effects revealed in judgments of typical color need not, of course, be the same as memory effects revealed in delayed matching to specific examples, familiar or otherwise. Amano, Uchikawa, and Kuriki (2002) presented subjects with images of natural scenes on an RGB monitor and then tested their recall 30 s later. They detected increases in contrast in the recalled images less well than decreases in contrast, suggesting that, as with typical color, the chroma or chromaticness of pictures is enhanced in memory.

An experiment similar to that of Siple and Springer (1983) was performed by Olkkonen, Hansen, and Gegenfurtner (2008) with images of fruits and vegetables presented on an RGB monitor but they also asked subjects to make achromatic settings of the stimuli. Stimuli were presented against a neutral background, and data for different illuminants were collected in different sessions. Illuminant changes were approximated by uniform chromatic shifts of the stimuli (Section 4.8). Subjects' settings were biased away from neutral towards the opposite direction of the typical color, with the strength of the effect decreasing with decreasing naturalness of the stimuli, unlike the effects reported by Siple and Springer (1983). As the authors noted, subjects could simply have matched the test stimulus to the neutral background field, but shifts were obtained only with the fruit stimuli, not with disks (Olkkonen et al., 2008).

Semantic content and familiarity were absent in a successive asymmetric color-matching experiment by Jin and Shevell (1996) using abstract patterns presented on an RGB monitor. Subjects learned the color of a central patch surrounded by an array of other colored patches. After a 10-min delay, they made a match using a pattern with a variable center patch. Subjects were encouraged to think of the patches as papers. Matches were good, and were consistent with subjects' remembering surface color. By contrast with some experiments in simultaneous asymmetric color matching, replacing the complex surround by a gray background produced matches consistent with subjects' remembering hue, saturation, and brightness (as did removing the background altogether). Similar experiments were performed by Uchikawa, Kuriki, and Tone (1998) and by Ling and Hurlbert (2008) in which subjects memorized a colored chip or paper sample under one illumination spectrum and, after a delay, matched it by memory under a different illumination spectrum. To separate general memory effects from constancy effects, Ling and Hurlbert (2008) subtracted memory matches without an illuminant change from those with an illuminant change (cf. Jin & Shevell, 1996; Nieves, Romero, García, & Hita,

[16] Arend and Reeves (1986), Valberg and Lange-Malecki (1990), and Arend et al. (1991) came to the same conclusion, but it was limited by a potential confound (Sections 4.8 and 5.1).

2000). The resulting Brunswik ratios were on average 0.92, similar to those for other measures of constancy with adaptation but without a memory component (Table 1). Lower Brunswik ratios were obtained by Hedrich et al. (2009) using illuminated papers and objects, and a correction for memory effects, although direct comparison is, again, problematic (Section 4.6).

4.8. Methodological issues

Both Arend and Reeves (1986) and Maloney (1999) drew attention to methodological difficulties afflicting color-constancy experiments. Several potential confounds have already been mentioned. These include the inadequate control of adaptational and inferential contributions to observed performance (Section 4.2); undifferentiated color judgments, i.e. hue and saturation vs. perceived surface color (Section 3.1); and the inclusion of matching surfaces in the stimulus scene that allow the experimental task to be circumvented, both in asymmetric color matching (Section 4.2) and in achromatic adjustment (Section 4.4).[17] The inclusion of a gray surface as a background or surround to the test stimulus, thereby providing a direct cue to the illumination spectrum, has been surprisingly common (Table 1). Whether subjects are able to exploit this cue, consciously or otherwise, is another matter. Nevertheless, it seems prudent to exclude the possibility.

Another methodological difficulty concerns the generalization from local to global measurements. Asymmetric color matching, color naming, achromatic adjustment, and discriminating illuminant from reflectance changes are typically local measurements. As noted in Section 4.2, to measure the effects of a change in illuminant, Arend and Reeves (1986) ensured that the corresponding fixed surfaces in the test and match patterns had identical spectral reflectances. If these surfaces had been changed, then there may or may not have been an effect on the match, but the absence of an effect would not have implied that subjects were insensitive to these changes (Section 4.4). To measure non-local changes, a global measure is needed such as discriminating illuminant from reflectance changes that affect the whole field (e.g. Craven & Foster, 1992).

A different kind of methodological difficulty occurs when the change in cone responses to light reflected from a particular surface undergoing an illuminant change is approximated by the corresponding change in cone responses to the illuminant. That is, if (l, m, s) and (l_0, m_0, s_0) are the long-, medium-, and short-wavelength-sensitive cone excitations in response to the reflected light and light from the illuminant, respectively, and (l', m', s') and (l'_0, m'_0, s'_0) are the corresponding values with a new illuminant, then (l', m', s') is approximated by $(ll'_0/l_0, mm'_0/m_0, ss'_0/s_0)$. A subject using von Kries' coefficient rule would then show perfect constancy (Section 5.2). Because these and other reflected-light approximations produced by uniform chromatic shifts exclude metamerism, estimates of color-constancy performance are likely to be elevated with respect to those obtained with real illuminant changes.

Predictably, higher indices may also be obtained by restricting the degrees of freedom of the subject's match. In simultaneous and sequential asymmetric color matching (Foster, Amano et al., 2001), matches based on chromaticity settings alone rather than on chromaticity and luminance settings were found to give an increase in the mean constancy index of about 9%.

Where appropriate, the studies listed in Table 1 indicate the inclusion of cues such as a gray background and the use of reflected-light approximations.

[17] A gray background whose reflected light does not covary with the illuminant provides no such cue (van Es et al., 2007).

5. What physical scene properties are relevant?

The theoretical approaches to color constancy summarized in Section 2 depend on assumptions about surface spectral reflectances and illuminants. These assumptions are important in their own right but also have implications for other color-constancy phenomena and the nature of the judgments made by observers, as indicated in Section 3. Some scene properties have less behavioral significance than expected, others a more pervasive impact, and still others a role that is incompletely elucidated.

5.1. Illumination spectrum

The requirement that an estimate of the scene illumination must first be obtained for a surface-color description to be retrieved (e.g. Buchsbaum, 1980) is sometimes known as the albedo hypothesis (Beck, 1972, p. 99) or the illuminant-estimation hypothesis (Maloney & Yang, 2003). Logically, illuminant estimation is not a prerequisite for surface-color estimation any more than surface-color estimation is a prerequisite for illuminant estimation. Given the one estimate, the other is also implicitly available. Nevertheless, as illustrated in Section 2.2, several constancy models have been predicated on obtaining independent illuminant estimates. Yet since Beck's (1972) work in the lightness domain, there has been an accumulation of evidence against the direct use of illuminant estimates in observers' constancy judgments. The evidence is of several kinds, as follows.

First, notwithstanding the fact that in some conditions observers can make precise estimates of the illumination spectrum, they may be remarkably insensitive to its variation. Thus, estimates from achromatic adjustment in nearly natural scenes (Brainard, 1998) are at least as good as asymmetric color matches (Table 1), and differences in illuminant can also be inferred reasonably well, even over different scenes, on the condition that the scenes retain some regularity. For example, in one experiment by Linnell and Foster (2002), subjects were able to detect a change in daylight from a correlated color temperature of the order of 6000 K to one of 4000 K over two different Mondrian patterns, presented sequentially on an RGB monitor, providing that the patterns had sufficiently many different surfaces. But illuminant variation within a scene, especially one containing an irregular population of objects, is much more difficult to detect. In an experiment by de Almeida and Nascimento (2009), in which subjects were presented binocularly with complex real three-dimensional scenes under spatially smooth color gradients in illumination, a variation in correlated color temperature of 4000 K to 25,000 K remained undetectable in the absence of duplicated objects in the scene. Only with a still larger color gradient, from 3300 K to 25,000 K, was detection possible. Although subjects have been shown to be exquisitely sensitive to changes in illuminant position, by as little as 4 degrees elevation (Ruppertsberg, Bloj, & Hurlbert, 2008), the detectability of these changes may have been more through their effects on the luminance distribution of the reflected light than on its chromaticity.

Second, illuminant estimates and spectral-reflectance estimates may be incompatible. In a test of the illuminant-estimation hypothesis by Rutherford and Brainard (2002), restricted to the lightness domain, subjects matched the illumination in one experimental chamber to that in another chamber and then a test patch in one chamber to a patch in the other. The results were inconsistent with the illuminant-estimation hypothesis. In a broader test of the hypothesis by Granzier et al. (2009a), four lamps of differing chromaticity were used to illuminate a three-dimensional scene. Subjects' judgments of the scene illuminant based on reflected light were much poorer than their judgments of the spectral reflectances of the surfaces in the scene, where the mean Brunswik ratio reached 0.84.

Third, the use of illuminant estimates appears not to be a question of salience. It might be argued that only when illuminant cues are particularly evident, as with specular highlights (Section 2.2), are they available to observers. The influence of specular highlights and other cues to the illuminant was tested in an experiment by Yang and Maloney (2001) with three-dimensional scenes simulated on stereo RGB monitors. The illumination chromaticity signaled by each candidate cue was perturbed to see whether there was an effect on subjects' achromatic settings of a small test patch embedded on a test object. The specular-highlight cue did have a significant influence, but the sensitivity of the test patch to each cue depended on its location (Section 4.8). In fact, there were abundant specular highlights available in the experiment by Granzier et al. (2009a) and they seem not to have been used by observers. In another test of illuminant saliency by Amano, Foster, and Nascimento (2006), subjects discriminated illuminant changes from material changes in hyperspectral images of natural scenes rendered on an RGB monitor. In one of those images, the sky illuminating the scene was directly visible to the subject and in another image a large gray sphere reflecting light from the sun and sky was inserted prominently in the field of view. There was no reliable effect of these illuminant cues on color constancy. All this is not to say that illuminants are ignored. Yang and Shevell (2003) showed that by adding a second illuminant to a three-dimensional scene simulated on stereo RGB monitors, the presence of two different lights illuminating part of the scene actually degraded asymmetric color-matching performance.

Fourth, and last, effectively removing the cues from a stimulus that make illuminant estimation possible appears to have little effect, at least in matching tasks. This cue removal can be achieved with flat patterns comprising just two surfaces, for then neither space-average color nor the brighter surface gives a reliable illuminant estimate (Section 4.6). Unfortunately for this purpose, many of the early experiments with patterns of two surfaces were, as noted earlier, center-surround arrangements in which the surround was spectrally neutral (Arend & Reeves, 1986; Arend et al., 1991; Tiplitz Blackwell & Buchsbaum, 1988b; Valberg & Lange-Malecki, 1990) and which could therefore have afforded a direct estimate of the illuminant (Section 4.8). Simultaneous asymmetric color matching with patterns of two Munsell surfaces where neither was gray was performed by Amano et al. (2005) and compared with matching with Mondrian patterns of 49 Munsell surfaces of the same size in a 7×7 array. As reported in Section 4.6, the mean constancy index was 0.72 with the two surfaces and 0.73 with the 49 surfaces.

Given the weight of evidence against it, it seems unlikely that the illuminant-estimation hypothesis is correct. Additionally, much of the same evidence makes it difficult to conclude that the purpose of color constancy is to provide illuminant estimates (Section 2). But this pessimistic conclusion does depend on what accuracy is required in practice for illuminant estimates, and indeed surface-color estimates, to be useful: that is, what is good enough for the task in hand (Brill, 2008; see also Smithson, 2005). This issue is taken up in Section 8.

5.2. Von Kries' rule

Originally, the coefficient rule of von Kries (1902, 1905) was formulated to describe the adaptation of the eye to colored lights, albeit not necessarily globally (Ives, 1912; Smithson & Zaidi, 2004). That is, a triplet of long-, medium-, and short-wavelength-sensitive cone excitations (l, m, s) is scaled to a triplet $(k_L l, k_M m, k_S s)$ in the presence of one adapting light and to another triplet $(k'_L l, k'_M m, k'_S s)$ in the presence of another adapting light, where the coefficients k_L, k_M, k_S and k'_L, k'_M, k'_S depend only on the activity within the corresponding cone class. This constraint is manifested in lightness algorithms (Section 2.1) as the independence of processing within spectral channels. Crucially, von Kries' rule leaves unspecified precisely how the adapting light determines the coefficients (see e.g. Troost, Wei, & de Weert, 1992; Worthey, 1985) or equivalently the observer's local white point. Formally, von Kries adaptation constitutes a diagonal matrix transformation of cone responses (Terstiege, 1972).

In conjunction with a method for determining the scaling coefficients, the coefficient rule offers a way of producing an approximately invariant response to a given scene, but only up to an equivalence, depending on how the coefficients are normalized (Sections 2.2 and 4.2). Despite limited experimental evidence, many models of color constancy, including the multiplication rule in Land's Retinex models (Land, 1983, 1986; Land & McCann, 1971), have assumed implicitly or explicitly that the formalism of the adaptation rule applied to the light reflected from surfaces, although the dimensional limits on the efficacy of von Kries adaptation were recognized early on (Brill & West, 1981, 1986; West, 1979). As Worthey and Brill (1986) noted, the overlap of the cone spectral sensitivities restricts the accuracy of von Kries adaptation since it introduces nonzero off-diagonal elements in the transformation matrix.

As a generalized adaptational mechanism, von Kries' rule was tested by Dannemiller (1993) in a computational study of the rank orderings of cone responses to light from 337 surfaces in the Krinov set (Krinov, 1947) of reflectances under daylight and tungsten illuminants.[18] Within the limits of these reflectances, rank orderings were found to be approximately preserved, consistent with cone adaptation modeled as a multiplicative, subtractive, or monotonic nonlinear process.

With the eye in constant motion over a scene, the time course of chromatic adaptation is critical in determining the coefficients k_L, k_M, k_S. Some features of the time course have been established by varying the period of adaptation to daylight or tungsten illuminants and measuring its effect on observers' achromatic adjustments (Fairchild & Lennie, 1992) or setting unique hues (Arend, 1993). Detailed data on chromatic adaptation with a range of different adapting colors at constant luminance were reported by Fairchild and Reniff (1995), again using achromatic adjustment. They found two components of adaptation: a slow one with a time course of about 40–50 s and a fast one of about 1 s. A more fine-grained study by Rinner and Gegenfurtner (2000), also using achromatic adjustment, obtained similar slow and faster time constants of 20 s and 40–70 ms, but also a very fast component with a time constant of <10 ms. A complication in this analysis is that some of the later aspects of the time course of chromatic adaptation may be affected by the structure of the adapting field. With simple and complex adapting fields and a red-green hue-cancellation technique, Shevell (2001) showed that adding chromatic context had a proportionately greater effect on color appearance after several minutes of adaptation.

As well as these three components, another component with an extremely long time course (e.g. Belmore & Shevell, 2008) has been considered in relation to seasonal variations in the natural environment (Juricevic & Webster, 2009; Webster et al., 2007).

[18] Krinov recorded reflected spectra from a distance, effectively averaging spectral reflectances from a mixture of sources, including foliage, tree branches and trunks, and some soil, which may have led to a smoothing of the spectra and some chromatic bias; see Penndorf (1956).

5.3. Spatial ratios of cone excitations

As indicated in Section 3.2, spatial ratios of cone excitations[19] may furnish a physical substrate for relational color constancy and the discriminations which depend on it (Sections 4.2 and 4.5). Unlike the constancy associated with absolute judgments of surface color, relational color constancy does not require even implicitly an estimate of the spectrum of the illumination on the scene. Given long-, medium-, and short-wavelength-sensitive cone excitations (l_1, m_1, s_1) and (l_2, m_2, s_2) generated in response to light reflected from a pair of surfaces (or groups of surfaces) 1 and 2, their spatial ratios $(l_1/l_2, m_1/m_2, s_1/s_2)$ show extensive stability under changes in illuminant. In numerical simulations, ratios were found to be invariant to within 4% both for the Munsell set of reflectances under randomly sampled daylights (Foster & Nascimento, 1994) and for 640,000 reflectance spectra from each of 30 hyperspectral images of natural scenes under daylights with correlated color temperatures 4300 K and 25,000 K (Nascimento, Ferreira, & Foster, 2002). Fig. 6 shows by cone class spatial ratios of cone excitations from randomly chosen natural surfaces under daylights e and e' with correlated color temperatures 4000 K and 25,000 K, respectively, over a four-decade range: (a) l_1/l_2 vs. l'_1/l'_2, (b) m_1/m_2 vs. m'_1/m'_2, and (c) s_1/s_2 vs. s'_1/s'_2. The near invariance of spatial cone-excitation ratios propagates through to the corresponding invariances in affine combinations of cone signals (Zaidi, 1998) and in isoluminant and achromatic images (Nascimento & Foster, 2000). Notice that the near invariance of spatial cone-excitation ratios, unlike the near invariance of cone responses from von Kries adaptation (Section 5.2), does not depend on specific normalizing assumptions, i.e. the determination of the coefficients k_L, k_M, k_S.

Spatial cone-excitation ratios provide a compelling cue to observers trying to distinguish between illuminant and reflectance changes in scenes (Section 4.5). The significance of this cue was demonstrated in a psychophysical experiment (Nascimento & Foster, 1997) that made use of the small natural departures of ratios of excitations from perfect invariance, mentioned earlier. Mondrian patterns of Munsell surfaces were simulated on an RGB monitor in a two-interval experimental design: in one interval, the surfaces of the pattern underwent an illuminant change; in the other interval, the surfaces underwent the same change but the images were then corrected so that, for each cone class, ratios of excitations were preserved exactly. Subjects had to report which interval contained the natural illuminant change. The intervals with corrected images corresponded individually to highly improbable natural events, yet subjects systematically misidentified them as containing the natural illuminant changes. For the range of illuminants and surfaces tested, subjects' sensitivity to violations of invariance was found to depend on cone class: greatest for long-wavelength-sensitive cones and least for short-wavelength-sensitive cones.

Although ratios may act as indicators of scene stability under illuminant changes, they need to be combined in some way for individual objects to be indentified within a larger environment. Funt and Finlayson (1995) proposed forming histograms of ratios to characterize objects in an illuminant-invariant way, a procedure called color-constant color indexing. In simulations with a large set of Mondrian patterns as a test set, they found that the broadness of the cone absorption spectra impaired indexing performance. But by transforming cone responses so that they were spectrally narrower or sharper (Finlayson, Drew, & Funt, 1994a, 1994b), almost perfect identification was achieved (Funt & Finlayson, 1995). The sharpening transformations they employed corresponded well (Finlayson et al., 1994b) to the cone-opponent interactions used to model sharpening of test and field spectral sensitivities in increment-threshold measurements (Foster & Snelgar, 1983; Sperling & Harwerth, 1971).

The near invariance of spatial cone-excitation ratios—and of cone-opponent ratios (Nascimento & Foster, 2000) and affine combinations (Zaidi, Spehar, & DeBonet, 1997)—has been used to account for a variety of phenomena related to surface-color perception in addition to relational color constancy and the discrimination of illuminant and reflectance changes. They have been used to explain the properties of visual transparency, i.e. the perception of multi-colored surfaces through transparent colored filters (Khang & Zaidi, 2002a, 2002b; Ripamonti & Westland, 2003; Westland & Ripamonti, 2000), where ratios were found to give a better account than convergence models (Ripamonti, Westland, & Da Pos, 2004). They have also been used to explain the effect of background articulation on visual search (Plet & Gerbino, 2001); the identification of a spotlight across disparate brightly lit variegated scenes (Khang & Zaidi, 2004); the invariance of asymmetric color matching to test-patch position and background permutation in Mondrian patterns, where ratios were computed between the test patches and a spatial average over the whole pattern (Amano & Foster, 2004); the properties of asymmetric color matching with individual surfaces in three-dimensional tableaux (Nascimento, de Almeida, Fiadeiro, & Foster, 2004); the compatibility of chromatic shadows in the chromatic wall-of-blocks illusion (Heckman, Muday, & Schirillo, 2005); the effect of natural scene structure on detecting material changes in a test probe (Foster, Amano, & Nascimento, 2006); the variability in symmetric and asymmetric color matching (Brenner, Granzier, & Smeets, 2007); and the role of wavelength ratios in defining lines of constant spectral and non-spectral hue across illuminant changes (Pridmore, 2008, 2010).

Spatial ratios of cone excitations or of opponent combinations may also be calculated temporally rather than spatially. Eye movements back and forth across an edge separating different surfaces define a ratio signal over recent time that is nearly invariant under changes in illuminant, and which may assist surface-color judgments (Cornelissen & Brenner, 1995). Temporal variations in spatial ratios of cone excitations appear to have an added salience and may be exploited in a similar way (Section 4.2). Thus evidence has been obtained of a low-level transient signal which is generated in response to rapid changes in scene reflectance and therefore presumably in spatial ratios of cone excitations and which is progressively attenuated as these changes occur more and more gradually (Linnell & Foster, 1996). The rate constant associated with this signal was found to be approximately 2.5 s^{-1}.

5.4. Spatial structure

The physical properties of scenes can influence color constancy through more than just illuminant and reflectance spectra. Spatial properties of scenes are also relevant, not least in determining which elements contribute to illuminant estimation (Section 2.2).

Most data on the effects of scene structure on constancy judgments have come from classical color-contrast or chromatic-induction effects, in which the hue of a test stimulus is normally shifted away from the hue of an immediate surround or background field. The relationship of these effects to color constancy is uncertain. One interpretation is that they facilitate color constancy or signal the error in achieving it (e.g. Shepherd, 1992; Tiplitz Blackwell & Buchsbaum, 1988a; Walraven, Benzschawel, & Rogowitz, 1987), specifically, as local illuminant compensation (Hurlbert & Wolf, 2004). The effects have been found to be spatially limited, to about 1 degree of visual angle from the point of gaze (Brenner & Cornelissen, 1991; Brenner, Ruiz, Herráiz, Cornelissen, & Smeets, 2003; Monnier & Shevell, 2003; see also Barbur et al., 2004), with significant modulatory effects extending to 10 degrees (Wachtler et al.,

[19] The term cone contrast is sometimes preferred, but since that term may also be taken to refer to a perceptual enhancement of the physical quantity (Gilchrist, 2006, pp. 8–9), it is not used here.

2001; Walraven, 1973). Both the average chromaticity (e.g. Brenner & Cornelissen, 1998; Brenner et al., 2003) and chromatic inhomogeneity or chromatic variability of the surround (e.g. Brenner & Cornelissen, 2002; Brenner et al., 2003; Brown & MacLeod, 1997; Jenness & Shevell, 1995; Shevell & Wei, 1998) can affect color appearance. In these manipulations, however, the physical origin of the relationship between test and surround stimuli is generally not the primary concern, whereas it is central to color constancy.

For color constancy, the surround should have an effect only if it and the test stimulus, both understood as surfaces, are seen to be under the same illumination; otherwise, the chromatic properties of the surround are irrelevant to making surface-color estimates—an exception is when the same surfaces appear under different illuminants; see e.g. D'Zmura (1992). The conclusion that there is a shared illuminant obviously depends on how the surfaces of the scene are visually segmented. Motion and depth between a test and background field might both be expected to increase segmentation and reduce the effect of the background. But experimental results have varied according to whether chromatic induction or color constancy was being assessed. Relative motion was found not to disrupt chromatic induction measured by nulling (Hurlbert & Wolf, 2004) but to improve constancy by achromatic adjustment (Werner, 2007). Relative depth was reported to have no effect on chromatic contrast induction (Hurlbert & Wolf, 2004), a weak effect on red-green equilibria (Shevell & Miller, 1996), and a strong effect on constancy by achromatic adjustment (Werner, 2006). In the last experiment, changes in the spectrum of illumination on the test and background areas were used that were either consistent or inconsistent with each other. Werner (2006) found that with consistent illuminant changes, color constancy was reduced when the test and background were separated in depth; conversely, with an inconsistent illuminant change, constancy was reduced when the test and background were in the same depth plane, but not if they were in different depth planes.

6. What neural mechanisms support color constancy?

Given the variety of cues to color constancy and the several forms it can manifest, it is not surprising that multiple neural mechanisms have been identified, operating at different levels in the visual system. Broadly speaking, evidence has been reported of three kinds of activity contributing to color constancy: within-class cone adaptation, spatial comparisons of cone and cone-opponent signals, and invariant cell responses. How these separate signals are related to each other and combined with other nonchromatic signals is unclear.

6.1. Cone adaptation

Substantial chromatic adaptation takes place in the retina, although it is incomplete. Recordings from horizontal cells in monkey have shown that adaptation is cone-specific at moderate light levels and spatially local (Lee, Dacey, Smith, & Pokorny, 1999), and consistent with von Kries' rule. Changes in chromaticity of the illumination, as distinct from changes in luminance, cause correspondingly smaller changes in adaptation. Normalizing shifts in chromatic sensitivity have also been found in monkey parvocellular lateral geniculate neurons and their retinal afferents (Creutzfeldt, Crook, Kastner, Li, & Pei, 1991; Creutzfeldt, Kastner, Pei, & Valberg, 1991). Recordings from goldfish, which can make color-constant judgments (Neumeyer, Dörr, Fritsch, & Kardelky, 2002), have indicated that cone synaptic gains are modulated by the horizontal cell network in such a way that the ratios of cone outputs are almost invariant with the illumination spectrum (Kamermans, Kraaij, & Spekreijse, 1998; Kraaij, Kamermans, & Spekreijse, 1998). A retinal contribution linked to Land's Retinex models (Section 2.1) has also been proposed on the basis of subjective reports of color-induction in hemianopia (Pöppel, 1986).

6.2. Spatial comparisons

Experiments with dichoptically presented stimuli have also pointed to spatial chromatic comparisons taking place retinally or in the lateral geniculate nucleus or the monocularly driven part of V1. In a psychophysical experiment by Moutoussis and Zeki (2000), subjects were presented with a Mondrian pattern through one eye and an isolated patch from the pattern through the other, with the result that the patch had the appearance of the void color, rather than its normal appearance when both it and the Mondrian surround were viewed monocularly or binocularly. In another psychophysical experiment by Nascimento and Foster (2001) with dichoptically presented simulations of two Munsell surfaces undergoing an illuminant change, it was found that the detection of a small change in reflectance was poorer than when the pair was presented binocularly (see also Barbur et al., 2004). Performance in asymmetric color matching by patients with lesions higher than V1 has also led to the suggestion that chromatic-con-

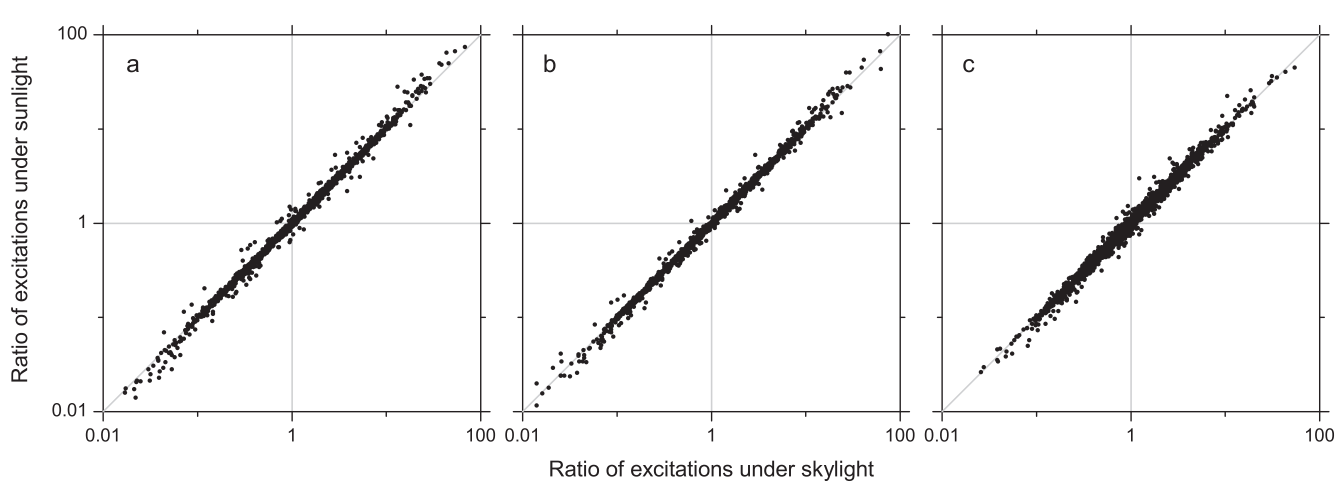

Fig. 6. Spatial cone-excitation ratios from a set of natural scenes for (a) long-, (b) medium-, and (c) short-wavelength-sensitive cones. Each point in each graph represents a pair of ratios of excitations produced by light reflected from two randomly chosen surface elements illuminated in turn by sunlight and skylight with respective correlated color temperatures 4000 K and 25,000 K (from hyperspectral data in Foster, Amano, Nascimento, and Foster (2006), after Nascimento et al. (2002)).

trast effects are calculated in area V1 or lower (Barbur et al., 2004; Hurlbert, Bramwell, Heywood, & Cowey, 1998; Kentridge, Heywood, & Cowey, 2004).

Relatively early in the visual system, therefore, there is significant local normalization to the illumination spectrum and signaling of invariants associated with surface spectral reflectance. But it is not until the cortex that signals from more spatially extended regions are available which can better capture the non-local effects of the prevailing illumination.

6.3. Invariant cell responses

Recordings from single neurons in primary visual cortex (area V1) of awake monkeys to chromatic stimuli presented on large neutral and colored backgrounds have shown changes in tuning consistent with sensitivity to chromatic contrast between stimulus and background (Wachtler, Sejnowski, & Albright, 2003). Average response changes matched psychophysical color-induction effects in human subjects under corresponding stimulus conditions, and the presence of colored stimuli remote from the test stimuli produced changes that paralleled some human chromatic-induction effects (Shevell & Wei, 1998; Wachtler et al., 2001; Wesner & Shevell, 1992). These cells in V1, with their extended spatial sensitivities, probably differ from the double-opponent cells in area V1 of macaque which have been reported by Conway and Livingstone (Conway, 2001; Conway & Livingstone, 2006). These have circular receptive-field centers and crescent-shaped surrounds with opposite chromatic tuning. With their sharpened spectral responses, they offer greater invariance to changes in illumination spectrum (Section 5.3), albeit over a more constrained extent, of the order of 1 degree (Conway & Livingstone, 2006) (see commentary by Hurlbert, 2003).

Color-selective influences from far beyond the classical receptive field have been demonstrated not only in V1 but also in area V4 (Schein & Desimone, 1990). In these V4 cells, although stimulation of the surround by itself did not cause any response, it could completely suppress the response to even optimally colored stimuli in the receptive field. More direct comparisons with human behavioral experiments analogous to those of Land and McCann (1971) have come from studies of neurons in V4 of awake and anesthetized cynomolgus and rhesus monkeys. Single-cell recordings by Zeki (1980, 1983) in anesthetized monkeys revealed responses to the surface color of a stimulus irrespective of its local spectral composition. In related experiments by Kusunoki, Moutoussis, and Zeki (2006), stimuli of different colors were presented within a pattern background and the illumination of the background then varied. The majority of V4 neurons shifted their color tuning in the direction of the chromatic shift of the background, exactly what would be needed to achieve a color-constant response. Kusunoki et al. confirmed the behavioral significance of these shifts by recording single-unit activity while the animals performed a color-discrimination task against different-colored backgrounds: the shift in monkey psychometric function was similar to the shift in the cell population response.

The importance of V4 for color constancy was tested in lesion studies by Wild et al. (1985) and Walsh, Carden, Butler, and Kulikowski (1993). They showed that intact rhesus monkeys could identify a colored patch under different illuminations but this ability was impaired in animals with V4 lesions, with little if any loss in color discrimination.

Interestingly, cells in inferior temporal cortex, which have chromaticity tuning similar to that in V4 though with even larger receptive fields (Komatsu, Ideura, Kaji, & Yamane, 1992), have been reported as changing activity depending on the visual task, either color categorization or color discrimination (Section 3.1) (Koida & Komatsu, 2007). These cells appear not to have been tested specifically for color constancy.

Evidence from humans with natural lesions is complicated. Patients with lesions of the lingual and fusiform gyri, which include part of the putative human area V4, show impaired color naming and asymmetric color matching (Clarke, Walsh, Schoppig, Assal, & Cowey, 1998; Kennard, Lawden, Morland, & Ruddock, 1995), and patients with lesions of the parieto-temporal cortex show impaired constancy in making achromatic settings although having normal color discrimination (Rüttiger et al., 1999). Additionally, fMRI experiments in humans have shown that instantaneous color constancy (Section 4.2) involves strong activation in both the fusiform color area and V1, along with significant activity in V2 and V3 (Barbur & Spang, 2008).

7. Are natural scenes and surfaces special?

As made clear in Section 4.6, there seems to be little difference between levels of color constancy recorded in the laboratory with natural or manufactured three-dimensional stimuli and with flat, coplanar, geometric scenes generated on an RGB monitor. Yet these comparisons, including those with rendered natural scenes, did not really address the question of whether the surfaces and objects of the natural environment present a special problem for color constancy. Not only do natural scenes contain color gamuts very different from those of typical laboratory stimuli, they contain very different spectral and spatial structures.

7.1. Colors of natural surfaces

As expected, the colors of natural scenes are dominated by browns, greens, and blues, from earth, vegetation, and sky. Although there had been earlier direct measurements of the colors of the surfaces in natural scenes (grass, soil, foliage, etc.) by Hendley and Hecht (1949), it was not until the work of Burton and Moorhead (1987) using photographic colorimetry that data became available on the detailed spatial chromatic variation of natural scenes, their color statistics, and spatial-frequency content. Subsequent imaging studies yielded more comprehensive data which have been important in the interpretation of the evolution of trichromacy and the trade-off between chromatic and luminance vision (e.g. Osorio & Bossomaier, 1992; Párraga, Brelstaff, Troscianko, & Moorehead, 1998) and in analyzing the effects of different adaptational mechanisms (Juricevic & Webster, 2009; Webster & Mollon, 1997), the chromatic diversity of natural scenes (Linhares et al., 2008; Nascimento et al., 2002), and the limitations of color for identifying surfaces in natural scenes (Foster et al., 2009).

Despite their selective color gamuts, the reflectance spectra of surfaces in natural scenes may present a stronger test of color constancy than the spectra of the Munsell set used in laboratory measurements. The Munsell pigments are low in metamerism (Worthey, 1985), whereas the chlorophylls and carotenoids in foliage have multiple absorbance peaks, which should lead to a higher metamerism of the surfaces containing them (Wyszecki & Stiles, 1982). As an indicator, the relative frequency of natural metamers has been found to be higher in vegetated natural scenes than in non-vegetated ones (Foster, Amano, Nascimento, & Foster, 2006), a property that may be linked to variations in observers' levels of color constancy over different scenes (Foster, Amano, & Nascimento, 2006). For surfaces and objects within the mid-to-near visual field, however, a more influential feature of natural scenes may be surfaces with high chroma, which empirically and from theoretical considerations (Morovič & Morovič, 2005; Nascimento et al., 2004) can markedly attenuate constancy levels.

7.2. Spectral basis functions for surfaces

The linear models of color constancy discussed in Section 2.3 required surface reflectance and illuminant spectra to be represented as a linear combination of a few spectral basis functions (D'Zmura & Lennie, 1986; Maloney & Wandell, 1986; Wandell, 1989; see also Brainard & Freeman, 1997; Brainard et al., 2006). The dimensionality of the spectral representation of surfaces is particularly problematic.[20] Some theoretical analyses have suggested that for sets of artificial surfaces such as the 1269 samples in the Munsell matte set, 3–8 basis functions would suffice, depending on the criterion of fit and whether part or all of the set was used (Cohen, 1964; Parkkinen, Hallikainen, & Jaaskelainen, 1989). A decisive factor is the method of approximation. Maloney (1986) repeated Cohen's (1964) analysis with the full Munsell set and 337 samples from one of the few sets of natural spectra available at the time, namely, the Krinov set. He concluded that 5–7 basis functions were needed numerically, but 3–4 would suffice if the luminous efficiency function were taken into account. Even so, the adequacy of the approximation may be context dependent, for as Maloney (1986) noted, when natural spectral reflectances are sampled, each spectral reflectance is weighted by its prevalence. Threshold numbers of basis functions need to be determined behaviorally.

To this end, Nascimento, Foster, and Amano (2005) undertook a psychophysical study with hyperspectral images of 20 outdoor scenes with reflectance spectra approximated with a variable number of basis functions. The scenes were rendered under a common daylight on an RGB monitor, and subjects had to discriminate the approximated images from the originals. In theory, an average of five basis functions should have made the two indistinguishable, with respect to a standard threshold color difference, but the original images were visually indistinguishable from their approximations only if there were at least eight basis functions. A separate psychophysical study by Oxtoby and Foster (2005) using Mondrian patterns of Munsell surfaces simulated on an RGB monitor showed that it mattered little whether the approximations were produced by PCA, by independent component analysis, or by artificial neural networks, and that at least five basis functions were needed for discrimination to be at chance levels. The differences between approximations to natural scenes and to Munsell surfaces may be due to the differing frequencies of metamerism (Section 7.1).

7.3. Spectral basis functions for illuminants

The spectral representation of natural illuminants is a little less demanding than that of natural surfaces. A characteristic-vector analysis of 622 samples of daylight by Judd, MacAdam, and Wyszecki (1964) showed that three basis functions sufficed numerically. A principal component analysis of over 1500 skylight spectra measured with a narrow field of view during a 7-month period in Granada, Spain, by Hernández-Andrés et al. (Hernández-Andrés, Romero, & Lee, 2001; Hernández-Andrés, Romero, & Nieves, 2001) showed also that, although six basis functions were necessary for spectral analysis, three were sufficient for accurate clear-sky colorimetry. Correlated color temperatures ranged from 3800 K to infinity. In fact, the distribution of the Granada spectra in the CIE (x, y) chromaticity diagram revealed marked departures from the CIE daylight locus at high correlated color temperatures.

Forests exhibit more spectral variation depending on the distribution of foliage and gaps in the canopy (Endler, 1993). A principal component analysis performed on illumination spectra in forest scenes by Chiao, Osorio, Vorobyev, and Cronin (2000) yielded good numerical approximations with three basis functions, but no behavioral discrimination experiments seem to have been performed with spectra based on these approximations. Given the greater degrees of freedom of forest illumination spectra, it might be expected that more than the three basis functions of Judd et al. (1964) would be needed.

[20] See Brill (2003) for remarks on how basis functions may be counted.

Fig. 7. Plant structures whose color appearance depends on complicated optics (Arend, 2001, Fig. 9; original image supplied by L.E. Arend).

7.4. Color constancy and the spatial variation of natural scenes

As Arend (2001) emphasized, human color constancy is fundamentally about perception in the natural world. Yet constancy experiments concerned with naturalistic stimuli have, in one way or another, been limited, variously using sets of simple real objects or surfaces in laboratory tableaux (e.g. de Almeida et al., 2010; Kraft & Brainard, 1999; Zaidi, 2001) or simulated on an RGB monitor with physics-based rendering (e.g. Yang & Maloney, 2001); colored papers moved between indoor and outdoor locations (e.g. Granzier, Brenner, & Smeets, 2009b); and hyperspectral images of natural scenes rendered with uniform illuminant changes on an RGB monitor (e.g. Foster, Amano, & Nascimento, 2006).

In many natural environments, however, surfaces and objects are complex and spatially inhomogeneous: spectral reflectance varies over surfaces and uneven geometry alters reflections and casts local shadows; and microscopic specular reflections as well as pigment effects may vary with depth, as illustrated in Fig. 7 (Arend, 2001). In this context, the notion of a unique surface spectral reflectance to be recovered visually is not well defined. Color constancy for these surfaces—which are the norm rather than the exception for vision—needs to be defined in such a way that it is congruent with their natural statistical variation.

8. Current state and significant problems

The understanding of human color constancy has advanced considerably since Wyszecki and Stiles' bleak assessment of 1982. As shown here, the theoretical requirements for constancy have been better delineated; the nature of visual judgments has been clarified; the range of experimental techniques has been greatly expanded; novel invariant properties of images and a variety of neural mechanisms have been identified; and increasing recognition has been given to the relevance of natural surfaces and scenes as laboratory stimuli. Even so, the questions posed in Section 1 have been answered at best only in part, and there remain significant uncertainties and new challenges:

(1) Although the local illumination at a point or region in a scene can be accurately estimated, it seems to be largely irrelevant to judgments about surface spectral reflectances, especially under spatially varying illumination. In such circumstances, it is unclear what sets an observer's white point (Section 5).
(2) The precise relationship between color constancy and chromatic induction remains to be determined (Section 5).
(3) Multiple mechanisms underlie constancy judgments, each providing cues to the state and stability of the observed surface, object, or scene (Section 6). Which surface-color attribute is given perceptual prominence may depend simply on the task at hand, but at present it is not possible to identify uniquely either the neural substrate for these attributes or how they are combined with other non-chromatic attributes to determine surface-color appearance.
(4) The levels of color constancy recorded experimentally vary with observers' adaptational state and decision criteria and with experimental method (Section 3 and 4, Table 1). But what degree of color constancy is good enough in practice is not known; that is, how color constant surfaces really have to be in order for objects to be reliably recognized.
(5) More generally, color constancy needs to be formulated in such a way that it can apply to the recovery of surfaces that are not spectrally uniform, or uniformly lit, or consistently structured (Section 7).

Of these, the last two challenges are potentially the most demanding. The developments of the last quarter of a century have been founded on analytical experiments with stimuli that, with a few exceptions, have been laboratory-based, deterministic, and relatively simple. A proper account of color constancy would deal with the intrinsically variable nature of real surfaces and illuminations, not by compiling summary measures involving prior probabilities but by properly capturing the complex statistical structure of natural environments. Within such an account, it may then be feasible to determine the extent of color constancy necessary for an observer's reliable interaction with the real world.

Acknowledgments

This work was supported by the Engineering and Physical Sciences Research Council (Grant number EP/F023669/1). I thank L.E. Arend, J. Broackes, B. Funt, and A.J. Reeves for useful advice; K. Amano and S.M.C. Nascimento for hyperspectral data and help with the figures; and K. Amano, B. Funt, K.R. Gegenfurtner, L.T. Maloney, M.S. Mould, S.M.C. Nascimento, A.J. Reeves, H.E. Smithson, T. Wachtler, and A. Werner for critical comments on the manuscript.

References

Amano, K., & Foster, D. H. (2004). Colour constancy under simultaneous changes in surface position and illuminant. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 271*, 2319–2326.

Amano, K., Foster, D. H., & Nascimento, S. M. C. (2005). Minimalist surface-colour matching. *Perception, 34*, 1009–1013.

Amano, K., Foster, D. H., & Nascimento, S. M. C. (2006). Color constancy in natural scenes with and without an explicit illuminant cue. *Visual Neuroscience, 23*, 351–356.

Amano, K., Uchikawa, K., & Kuriki, I. (2002). Characteristics of color memory for natural scenes. *Journal of the Optical Society of America A – Optics Image Science and Vision, 19*, 1501–1514.

Angelopoulou, E., & Poger, S. (2003). The color of specular highlights. In B. E. Rogowitz & T. N. Pappas (Eds.). *Human vision and electronic imaging VIII. Proceedings of SPIE–IS&T Electronic Imaging, SPIE* (Vol. 5007, pp. 298–309). Santa Clara, CA: SPIE-IS&T.

Arend, L. E. (1993). How much does illuminant color affect unattributed colors? *Journal of the Optical Society of America A – Optics Image Science and Vision, 10*, 2134–2147.

Arend, L. (2001). Environmental challenges to color constancy. In B. E. Rogowitz & T. N. Pappas (Eds.). *Human vision and electronic imaging VI. Proceedings of SPIE* (Vol. 4299, pp. 392–399). Bellingham, WA: SPIE.

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

Arend, L. E., Jr, Reeves, A., Schirillo, J., & Goldstein, R. (1991). Simultaneous color constancy: Papers with diverse Munsell values. *Journal of the Optical Society of America A – Optics Image Science and Vision, 8*, 661–672.

Baraas, R. C., Foster, D. H., Amano, K., & Nascimento, S. M. C. (2010). Color constancy of red–green dichromats and anomalous trichromats. *Investigative Ophthalmology & Visual Science, 51*, 2286–2293.

Barbur, J. L., de Cunha, D., Williams, C. B., & Plant, G. (2002). Experimental studies of instantaneous colour constancy – Dynamic colour matching under rapid changes of illuminant. In B. E. Rogowitz & T. N. Pappas (Eds.). *Human vision and electronic imaging VII. Proceedings of SPIE* (Vol. 4662, pp. 298–314). Bellingham, WA: SPIE.

Barbur, J. L., de Cunha, D., Williams, C. B., & Plant, G. (2004). Study of instantaneous color constancy mechanisms in human vision. *Journal of Electronic Imaging, 13*, 15–28.

Barbur, J. L., & Spang, K. (2008). Colour constancy and conscious perception of changes of illuminant. *Neuropsychologia, 46*, 853–863.

Barnard, K., Cardei, V., & Funt, B. (2002). A comparison of computational color constancy algorithms – Part I: Methodology and experiments with synthesized data. *IEEE Transactions on Image Processing, 11*, 972–984.

Bäuml, K.-H. (1995). Illuminant changes under different surface collections: Examining some principles of color appearance. *Journal of the Optical Society of America A – Optics Image Science and Vision, 12*, 261–271.

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

Beck, J. (1972). *Surface color perception*. Ithaca, NY: Cornell University Press.

Belmore, S. C., & Shevell, S. K. (2008). Very-long-term chromatic adaptation: Test of gain theory and a new method. *Visual Neuroscience, 25*, 411–414.

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

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

Boyaci, H., Doerschner, K., Snyder, J. L., & Maloney, L. T. (2006). Surface color perception in three-dimensional scenes. *Visual Neuroscience, 23*, 311–321.

Boynton, R. M. (1979). *Human color vision*. New York: Holt, Rinehart and Winston.

Boynton, R. M., & Olson, C. X. (1987). Locating basic colors in the OSA space. *Color Research and Application, 12*, 94–105.

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

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

Brainard, D. H., & Freeman, W. T. (1997). Bayesian color constancy. *Journal of the Optical Society of America A – Optics Image Science and Vision, 14*, 1393–1411.

Brainard, D. H., Longère, P., Delahunt, P. B., Freeman, W. T., Kraft, J. M., & Xiao, B. (2006). Bayesian model of human color constancy. *Journal of Vision, 6*, 1267–1281.

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

Brainard, D. H., & Wandell, B. A. (1992). Asymmetric color matching: How color appearance depends on the illuminant. *Journal of the Optical Society of America A – Optics Image Science and Vision, 9*, 1433–1448.

Bramwell, D. I., & Hurlbert, A. C. (1996). Measurements of colour constancy by using a forced-choice matching technique. *Perception, 25*, 229–241.

Brenner, E., & Cornelissen, F. W. (1991). Spatial interactions in color vision depend on distances between boundaries. *Naturwissenschaften, 78*, 70–73.

Brenner, E., & Cornelissen, F. W. (1998). When is a background equivalent? Sparse chromatic context revisited. *Vision Research, 38*, 1789–1793.

Brenner, E., & Cornelissen, F. W. (2002). The influence of chromatic and achromatic variability on chromatic induction and perceived colour. *Perception, 31*, 225–232.

Brenner, E., Granzier, J. J. M., & Smeets, J. B. J. (2007). Combining local and global contributions to perceived colour: An analysis of the variability in symmetric and asymmetric colour matching. *Vision Research, 47*, 114–125.

Brenner, E., Ruiz, J. S., Herráiz, E. M., Cornelissen, F. W., & Smeets, J. B. J. (2003). Chromatic induction and the layout of colours within a complex scene. *Vision Research, 43*, 1413–1421.

Brill, M. H. (1978). A device performing illuminant-invariant assessment of chromatic relations. *Journal of Theoretical Biology, 71*, 473–478.

Brill, M. H. (1979). Further features of the illuminant-invariant trichromatic photosensor. *Journal of Theoretical Biology, 78*, 305–308.

Brill, M. H. (2003). A non-PC look at principal components. *Color Research and Application, 28*, 69–71.

Brill, M. H. (2008). Minimal Von Kries illuminant invariance. *Color Research and Application, 33*, 320–323.

Brill, M., & West, G. (1981). Contributions to the theory of invariance of color under the condition of varying illumination. *Journal of Mathematical Biology, 11*, 337–350.

Brill, M. H., & West, G. (1986). Chromatic adaptation and color constancy: A possible dichotomy. *Color Research and Application, 11*, 196–204.

Broackes, J. (2010). What do the color-blind see? In J. Cohen & M. Matthen (Eds.), *Color ontology and color science*. Cambridge MA: MIT Press. pp. 291–405.

Brown, R. O., & 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.

Burton, G. J., & Moorhead, I. R. (1987). Color and spatial structure in natural scenes. *Applied Optics, 26*, 157–170.

Chiao, C.-C., Osorio, D., Vorobyev, M., & Cronin, T. W. (2000). Characterization of natural illuminants in forests and the use of digital video data to reconstruct illuminant spectra. *Journal of the Optical Society of America A – Optics Image Science and Vision, 17*, 1713–1721.

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

CIE (2004). *Colorimetry* (3rd ed.). Vienna: CIE Central Bureau.

Ciurea, F., & Funt, B. (2004). Failure of luminance-redness correlation for illuminant estimation. *Twelfth color imaging conference: Color science and engineering systems, technologies, and applications* (Vol. 12, pp. 42–46). Scottsdale, AZ: Society for Imaging Science and Technology.

Clarke, S., Walsh, V., Schoppig, A., Assal, G., & Cowey, A. (1998). Colour constancy impairments in patients with lesions of the prestriate cortex. *Experimental Brain Research, 123*, 154–158.

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

Conway, B. R. (2001). Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (V-1). *Journal of Neuroscience, 21*, 2768–2783.

Conway, B. R., & Livingstone, M. S. (2006). Spatial and temporal properties of cone signals in alert macaque primary visual cortex. *Journal of Neuroscience, 26*, 10826–10846.

Cornelissen, F. W., & Brenner, E. (1995). Simultaneous colour constancy revisited: An analysis of viewing strategies. *Vision Research, 35*, 2431–2448.

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

Creutzfeldt, O. D., Crook, J. M., Kastner, S., Li, C.-Y., & Pei, X. (1991). The neurophysiological correlates of colour and brightness contrast in lateral geniculate neurons: 1. Population analysis. *Experimental Brain Research, 87*, 3–21.

Creutzfeldt, O. D., Kastner, S., Pei, X., & Valberg, A. (1991). The neurophysiological correlates of colour and brightness contrast in lateral geniculate neurons: II. Adaptation and surround effects. *Experimental Brain Research, 87*, 22–45.

Dalton, J. (1794). Letter to Elihu Robinson, February, 1794. In: *Manchester Literary and Philosophical Society 1781–1851, published in Manchester Memoirs, 68* (1924) (9) (p. 115). Manchester.

Dannemiller, J. L. (1991). Lightness is not illuminant invariant: Reply to Troost and de Weert (1990). *Psychological Review, 98*, 146–148.

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

de Almeida, V. M. N., Fiadeiro, P. T., & Nascimento, S. M. C. (2004). Color constancy by asymmetric color matching with real objects in three-dimensional scenes. *Visual Neuroscience, 21*, 341–345.

de Almeida, V. M. N., Fiadeiro, P. T., & Nascimento, S. M. C. (2010). Effect of scene dimensionality on colour constancy with real three-dimensional scenes and objects. *Perception, 39*, 770–779.

de Almeida, V. M. N., & Nascimento, S. M. C. (2009). Perception of illuminant colour changes across real scenes. *Perception, 38*, 1109–1117.

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

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

D'Zmura, M. (1992). Color constancy: Surface color from changing illumination. *Journal of the Optical Society of America A – Optics Image Science and Vision, 9*, 490–493.

D'Zmura, M., & Iverson, G. (1993a). 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 – Optics Image Science and Vision, 10*, 2148–2165.

D'Zmura, M., & Iverson, G. (1993b). Color constancy: II. Results for two-stage linear recovery of spectral descriptions for lights and surfaces. *Journal of the Optical Society of America A – Optics Image Science and Vision, 10*, 2166–2180.

D'Zmura, M., & Iverson, G. (1994). Color constancy: III. General linear recovery of spectral descriptions for lights and surfaces. *Journal of the Optical Society of America A – Optics Image Science and Vision, 11*, 2389–2400.

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

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

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

Evans, R. M. (1946/1951). *Method for correcting photographic color prints*. USA: Eastman Kodak Company.

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

Fairchild, M. D., & Lennie, P. (1992). Chromatic adaptation to natural and incandescent illuminants. *Vision Research, 32*, 2077–2085.

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

Fine, I., MacLeod, D. I. A., & Boynton, G. M. (2003). Surface segmentation based on the luminance and color statistics of natural scenes. *Journal of the Optical Society of America A – Optics Image Science and Vision, 20*, 1283–1291.

Finlayson, G. D., Drew, M. S., & Funt, B. V. (1994a). Color constancy: Generalized diagonal transforms suffice. *Journal of the Optical Society of America A – Optics Image Science and Vision, 11*, 3011–3019.

Finlayson, G. D., Drew, M. S., & Funt, B. V. (1994b). Spectral sharpening: Sensor transformations for improved color constancy. *Journal of the Optical Society of America A – Optics Image Science and Vision, 11*, 1553–1563.

Finlayson, G. D., Hordley, S. D., & Hubel, P. M. (2001). Unifying color constancy. *Journal of Imaging Science and Technology, 45*, 107–116.

Finlayson, G. D., & Trezzi, E. (2004). Shades of gray and colour constancy. *Twelfth color imaging conference: Color science and engineering systems, technologies, and applications* (Vol. 12, pp. 37–41). Scottsdale, AZ: Society for Imaging Science and Technology.

Forsyth, D. A., Haddon, J., & Ioffe, S. (2001). The joy of sampling. *International Journal of Computer Vision, 41*, 109–134.

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

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

Foster, D. H., Amano, K., & Nascimento, S. M. C. (2003). Tritanopic colour constancy under daylight changes? In J. D. Mollon, J. Pokorny, & K. Knoblauch (Eds.), *Normal & Defective Colour Vision* (pp. 218–224). Oxford: Oxford University Press.

Foster, D. H., Amano, K., & Nascimento, S. M. C. (2006). Color constancy in natural scenes explained by global image statistics. *Visual Neuroscience, 23*, 341–349.

Foster, D. H., Amano, K., Nascimento, S. M. C., & Foster, M. J. (2006). Frequency of metamerism in natural scenes. *Journal of the Optical Society of America A – Optics Image Science and Vision, 23*, 2359–2372.

Foster, D. H., Craven, B. J., & Sale, E. R. H. (1992). Immediate colour constancy. *Ophthalmic and Physiological Optics, 12*, 157–160.

Foster, D. H., Marín-Franch, I., Amano, K., & Nascimento, S. M. C. (2009). Approaching ideal observer efficiency in using color to retrieve information from natural scenes. *Journal of the Optical Society of America A – Optics Image Science and Vision, 26*, B14–B24.

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

Foster, D. H., Nascimento, S. M. C., Amano, K., Arend, L., Linnell, K. J., Nieves, J. L., et al. (2001). Parallel detection of violations of color constancy. *Proceedings of the National Academy of Sciences of the United States of America, 98*, 8151–8156.

Foster, D. H., & Snelgar, R. S. (1983). Test and field spectral sensitivities of colour mechanisms obtained on small white backgrounds: Action of unitary opponent-colour processes? *Vision Research, 23*, 787–797.

Funt, B., Ciurea, F., & McCann, J. (2004). Retinex in MATLAB (TM). *Journal of Electronic Imaging, 13*, 48–57.

Funt, B. V., Drew, M. S., & Ho, J. (1991). Color constancy from mutual reflection. *International Journal of Computer Vision, 6*, 5–24.

Funt, B. V., & Finlayson, G. D. (1995). Color constant color indexing. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 17*, 522–529.

Funt, B., & Ho, J. (1989). Color from black and white. *International Journal of Computer Vision, 3*, 109–117.

Gijsenij, A., Gevers, T., & Lucassen, M. P. (2009). Perceptual analysis of distance measures for color constancy algorithms. *Journal of the Optical Society of America A – Optics Image Science and Vision, 26*, 2243–2256.

Gijsenij, A., Gevers, T., & van de Weijer, J. (2010). Generalized gamut mapping using image derivative structures for color constancy. *International Journal of Computer Vision, 86*, 127–139.

Gilchrist, A. (2006). *Seeing black and white*. New York: Oxford University Press.

Golz, J. (2008). The role of chromatic scene statistics in color constancy: Spatial integration. *Journal of Vision, 8*(13): 6, 1–16.

Golz, J. (2010). Colour constancy: Influence of viewing behaviour on grey settings. *Perception, 39*, 606–619.

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

Granzier, J. J. M., Brenner, E., Cornelissen, F. W., & Smeets, J. B. J. (2005). Luminance-color correlation is not used to estimate the color of the illumination. *Journal of Vision, 5*, 20–27.

Granzier, J. J. M., Brenner, E., & Smeets, J. B. J. (2009a). Can illumination estimates provide the basis for color constancy? *Journal of Vision, 9*(3): 18, 1–11.

Granzier, J. J. M., Brenner, E., & Smeets, J. B. J. (2009b). Reliable identification by color under natural conditions. *Journal of Vision, 9*(1): 39, 1–8.

Hagedorn, J., & D'Zmura, M. (2000). Color appearance of surfaces viewed through fog. *Perception, 29*, 1169–1184.

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

Hansen, T., Walter, S., & Gegenfurtner, K. R. (2007). Effects of spatial and temporal context on color categories and color constancy. *Journal of Vision, 7*(4): 2, 1–15.

Heckman, G. M., Muday, J. A., & Schirillo, J. A. (2005). Chromatic shadow compatibility and cone-excitation ratios. *Journal of the Optical Society of America A – Optics Image Science and Vision, 22*, 401–415.

Hedrich, M., Bloj, M., & Ruppertsberg, A. I. (2009). Color constancy improves for real 3D objects. *Journal of Vision, 9*(4): 16, 1–16.

Helson, H. (1943). Some factors and implications of color constancy. *Journal of the Optical Society of America, 33*, 555–567.

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

Hendley, C. M., & Hecht, S. (1949). The colors of natural objects and terrains, and their relation to visual color deficiency. *Journal of the Optical Society of America, 39*, 870–873.

Hering, E. (1920). *Grundzüge der Lehre vom Lichtsinn*. Berlin: Springer (Trans. Hurvich, L. M. & Jameson, D., 1964, *Outlines of a theory of the light sense*, Cambridge MA: Harvard University Press).

Hernández-Andrés, J., Romero, J., & Lee, R. L. Jr., (2001). Colorimetric and spectroradiometric characteristics of narrow-field-of-view clear skylight in Granada, Spain. *Journal of the Optical Society of America A – Optics Image Science and Vision, 18*, 412–420.

Hernández-Andrés, J., Romero, J., & Nieves, J. L. (2001). Color and spectral analysis of daylight in southern Europe. *Journal of the Optical Society of America A – Optics Image Science and Vision, 18*, 1325–1335.

Hordley, S. D. (2006). Scene illuminant estimation: Past, present, and future. *Color Research and Application, 31*, 303–314.

Hordley, S. D., & Finlayson, G. D. (2006). Reevaluation of color constancy algorithm performance. *Journal of the Optical Society of America A – Optics Image Science and Vision, 23*, 1008–1020.

Howe, C. Q., Lotto, R. B., & Purves, D. (2006). Comparison of Bayesian and empirical ranking approaches to visual perception. *Journal of Theoretical Biology, 241*, 866–875.

Hubel, P. M. (2000). The perception of color at dawn and dusk. *Journal of Imaging Science and Technology, 44*, 371–375, 387.

Hunt, D. M., Dulai, K. S., Bowmaker, J. K., & Mollon, J. D. (1995). The chemistry of John Dalton's color-blindness. *Science, 267*, 984–988.

Hurlbert, A. (1986). Formal connections between lightness algorithms. *Journal of the Optical Society of America A – Optics Image Science and Vision, 3*, 1684–1693.

Hurlbert, A. C. (1998). Computational models of color constancy. In V. Walsh & J. Kulikowski (Eds.), *Perceptual constancy: Why things look as they do* (pp. 283–322). Cambridge: Cambridge University Press.

Hurlbert, A. (2003). Colour vision: Primary visual cortex shows its influence. *Current Biology, 13*, R270–R272.

Hurlbert, A. C., Bramwell, D. I., Heywood, C., & Cowey, A. (1998). Discrimination of cone contrast changes as evidence for colour constancy in cerebral achromatopsia. *Experimental Brain Research, 123*, 136–144.

Hurlbert, A., & Wolf, K. (2004). Color contrast: A contributory mechanism to color constancy. *Progress in Brain Research, 144*, 147–160.

Ives, H. E. (1912). The relation between the color of the illuminant and the color of the illuminated object. *Transactions of Illuminating Engineering Society, 7*, 62–72.

Jameson, D. (1983). Some misunderstandings about color perception, color mixture and color measurement. *Leonardo, 16*, 41–42.

Jameson, D., & Hurvich, L. M. (1989). Essay concerning color constancy. *Annual Review of Psychology, 40*, 1–22.

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

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

Joost, U., Lee, B. B., & Zaidi, Q. (2002). Lichtenberg's letter to Goethe on "Färbige Schatten" - Commentary. *Color Research and Application, 27*, 300–303.

Judd, D. B. (1940). Hue saturation and lightness of surface colors with chromatic illumination. *Journal of the Optical Society of America, 30*, 2–32.

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

Juricevic, I., & Webster, M. A. (2009). Variations in normal color vision: V. Simulations of adaptation to natural color environments. *Visual Neuroscience, 26*, 133–145.

Kamermans, M., Kraaij, D. A., & Spekreijse, H. (1998). The cone/horizontal cell network: A possible site for color constancy. *Visual Neuroscience, 15*, 787–797.

Katz, D. (1935). *The world of colour*. London: Kegan Paul, Trench, Trübner & Co., Ltd.

Kennard, C., Lawden, M., Morland, A. B., & Ruddock, K. H. (1995). Colour identification and colour constancy are impaired in a patient with incomplete achromatopsia associated with prestriate cortical lesions. *Proceedings of the Royal Society of London Series B – Biological Sciences, 260*, 169–175.

Kentridge, R. W., Heywood, C. A., & Cowey, A. (2004). Chromatic edges, surfaces and constancies in cerebral achromatopsia. *Neuropsychologia, 42*, 821–830.

Khang, B.-G., & Zaidi, Q. (2002a). Accuracy of color scission for spectral transparencies. *Journal of Vision, 2*, 451–466.

Khang, B.-G., & Zaidi, Q. (2002b). Cues and strategies for color constancy: Perceptual scission, image junctions and transformational color matching. *Vision Research, 42*, 211–226.

Khang, B.-G., & Zaidi, Q. (2004). Illuminant color perception of spectrally filtered spotlights. *Journal of Vision, 4*, 680–692.

Koida, K., & Komatsu, H. (2007). Effects of task demands on the responses of color-selective neurons in the inferior temporal cortex. *Nature Neuroscience, 10*, 108–116.

Komatsu, H., Ideura, Y., Kaji, S., & Yamane, S. (1992). Color selectivity of neurons in the inferior temporal cortex of the awake macaque monkey. *Journal of Neuroscience, 12*, 408–424.

Kraaij, D. A., Kamermans, M., & Spekreijse, H. (1998). Spectral sensitivity of the feedback signal from horizontal cells to cones in goldfish retina. *Visual Neuroscience, 15*, 799–808.

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

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

Krinov, E. L. (1947). *Spektralnaja otrazhatelnaja sposobnost prirodnyh obrazovanij*. Moscow: Laboratorija Aerometodov Akademiia Nauk SSSR (Trans. Belkov, G. *Spectral reflectance properties of natural formations*, Ottawa: National Research Council of Canada) (Technical translation: TT-1439).

Kuehni, R. G. (1997). Memoir concerning certain phenomena of vision, by M. Monge – Commentary and translation. *Color Research and Application, 22*, 199–203.

Kulikowski, J. J., & Vaitkevicius, H. (1997). Colour constancy as a function of hue. *Acta Psychologica, 97*, 25–35.

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

Kusunoki, M., Moutoussis, K., & Zeki, S. (2006). Effect of background colors on the tuning of color-selective cells in monkey area V4. *Journal of Neurophysiology, 95*, 3047–3059.

Land, E. H. (1983). Recent advances in retinex theory and some implications for cortical computations: Color vision and the natural image. *Proceedings of the National Academy of Sciences of the United States of America, 80*, 5163–5169.

Land, E. H. (1986). Recent advances in Retinex theory. *Vision Research, 26*, 7–21.

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

Lee, H.-C. (1986). Method for computing the scene-illuminant chromaticity from specular highlights. *Journal of the Optical Society of America A – Optics Image Science and Vision, 3*, 1694–1699.

Lee, B. B., Dacey, D. M., Smith, V. C., & Pokorny, J. (1999). Horizontal cells reveal cone type-specific adaptation in primate retina. *Proceedings of the National Academy of Sciences of the United States of America, 96*, 14611–14616.

Lichtenberg, G.-C. (1793). Letter to. J. W. von Goethe, Translation and commentary by Joost et al., 2002, p. 302.

Ling, Y.-Z., & Hurlbert, A. (2008). Role of color memory in successive color constancy. *Journal of the Optical Society of America A – Optics Image Science and Vision, 25*, 1215–1226.

Linhares, J. M. M., Pinto, P. D., & Nascimento, S. M. C. (2008). The number of discernible colors in natural scenes. *Journal of the Optical Society of America A – Optics Image Science and Vision, 25*, 2918–2924.

Linnell, K. J., & Foster, D. H. (1996). Dependence of relational colour constancy on the extraction of a transient signal. *Perception, 25*, 221–228.

Linnell, K. J., & Foster, D. H. (2002). Scene articulation: Dependence of illuminant estimates on number of surfaces. *Perception, 31*, 151–159.

Long, F., & Purves, D. (2003). Natural scene statistics as the universal basis of color context effects. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 15190–15193.

Lucassen, M. P., & Walraven, J. (1996). Color constancy under natural and artificial illumination. *Vision Research, 36*, 2699–2711.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

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 – Optics Image Science and Vision, 3*, 1673–1683.

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

Maloney, L. T. (2002). Illuminant estimation as cue combination. *Journal of Vision, 2*, 493–504.

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

Maloney, L. T., & Yang, J. N. (2003). The illuminant estimation hypothesis and surface colour perception. In R. Mausfeld & D. Heyer (Eds.), *Colour perception: Mind and the physical world* (pp. 336–358). Oxford: Oxford University Press.

Mausfeld, R. (2003). 'Colour' as part of the format of different perceptual primitives: The dual coding of colour. In R. Mausfeld & D. Heyer (Eds.), *Colour perception: Mind and the physical world* (pp. 381–430). Oxford: Oxford University Press.

McCann, J. J., McKee, S. P., & Taylor, T. H. (1976). Quantitative studies in Retinex theory. A comparison between theoretical predictions and observer responses to the "Color Mondrian" experiments. *Vision Research, 16*, 445–458.

Mollon, J. (2006). Monge – The Verriest Lecture, Lyon, July 2005. *Visual Neuroscience, 23*, 297–309.

Monge, G. (1789). Memoire sur quelques phénomènes de la vision. *Annales de Chimie, 3*, 131–147 (Commentary and Translation by R. G. Kuehni, 1997, *Color Research and Application, 22*, 199–203).

Monnier, P., & Shevell, S. K. (2003). Large shifts in color appearance from patterned chromatic backgrounds. *Nature Neuroscience, 6*, 801–802.

Morovič, J., & Morovič, P. (2005). Can highly chromatic stimuli have a low color inconstancy index. *Thirteenth color imaging conference: Color science and engineering systems, technologies applications* (Vol. 13, pp. 321–325). Scottsdale, AZ: Society for Imaging Science and Technology.

Mosny, M., & Funt, B. (2007). Multispectral color constancy: Real image tests. In B.E. Rogowitz, T.N. Pappas, & S.J. Daly (Eds.), *Human vision and electronic imaging XII. Proceedings of SPIE-IS&T Electronic Imaging, SPIE* (Vol. 6492, 64920S, pp. 1–6). Bellingham, WA: SPIE-IS&T.

Moutoussis, K., & Zeki, S. (2000). A psychophysical dissection of the brain sites involved in color-generating comparisons. *Proceedings of the National Academy of Sciences of the United States of America, 97*, 8069–8074.

Murray, I. J., Daugirdiene, A., Vaitkevicius, H., Kulikowski, J. J., & Stanikunas, R. (2006). Almost complete colour constancy achieved with full-field adaptation. *Vision Research, 46*, 3067–3078.

Nascimento, S. M. C., de Almeida, V. M. N., Fiadeiro, P. T., & Foster, D. H. (2004). Minimum-variance cone-excitation ratios and the limits of relational color constancy. *Visual Neuroscience, 21*, 337–340.

Nascimento, S. M. C., de Almeida, V. M. N., Fiadeiro, P. T., & Foster, D. H. (2005). Effect of scene complexity on colour constancy with real three-dimensional scenes and objects. *Perception, 34*, 947–950.

Nascimento, S. M. C., Ferreira, F. P., & Foster, D. H. (2002). Statistics of spatial cone-excitation ratios in natural scenes. *Journal of the Optical Society of America A – Optics Image Science and Vision, 19*, 1484–1490.

Nascimento, S. M. C., & Foster, D. H. (1997). Detecting natural changes of cone-excitation ratios in simple and complex coloured images. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 264*, 1395–1402.

Nascimento, S. M. C., & Foster, D. H. (2000). Relational color constancy in achromatic and isoluminant images. *Journal of the Optical Society of America A – Optics Image Science and Vision, 17*, 225–231.

Nascimento, S. M. C., & Foster, D. H. (2001). Detecting changes of spatial cone-excitation ratios in dichoptic viewing. *Vision Research, 41*, 2601–2606.

Nascimento, S. M. C., Foster, D. H., & Amano, K. (2005). Psychophysical estimates of the number of spectral-reflectance basis functions needed to reproduce natural scenes. *Journal of the Optical Society of America A – Optics Image Science and Vision, 22*, 1017–1022.

Neumeyer, C., Dörr, S., Fritsch, J., & Kardelky, C. (2002). Colour constancy in goldfish and man: Influence of surround size and lightness. *Perception, 31*, 171–187.

Nieves, J. L., García-Beltrán, A., & Romero, J. (2000). Response of the human visual system to variable illuminant conditions: An analysis of opponent-colour mechanisms in colour constancy. *Ophthalmic and Physiological Optics, 20*, 44–58.

Nieves, J. L., Romero, J., García, J. A., & Hita, E. (2000). Visual system's adjustments to illuminant changes: Heuristic-based model revisited. *Vision Research, 40*, 391–399.

Olkkonen, M., Hansen, T., & Gegenfurtner, K. R. (2008). Color appearance of familiar objects: Effects of object shape, texture, and illumination changes. *Journal of Vision, 8*(5): 13, 1–16.

Olkkonen, M., Hansen, T., & Gegenfurtner, K. R. (2009). Categorical color constancy for simulated surfaces. *Journal of Vision, 9*(12): 6, 1–18.

Osorio, D., & Bossomaier, T. R. J. (1992). Human cone-pigment spectral sensitivities and the reflectances of natural surfaces. *Biological Cybernetics, 67*, 217–222.

Oxtoby, E. K., & Foster, D. H. (2005). Perceptual limits on low-dimensional models of Munsell reflectance spectra. *Perception, 34*, 961–966.

Parkkinen, J. P. S., Hallikainen, J., & Jaaskelainen, T. (1989). Characteristic spectra of Munsell colors. *Journal of the Optical Society of America A – Optics Image Science and Vision, 6*, 318–322.

Párraga, C. A., Brelstaff, G., Troscianko, T., & Moorehead, I. R. (1998). Color and luminance information in natural scenes. *Journal of the Optical Society of America A – Optics Image Science and Vision, 15*, 563–569.

Penndorf, R. (1956). Colors of natural objects. *Journal of the Optical Society of America, 46*, 180–182.

Plet, S., & Gerbino, W. (2001). Background articulation and relational colour constancy. *Color Research and Application, 26*, S201–S204.

Pointer, M. R., & Attridge, G. G. (1998). The number of discernible colours. *Color Research and Application, 23*, 52–54.

Pöppel, E. (1986). Long-range color-generating interactions across the retina. *Nature, 320*, 523–525.

Pridmore, R. W. (2008). Color constancy from invariant wavelength ratios: I. The empirical spectral mechanism. *Color Research and Application, 33*, 238–249.

Pridmore, R. W. (2010). Color constancy from invariant wavelength ratios: II. The nonspectral and global mechanisms. *Color Research and Application, 35*, 134–144.

Provenzi, E., De Carli, L., Rizzi, A., & Marini, D. (2005). Mathematical definition and analysis of the Retinex algorithm. *Journal of the Optical Society of America A – Optics Image Science and Vision, 22*, 2613–2621.

Reeves, A. (1992). Areas of ignorance and confusion in color science. *Behavioral and Brain Sciences, 15*, 49–50.

Reeves, A. J., Amano, K., & Foster, D. H. (2008). Color constancy: Phenomenal or projective? *Perception & Psychophysics, 70*, 219–228.

Rinner, O., & Gegenfurtner, K. R. (2000). Time course of chromatic adaptation for color appearance and discrimination. *Vision Research, 40*, 1813–1826.

Ripamonti, C., & Westland, S. (2003). Prediction of transparency perception based on cone-excitation ratios. *Journal of the Optical Society of America A – Optics Image Science and Vision, 20*, 1673–1680.

Ripamonti, C., Westland, S., & Da Pos, O. (2004). Conditions for perceptual transparency. *Journal of Electronic Imaging, 13*, 29–35.

Romero, J., Hernández-Andrés, J., Nieves, J. L., & García, J. A. (2003). Color coordinates of objects with daylight changes. *Color Research and Application, 28*, 25–35.

Ruppertsberg, A. I., Bloj, M., & Hurlbert, A. (2008). Sensitivity to luminance and chromaticity gradients in a complex scene. *Journal of Vision, 8*(9): 3, 1–16.

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

Rüttiger, L., Braun, D. I., Gegenfurtner, K. R., Petersen, D., Schönle, P., & Sharpe, L. T. (1999). Selective colour constancy deficits after circumscribed unilateral brain lesions. *Journal of Neuroscience, 19*, 3094–3106.

Schein, S. J., & Desimone, R. (1990). Spectral properties of V4 neurons in the macaque. *Journal of Neuroscience, 10*, 3369–3389.

Schultz, S., Doerschner, K., & Maloney, L. T. (2006). Color constancy and hue scaling. *Journal of Vision, 6*, 1102–1116.

Shapley, R. (1986). The importance of contrast for the activity of single neurons, the VEP and perception. *Vision Research, 26*, 45–61.

Shepherd, A. J. (1992). Coloured shadows described in a cone contrast space. *Ophthalmic and Physiological Optics, 12*, 220–224.

Shevell, S. K. (2001). The time course of chromatic adaptation. *Color Research and Application, 26*, S170–S173.

Shevell, S. K., & Kingdom, F. A. A. (2008). Color in complex scenes. *Annual Review of Psychology, 59*, 143–166.

Shevell, S. K., & Miller, P. R. (1996). Color perception with test and adapting lights perceived in different depth planes. *Vision Research, 36*, 949–954.

Shevell, S. K., & Wei, J. (1998). Chromatic induction: Border contrast or adaptation to surrounding light? *Vision Research, 38*, 1561–1566.

Siple, P., & Springer, R. M. (1983). Memory and preference for the colors of objects. *Perception & Psychophysics, 34*, 363–370.

Smithson, H. E. (2005). Sensory, computational and cognitive components of human colour constancy. *Philosophical Transactions of the Royal Society of London Series B – Biological Sciences, 360*, 1329–1346.

Smithson, H., & Zaidi, Q. (2004). Colour constancy in context: Roles for local adaptation and levels of reference. *Journal of Vision, 4*, 693–710.

Speigle, J. M., & Brainard, D. H. (1996). Is color constancy task independent? *Fourth color imaging conference: Color science, systems, and applications* (Vol. 4, pp. 167–172). Scottsdale, AZ: Society for Imaging Science and Technology.

Sperling, H. G., & Harwerth, R. S. (1971). Red–green cone interactions in the increment-threshold spectral sensitivity of primates. *Science, 172*, 180–184.

Terstiege, H. (1972). Chromatic adaptation: A state-of-the-art report. *Journal of Color Appearance, 1*, 19–23 [cont. 40].

Tiplitz Blackwell, K., & Buchsbaum, G. (1988a). The effect of spatial and chromatic parameters on chromatic induction. *Color Research and Application, 13*, 166–173.

Tiplitz Blackwell, K., & Buchsbaum, G. (1988b). Quantitative studies of color constancy. *Journal of the Optical Society of America A – Optics Image Science and Vision, 5*, 1772–1780.

Tokunaga, R., & Logvinenko, A. D. (2010). Material and lighting dimensions of object colour. *Vision Research, 50*, 1740–1747.

Tominaga, S. (1991). Surface identification using the dichromatic reflection model. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13*, 658–670.

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

Troost, J. M., Wei, L., & de Weert, C. M. M. (1992). Binocular measurements of chromatic adaptation. *Vision Research, 32*, 1987–1997.

Uchikawa, K., Kuriki, I., & Tone, Y. (1998). Measurement of color constancy by color memory matching. *Optical Review, 5*, 59–63.

Valberg, A. (1971). Method for precise determination of achromatic colours including white. *Vision Research, 11*, 157–160.

Valberg, A., & Lange-Malecki, B. (1990). "Colour constancy" in Mondrian patterns: A partial cancellation of physical chromaticity shifts by simultaneous contrast. *Vision Research, 30*, 371–380.

van Es, J. J., Vladusich, T., & Cornelissen, F. W. (2007). Local and relational judgements of surface colour: Constancy indices and discrimination performance. *Spatial Vision, 20*, 139–154.

Vazquez-Corral, J., Parrága, C. A., Vanrell, M., & Baldrich, R. (2009). Color constancy algorithms: Psychophysical evaluation on a new dataset. *Journal of Imaging Science and Technology, 53*(3): 031105, 1–9.

von Helmholtz, H. (1867). Handbuch der Physiologischen Optik, Vol. II, 1st Edition, published by Leopold Voss, Leipzig. In J.P.C. Southall (Ed.), Translation 3rd Edition (Helmholtz's Treatise on Physiological Optics, 1909), Optical Society of America, Washington, D.C., 1924, pp. 286–287. Republished by Dover Publications, New York, 1962.

von Kries, J. (1902). Theoretische Studien über die Umstimmung des Sehorgans. In *Festschrift der Albrecht-Ludwigs-Universität* (pp. 143–158). Freiburg.

von Kries, J. (1905). Die Gesichtsempfindungen. In W. Nagel (Ed.), *Handbuch der Physiologie des Menschen* (pp. 211–212). Braunschweig: Vieweg und Sohn.

Wachtler, T., Albright, T. D., & Sejnowski, T. J. (2001). Nonlocal interactions in color perception: Nonlinear processing of chromatic signals from remote inducers. *Vision Research, 41*, 1535–1546.

Wachtler, T., Sejnowski, T. J., & Albright, T. D. (2003). Representation of color stimuli in awake macaque primary visual cortex. *Neuron, 37*, 681–691.

Walraven, J. (1973). Spatial characteristics of chromatic induction; the segregation of lateral effects from straylight artefacts. *Vision Research, 13*, 1739–1753.

Walraven, J., Benzschawel, T. L., & Rogowitz, B. E. (1987). Color-constancy interpretation of chromatic induction. *Die Farbe, 34*, 269–273.

Walsh, V., Carden, D., Butler, S. R., & Kulikowski, J. J. (1993). The effects of V4 lesions on the visual abilities of macaques: Hue discrimination and colour constancy. *Behavioural Brain Research, 53*, 51–62.

Wandell, B. A. (1989). Color constancy and the natural image. *Physica Scripta, 39*, 187–192.

Webster, M. A., & Leonard, D. (2008). Adaptation and perceptual norms in color vision. *Journal of the Optical Society of America A – Optics Image Science and Vision, 25*, 2817–2825.
Webster, M. A., Mizokami, Y., & Webster, S. M. (2007). Seasonal variations in the color statistics of natural images. *Network: Computation in Neural Systems, 18*, 213–233.
Webster, M. A., & Mollon, J. D. (1997). Adaptation and the color statistics of natural images. *Vision Research, 37*, 3283–3298.
Werner, A. (2006). The influence of depth segmentation on colour constancy. *Perception, 35*, 1171–1184.
Werner, A. (2007). Color constancy improves, when an object moves: High-level motion influences color perception. *Journal of Vision*, *7*(14): 19, 1–14.
Werner, J. S., & Walraven, J. (1982). Effect of chromatic adaptation on the achromatic locus: The role of contrast, luminance and background color. *Vision Research, 22*, 929–943.
Wesner, M. F., & Shevell, S. K. (1992). Color perception within a chromatic context: Changes in red/green equilibria caused by noncontiguous light. *Vision Research, 32*, 1623–1634.
West, G. (1979). Color perception and the limits of color constancy. *Journal of Mathematical Biology, 8*, 47–53.
Westland, S., & Ripamonti, C. (2000). Invariant cone-excitation ratios may predict transparency. *Journal of the Optical Society of America A – Optics Image Science and Vision, 17*, 255–264.
Wild, H. M., Butler, S. R., Carden, D., & Kulikowski, J. J. (1985). Primate cortical area V4 important for colour constancy but not wavelength discrimination. *Nature, 313*, 133–135.
Worthey, J. A. (1985). Limitations of color constancy. *Journal of the Optical Society of America A – Optics Image Science and Vision, 2*, 1014–1026.
Worthey, J. A., & Brill, M. H. (1986). Heuristic analysis of von Kries color constancy. *Journal of the Optical Society of America A – Optics Image Science and Vision, 3*, 1708–1712.
Wyszecki, G., & Stiles, W. S. (1982). *Color science: Concepts and methods, quantitative data and formulae*. New York: John Wiley & Sons.
Yang, J. N., & Maloney, L. T. (2001). Illuminant cues in surface color perception: Tests of three candidate cues. *Vision Research, 41*, 2581–2600.
Yang, J. N., & Shevell, S. K. (2002). Stereo disparity improves color constancy. *Vision Research, 42*, 1979–1989.
Yang, J. N., & Shevell, S. K. (2003). Surface color perception under two illuminants: The second illuminant reduces color constancy. *Journal of Vision, 3*, 369–379.
Young, T. (1807). *A course of lectures on natural philosophy and the mechanical arts (Vol. I)*. London: Joseph Johnson [Lecture XXXVIII].
Zaidi, Q. (1998). Identification of illuminant and object colors: Heuristic-based algorithms. *Journal of the Optical Society of America A – Optics Image Science and Vision, 15*, 1767–1776.
Zaidi, Q. (2001). Color constancy in a rough world. *Color Research and Application, 26*, S192–S200.
Zaidi, Q., & Bostic, M. (2008). Color strategies for object identification. *Vision Research, 48*, 2673–2681.
Zaidi, Q., Spehar, B., & DeBonet, J. (1997). Color constancy in variegated scenes: Role of low-level mechanisms in discounting illumination changes. *Journal of the Optical Society of America A – Optics Image Science and Vision, 14*, 2608–2621.
Zeki, S. (1980). The representation of colours in the cerebral cortex. *Nature, 284*, 412–418.
Zeki, S. (1983). Colour coding in the cerebral cortex: The reaction of cells in monkey visual cortex to wavelengths and colours. *Neuroscience, 9*, 741–765.
Zeki, S. (1993). *A vision of the brain*. Oxford: Blackwell Science Ltd.