Colour constancy in context: Roles for local adaptation and levels of reference

Hannah Smithson Institute of Ophthalmology, University College London, London, UK

Qasim Zaidi SUNY State College of Optometry, New York, NY, USA

By determining the locations of boundaries between colour categories, we measured changes in the colour appearance of test-reflectances as a function of the simulated illumination. Test-reflectances were displayed against a variegated background of reflectance samples. Under prolonged adaptation to each illuminant, observers demonstrated a high degree of appearance-based colour constancy. By using backgrounds that consisted of chromatically biased sets of reflectances, we tested whether this stability depends on estimates of the illuminant's cone-coordinates based on simple scene statistics. The chromatic bias of the background had only a small effect on the classification of test materials. To compare the roles of spatially local and spatially extended estimation processes, we then (unknown to the observer) simulated different illuminants on the test and on the background. Observers continued to demonstrate reasonable colour constancy. To examine the relative roles of automatic adaptation and perceptual strategies, we reduced the duration of exposure to the test compared to exposure to the background (under the conflicting illuminant). The results suggest that mechanisms that preserve information across successive test-presentations (e.g. spatially local adaptation with a time course of a few seconds, and perceptual adjustments to levels of reference) are key determinants of the stability of colour appearance.

Keywords: colour constancy, local adaptation, levels of reference, temporal context, spatial context

Introduction

The term colour constancy describes the extent to which the colour of an object appears unchanging despite changes in the spectral composition of the light reflected from that object to the eye (Helson, Judd, & Warren, 1952; Land & McCann, 1971; Land, 1983; Brainard, 1998; Foster, 2003). In the present paper we consider colour constancy under a change in illuminant from sunlight to skylight, although in general the light reflected to the eye from a particular object can change for a number of reasons (e.g. occlusion or filtering of one or multiple light sources, or other changes in the geometry of the scene). With environmental reflectance spectra, the "colour conversion" (Helson, 1938) between two illuminant conditions has a simple form when expressed in terms of cone-coordinates. Our experiments were aimed towards identifying the receptoral and post-receptoral neural processes that undo this colour conversion and "transform" (Helson, 1938) the perceived colours of objects under a test illuminant towards the colours of objects under a reference illuminant.

To quantify colour constancy, we assessed changes in colour appearance under different illuminants. Our stimulus displays consisted of a square test patch presented on a variegated background of randomly oriented elliptical patches. Examples of these displays are given in Figure 1A & B. Each patch was assigned a reflectance spectrum and rendered under a particular illuminant. Reflectance spectra were chosen from measurements of natural and man-made

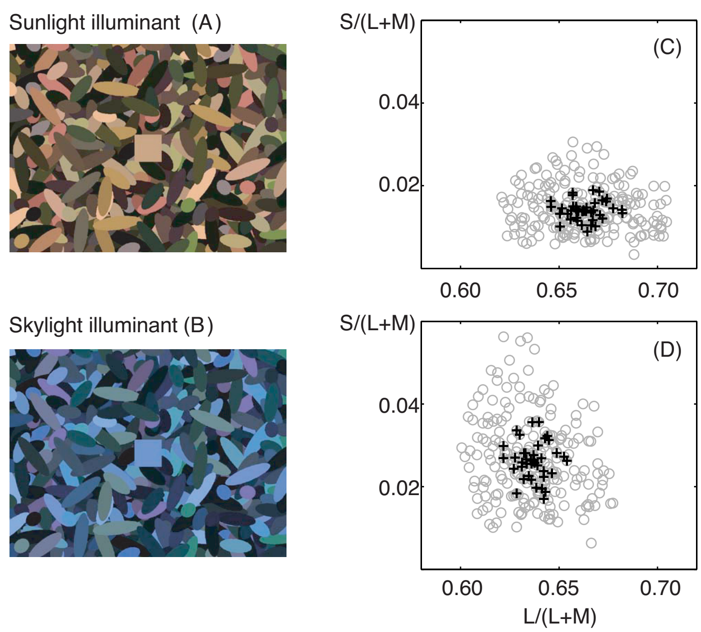

Figure 1. Left-hand panels (A & B) show examples of stimuli used in Experiment 1. On each trial, a square test patch was presented on a variegated background of randomly oriented elliptical patches. Right-hand panels show the MacLeod-Boynton chromaticity coordinates of our stimuli, rendered under sunlight (C) and skylight (D). Open circles show the complete set of 280 test-materials. The materials indicated with black plus-symbols were used to generate chromatically "balanced" backgrounds.

doi:10.1167/4.9.3 Received February 5, 2004; August 26, 2004 ISSN 1534-7362 © 2004 ARVO

objects, and we used the spectra of direct sunlight and of zenith skylight as illuminants. The observer's task was to classify the appearance of sequentially presented test-patches as either red or green in one set of trials, and as either yellow or blue in a second set (Chichilnisky & Wandell, 1999). We thus obtained a locus of test-patches that appeared neither red nor green, and a second locus that appeared neither yellow nor blue. If we assume that colour boundaries measured under different conditions describe a set of stimuli that generate equivalent signals at the decision stage, then shifts in the locations of colour boundaries provide a measure of the neural transformations performed under different conditions of observing.

In a series of experiments we performed critical tests of whether these neural transformations depend on information that is distributed over space, or on information that is spatially localized but distributed over time. In addition, we ask, are judgements of colour appearance under different conditions well predicted by differences in early adaptation, or do they reflect higher-level perceptual mechanisms?

In order to identify the neural transformations required for colour constancy we must first consider the nature of the colour conversion due to changes in the illuminant. For sets of everyday objects, and natural and man-made illuminants, when the L- (or M-, or S-) cone-coordinate for each object under one illuminant is plotted against the L- (or M-, or S-) cone-coordinate for that object under a different illuminant, the points fall close to a straight line through the origin (Dannemiller, 1993; Foster & Nascimento, 1994; Zaidi, Spehar, & DeBonet, 1997). For the object reflectances used in this study, such plots are shown in the left-hand panels of Figure 2. Within each cone class, the effect of a change in the spectrum of the illuminant is to scale the cone-coordinate by approximately the same multiplicative constant for each object. Cone-excitation ranks across a set of objects are thus approximately invariant under an illuminant change. The Macleod-Boynton (1979) chromaticity axes (L/(L+M), S/(L+M) provide a good representation of the post-receptoral colour signals that are transmitted to the cortex (Derrington, Krauskopf, & Lennie, 1984). Zaidi et al. (1997) showed that when the effects of changes in illuminant spectrum are transformed to Macleod-Boynton coordinates, the L/(L+M) chromaticities are shifted by an additive constant, whereas the S/(L+M) chromaticities are shifted by a multiplicative constant (see right-hand panels of Figure 2). Nascimento & Foster (1997) showed that multiplicative scaling of cone-signals provides a compelling cue to observers trying to distinguish between illuminant and reflectance changes in scenes, even when such scaling corresponds to highly unlikely natural events.

Identifying the type of transformation required to undo a colour conversion is the first stage in specifying a model of colour constancy. Determining how the parameters of the transformation might be set by the image is the second (e.g. Stiles, 1961; Brainard, 2004). Any complete model of colour constancy must additionally include a third component that specifies where in our perceptual apparatus these transformations are implemented. The highly systematic nature of colour conversions under a change in illuminant implies that colour constancy could be supported by simple neural mechanisms that could in principle range from automatic to volitional and from peripheral to central. The present study is aimed at elucidating the second and third components of a model of colour constancy.

Von Kries (1878, 1905) suggested that the invariance of colour metamers to adaptation level, might be due to multiplicative gain control at the photoreceptor level, and that these gains are set independently within each class of photoreceptor in inverse proportion to the local stimulation. Ives (1912) may have been the first to suggest an explicit mechanism for constancy under an illuminant change. He showed that the multiplicative factors that transform the *illuminant*'s cone-coordinates to those of an

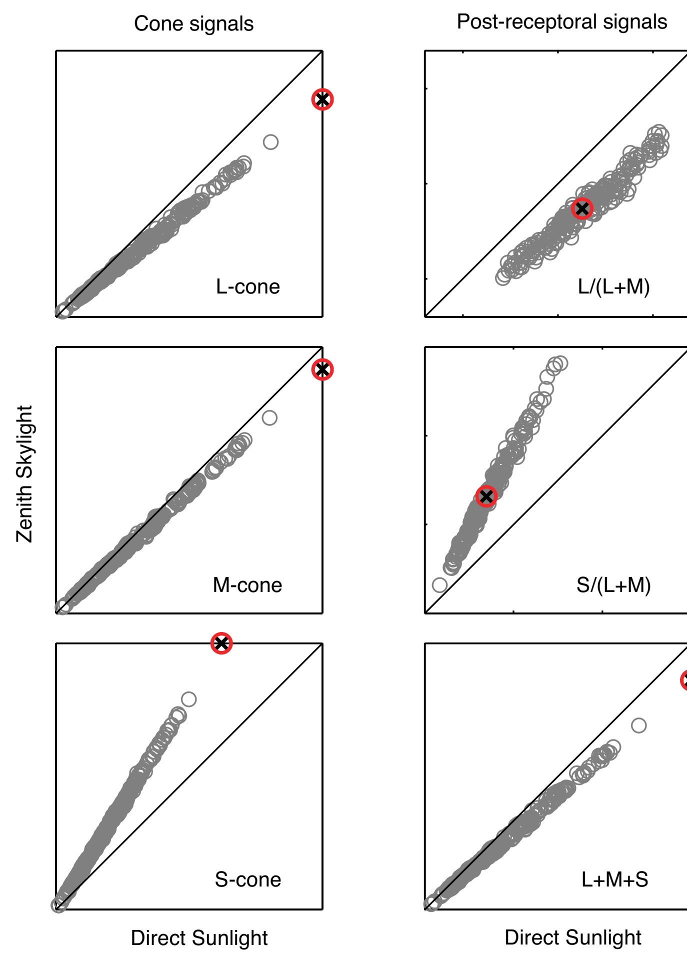

Figure 2. Left-hand plots show excitations of L-, M- and S-cones from each of the 280 reflectance spectra used in this study rendered under two illuminants: zenith skylight on the ordinate and direct sunlight on the abscissa. Right-hand plots show excitations of second-stage chromatic and luminance mechanisms (L/(L+M), S/(L+M) and L+M+S) from the same stimuli. In each plot, the black cross within a red circle represents the object of uniform spectral reflectance.

equal energy illuminant, also transform the cone-coordinates of *surfaces* to approximately their cone-coordinates under the equal-energy illuminant. The left-hand panels of Figure 2 help to illustrate why this simple transform will work. The illuminant (indicated by a black cross within a red circle) plots at the extreme end of the line of reflectances. Multiplying each cone-coordinate by the ratio of the illuminant cone-coordinates will transform most cone-coordinates to the unit diagonal, thus equating neural signals under the two illuminants. Mathematically, the Ives transform consists of multiplying all cone-coordinates by the same diagonal matrix and has been widely analyzed in the computer vision literature where it is misnamed the Von Kries transform. Von Kries' original transform multiplies each local cone-coordinate by a scalar depending only on its *local* magnitude, and thus shifts all colours towards a neutral colour (Vimal, Pokorny, & Smith, 1987; Webster, 1996) rather than achieving the required transformation to an equal energy illuminant.

The Ives transformation relies on the visual system's ability to estimate the cone-coordinates of the illuminant. Since the illuminant itself is often not in the field of view, its cone-coordinates have to be estimated from the visual scene. The most common suggestion for the estimate involves taking the mean cone-coordinates of the scene (Buchsbaum, 1980) under the assumption that the mean surface reflectance is likely to have uniform spectral reflectance (the "grey-world" hypothesis). This assumption is unlikely to be true for most scenes (Brown, 1994; Brown & MacLeod, 1997; Webster & Mollon, 1997; Webster, Malkoc, Bilson, & Webster, 2002), so Golz & MacLeod (2002) have suggested that luminance-chromaticity correlations may provide estimates that are less influenced by the set of reflectances available. Tominaga, Ebisui & Wandell (2001) argue that it is better to use just the brightest objects to make the illuminant estimate, since darker surfaces in the scene contribute more noise than signal to the estimate. Specular highlights are the extreme example of bright objects, and several authors have suggested using these to derive the illuminant estimate (D'Zmura & Lennie, 1986; Lee, 1986; Lehmann & Palm, 2001; Yang & Maloney, 2001).

A neural mechanism that integrated over a large spatial area could in principle extract the mean chromaticity. If the outputs of local subunits of such a mechanism were subjected to accelerating nonlinearities before integration, then this mechanism would estimate the illuminant by weighting scene chromaticities as an increasing function of their brightness. Psychophysical measurements, however, indicate that early adaptation mechanisms are extremely local in their spatial properties (MacLeod, Williams, & Makous, 1992; MacLeod & He, 1993; He & MacLeod, 1998). Local mechanisms could estimate the illuminant from an extended scene by using eye movements to convert spatial variations into temporal variations (D'Zmura & Lennie, 1986; Fairchild & Lennie, 1992).

Early adaptation is not the only neural transformation that could use estimated illuminant cone-coordinates. Later perceptual mechanisms could use these estimates to adjust for colour conversions (Adelson & Pentland, 1996), without losing information about the illuminant colour (Zaidi, 1998). Such mechanisms are particularly salient when the geometrical properties of the scene promote colour scission, i.e. separation of the colours of the scene into material colours and the colours of illuminants or transparencies (Hagedorn & D'Zmura, 2000). Khang & Zaidi (2002) showed that observers were able to identify like versus unlike filters across illuminants based on the similarity between colour-shifts of backgrounds and the colour-shifts of tests.

A different class of transformation mechanism involves the concept of "level of reference" or "anchoring" (Rogers, 1941; Helson, 1947). Thomas & Jones (1962) showed that matches to a reference colour were biased by the distribution of possible matching colours. In its extreme form, if perceived colours in a scene were determined entirely by rank-orders of cone-coordinates, good colour constancy would be the result because, as shown in Figure 2, colour conversions do not disturb rank-orders of cone-coordinates. This mechanism would not need an estimate for the illuminant but would, like adaptation to the mean, lead to inconstancy if the set of available materials changed.

In this study we have tried to distinguish between different types of neural transformation and the ways in which they are driven by properties of the scene. Our observers were not asked to make inferences about objects in the world. They were simply asked to judge the appearance of a test-patch displayed in the centre of a variegated image. These images were constructed by rendering a set of materials (reflectance spectra) under a particular illuminant. In the first experiment, we determined boundaries between colour categories as a function of the illuminant. Under prolonged adaptation to a single illuminant, observers demonstrated a high degree of phenomenological (appearance-based) colour constancy. The *chromaticity* that elicited the percept of neither red nor green (or neither yellow nor blue) was substantially different for the two illuminant-conditions, while the classification of *materials* was largely unaffected.

In the first experiment, the set of object reflectances was balanced so that the mean chromaticity was a reasonable estimate of the illuminant chromaticity. In the second experiment, we used sets of background reflectances whose means were significantly biased, yet this had only a small effect on the classification of test materials. Khang and Zaidi (2004) showed that on biased backgrounds, the perceived colour of the illuminant is close to that of the mean chromaticity of the scene. The high levels of constancy we observe with biased backgrounds suggest that the colour constancy transformation is not based on the simple spatial integration that seems to set the perceived colour of the illuminant. However, our results with biased backgrounds do not rule out all spatially extended constancy mechanisms since it is always possible that there exists some scene-statistic that could appropriately set the parameters of a constancy transformation, even for biased scenes.

In our third experiment, we performed a critical manipulation. We used one illuminant for the test and a different illuminant for the background. Under these conditions, the spatial context provides information only about the background illuminant, and so any spatially extended illumination mechanism would estimate the wrong illuminant for the test, and constancy would be low. In this experiment, information about the test illuminant is available only by collating local information over successive trials. Observers continued to demonstrate reasonable colour constancy for reflectances presented under the test-illuminant.

Our final experiment was designed to separate purely automatic, adaptation mechanisms from higher-level perceptual mechanisms. Test-patches, rendered under one illuminant, were briefly presented within a background rendered under a conflicting illuminant. If the test illuminant influences observers' judgements in disproportion to the relative exposure time to the two illuminants, we have evidence that contextual information about the test is tracked by higher-level mechanisms that can collate information about the test independently from information about the background.

General Methods

Equipment

Stimulus presentation and data collection were computer controlled. Stimuli were displayed on the 36° × 27° screen (1024 × 768 pixels at a viewing distance of 60 cm) of a Sony Trinitron Multiscan GDM-FW900 colour monitor with a refresh rate of 100 Hz. Images were generated using a Cambridge Research Systems (CRS) Visual Stimulus Generator (VSG 2/5) running in a 400 MHz Pentium II based system.

Gamma correction was performed using a CRS ColorCAL system. A Spectra-Scan PR-704 photospectroradiometer was used to measure complete spectra for the three phosphors. Cone absorptions were calculated for the phosphors using the Smith & Pokorny (1975) spectral sensitivities, and the resulting matrix was used to transform cone absorptions for rendered materials, to gun-values for display.

A cardboard-box (60 × 60 × 60 cm) with a rectangular window the same size as the display abutted the colour monitor, and observers viewed the display by looking through eye-holes in the opposite side of the box. The inside of the box was painted matt-black. The experiment took place in a dark room.

Observers

Three observers participated in these experiments, all of whom had normal or corrected to normal visual acuity and normal colour vision, as assessed by the Ishihara colour test and the SP II test for acquired deficiencies. Observer HES, the first author, was aware of the nature and purpose of the experiments; the other observers were not informed until after the conclusion of the experiments. Observers HES and HS are experienced psychophysical observers; JEM was not. An additional five naïve observers participated in preliminary versions of Experiments 1 and 2, and gave very similar results to those presented here.

Stimuli

Our stimulus displays consisted of a square test patch (3° × 3°) presented on a variegated background of randomly oriented elliptical patches (minor axis 1.8°, major axis 2.2° to 6.6°). Examples of these displays are given in Figure 1A & B. During the experiments we used a total of 280 simulated materials. Reflectance spectra were chosen from measurements of natural and man-made objects (Chittka, Shmida, Troje, & Menzel, 1994; Vrhel, Gershon, & Iwan, 1994; Hiltunen, 1996; Marshall, 2000) so as to obtain an even coverage of colour space within the gamut afforded by our monitor. Materials were rendered under one of two illuminants: the spectrum of direct sunlight or of zenith skylight, as measured by Taylor & Kerr (1941). The open circles in Figure 1C show the MacLeod-Boynton chromaticity coordinates of the 280 materials rendered under sunlight. Open circles in Figure 1D show the same 280 materials rendered under skylight. The rendered stimuli differ in luminance although in these figures they are shown collapsed on to the equiluminant plane. Our background patterns were coloured with subsets of 40 reflectance spectra. The subset of reflectances chosen to colour the background in Experiment 1 is indicated with black plus-symbols in plots 1C & D. The selection of background materials was different for different experiments, and is described below. Selection of the test materials is described in "Preliminary Data Collection and Analysis".

Procedure

Each trial consisted of a single judgment about a single rendered material. The observer's task was to classify the appearance of test-patches as either red or green in one set of trials, and as either yellow or blue in another set. The chromaticity of the test square was specified by the test reflectance and current illuminant condition. Presentation-order of test stimuli was random.

Ten spatially different backgrounds were generated at the start of the session and used in random sequence for the series of trials. The average chromatic properties of the background were kept constant in a given condition since the background ellipses were coloured with a predetermined set of 40 material-reflectances.

The duration of each trial was fixed at 1500 msec, and stimuli were automatically updated. If no response was received in this time, the computer stored the stimulus properties, and the trial was repeated at a later stage. The session continued until a full-set of responses had been gathered. Observers were asked not to rush, and were reassured that missed trials would be repeated. They generally found it easy to make decisions in the time allowed.

A single session lasted approximately 10 minutes. Observers first adapted to a 2-minute sequence of 80 pseudo-trials, which had all the properties of real trials for a particular session, but which did not require a response.

Preliminary data collection and analysis

Initially, a single classification was obtained for each of the 280 test materials, for each condition of Experiments 1 and 2 (see below). Our aim was to identify, for each observer, a subset of 80 test-materials (from the complete set of 280) that were close to either the red-green or the blue-yellow colour-boundary under particular conditions of observation. Shifts in classification boundaries are most clearly revealed by judgements of materials that lie close to classification boundaries. We therefore wanted to concentrate our measurements in the main phases of Experiments 1, 2, 3, and 4 on these materials of interest.

In order to formally identify a subset of test materials near a classification boundary, we submitted our preliminary data to discriminant function analysis. The goal is to predict group membership (e.g. red versus green) based on a linear combination of interval variables (i.e. luminance and chromaticity coordinates of the test materials). We performed multiple logistic regression with response (green = 0 and red = 1, or yellow = 0 and blue = 1) transformed to the "log odds ratio," and used chromaticity coordinates, and powers of chromaticity coordinates, as interval variables. The resulting discriminant functions provide a convenient description of our data, but we make no attempt to interpret them in terms of visual mechanisms.

Our preliminary data showed that the set of materials near the boundary was very similar in all conditions of Experiments 1 and 2. Additional data for HES revealed that the boundary obtained under the conditions of Experiment 3 was also defined by a similar set of materials. We therefore used preliminary data for Experiments 1 and 2 to choose a single set of 80 test-materials that would be used for all subsequent red-green judgements, and a second single set that would be used for all subsequent blue-yellow judgements. This guarantees that shifts in the locations of classification boundaries reflect differences in experimental conditions, rather than differences in the sampling of reflectance space. The discriminant function was evaluated for all materials, and those that fell within a specified range (symmetric around a classification probability of 0.5) were chosen. The range was extended until the number of chosen materials (i.e. the union for all conditions) reached 80.

Final data collection and analysis

To improve the accuracy of our data, the 80 test-materials near the red-green boundary and the 80 test-materials near the blue-yellow boundary were classified repeatedly by our observers under each condition of Experiments 1, 2, 3 and 4. The final estimate of each classification boundary is derived from 8 classifications of each of the 80 test materials presented in random order across two sessions (i.e. 320 judgements per session). Each session was confined to one illuminant and one background condition for a single experiment. Sessions for Experiments 1, 2, 3 and 4 were interleaved and counter-balanced. Observers were unaware of the experiment or type of session they were running.

Repeated classifications for each material provided psychometric functions relating stimulus chromaticity to classification-probability i.e. the percentage of times the stimulus is classified as red (versus green), or yellow (versus blue). Multiple linear least squares regression was then used to determine the best fitting function (3-dimensional, second- or third-order polynomial) relating chromaticity and classification-probability. Logistic regression gave similar results. Fits were accepted if they accounted for a significant proportion of the variance and if there was no obvious structure to the residuals (assessed by inspecting animated 3-dimensional plots of the data points and the best-fitting surface, and 3-dimensional plots of the residuals). We assume that the surface defined by a classification-probability of 0.5 is the surface that best divides chromaticity space into red versus green, or yellow versus blue, and thus represents the classification boundary we are seeking.

Experiment 1

Introduction

We assessed the extent of appearance-based colour constancy under a (simulated) global illuminant change. We obtained classification boundaries for two conditions, one in which scenes were rendered under direct sunlight and the other in which scenes were rendered under zenith skylight (Figure 1A & B). Background patterns were coloured with 40 reflectance spectra chosen to form a balanced set with an approximately uniform mean reflectance.

Methods

The open circles in Figure 1C & D show the MacLeod-Boynton chromaticity coordinates of the 280 materials rendered under sunlight and skylight respectively. The subset of reflectances chosen to colour the background is indicated with black plus-symbols. The set was chosen to have an approximately uniform mean reflectance, and so is centred on the chromaticity of the illuminant.

As described above, preliminary measurements allowed us to identify, for each observer, subsets of materials that fall close to classification boundaries. Figure 3A-D shows preliminary classifications from one observer (HES). Symbols are colour-coded on the basis of the observer's response and, even with only one response per point, chromaticity space segments lawfully into colour categories. The grey circles in these plots identify the 80 test locations that were closest to the red/green, or yellow/blue boundaries. Our final estimates of classification boundaries for this observer were derived from repeated classifications of each of these materials.

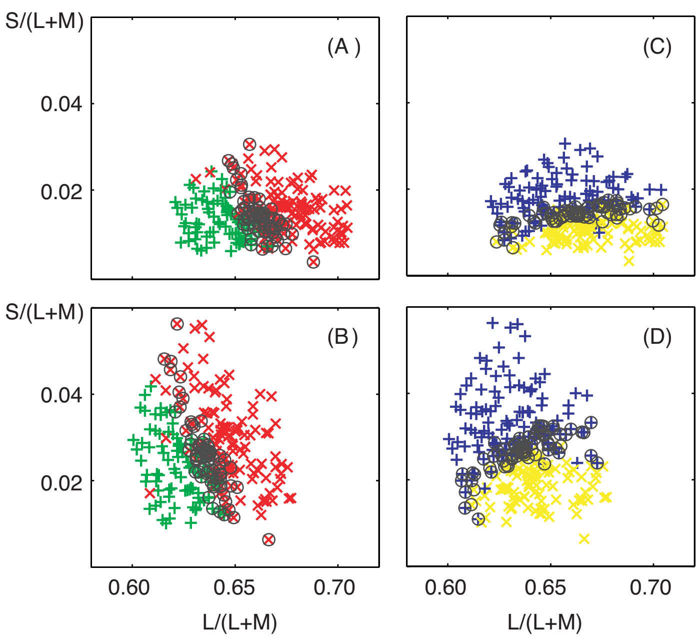

Figure 3. The 4 panels show the MacLeod-Boynton chromaticity coordinates of our stimuli, rendered under sunlight (A & C) and skylight (B & D), and colour-coded according to preliminary classifications obtained from one observer (HES). Coloured symbols indicate "red (×) / green (+)" (A & B) and "yellow (×) / blue (+)" (C & D) classifications. Open circles indicate those materials that were identified from preliminary measurements (under all conditions of Experiments 1 and 2) as being close to the colour boundaries for this observer, and that were used as test-materials in Experiments 1, 2, 3 and 4.

Results

Classification boundaries

The classification boundary that divides colour space into red versus green (or the boundary that divides yellow from blue) forms a surface in 3-dimensional colour space. The panels on the left of Figure 4 show the lines of intersection of these surfaces with the mean luminance (15 cd/m^2) equiluminant plane of the MacLeod-Boynton chromaticity diagram. Traces of classification boundaries are plotted for three observers. Red lines represent boundaries obtained under sunlight illumination; blue lines represent boundaries obtained under skylight illumination. Clearly, the illuminant has a large effect on the location of red/green and yellow/blue boundaries in chromaticity space.

Panels on the right show classification boundaries evaluated in reflectance space (i.e. material chromaticities plotted as if rendered under a spectrally uniform illuminant). We now see that classification boundaries obtained under the two illuminant conditions partition reflectance space in the same way i.e. the grouping of test-materials into colour categories is largely unaffected by the illuminant under which they are rendered. This is a demonstration of classical appearance-based colour constancy.

Constancy Index

One common way to assess the extent of colour constancy is to calculate a colour constancy index. These indi-

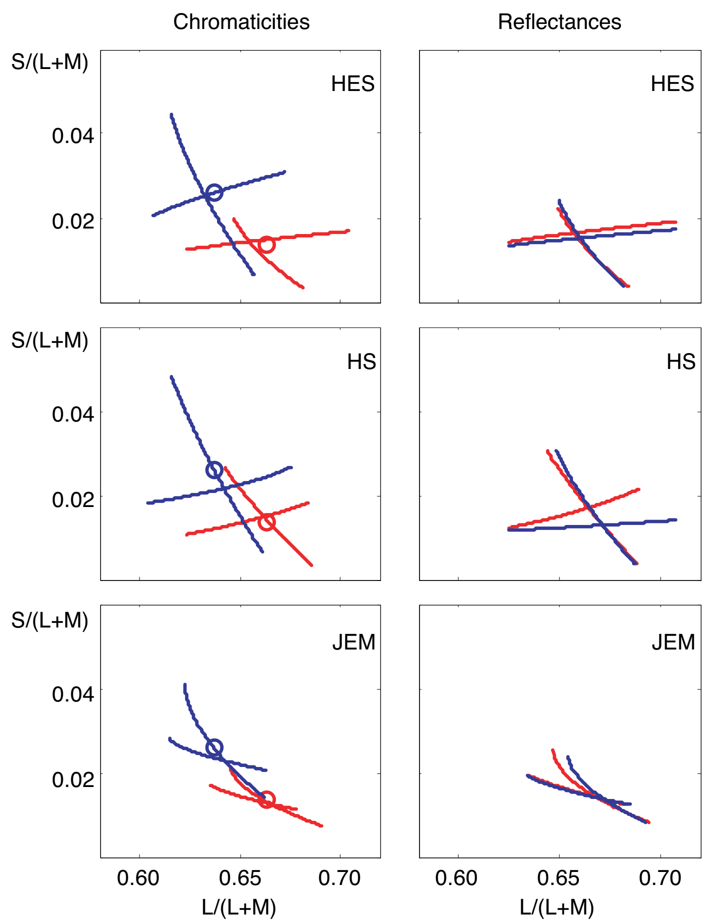

Figure 4. Data for the three observers from Experiment 1, with a global illuminant change on balanced backgrounds. Panels on the left show traces of classification boundaries in chromaticity space evaluated at a luminance of 15 cd/m^2 (i.e. line of intersection of the surface in 3-D colour space that is defined by a classification-probability of 0.5, and the 15 cd/m^2 equiluminant plane). Red lines show boundaries obtained under sunlight; blue lines show boundaries obtained under skylight. Red and blue open-circles show the corresponding illuminant chromaticities. Panels on the right show the same boundaries represented in reflectance space (i.e. as if materials were rendered under an equal energy illuminant).

ces typically relate the measured shift in the location of the achromatic point to the shift in the chromaticity of a material of uniform spectral reflectance that is caused by an illuminant change (Brainard, 1998). As illustrated in Figure 2, the effect of an illuminant change on cone-excitations can be well summarised by multiplicative scaling, and at the opponent-stage by multiplicative scaling of the S-opponent signal, and translational (additive) scaling of the L/M-opponent signal. We have defined two colour constancy indices, one appropriate for dimensions undergoing multiplicative change, and the other for dimensions undergoing additive change.

If b_1 is the coordinate of Illuminant 1, b_2 is the coordinate of Illuminant 2, and a_1 and a_2 are the respective achromatic settings, then our multiplicative constancy index is defined as

$$C = (\log(a_1/a_2))/(\log(b_1/b_2)) \quad (1)$$

The value (a_1/a_2), which is derived from the achromatic settings, reveals the scaling factor used by the putative multiplicative neural transformation; b_1/b_2 provides a summary of the colour conversion imposed by the illuminant change. For perfect constancy $(a_1/a_2) = (b_1/b_2)$ and the index is equal to one. If there is no neural transformation due to the illuminant, the achromatic setting is determined by the cone-coordinates, therefore $a_1 = a_2$, and the index is zero.

For dimensions undergoing translational scaling, e.g. the L/M-opponent axis of MacLeod-Boynton space, we use an index defined as

$$C = |a_1\text{-}a_2|/|b_1\text{-}b_2| \quad (2)$$

This is equivalent to the index proposed by Yang & Shevell (2002). Again, $C = 0$ indicates no constancy, and $C = 1$ indicates perfect constancy. However, since the mapping between chromaticity space and so-called uniform colour space is not yet known, and is likely to be nonlinear and depend on adaptation state, no constancy index can provide an absolute measure of how steady a material will appear under an illuminant change.

We derived achromatic points from our data by calculating the point of intersection of the red/green and the blue/yellow classification boundaries (i.e. the intersection of the lines in the left-hand panels of Figure 4). Figure 5 shows constancy indices evaluated at 15 cd/m^2, and expressed as percentages, for the 3 observers, for the L-, M- and S-cone signals and for the L/(L+M) and S/(L+M) opponent signals. Indices for achromatic points evaluated at luminances between 10 and 20 cd/m^2 vary by less than 6%. The data presented here show high levels of constancy, with indices ranging from 58% to 94%. For the three observers, HES, HS and JEM, mean constancy indices were 87%, 72% and 94% calculated from cone signals, and 87%, 68% and 93% calculated from opponent signals.

Discussion

The scenes used in our experiments contained a rich variety of surface reflectances, but were impoverished compared to three-dimensional real-world scenes (for example, they were devoid of specular highlights, shadows, binocular disparity and mutual reflections, which Kraft & Brainard, 1999 identified as useful cues). Yet our observers demonstrated reliable colour constancy.

Recovering surface-reflectance under an unknown illuminant is a mathematically under-determined problem. But cues to the illuminant *are* available from the statistical properties of the sample of chromaticities presented to the observer. The mean chromaticity of a scene has been suggested as a cue to the colour of the illuminant because, for a given scene, this statistic varies systematically with changes in illumination. In Experiment 1, the mean chromaticity of the scene provides a reliable estimate of the cone-coordinates of the illuminant, so good colour constancy would be predicted by spatially extended adaptation or alternatively by a high-level mechanism that used the mean to derive an illuminant estimate.

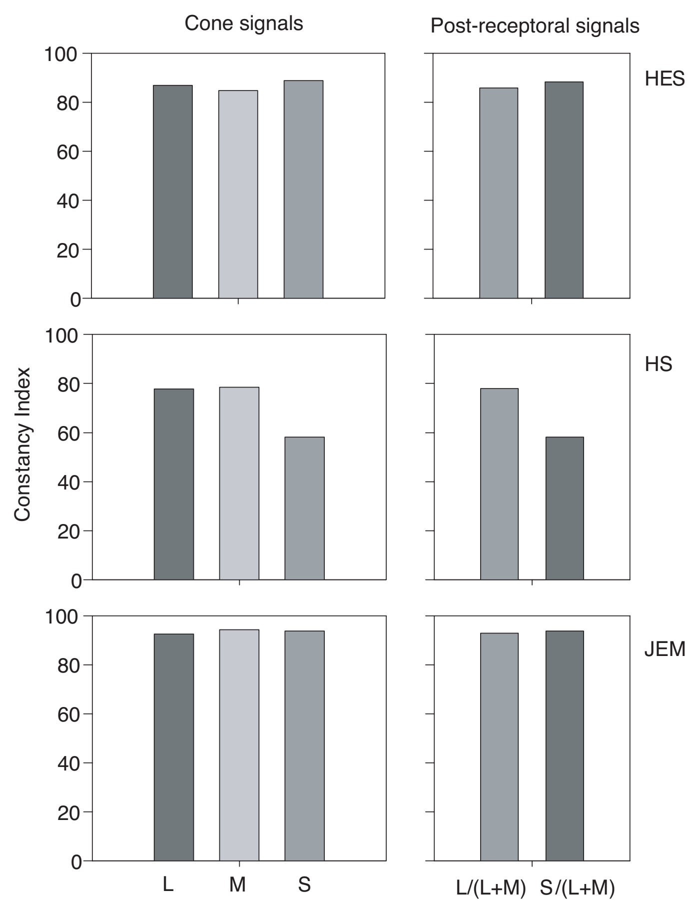

Figure 5. Constancy indices (expressed as percentages) obtained in Experiment 1 for the three observers. Panels on the left show indices calculated with cone-signals; panels on the right show indices calculated with signals in the L/(L+M) and S/(L+M) chromatic mechanisms. See text for definition of constancy indices.

In more realistic situations, however, the group of objects in a scene might have a chromatic bias (Brown, 1994; Webster et al., 2002), and might differ under the two illuminants. In Experiment 2 we simulated conditions where the mean chromaticity of the scene does not provide a good estimate of the cone-coordinates of the illuminant.

Experiment 2

Introduction

We obtained classification boundaries under four additional conditions (Figure 6A-D). We used two illuminants (sunlight and skylight) and two biased sets of reflectances for the background (red-blue biased, and green-yellow biased).

Methods

The grey circles in Figure 6E & F show the MacLeod-Boynton chromaticity coordinates of the 280 materials rendered under sunlight (upper panel) and skylight (lower panel). The biased subsets used in Experiment 2 are drawn from the +S / +L (red-blue) and the –S / –L (green-yellow) quadrants of MacLeod-Boynton space and are indicated with pink plus-symbols and lime-green crosses respectively. Our final estimates of classification boundaries were derived from repeated classification of the 80 test materials that were identified as lying close to classification boundaries from preliminary measurements for Experiments 1 and 2.

Results

Classification boundaries

Figure 7A & B shows classification boundaries for red-blue and green-yellow biased backgrounds respectively. The red lines represent colour boundaries under sunlight and the blue lines represent boundaries under skylight. Again the illuminant has a large effect on the locations of red/green and yellow/blue boundaries in chromaticity space, but the locations of the boundaries in reflectance space are largely unaltered by the illuminant condition.

Figure 8 allows direct comparison of the effect of the chromatic bias of the background. Each panel shows data for one observer, and for one illuminant condition. The three traces (black, pink and lime-green lines) show the boundaries obtained with three different backgrounds (balanced, red-blue and green-yellow respectively). Colour-coded plus-symbols indicate the location of the mean chromaticities of each of the three backgrounds. Large changes in the mean chromaticity of the scene produce small but systematic shifts in the locations of the classification boundaries. Achromatic points determined in the presence of biased backgrounds are slightly displaced towards the mean chromaticity of the background.

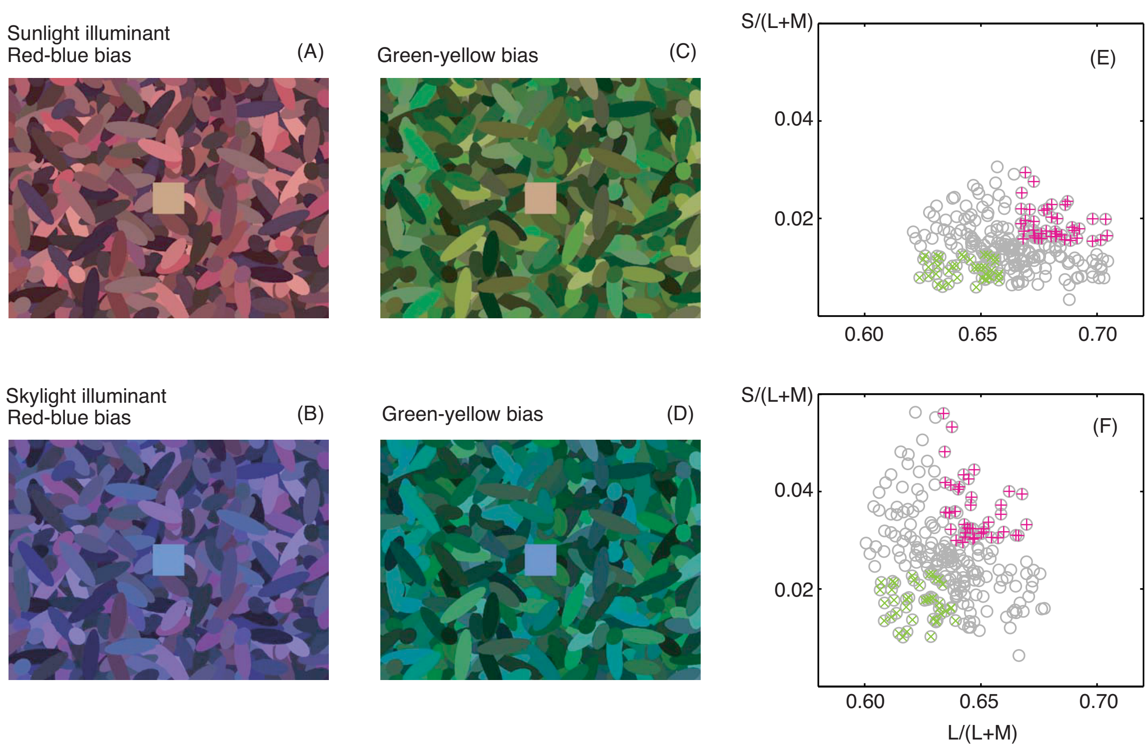

Figure 6. Left-hand panels (A, B, C & D) show examples of stimuli used in Experiment 2. The top row shows stimuli rendered under sunlight; the bottom row shows stimuli rendered under skylight. The subset of materials used to generate red-blue biased backgrounds is indicated with pink plus-symbols in the panels on the right (E & F). The subset used to generate green-yellow biased backgrounds is indicated with lime-green crosses.

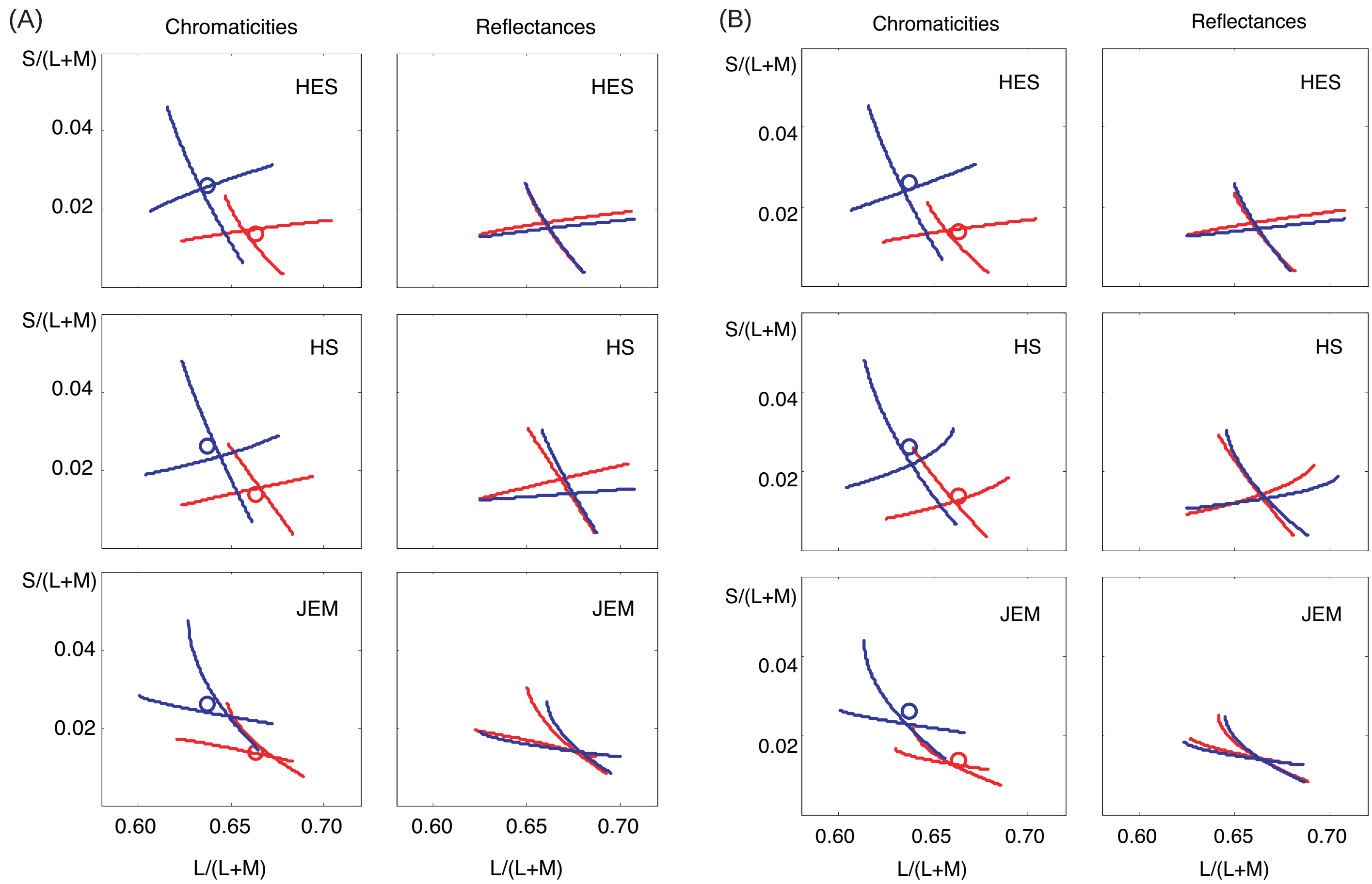

Figure 7. Data for the three observers from Experiment 2, with a global illuminant change on (A) red-blue biased backgrounds, (B) green-yellow biased backgrounds. Panels on the left show trace of classification boundaries in chromaticity space, evaluated at a luminance of 15 cd/m^2, and obtained under sunlight (red lines) or skylight (blue lines). Open-circles show the corresponding illuminant chromaticities. Panels on the right show the same boundaries represented in reflectance space (i.e. as if materials were rendered under an equal energy illuminant).

Constancy indices

Figure 9 shows, for the 3 observers, constancy indices evaluated at 15 cd/m^2 along the L/M-opponent and S-opponent axes of MacLeod-Boynton space. Constancy indices obtained with balanced backgrounds are re-plotted for comparison (black bars). Indices obtained with red-blue and green-yellow biased backgrounds are plotted as light and dark grey bars respectively. Observers demonstrate high levels of constancy in all conditions.

One strategy by which the visual system could achieve stability of the colour appearance of materials is by separating the signal reaching the eye into a component that depends on the illumination and a component that depends on the material reflectance. Given that the mean chromaticity of a scene varies systematically with changes in illumination, it seems appropriate to ask whether performance with biased backgrounds is consistent with misattribution of the bias in reflectances to a bias in illumination. We can use modified versions of the constancy indices defined above (Equations 1 and 2). Now b_1 and b_2 represent the coordinates of the mean chromaticities of the backgrounds, and a_1 and a_2 represent the corresponding achromatic settings. Attributing the change in mean chromaticity of the background to a change in illuminant is the wrong assumption, so here the constancy index is zero for perfect constancy, and 1.0 for no constancy. For observers HES and HS, constancy indices evaluated along both axes of MacLeod-Boynton colour space are all less than 0.2 indicating good constancy. Indices for observer JEM are less than 0.2 for the S/(L+M) axis but around 0.4 for the L/(L+M) axis. Under the conditions of our experiment, colour appearance is relatively little affected by a change in the mean chromaticity of the background, and a bias in the set of reflectances available is largely not misattributed to a bias in the spectrum of the illumination.

Discussion

Clearly, the illuminant has a large effect on the location of red/green and yellow/blue boundaries in chromaticity space, while the chromatic bias of the background has a small effect. The data from Experiments 1 and 2 show that, under prolonged adaptation to a single illuminant,

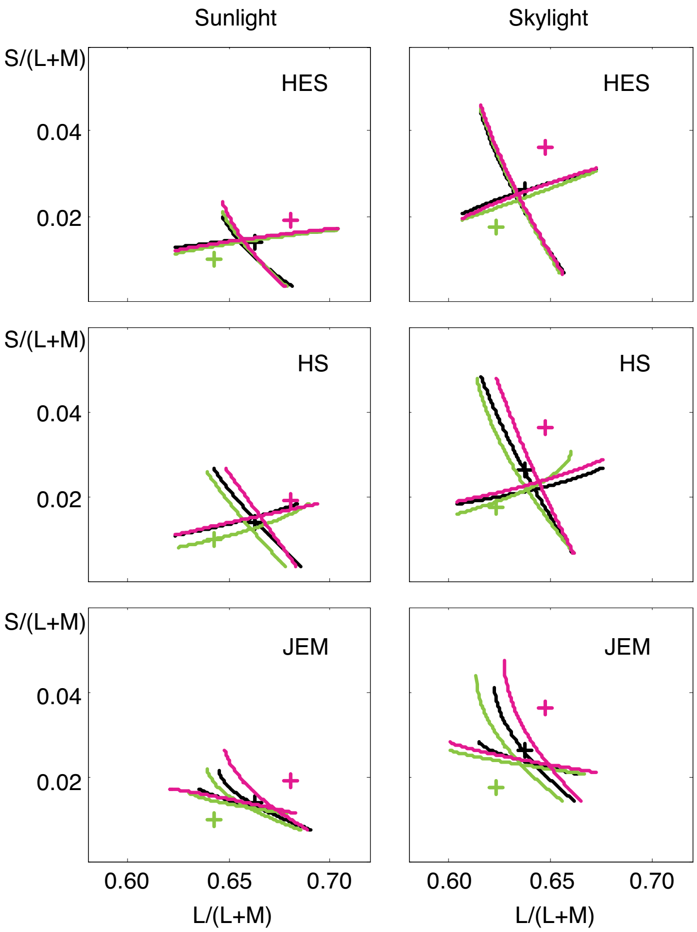

Figure 8. Data from Experiments 1 and 2 showing the effect of mean background chromaticity. Each panel shows classification boundaries obtained with balanced backgrounds (black lines), red-blue biased backgrounds (pink lines) and green-yellow biased backgrounds (lime-green lines). Plus-symbols show the mean chromaticities of the three backgrounds, and are colour-coded accordingly.

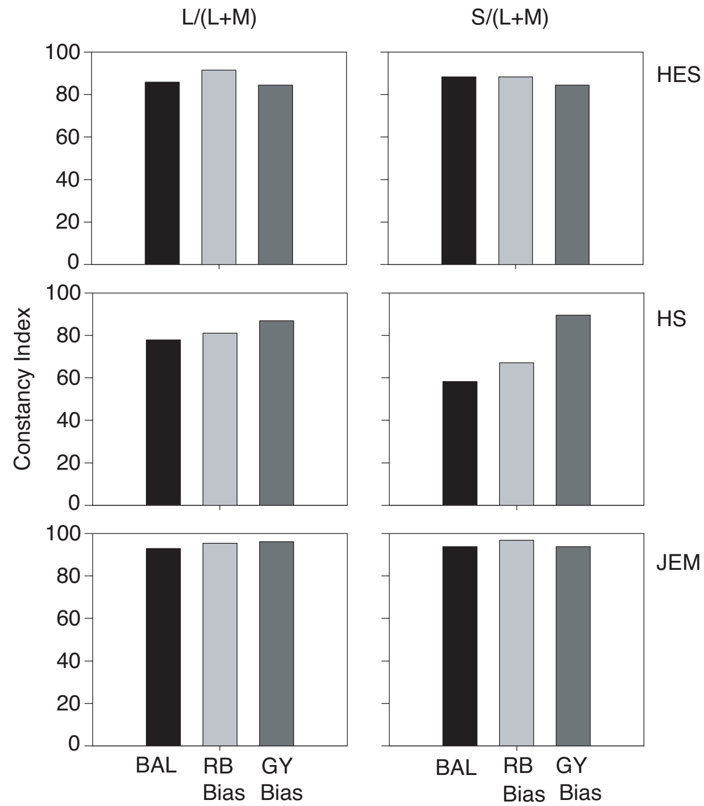

Figure 9. Constancy indices obtained in Experiment 2 (light-grey and mid-grey bars) plotted with indices from Experiment 1 (black bars) for comparison. Indices calculated with signals in the L/(L+M) (left panels) and S/(L+M) (right panels) chromatic mechanisms. See text for definition of constancy indices.

observers demonstrate a high level of appearance-based colour constancy, even across scenes that differ in mean reflectance. Performance in Experiment 2 cannot be explained by normalization to the mean chromaticity of the scene, since measured achromatic loci did not coincide with the mean chromaticities. Moreover, colour appearance is not set relative to the mean, since the shifts in achromatic loci were not comparable to the shifts in mean chromaticity.

It is well known that colour appearance can be influenced by neighbouring colours. One might ask why we observe such small chromatic contrast effects with our stimuli. Zaidi et al. (1992; 1999) showed that chromatic induction is reduced by the presence of high spatial-frequency chromatic variation in the inducing stimulus. To test whether this is also true for our conditions, we repeated Experiment 2 but replaced the variegated background with a uniform background of the same mean chromaticity. We additionally replaced the 3-degree square test-patch with a small test-annulus, with inner radius of 0.8 degrees and outer radius of 1.3 degrees (the same sized test-stimulus was used by Shevell, 1982 in studying chromatic adaptation to a background). With this spatial configuration of test and background we observed large changes in the location of the achromatic loci as a function of the background chromaticity. On the red-blue background, a higher number of test materials were classified as green (versus red) and as yellow (versus blue) compared to classifications obtained on the green-yellow background.

The data from Experiment 2 confirm that the colour appearance of our test-materials is not set by the mean chromaticity of the global scene. But there are other global scene statistics that could be used to disentangle the set of reflectances in the scene from the illumination. Golz & MacLeod (2002) have suggested that, rather than using the mean chromaticity, an illuminant estimation mechanism might make use of higher-order correlations, for example between redness and luminance within the image. Tominaga, Ebisui & Wandell (2001) showed that it is better to use just the brightest objects in the scene to estimate the illuminant, since darker surfaces may contribute more noise than signal to the estimate. We simulated such estimates by taking the average of the chromaticity of each material weighted by its luminance raised to various positive powers (Khang & Zaidi, 2004). The resultant illuminant estimates were considerably better than those obtained from a simple mean. To assess whether performance could be explained by any spatially extended illuminant estimate we performed Experiment 3.

Experiment 3

Introduction

In our third experiment we performed a critical manipulation. We simulated one illuminant for the test and a different illuminant for the background (Figure 10). Under these conditions, the spatial context provides information only about the background illuminant, so any *global* mechanism would estimate the wrong illuminant for the test, and constancy would be low. In a single trial, the observer has no information about the test-illuminant, since it falls only on a single material and there are no statistical cues to disentangle the material reflectance and the illuminant spectrum. Under this manipulation, information about the test illuminant is available only by collating information over successive trials. We ask whether the classification of test-materials in the inconsistent illuminant conditions will follow that predicted by the background illuminant, or that predicted by the test illuminant.

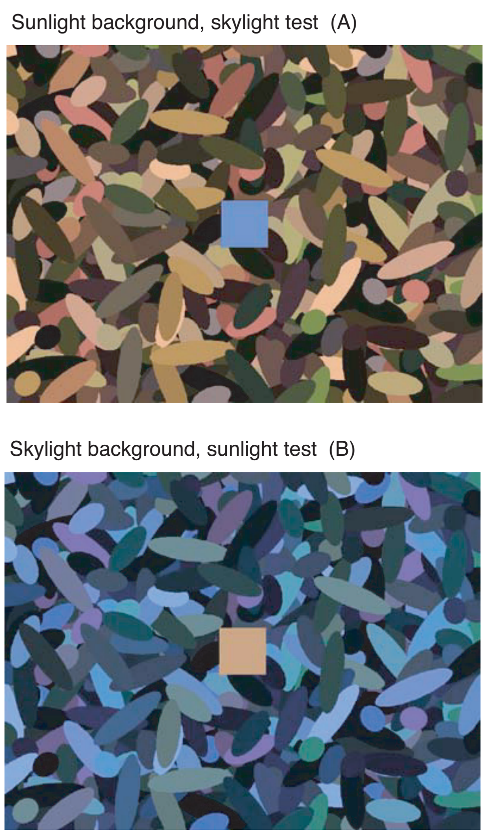

Figure 10. Examples of stimuli used in Experiments 3 and 4 with conflicting illuminants on test and background. Panel A shows the test material rendered under skylight and the background under sunlight; panel B shows the test material under sunlight and the background under skylight. Information about the test-illuminant is available only by collating information over successive trials.

Methods

Again, stimulus displays comprised a central square test-patch within a variegated background of elliptical patches. As for Experiment 1, we used the balanced set of background materials, and the standard set of test-materials (see Figure 1). But now, rather than using a global illuminant for the whole scene, we rendered either the test-material under sunlight and the background materials under skylight, or the test-material under skylight and the background materials under sunlight. Since the stimuli to be classified were held constant as surfaces, the locus of test-chromaticities shifted with the test-illuminant.

Results

Classification boundaries

The dotted lines in Figure 11 show classification boundaries re-plotted from Experiment 1. They are colour-coded according to the global illumination used: red for sunlight and blue for skylight. The solid lines show classification boundaries for the inconsistent illumination conditions of Experiment 3, and are colour coded according to the illuminant falling on the test-patch. So, red lines show performance with sunlight on the test and skylight on the background, and blue lines show performance with skylight on the test and sunlight on the background. For observer HS, for the condition with sunlight on the test and skylight on the background, the blue-yellow boundary fell at the extreme edge of our 3-dimensional cloud of test stimuli and therefore could not be estimated reliably. The locations of the boundaries obtained in Experiment 3 are slightly displaced relative to the locations of the boundaries obtained in Experiment 1, but performance in the inconsistent illumination conditions is more closely predicted by the illuminant falling on the test-patch than by the illuminant falling on the background.

Constancy indices

In Experiment 1, we calculated constancy indices for a global illuminant change. Here we calculate constancy indices for an illuminant change on the test-material only. We use the achromatic loci obtained in Experiments 1 and 3 to obtain two new estimates of constancy under a change in the illumination on the test patch from sunlight to skylight. One value describes performance in the presence of a sunlight-illuminated background (i.e. global sunlight compared with skylight test and sunlight background – Figure 1A compared with 10A) and one describes performance in the presence of a skylight-illuminated background (i.e. sunlight test and skylight background compared with global skylight – Figure 10B compared with 1B).

Constancy indices are plotted in Figure 12. Dotted lines show constancy indices achieved in Experiment 1,

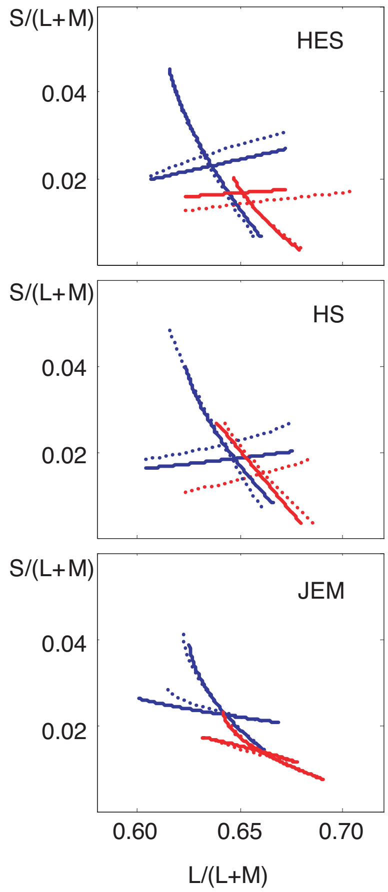

Figure 11. Data from Experiment 3, with different illuminants on the test and background. Solid red lines indicate classification boundaries obtained with sunlight illumination on the test, and skylight on the background. Solid blue lines indicate classification boundaries obtained with skylight on the test, and sunlight on the background. Dotted red and blue lines show boundaries obtained in Experiment 1 with a global illuminant of sunlight or skylight respectively, and are thus colour-coded to predict boundary locations based on the test illuminant. Discriminant functions for Observer HS, with sunlight on the test and skylight on the background, were out of range.

under a global illuminant change. This indicates the maximum level of performance under our experimental conditions when both the spatial and temporal contexts give appropriate cues to the illuminant. Light grey and dark grey bars show constancy indices obtained under a change in the illuminant on the test, i.e. a change in the test-illuminant coordinates from b_1 to b_2 in Equations 1 and 2. In this analysis, a change in the illuminant on the test is not accompanied by a corresponding change in the illuminant on the background, so the spatial context provides no cues to the illuminant change, signalling instead either steady sunlight (light bars) or skylight (dark bars). So, if performance were determined by the spatial context, the coordinates of the achromatic point (a_1 and a_2 in Equations 1 and 2) should be identical, and we should measure constancy indices equal to zero. If however performance were

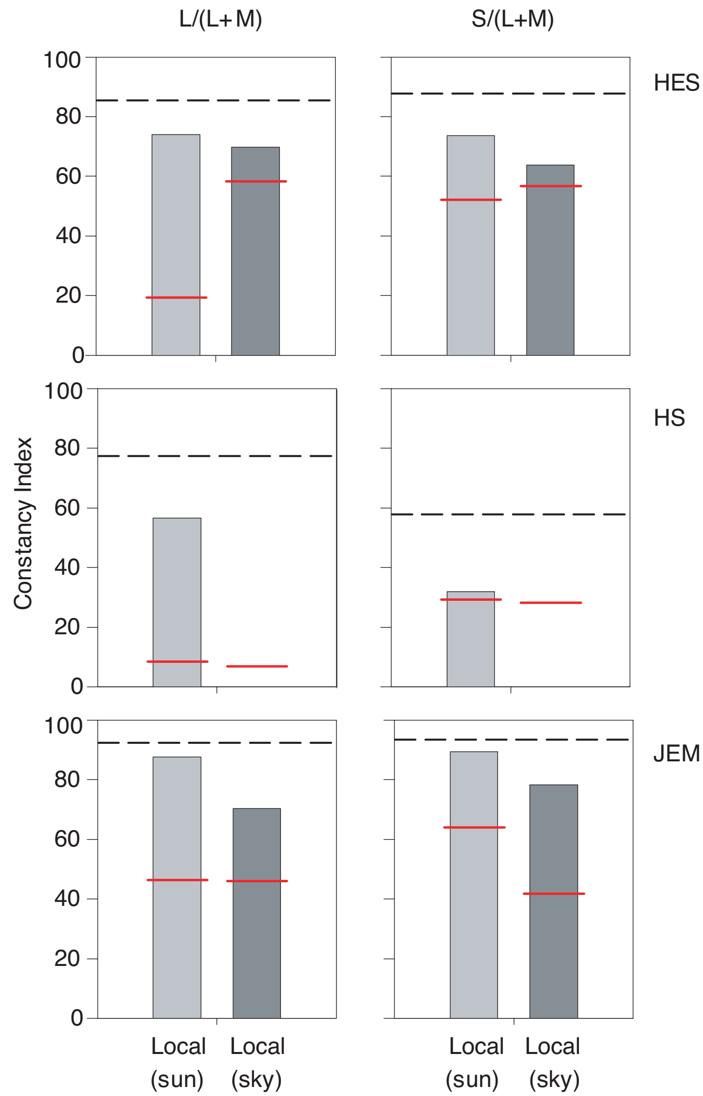

Figure 12. Constancy indices obtained in Experiment 3, for a change from sunlight to skylight on the test, with either steady sunlight (light-grey bars) or steady skylight (mid-grey bars) on the background. Indices calculated with signals in the L/(L+M) (left panels) and S/(L+M) (right panels) chromatic mechanisms. Dotted black lines show constancy indices obtained in Experiment One, and therefore indicate the maximum level we would expect. Red lines show the minimum constancy indices we could measure given our fixed sets of test-materials. For Observer HS, constancy indices with skylight on the background were out-of-range and fall below the red lines. See text for definition of constancy indices.

determined by the test illuminant, constancy indices should approach one (or rather the value obtained in Experiment 1 for a global illuminant change). Since we used the same set of test-materials for all conditions, the achromatic point predicted by the background illuminant was generally outside the range of boundaries we could measure. Correspondingly, the short red lines in Figure 12 indicate the lower limit of constancy indices that we could measure. For the three conditions where constancy indices are not plotted, we can say only that they fall somewhere below the red lines. Constancy indices obtained in Experiment 3 are in all cases slightly lower than those obtained in Experiment 1, but constancy is far from abolished.

Discussion

Observers' performances in Experiment 3 cannot be wholly explained by any spatially extended process, since the spatial context provided cues to the wrong illuminant. In a single trial of Experiment 3, observers had no information about the test illuminant, and yet they still achieved fairly high levels of constancy. Information about the test illuminant was available only by collating local information over successive trials. As each new test-material was presented it provided the observer with a reflected sample of the illuminant, and the statistical properties of successive samples, collated over time, could in principle be used to disentangle the properties of the test-reflectances and the properties of the illuminant.

The type of neural mechanism that could perform this temporal collation process could be peripheral or central. A process of temporal adaptation, with long time-constants of the order of a few seconds, would converge on the mean chromaticity of the test-materials. Since the mean chromaticity of our test materials was balanced, this would be sufficient to support reasonable constancy. So observers' performance in Experiment 3 could be accounted for, to a large extent, by spatially localised adaptation, which may even be retinal.

However, as noted above, mean chromaticity is not a perfect cue to the illuminant, especially if we allow that the set of materials may change. In the main experiments reported here, we always used the same set of material-reflectances for the test-patches. However, ancillary sessions, in which the sets of materials used for the biased-backgrounds were used for the test-patches, suggested that category boundaries are not well predicted by the temporal mean of the set of chromaticities. In fact, in the red-green classification task, observers classified all but one or two materials from the red-blue biased set as red, and all but one or two materials from the green-yellow biased set as green (see Figure 6E & F for the chromatic loci of the biased sets under the two illuminants). So, it is possible that the visual system is able to use other cues to disentangle a change in the set of temporally distributed reflectances from a change in the illuminant.

Our final experiment was designed to determine whether the neural processes that collate information about a temporally extended sample of illuminated spectral reflectances are automatic and adaptational, or whether they are based on perceptual processes that are selective for information about the test stimuli.

Experiment 4

Introduction

As in Experiment 3, we used conflicting illuminants for test and background (Figure 10) but now reduced the amount of time observers were exposed to the test-patches, relative to the amount of time they were exposed to the background. Any automatic adaptation process must collate information indiscriminately from test and background. If colour constancy is achieved for patches presented under the test illuminant it must be based on a selective mechanism that collates only the test samples.

Methods

Stimuli were as for Experiment 3. We used the balanced set of background materials, the standard set of test-materials, and different illuminants for test and background. As for all experiments reported here, trial duration was fixed at 1500 msec. But now the test-materials were presented only for 200 msec, after which the display reverted to the background pattern. Exposure to materials illuminated by the test illuminant was thus reduced to a fraction (0.13) of the duration of exposure to materials illuminated by the background illuminant, and successive test-presentations were separated in time.

Results

Classification boundaries

The format of Figure 13 is directly analogous to that of Figure 11. Classification boundaries obtained in Experiment 1 under global illumination are re-plotted with dotted lines. Classification boundaries obtained in Experiment 4, under inconsistent illumination and with brief test presentations, are plotted with solid lines. Colour coding relates to the test illuminant (red lines indicate sunlight conditions and blue lines indicate skylight conditions).

Several classification boundaries were not well constrained by our data, and are not plotted. In general, these boundaries had shifted towards the location of boundaries predicted by the background illuminant. The three observers were differently influenced by the reduction in exposure to the test illuminant. For observers HES and HS, the red-green boundaries are consistent with the background illuminant. However, for HS and JEM, the blue-yellow boundaries obtained with skylight on the test and sunlight on the background are well predicted by the test illuminant.

Constancy indices

Figure 14 shows constancy indices for a change in test illumination from sunlight to skylight. Dotted lines show constancy indices from Experiment 1, with a global illuminant change. Light grey bars show indices obtained with

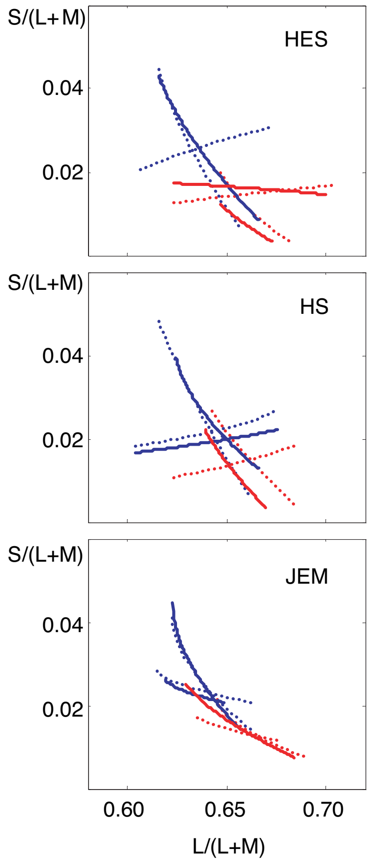

Figure 13. Data from Experiment 4, with different illuminants on the test and background and brief test presentation. Solid red lines indicate classification boundaries obtained with sunlight illumination on the test, and skylight on the background. Solid blue lines indicate classification boundaries obtained with skylight on the test, and sunlight on the background. Dotted red and blue lines show boundaries obtained in Experiment 1 with a global illuminant of sunlight or skylight respectively, and are thus colour-coded to predict boundary locations based on the test illuminant. Boundaries are missing where discriminant functions were out-of-range.

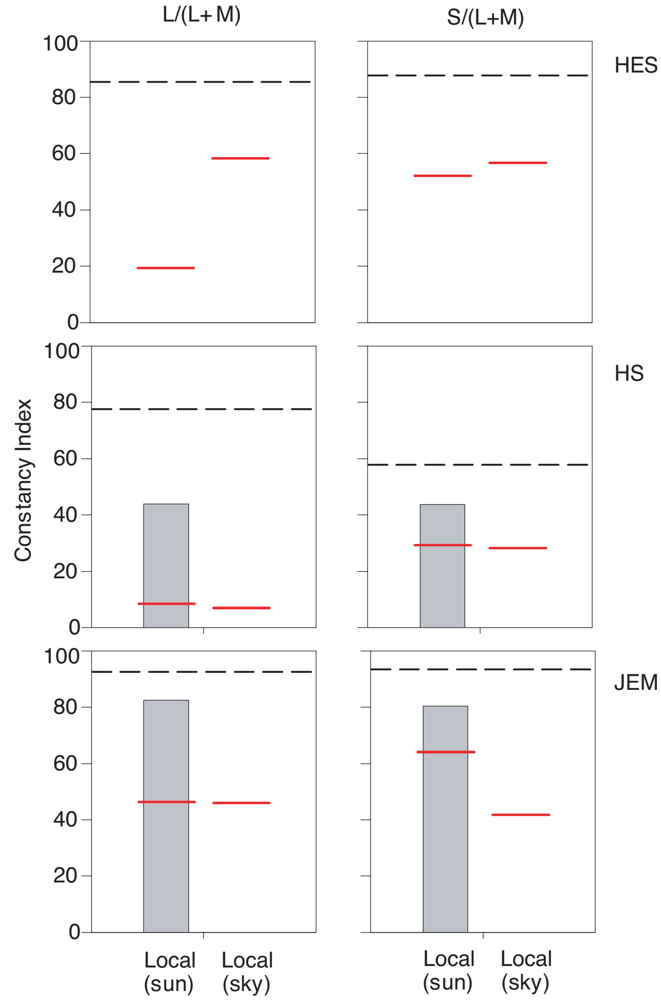

Figure 14. Constancy indices obtained in Experiment 4, for a change from sunlight to skylight on the test, with either steady sunlight (light-grey bars) or steady skylight (mid-grey bars) on the background. Indices calculated with signals in the L/(L+M) (left panels) and S/(L+M) (right panels) chromatic mechanisms. Dotted black lines show constancy indices obtained in Experiment 1, and therefore indicate the maximum level we would expect. Red lines show the minimum constancy index we could measure given our fixed sets of test-materials. Constancy indices with skylight on the background were out-of-range for all observers, and all constancy indices were out-of-range for observer HES. Constancy indices for these conditions lie below the red lines. See text for definition of constancy indices.

brief test-presentations interspersed amongst presentations of sunlight-illuminated background patterns. Values are missing for conditions in which our data did not allow reliable estimation of the achromatic points (see Figure 13). For these conditions we assume the achromatic points lie outside the loci of test-materials and that constancy indices lie below the short red lines in Figure 14. Constancy is below the measurable range for HES, but remains reasonably high in at least one condition for HS and JEM.

Discussion

The mixed performance in Experiment 4 cannot be completely explained by an automatic neural process that acts upon incoming chromatic signals to discount the illuminant. Observers were exposed to the test illuminant for 13% of the duration of each trial, and to the background illuminant for the remaining 87%. In addition, the spatial context provided cues only to the background illuminant. Under these conditions the output of any automatic process is predicted to be inescapably dominated by the properties of the background illuminant. While this does seem to be the case in some conditions, in other conditions, both HS and JEM make colour appearance judgements that are dominated by the test-illuminant. Classification boundaries in these conditions shift towards the boundaries predicted by the background illuminant, but constancy indices confirm that this shift is small and not in proportion to the relative exposure times to the two illuminants.

So in some conditions, observers' judgements are consistent with the hypothesis that the test-materials are illuminated by a different illuminant from the background. In a single trial there is no information about the test illuminant (since this falls only on a single material) so information about the properties of the test illuminant can only be obtained by collating information from successive trials. Moreover, successive presentations of materials illuminated by the test-illuminant are interspersed with presentations of materials illuminated by the background illuminant. To collate the properties of the test illuminant separately from the properties of the background illuminant requires a process that tracks the chromatic statistics of the task-relevant test-squares. Such a process could be a mechanism of adaptation gated by attention, or it may be a perceptual "level of reference" or "anchoring" mechanism (Rogers, 1941; Helson, 1947) that segregates test and background presentations. Several cues distinguish test from background. The most obvious is that the test-squares require a judgement while the background ellipses do not. A more subtle cue is highlighted in Forsyth's constancy algorithm (Forsyth, 1990). Since the illuminant limits the gamut of spectra reaching the eye, it is possible that (due to the conflicting illuminant) the colours of the test were very unlikely under the background illuminant, and this might provide a cue for the visual system to estimate the illuminant separately for the test-patches.

General Discussion

Our method allowed us to accurately locate perceptual colour boundaries under different conditions of observing. By relying on internal criteria we were able to assess colour constancy in scenes exposed to only one illuminant. Psychometric functions, obtained from repeated classification of the colour appearance of rendered test-materials, were steep and reproducible. Although we used only the achromatic point in deriving our constancy indices, the boundary plots provide a graphical representation of the transformations of the two perceptually significant colour axes (unique green to unique red, and unique yellow to unique blue). We should be cautious however in interpreting changes in the boundary locations at the extremes of the range, since these may be influenced by the properties of the curve fit. We also cannot draw conclusions about the transformations of off-axis colours.

In this study, we chose not to introduce observers to the concepts of materials and illuminants. Instead, observers were simply shown coloured test-patches, set within a variegated background of coloured ellipses, and were asked to judge the appearance of the test-patch. The two-dimensional, rendered images in our experiments, unlike three-dimensional, real-world scenes, did not contain image features that signal illuminants or objects, nor did they contain direct cues to the illuminant colour (e.g. specular highlights, shadows, and mutual reflections), but they did contain a rich sample of surface reflectances. In Experiments 1 and 2, observers demonstrated high levels of phenomenological colour constancy.

The majority of experiments on colour constancy have focused on the spatial information available in a scene. While it is generally assumed that observers are able to collate cues to the illuminant over space, it is not asked whether observers are able to collate the same information over time. In Experiment 3, we pitted temporal and spatial cues to the illuminant against one another. Our critical manipulation was to use different illuminants for the test-patch and for the variegated background. In a single trial, the observer had no information about the test-illuminant. This information could only be gathered by collating cues from successive test-presentations. Under the conditions of our main experiments, the illuminant used for the test-patch had a stronger effect on colour appearance than the illuminant used for the background i.e. for our, relatively large and well segregated, 3-degree test-patch, performance was best predicted by the temporally extended sample of materials presented under the test-illuminant, than by the spatially extended sample of materials presented under the background illuminant. The primary message from this study is that the stability of colour appearance can be set by local mechanisms that collate information over time. Observers in our experiments were given no explicit instructions regarding fixation, but it is likely that they were deterred from making eye-movements since the response period was limited and the test patch was always presented in the same position.

The phenomenon of colour constancy, and any attempt to study it, is inherently linked to the way in which we make judgements. In psychology, it has long been accepted that some sort of "level of reference" underlies all judgements, from aesthetic and social judgements to judgements of the perceptual qualities of objects. Helson (1947; 1964) attempted to formalise level of reference for prediction of psychophysical data. He states, "Fundamental to the theory is the assumption that effects of stimulation form a spatio-temporal configuration in which order prevails. For every excitation-response configuration there is assumed to be a stimulus which represents the pooled effect of all stimuli and to which the organism may be said to be attuned or adapted." The important point for the current experiments is that, while Helson's level of reference may be influenced by peripheral mechanisms, it is stored centrally and used as a reference point for all judgements.

The colour-appearance judgements obtained in Experiment 3 would be predicted equally well by mechanisms that are selective for the test and by non-selective, automatic mechanisms that are spatially local in extent. We therefore cannot be sure from this experiment whether temporal context acts centrally to modify the observer's level of reference, or whether the information reaching the decision-stage is modified automatically by peripheral mechanisms. Performance in Experiment 4, however, is complicated to explain. In certain conditions and for certain observers, the drastic decrease in constancy from Experiment 3 to Experiment 4 is consistent with an automatic adaptation mechanism with a long time constant. However, in other conditions there is little decrease in constancy, indicating the use of a central value that stores contextual information about the stimuli requiring judgement.

A sceptic might argue that observers were simply equating the number of "red" responses and "green" responses, but we have evidence that this is not the case. For we obtained data from each observer under conditions in which they were prepared to classify 95% to 100% of presentations into one of the two categories (for example in the auxiliary experiments with uniform backgrounds and small annular test-stimuli, or with biased sets of test-materials).

The set of materials to be classified has a substantial effect on the location of classification boundaries in chromaticity space. However, the way in which this effect depends on the statistical properties of the set of stimuli to be classified is at present uncertain and needs to be addressed in further parametric studies. Data from Experiments 3 and 4 suggest that information collated over time is important in maintaining colour constancy, and that both central and peripheral mechanisms can contribute to achieving this collation.

In the Introduction we identified three necessary components of a model of colour constancy: what is the nature of the required transformation, how are the parameters of the transformation set by the scene, and at what neural stage is the transformation performed? The invariance of cone-excitation ratios under an illuminant change suggests that the likely key to models of colour constancy is multiplicative scaling of cone-signals, or an operation at a later stage that achieves the same effect on transformed cone-signals. Many elements of the visual environment have been proposed as important in setting the level of this normalization factor. Data from the present study suggest that the recent history of reflectances sampled by the observer is a primary contributor. The history of reflectances might be taken from successive presentations (as in this study), or from successive fixations within a steady image. Spatially distributed cues to the illuminant are carefully specified in studies and models of colour constancy. The data presented here highlight the importance of also specifying temporally distributed cues. Our final point is that, in addition to local, automatic, adaptation mechanisms, central mechanisms are important in tracking chromatic context. The stability of the visual world is a complex phenomenon, which may in part be task-dependent, and judgements of the colour-appearance of objects are not made in isolation.

Acknowledgements

We would like to thank our observers Hao Sun and James Mead. This work was supported by National Eye Institute Grant EY07556 to Qasim Zaidi.

Commercial relationships: none.
Corresponding author: Hannah Smithson.
Email: h.smithson@ucl.ac.uk.
Address: Institute of Ophthalmology, University College London, 11-43 Bath Street, London, EC1V 9EL, UK.

References

Adelson, E. H., & Pentland, A. P. (1996). The perception of shading and reflectance. In D. C. Knill & W. Richards (Eds.), *Perception as Baysian inference* (pp. 409-423). Cambridge: Cambridge University Press.

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

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

Brainard, D. H., & Wandell, B. A. (1992). Asymmetric color matching: How color appearance depends on the illuminant. *Journal of the Optical Society of America A, 9*(9), 1433-1448. [PubMed]

Brown, R. O. (1994). The world is not gray [Abstract]. *Investigative Ophthalmology & Visual Science, 35*(4), 2165.

Brown, R. O., & MacLeod, D. I. (1997). Color appearance depends on the variance of surround colors. *Current Biology, 7*(11), 844-849. [PubMed]

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

Chichilnisky, E. J., & Wandell, B. A. (1999). Trichromatic opponent color classification. *Vision Research, 39*(20), 3444-3458. [PubMed]

Chittka, L., Shmida, A., Troje, N., & Menzel, R. (1994). Ultraviolet as a component of flower reflections, and the colour perception of Hymenoptera. *Vision Research, 34*(11), 1489-1508. [PubMed]

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

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

D'Zmura, M., & Lennie, P. (1986). Mechanisms of color constancy. *Journal of the Optical Society of America A, 3*(10), 1662-1672. [PubMed]

Fairchild, M. D., & Lennie, P. (1992). Chromatic adaptation to natural and incandescent illuminants. *Vision Research, 32*(11), 2077-2085. [PubMed]

Forsyth, D. (1990). A novel algorithm for color constancy. *International Journal of Computer Vision, 30*, 5-36.

Foster, D. H. (2003). Does colour constancy exist? *Trends in Cognitive Sciences*, *7*(10), 439-443. [PubMed]

Foster, D. H., & Nascimento, S. M. (1994). Relational colour constancy from invariant cone-excitation ratios. *Proceedings of the Royal Society of London B Biological Sciences, 257*(1349), 115-121. [PubMed]

Golz, J., & MacLeod, D. I. (2002). Influence of scene statistics on colour constancy. *Nature, 415*(6872), 637-640. [PubMed]

Hagedorn, J., & D'Zmura, M. (2000). Color appearance of surfaces viewed through fog. *Perception, 29*(10), 1169-1184. [PubMed]

He, S., & MacLeod, D. I. (1998). Local nonlinearity in S-cones and their estimated light-collecting apertures. *Vision Research, 38*(7), 1001-1006. [PubMed]

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

Helson, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *Journal of the Optical Society of America, 60*, 1-29.

Helson, H. (1964). *Adaptation-level theory; an experimental and systematic approach to behavior*. New York: Harper & Row.

Helson, H., Judd, D. B., & Warren, M. H. (1952). Object-color changes from daylight to incandescent filament illumination. *Illuminating Engineering, 47*, 221-223.

Hiltunen, J. (1996). *Munsell colors matt (spectrophotometer measurements by Hiltunen)*. Retrieved September 10, 1999, from http://www.it.lut.fi/ip/research/color/database/download.html#munsell_spec_matt

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

Khang, B. G., & Zaidi, Q. (2002). Cues and strategies for color constancy: Perceptual scission, image junctions and transformational color matching. *Vision Research, 42*(2), 211-226. [PubMed]

Khang, B. G., & Zaidi, Q. (2004). Illuminant color perception of spectrally filtered spotlights. *Journal of Vision, 4*(9), 680-692, http://journalofvision.org/4/9/2/, doi:10.1167/4.9.2. [PubMed][Article]

Kraft, J. M., & Brainard, D. H. (1999). Mechanisms of color constancy under nearly natural viewing. *Proceedings of the National Academy of Sciences USA, 96*(1), 307-312. [PubMed][Article]

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 USA, 80*(16), 5163-5169. [PubMed][Article]

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

Lee, H. C. (1986). Method for computing the scene-illuminant chromaticity from specular highlights. *Journal of the Optical Society of America A, 3*(10), 1694-1699. [PubMed]

Lehmann, T. M., & Palm, C. (2001). Color line search for illuminant estimation in real-world scenes. *Journal of the Optical Society of America A, 18*(11), 2679-2691. [PubMed]

MacLeod, D. I., & Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America, 69*(8), 1183-1186. [PubMed]

MacLeod, D. I., & He, S. (1993). Visible flicker from invisible patterns. *Nature, 361*(6409), 256-258. [PubMed]

MacLeod, D. I., Williams, D. R., & Makous, W. (1992). A visual nonlinearity fed by single cones. *Vision Research, 32*(2), 347-363. [PubMed]

Marshall, N. J. (2000). Communication and camouflage with the same 'bright' colours in reef fishes. *Philosophical Transactions of the Royal Society of London B Biological Sciences, 355*(1401), 1243-1248. [PubMed]

Nascimento, S. M., & Foster, D. H. (1997). Detecting natural changes of cone-excitation ratios in simple and complex coloured images. *Proceedings of the Royal Society of London B Biological Sciences, 264*(1386), 1395-1402. [PubMed]

Rogers, S. (1941). The anchoring of absolute judgments. *Archives of Psychology Columbia University, 261*, 42.

Shevell, S. K. (1982). Color perception under chromatic adaptation: Equilibrium yellow and long-wavelength adaptation. *Vision Research, 22*(2), 279-292. [PubMed]

Smith, V. C., & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research, 15*(2), 161-171. [PubMed]

Stiles, W. S. (1961). Adaptation, chromatic adaptation, colour transformation. *Anales De La Real Sociedad Espanola De Fisica Y Quimica: Seria A, 57*, 149-175.

Taylor, A. H., & Kerr, G. P. (1941). The distribution of energy in the visible spectrum of daylight. *Journal of the Optical Society of America, 31*(1), 3-8.

Thomas, D. R., & Jones, C. G. (1962). Stimulus generalization as a function of the frame of reference. *Journal of Experimental Psychology, 64*(1), 77-80. [PubMed]

Tominaga, S., Ebisui, S., & Wandell, B. A. (2001). Scene illuminant classification: Brighter is better. *Journal of the Optical Society of America A, 18*(1), 55-64. [PubMed]

Vimal, R. L., Pokorny, J., & Smith, V. C. (1987). Appearance of steadily viewed lights. *Vision Research, 27*(8), 1309-1318. [PubMed]

von Kries, J. (1878). Physiology of visual sensations. In D. L. MacAdam (Ed.), *Sources of Color Science* (pp. 101-108). Cambridge, MA: MIT Press.

von Kries, J. (1905). Die Gesichtsempfindungen. In W. Nagel (Ed.), *Handbuch der Physiologie des Menschen* (Vol. 3) Physiologie der Sinne. Braunschweig: Vieweg und Sohn.

Vrhel, M., Gershon, R., & Iwan, L. S. (1994). Measurement and analysis of object reflectance spectra. *Color Research and Application, 19*, 4-9.

Webster, M. A. (1996). Human colour perception and its adaptation. *Network: Computation in Neural Systems, 7*(4), 587-634.

Webster, M. A., Malkoc, G., Bilson, A. C., & Webster, S. M. (2002). Color contrast and contextual influences on color appearance. *Journal of Vision, 2*(6), 505-519, http://journalofvision.org/2/6/7/, doi:10.1167/2.6.7. [PubMed][Article]

Webster, M. A., & Mollon, J. D. (1997). Adaptation and the color statistics of natural images. *Vision Research, 37*(23), 3283-3298. [PubMed]

Yang, J. N., & Maloney, L. T. (2001). Illuminant cues in surface color perception: Tests of three candidate cues. Vision Research, *41*(20), 2581-2600. [PubMed]

Yang, J. N., & Shevell, S. K. (2002). Stereo disparity improves color constancy. *Vision Research, 42*(16), 1979-1989. [PubMed]

Zaidi, Q. (1998). Identification of illuminant and object colors: Heuristic-based algorithms. *Journal of the Optical Society of America A, 15*(7), 1767-1776. [PubMed]

Zaidi, Q. (1999). Color and brightness induction: From Mach bands to 3-D configurations. In K. Gegenfurtner & L. Sharpe (Eds.), *Color Vision: From Genes to Perception*. Cambridge: Cambridge University Press.

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, 14*(10), 2608-2621. [PubMed]

Zaidi, Q., Yoshimi, B., Flanigan, N., & Canova, A. (1992). Lateral interactions within color mechanisms in simultaneous induced contrast. *Vision Research, 32*(9), 1695-1707. [PubMed]