THE ROYAL
SOCIETY

Received 8 July 2004
Accepted 5 August 2004

Colour constancy under simultaneous changes in surface position and illuminant

Kinjiro Amano and David H. Foster*

Visual and Computational Neuroscience Group, University of Manchester Institute of Science and Technology, Manchester M60 1QD, UK

Two kinds of constancy underlie the everyday perception of surface colour: constancy under changes in illuminant and constancy under changes in surface position. Classically, these two constancies seem to place conflicting demands on the visual system: to both take into account the region surrounding a surface and also discount it. It is shown here, however, that the ability of observers to make surface-colour matches across simultaneous changes in test-surface position and illuminant in computer-generated 'Mondrian' patterns is almost as good as across changes in illuminant alone. Performance was no poorer when the surfaces surrounding the test surface were permuted, or when information from a potential comparison surface, the one with the highest luminance, was suppressed. Computer simulations of cone-photoreceptor activity showed that a reliable cue for making surface-colour matches in all experimental conditions was provided by the ratios of cone excitations between the test surfaces and a spatial average over the whole pattern.

Keywords: vision; colour perception; colour constancy; spatial cone-excitation ratios; colour contrast; grey world

1. INTRODUCTION

Imagine looking at a coloured object somewhere in a room and then looking at the same object somewhere else in the room where the lighting is different. The colour of the object generally appears the same. Yet the spectrum of the reflected light reaching the eye is different and the immediate surroundings of the object, which might be used to compensate for the change in spectrum of the illumination, are also different.

This everyday visual task involves two kinds of perceptual colour constancy that in the laboratory are normally considered independently of each other: illuminant colour constancy, which describes the invariance of perceived surface colour under changes in illuminant, and positional colour constancy, which describes the invariance of perceived surface colour under changes in surface position (Young 1807; von Helmholtz 1867). In principle, these two constancies seem to entail a paradox, first identified in studies of brightness matching (Whittle & Challands 1969). On the one hand, illuminant colour constancy requires the chromatic context of the surface to be taken into account (Shapley 1986; Kraft & Brainard 1999); otherwise, it would be impossible to separate the spectral properties of the illuminant from those of the surface. On the other hand, positional colour constancy requires the surround to be discounted; otherwise, perceived surface colour would be an accident of location.

The task of judging surface colour under simultaneous changes in surface position and illuminant should therefore represent a significant challenge for the visual system, ostensibly requiring different modes of processing to achieve different perceptual goals (Whittle & Challands 1969). The aim of the present experiments was to measure observers' performance in such a task, which required them to match test surfaces embedded in pairs of computer-generated 'Mondrian' patterns viewed under different illuminants. These simulated patterns were used in preference to real tableaux of physical objects (e.g. Kraft & Brainard 1999; de Almeida *et al.* 2004) so that their spatial and spectral properties could be easily randomized, and to avoid uncontrolled or irrelevant stimulus cues such as mutual reflections (Bloj *et al.* 1999) and specularities (Yang & Maloney 2001) and, more generally, any effects of past experience with specific objects. Surprisingly, it was found that the goodness of the matches was about the same independent of the displacement of the test surfaces within the patterns, and, furthermore, independent of the permutation of all the other surfaces in the patterns.

One possible explanation of this position-invariant performance is that a test surface, rather than being compared with its local chromatic context, is instead compared with a particular surface in the scene whose position is generally stable under illuminant changes, such as the surface having the highest luminance. To test this explanation, performance was measured when information from this surface was suppressed. The degree of constancy remained about the same whether or not the test surfaces were displaced.

In general, surface-colour matching under simple changes of illuminant can be interpreted in terms of signals such as the spatial ratios of cone excitations produced by light reflected from the test and surround surfaces within each pattern (Foster 2003). The particular variation in the stability of these ratios as the number of surround surfaces increases suggests that making matches under simultaneous changes in illuminant and surface position depends on the large-scale properties of scenes such as space-average colour.

---

* Author for correspondence (d.h.foster@manchester.ac.uk).

Proc. R. Soc. Lond. B (2004) **271**, 2319–2326
doi:10.1098/rspb.2004.2884

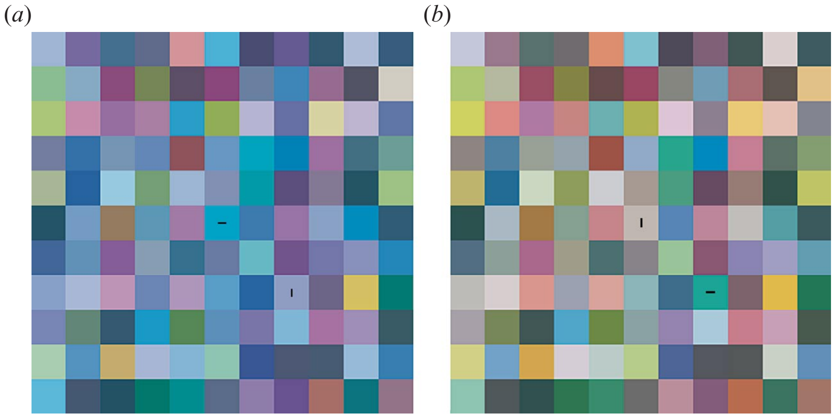

Figure 1. Stimulus with test surfaces transposed. Each pattern consisted of an array of randomly selected, uniformly coloured surfaces, each surface subtending $1^\circ \times 1^\circ$ visual angle, illuminated (*a*) by daylight of correlated colour temperature 25 000 K and (*b*) by daylight of correlated colour temperature 6700 K. The two test surfaces, marked with small horizontal and vertical black bars, were transposed in (*b*). Observers had to match the test surfaces across the two patterns.

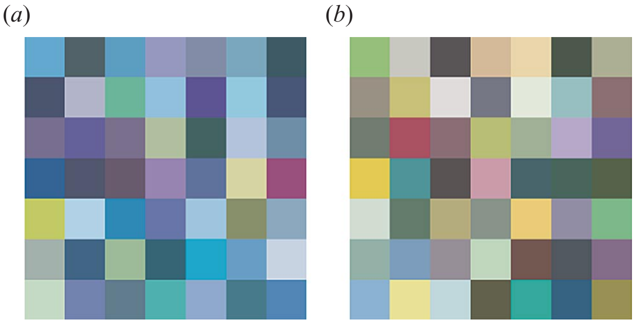

Figure 2. Stimulus with surround surfaces permuted. The patterns, which were similar to those in figure 1, were illuminated: (*a*) by daylight of correlated colour temperature 25 000 K and (*b*) by daylight of correlated colour temperature 6700 K, except that the position of the single test surface was fixed at the centre and the remaining surfaces were randomly permuted in (*b*).

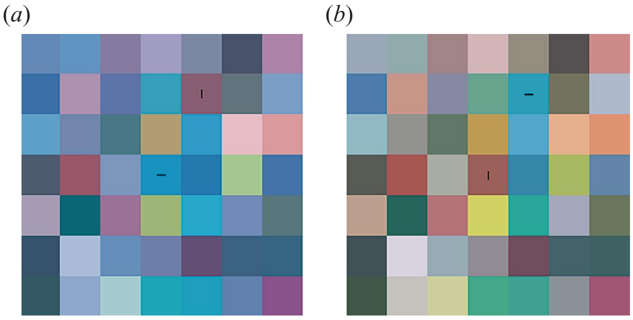

Figure 3. Stimulus with test surfaces transposed and altered highest-luminance surface. The patterns, which were similar to those in figure 1, were illuminated (*a*) by daylight of correlated colour temperature 25 000 K and (*b*) by daylight of correlated colour temperature 6700 K, except that the luminance of the highest-luminance surface in (*a*) (coordinates (6,4) from bottom left) was replaced by the luminance of another surface (coordinates (4,3)) in (*b*).

2. METHODS

(a) *Stimuli*

The stimuli were generated on a computer-controlled colour monitor. They were simulations of pairs of illuminated square coloured patterns, of side 11° or 7° visual angle, viewed simultaneously side-by-side in a dark surround with 1° gap between them, as illustrated in figure 1. Each pattern consisted of an array of 121 (11×11) or 49 (7×7) uniform, square, Lambertian coloured surfaces, of side 1°, drawn randomly from 1269 samples in the *Munsell book of color* (Munsell Color Corporation 1976). The random sampling producing each pattern was repeated, if necessary, to eliminate any accidental similarities between the illuminated test surfaces (patches) and the surrounding surfaces (Maloney 1999; Foster *et al.* 2001*a*).

There were two test surfaces in each pattern: one was the centre surface and the other was chosen randomly from the remaining surfaces in the pattern excluding those at the edge. The left-hand pattern was presented under a fixed spatially uniform daylight of correlated colour temperature of 25 000 K and luminance $50\ \mathrm{cd\,m^{-2}}$. The right-hand pattern was identical but presented under a fixed spatially uniform daylight of correlated colour temperature of 6700 K and luminance $50\ \mathrm{cd\,m^{-2}}$, except for the two test surfaces where the 6700 K daylight was replaced by a common, adjustable, spatially uniform local illuminant constructed from three daylight spectral basis functions (Judd *et al.* 1964), independent of the global illuminant. By varying the coefficients of these functions with a joy-pad input control to the computer, the observer could vary the chromaticity and luminance of the local illuminant, and therefore simultaneously that of the two test patches (see Foster *et al.* 2001*a*). For example, in figure 1, the brownish grey and greenish target patches in the right-hand pattern (marked by small vertical and horizontal black bars) could be made simultaneously bluer by about the same amount, or redder, and so on. This device of varying the colour of the local illuminant rather than the colour of the surfaces directly has certain technical advantages (Foster *et al.* 2001*a*). No observer reported difficulty in making these matches. The variation of the local illuminant was unconstrained within a convex region of the Commission Internationale de l'Eclairage (CIE) 1931 (x, y) chromaticity diagram $(0.250 \leqslant x \leqslant 0.400;\ 0.230 \leqslant y \leqslant 0.400)$ containing the coordinates (0.250, 0.255) and (0.310, 0.326) of the fixed 25 000 K and 6700 K illuminants. These two daylights are, respectively, typical of the north sky and of the sun and total sky (Wyszecki & Stiles 1982).

Two test surfaces were used, for displacing just one would leave a gap in the pattern (an experimental control with a single, fixed, test surface is described later). With the larger number of surfaces

Table 1. Judging surface colour across an illuminant change with and without simultaneous transposition of test surfaces (figure 1) in Mondrian patterns.
(Means (s.e.m.s) of colour-constancy indices were calculated over groups of 11 and 8 observers viewing, respectively, new and recycled samples of patterns of side either 11° or 7°. An index of unity corresponds to perfect constancy.)

		test-surface transposition	no transposition	difference
new Mondrians	121 surfaces	0.70 (0.04)	0.74 (0.05)	−0.041
	49 surfaces	0.68 (0.04)	0.71 (0.03)	−0.036
recycled Mondrians	49 surfaces	0.71 (0.05)	0.76 (0.05)	−0.046

of the 11° patterns, the test surfaces were initially darkened in each trial, in addition to being marked by two steady, small horizontal and vertical black bars (subtending 0.1° × 0.025°). The 11° and 7° patterns were viewed binocularly at 90 cm and 100 cm, respectively. By making measurements with two sizes of patterns, one a proper subset of the other, it was possible to test whether performance was constrained by the maximum displacement of the test surfaces and the gamut of colours available in each pattern.

(b) *Apparatus*

Stimuli were generated with an RGB colour-graphics system with nominal 15-bit intensity resolution on each gun (VSG 2/3F, Cambridge Research Systems Ltd, Rochester, Kent, UK), controlled by a laboratory computer and displayed on a 20-inch RGB monitor (GDM-20SE2T5, Sony Corporation, Tokyo, Japan). The screen resolution was 1024 pixels × 768 pixels. The screen refresh rate was *ca.* 100 Hz. A telespectroradiometer (SpectraColorimeter, PR-650; Photo Research Inc., Chatsworth, CA, USA) that had previously been calibrated by the National Physical Laboratory was used to calibrate the display system. Errors in the displayed CIE (x, y, Y) coordinates of a white test patch were less than 0.005 in (x, y) and less than 3% in Y (less than 5% at lower light levels).

(c) *Procedure*

The observer's task was to adjust the chromaticity and luminance of the local illuminant so that the patterns in each pair looked as if they were made up of exactly the same pieces of coloured paper, that is, to make a 'paper match' (Arend *et al.* 1991). Matches were made with and without transposition of the two test surfaces in the right-hand 6700 K pattern (see figure 1). Observers were allowed to move their eyes freely (Cornelissen & Brenner 1995), and were given unlimited time to make each setting.

(d) *Observers*

A pool of 15 observers participated in the experiments, 5 male and 10 female, aged 19–41 years, with normal colour vision (verified by Rayleigh and Moreland anomaloscopy) and visual acuity. All except co-author K.A., who also served as an observer, were unaware of the purpose of the experiment.

(e) *Performance measure*

As matches made by observers were effectively estimates of local illuminant colour, they could be represented as points in the CIE 1976 three-dimensional ($L^*u^*v^*$) colour space, and the degree of colour constancy then quantified with a standard index (Bäuml 1999; Foster *et al.* 2001*a*) introduced by Arend *et al.* (1991). Thus, in this space, let a be the error in the match, that is, the distance between the observer's setting and the 6700 K illuminant and let b be the scale factor, that is, the distance between the 25 000 K and 6700 K illuminants; then the constancy index is $1-a/b$. Perfect constancy therefore corresponds to an index of unity, and the greater the error—independent of its direction in colour space—the lower the index. The values reported here do not depend critically on the choice of colour space: similar values were obtained with the CIE 1976 two-dimensional (u', v') colour space. Some more general limitations on surface-colour matching as a measure of colour constancy have been discussed elsewhere (Foster 2003).

3. RESULTS AND COMMENT

(a) *Constancy with transposed test surfaces*

The first experiment tested whether interchanging the positions of the test surfaces in the Mondrian patterns affected surface-colour matches. New patterns (i.e. fresh samples from the Munsell set) were generated in each trial, and the matches made by each observer were averaged over 16 such trials. The degree of colour constancy was then calculated. The entries in table 1 (first and second rows) show mean colour-constancy indices calculated over a group of 11 observers for the 11° and 7° patterns. Despite the transposition of the test surfaces, the mean constancy indices of 0.70 and 0.68 (first column) were high and closely similar to previously reported values for asymmetric colour matching with fixed stimulus geometry and simulated Mondrian patterns (e.g. Bäuml 1999; Foster *et al.* 2001*a*) or real surfaces (Brainard *et al.* 1997). Without the transposition (second column), the mean indices were numerically slightly higher, by *ca.* 0.04 (third column).

A repeated-measures ANOVA revealed no significant effect of test-surface displacement ($F_{1,10} = 4.1$; $p = 0.07$) and no significant effect of pattern size ($F_{1,10} = 1.1$; $p = 0.3$). There was little or no correlation between the degree of constancy and the distance between the central and randomly located test surfaces, either for the 11° patterns (with and without transposition, respectively, $r = -0.03 \pm 0.11$ and -0.16 ± 0.06, $n = 11$) or the 7° patterns (with and without transposition, respectively, mean $r = -0.07 \pm 0.09$ and -0.02 ± 0.08, $n = 11$). Only for the 11° patterns without transposition was the correlation significant ($t_{10} = 2.61$; $p = 0.03$, two-tailed test).

Was this failure to find a significant effect of transposing the test surfaces an artefact of averaging observers' matches over different Mondrian patterns? It might be argued that the transposition had a strong effect which depended on the chromatic characteristics of each pattern but which was masked by being averaged over several patterns with different chromatic characteristics. This hypothesis was tested in two ways.

Table 2. Judging surface colour across an illuminant change with and without simultaneous permutation of surfaces surrounding the test surfaces in a 49-surface Mondrian pattern (figure 2).
(Means (s.e.m.s) of colour-constancy indices were calculated over a group of 11 observers viewing new samples of patterns of side 7°.)

surround permutation	no permutation	difference
0.75 (0.03)	0.73 (0.05)	0.017

First, within-observer standard deviations of matches were calculated. They were no larger when test surfaces were transposed than when they were not. Thus, with the 121-surface Mondrian patterns, the standard deviations of the chromaticity coordinates (u', v') were 0.014 and 0.017 with the transposition and 0.013 and 0.018 without. With the 49-surface Mondrian patterns, the standard deviations were 0.012 and 0.019 with the transposition and 0.013 and 0.020 without.

Second, the experiment was repeated, but rather than being generated afresh in each trial, Mondrian patterns were drawn cyclically from a set of six constant patterns, A, B, . . ., F, generated exactly as before. Because the previous experiment revealed no effect of pattern size, only the 7° patterns were used, which, with fewer surfaces, should have provided the stronger test of any pattern-specific effects. Measurements were made by a group of eight observers, four new and four who had performed the previous experiment. Matches made by each observer were averaged over 10 trials with pattern A and the constancy index calculated; similarly with pattern B; and so on. These index values were then averaged over patterns and observers. Notice the distinction between averaging indices from (i) matches averaged over different patterns presented once, as was done earlier, and (ii) matches averaged over the same patterns presented repeatedly, as was done here.

Table 1 (bottom row) shows the resulting mean colour-constancy indices for the recycled patterns. The mean index of 0.71 with the transposition was only 0.05 smaller than without the transposition. The mean indices over observers for each of the patterns A, B, . . ., F, were, respectively, 0.76, 0.71, 0.65, 0.74, 0.73, 0.68 with the transposition and, respectively, 0.82, 0.73, 0.66, 0.81, 0.75, 0.77 without. The earlier null result seems not to be an artefact of averaging matches over different patterns.

Taken together, these results suggest that surface-colour judgements are, in general, almost independent of the position of the test surfaces within the pattern. Is this true also for the surround surfaces?

(b) *Constancy with permuted surround surfaces*

In this second experiment, surface-colour matches were made with one fixed test surface at the centre of the Mondrian pattern and the remaining surfaces randomly permuted in one pattern relative to the other, as illustrated in figure 2. Only the 7° patterns were used, and new patterns were generated in each trial. The observers were the same as in the first experiment. Table 2 shows mean colour-constancy indices for 11 observers with and without transposition of the surround surfaces. There was no significant difference between the two mean indices ($t_{10} = 0.61; p > 0.5$, two-tailed test).

Surface-colour judgements therefore seem also to be independent of the positions of the surround surfaces. This is not to say of course that systematic manipulations of the surround surfaces may not affect colour appearance, particularly with chromatically sparse stimuli (Jenness & Shevell 1995; Monnier & Shevell 2003; cf. Brenner & Cornelissen 1998), even when remote from the test region (Shevell & Wei 1998; Wachtler *et al.* 2001; cf. Brenner *et al.* 2003). Where the present experiments differed, however, is that the stimuli were chromatically dense, the manipulations were random rather than systematic, and they preserved the spectral and spatial characteristics of the surround field.

(c) *Constancy with altered highest-luminance surface*

If judgements about surface colour do not generally depend on relative positions and therefore on local chromatic context, how else could they be achieved? One possibility is that observers compare the test surface with certain other surfaces in the pattern, and then attempt to establish that same relationship for the same surfaces in the second pattern. The ability to make accurate relational judgements independent of the illuminant has been attributed to relational colour constancy (Foster & Nascimento 1994); that is, the constancy under different illuminants of the perceived colour relations between surfaces, as distinct from their perceived colours. This perceptual invariant, which has been investigated operationally by measuring the discriminability of illuminant and surface-reflectance changes on scenes (Craven & Foster 1992; Nascimento & Foster 1997), was initially conceived for situations in which scenes differ only in illuminant, so that one surface could be directly related to another: the simplest possible comparison. It can also be extended to situations in which scene configuration changes if instead a surface is related to multiple other surfaces or to some average (possibly weighted) over the scene as a whole or to a distinguished surface, such as the one with the highest luminance (assumed not to be the surface itself). How these putative comparisons might be achieved is considered later. There are other more complex chromatic properties of patterns which are also invariant under spatial permutations of its surfaces and which might be used in this way (e.g. Maloney 2002; Golz & MacLeod 2002). Although space-average scene colour or the surface with the highest luminance can, in principle, be used to make an estimate of the illuminant (the 'grey-world' assumption (Buchsbaum 1980) and the 'bright-is-white' assumption (Land & McCann 1971; Gilchrist *et al.* 1999), respectively), the comparisons underlying relational colour constancy do not themselves require or provide such an estimate (Foster 2003).

Despite space-average scene colour being well defined under changes in illuminant, the surface with the highest luminance under one illuminant need not always be the same as under another illuminant. An analysis of the 800 patterns used in the first experiment showed that 70% of the patterns preserved the position of the highest-luminance surface during an illuminant change, but this proportion rose to 99% if luminance variations of 5% were ignored.

Table 3. Judging surface colour across an illuminant change with and without simultaneous transposition of test surfaces in a 49-surface Mondrian pattern where information provided by the highest-luminance surface was suppressed (see figure 3). (Means (s.e.m.s) of colour-constancy indices were calculated over groups of 11 and 8 observers viewing, respectively, new and recycled samples of patterns of side 7°.)

	test-surface transposition	no transposition	difference
new Mondrians	0.70 (0.05)	0.68 (0.06)	0.018
recycled Mondrians	0.74 (0.03)	0.80 (0.03)	−0.062

To test the extent to which positional colour constancy might depend on the highest-luminance surface, its luminance was manipulated in a third experiment that in all other respects duplicated the first experiment with the 7° patterns, which, with fewer surfaces than the 11° patterns, should have provided the stronger test (Linnell & Foster 2002). In each trial, the luminance of the surface in the right-hand pattern that in the left-hand pattern had the highest luminance under the 25 000 K illuminant was exchanged with the luminance of a randomly chosen surface in the right-hand pattern, but constrained to be lower than the second highest luminance of the remaining surfaces under the 6700 K illuminant (the chromaticity coordinates were unaltered). Figure 3 shows an example. If the transposed test surfaces were compared with the highest-luminance surface, then performance should have been markedly poorer. The observers were the same as in the first experiment.

The entries in table 3 (first row) show mean colour-constancy indices for 11 observers with and without transposition of the test surfaces. There was no significant difference between the two mean indices ($t_{10} = 0.61$; $p > 0.5$, two-tailed test). Some observers noticed the manipulation of the highest-luminance surfaces, and reported that it made colour matching more difficult, although this seemed not to be reflected in the mean scores.

To test whether taking the mean of observers' matches over Mondrian patterns with different chromatic characteristics might have masked the contribution of the highest-luminance surface, the experiment was repeated, but with patterns drawn cyclically from the same set of six constant patterns used in the control to the first experiment. Measurements were made by the same group of eight observers who participated in that earlier control. Matches by each observer were averaged over 10 trials with each of the patterns and constancy indices calculated as before. Table 3 (bottom row) shows the resulting mean constancy indices for the recycled patterns. With the transposition of the test surfaces, the mean index of 0.74 was 0.06 less than without the transposition, an effect which was significant ($t_7 = 4.6$; $p < 0.01$, two-tailed test), but the mean index was no worse than the corresponding index for the same observers with the same recycled Mondrian patterns in the first experiment without the luminance manipulation (table 1, bottom row).

Thus, observers might use information from the highest-luminance surface, as distinct from any other distinguished surface, but its possible contribution to surface-colour judgements is small in comparison with other sources (Linnell & Foster 2002).

4. DISCUSSION

Looking at objects in different locations under different lights is part of everyday experience, and it is reasonable that the visual system should have developed mechanisms for surface-colour judgement that are stable under these varying conditions. Such an inference is compatible with a general theory of surface-colour perception (Lotto & Purves 2000; Purves *et al.* 2001) in which chromatic percepts are assumed to be generated empirically according to past experience. Past experience, however, is not usually random: the colour signals presented to the eye produced by different combinations of spectral reflectances and illuminants are constrained by physical factors such as the locally uniformity of illumination, the integrity of objects under displacement, and so on. Moreover, if these combinations were random and unconstrained, then it would be difficult to distil from experience the underlying physical regularities that allow veridical interaction with the world.

So, what kinds of physical regularities could the visual system exploit for surface-colour perception? There are many potential signals, but for images in which the eye moves over differently illuminated regions, and has little chance to adapt, an important regularity is the spatial ratios of cone excitations—or ratios of related quantities such as cone-opponent signals—generated in response to light reflected from different illuminated surfaces, either individually or averaged over several surfaces. For a large class of pigmented surfaces, these ratios, defined within rather than between cone classes or classes of opponent signals, are almost invariant under changes in the illuminant, whether its spectrum is drawn from the sun and sky or from a Planckian radiator (Foster & Nascimento 1994). This stability is preserved—and is actually slightly better—with surfaces of natural scenes under changes in daylight (Nascimento *et al.* 2002).

Spatial cone-excitation ratios provide compelling evidence to observers about the origin of changes in scenes: during abrupt illuminant changes, natural deviations in ratios are interpreted as being a result of changes in surface reflectance even when they are actually a result of changes in illuminant (Nascimento & Foster 1997), and there is evidence that such deviations are processed efficiently and in a spatially parallel way over the visual field (Foster *et al.* 2001*b*). Cone-excitation ratios have been assumed to underlie the phenomenon of relational colour constancy referred to earlier (Craven & Foster 1992; Foster & Nascimento 1994), and they may explain the perceived transparency of different combinations of coloured filters placed over scenes (Westland & Ripamonti 2000; Ripamonti & Westland 2003). Minimizing the variance in ratios has been used to predict numerically (Nascimento *et al.* 2004) surface-colour matches in complex three-dimensional scenes (de Almeida *et al.* 2004). Possible neurophysiological mechanisms mediating cone-excitation ratio effects have been considered by Hurlbert & Wolf (2004). Nevertheless, ratios should not be assumed to be relevant to all surface-colour judgements; in particular, they are uninformative in the task of estimating illuminant colour from a scene (Kraft & Brainard 1999; Foster 2003).

Table 4. Mean relative deviations in spatial cone-excitation ratios.
(The unsigned difference in ratios divided by the smaller of the two was evaluated between test surfaces and various configurations of comparison surfaces in Mondrian patterns of 49 and 121 Munsell surfaces, under a change in illuminant from a daylight with a correlated colour temperature of 25 000 K to one of 6700 K. Where like configurations existed for the two sizes of patterns, results were averaged (recycled patterns with 121 surfaces were not used). Entries more than three standard deviations from the mean with 120 surround surfaces are indicated in bold.)

		new Mondrians		recycled Mondrians	
highest-luminance surface	comparison surfaces	test-surface transposition	no transposition	test-surface transposition	no transposition
unaltered	1 adjacent	**1.4**	**0.045**	**1.4**	0.035
	8 surround	**0.28**	0.026	**0.21**	0.031
	24 surround	**0.11**	0.025	**0.071**	0.032
	48 surround	0.025	0.025	0.029	0.029
	120 surround	0.027	0.025	—	—
	highest luminance	**0.045**	**0.045**	0.036	0.036
altered	48 surround	0.027	0.028	0.031	0.031
	highest luminance	**0.77**	**0.78**	**0.55**	**0.55**

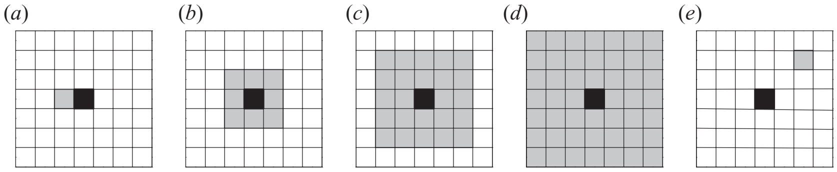

Figure 4. Comparison surfaces for calculation of relative deviations in spatial cone-excitation ratios summarized in table 4. Each matrix represents a 7 × 7 Mondrian pattern with a test surface at the centre. The comparison surfaces were: (*a*) one adjacent surface; (*b*) a surround of eight surfaces; (*c*) a surround of 24 surfaces; (*d*) a surround of 48 surfaces; and (*e*) a surface giving the highest luminance.

To test whether the invariance or otherwise of spatial cone-excitation ratios could provide a basis for making the illuminant- and position-invariant matches reported here, ratios were calculated for each of the experimental conditions. Various configurations of comparison surfaces were tested and the results are described in Appendix A. In brief, ratios between a test surface and an adjacent surface or a spatial average over a local context of 8 or 24 surfaces were not stable under changes in test surface position or permutation of the surround surfaces, but they were stable when evaluated between a test surface and an average over 48 surfaces, or more. They were also stable between a test surface and the highest-luminance surface of the pattern, although less so than with spatial averages. When the information provided by the highest-luminance surface was rendered ineffective, as in the last experiment, only ratios between a test surface and the spatial average remained stable.

To summarize, in determining surface-colour judgements, the visual system could avoid the paradoxical demands of taking into account the region surrounding a surface and simultaneously discounting it by using the larger-scale properties of surfaces in a scene. Relating a surface to space-average scene colour may be more important than making comparisons with specific high-luminance surfaces or local chromatic context.

This work was supported by the Engineering and Physical Sciences Research Council and the Biotechnology and Biological Sciences Research Council. We thank J. J. McCann for discussion, and R. C. Baraas, J. M. Kraft, R. B. Lotto, S. M. C. Nascimento and A. Reeves for comments on the manuscript.

APPENDIX A

The aim in this analysis was to test how the variation in spatial cone-excitation ratios calculated between a test surface and other surfaces in a pattern might explain observers' surface-colour matching performance. As noted in the main text, cone-excitation ratios remain largely invariant under changes in illuminant, and previous work has shown that several surface-colour phenomena can be interpreted in terms of these ratios. Their stability under simultaneous changes in test surface position and illuminant has not been previously documented.

For any two surfaces in a pattern under illuminant e, let r_1, r_2, and r_3 be the ratios of excitations in short-, medium- and long-wavelength-sensitive cones, respectively. Let r'_1, r'_2, and r'_3 be the corresponding ratios when illuminant e is replaced by illuminant e'. The difference between the triples $\boldsymbol{r} = (r_1,\ r_2,\ r_3)$ and $\boldsymbol{r}' = (r'_1,\ r'_2,\ r'_3)$ is small if the failures of invariance in the ratios are small, and large if the failures are large. A convenient summary measure of this difference is given by the quotient $|\boldsymbol{r} - \boldsymbol{r}'|/\min\{|\boldsymbol{r}|, |\boldsymbol{r}'|\}$, where the vertical bars signify the length of the vector: $|r| = (r_1^2 + r_2^2 + r_3^2)^{1/2}$. This relative deviation can be recast as a Michelson contrast and the terms weighted according to the different increment-threshold sensitivities of the three cone classes (Wyszecki & Stiles 1982) or recalculated as opponent combinations (Nascimento & Foster 2000); but in the present context the precise choice does not matter.

Table 4 shows mean relative deviations calculated between the test surfaces and representative configurations of comparison surfaces in a Mondrian pattern under daylights with correlated colour temperatures of 25 000 K and 6700 K. The configurations are indicated in figure 4. They consisted of one randomly selected adjacent surface, four rectangular surrounds of 8, 24, 48 and (not shown) 120 surfaces, and a surface giving the highest luminance under the 25 000 K daylight. For comparisons against multiple surfaces, unweighted cone excitations were spatially averaged over the pattern before ratios were calculated (a similar pattern of performance was obtained with averages of ratios instead of ratios of averages). Results in table 4 are shown with and without simultaneous transposition of the test surfaces (results for permuted surround surfaces were almost identical), and represent averages over the patterns actually used in the experiments: 176–272 new patterns and 6 recycled patterns. The position of the highest-luminance surface under one illuminant was treated as being unchanged if under the other illuminant its luminance was within 5% of maximum.

Entries in bold indicate mean relative deviations more than three standard deviations from the mean obtained with the largest 120-surface surrounds, that is, *ca.* 0.03, which characterizes the limit on the stability of cone-excitation ratios here (cf. Foster & Nascimento 1994). This value is slightly smaller than that reported in Nascimento *et al.* (2002, table 1, Munsell set), where a larger illuminant change of 25 000 K to 4300 K was used.

From table 4, it is clear that ratios between a test surface and an adjacent surface or surrounds of 8 and 24 surfaces would provide an unreliable guide to making surface-colour matches with transposed test surfaces. This is also true for ratios between a test surface and the highest-luminance surface when its luminance was manipulated, for which relative deviations exceeded 0.5. By contrast, ratios between a test surface and a surround of 48 surfaces produced relative deviations of less than 0.03, and would provide a reliable guide to matching, independent of whether the test surfaces or the surrounds were transposed. The extent to which such deviations in cone-excitation ratios predict observed constancy indices has been analysed elsewhere (Nascimento *et al.* 2004).

REFERENCES

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 Neurosci.* **21**, 341–345.

Arend Jr L. E., Reeves, A., Schirillo, J. & Goldstein, R. 1991 Simultaneous color constancy: papers with diverse Munsell values. *J. Opt. Soc. Am.* A **8**, 661–672.

Bäuml, K.-H. 1999 Simultaneous color constancy: how surface color perception varies with the illuminant. *Vision Res.* **39**, 1531–1550.

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

Brainard, D. H., Brunt, W. A. & Speigle, J. M. 1997 Color constancy in the nearly natural image. 1. Asymmetric matches. *J. Opt. Soc. Am.* A **14**, 2091–2110.

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

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 Res.* **43**, 1413–1421.

Buchsbaum, G. 1980 A spatial processor model for object colour perception. *J. Franklin Inst.* **310**, 1–26.

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

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

Foster, D. H. 2003 Does colour constancy exist? *Trends Cogn. Sci.* **7**, 439–443.

Foster, D. H. & Nascimento, S. M. C. 1994 Relational colour constancy from invariant cone-excitation ratios. *Proc. R. Soc. Lond.* B **257**, 115–121.

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

Foster, D. H., Nascimento, S. M. C., Amano, K., Arend, L., Linnell, K. J., Nieves, J. L., Plet, S. & Foster, J. S. 2001*b* Parallel detection of violations of color constancy. *Proc. Natl Acad. Sci. USA* **98**, 8151–8156.

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X. J., Spehar, B., Annan, V. & Economou, E. 1999 An anchoring theory of lightness perception. *Psychol. Rev.* **106**, 795–834.

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

von Helmholtz, H. 1867 *Handbuch der physiologischen optik*, vol. II, 1st edn, Leipzig: Leopold Voss. Translation of 3rd edn, *Helmholtz's treatise on physiological optics*, 1909 (ed. J. P. C. Southall), pp. 286–287. Washington, DC: Optical Society of America, 1924, republished by Dover Publications, New York, 1962.

Hurlbert, A. & Wolf, K. 2004 Color contrast: a contributory mechanism to color constancy. *Prog. Brain Res.* **144**, 147–160.

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

Judd, D. B., MacAdam, D. L. & Wyszecki, G. 1964 Spectral distribution of typical daylight as a function of correlated color temperature. *J. Opt. Soc. Am.* **54**, 1031–1040.

Kraft, J. M. & Brainard, D. H. 1999 Mechanisms of color constancy under nearly natural viewing. *Proc. Natl Acad. Sci. USA* **96**, 307–312.

Land, E. H. & McCann, J. J. 1971 Lightness and retinex theory. *J. Opt. Soc. Am.* **61**, 1–11.

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

Lotto, R. B. & Purves, D. 2000 An empirical explanation of color contrast. *Proc. Natl Acad. Sci. USA* **97**, 12 834–12 839.

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

Maloney, L. T. 2002 Illuminant estimation as cue combination. *J. Vision* **2**, 493–504.

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

Munsell Color Corporation 1976 *Munsell book of color: matte finish collection*. Baltimore, MD: Munsell Color Corporation.

Nascimento, S. M. C. & Foster, D. H. 1997 Detecting natural changes of cone-excitation ratios in simple and complex coloured images. *Proc. R. Soc. Lond.* B **264**, 1395–1402. (doi:10.1098/rspb.1997.0194)

Nascimento, S. M. C. & Foster, D. H. 2000 Relational color constancy in achromatic and isoluminant images. *J. Opt. Soc. Am.* A **17**, 225–231.

Nascimento, S. M. C., Ferreira, F. P. & Foster, D. H. 2002 Statistics of spatial cone-excitation ratios in natural scenes. *J. Opt. Soc. Am.* A **19**, 1484–1490.

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 Neurosci.* **21**, 337–340.

Purves, D., Lotto, R. B., Williams, S. M., Nundy, S. & Yang, Z. 2001 Why we see things the way we do: evidence for a wholly empirical strategy of vision. *Phil. Trans. R. Soc. Lond.* B **356**, 285–297. (doi:10.1098/rstb.2000.0772)

Ripamonti, C. & Westland, S. 2003 Prediction of transparency perception based on cone-excitation ratios. *J. Opt. Soc. Am.* A **20**, 1673–1680.

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

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

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

Westland, S. & Ripamonti, C. 2000 Invariant cone-excitation ratios may predict transparency. *J. Opt. Soc. Am.* A **17**, 255–264.

Whittle, P. & Challands, P. D. C. 1969 The effect of background luminance on the brightness of flashes. *Vision Res.* **9**, 1095–1110.

Wyszecki, G. & Stiles, W. S. 1982 *Color science: concepts and methods, quantitative data and formulae*. New York: Wiley.

Yang, J. N. & Maloney, L. T. 2001 Illuminant cues in surface color perception: tests of three candidate cues. *Vision Res.* **41**, 2581–2600.

Young, T. 1807 *A course of lectures on natural philosophy and the mechanical arts*, vol. I, lecture XXXVIII. London: Joseph Johnson.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.