CHAPTER 8.7

Space-average Scene Colour Used to Extract Illuminant Information

Karina J. Linnell and David H. Foster

8.7.1 INTRODUCTION

Human observers are able to determine whether changes in the colour appearance of a complex scene are consistent with a change in the spectral composition of the illuminating light or of the reflecting properties of its constituent materials (Craven and Foster, 1992; Foster *et al.*, 1992). When an illuminant change is sufficiently large, it is perceived as a change in the space-average colour of the scene; in other words, it appears as though a wash of a different colour had been applied uniformly across the scene. The two experiments reported here were designed to investigate the capacity of human observers both to extract information about space-average colour and to use it to provide information about illuminant colour.

8.7.2 EXPERIMENT 1

If information about space-average colour is extracted from a scene, it should be possible to detect a change in illuminant over *different* as well as *identical* scenes, providing that the surfaces in each scene are sampled sufficiently uniformly to ensure that, on average, they reflect approximately the same amounts of energy in all regions of the visible spectrum. This experiment tested observers' capacity to detect a change in illuminant over two different random samples of reflecting surfaces, and how that capacity depended on the chromatic uniformity of the samples as determined by the number of surfaces in each sample. A practical analogy is provided by our everyday experience of walking from one room into another: can we detect a difference in illuminant between the two rooms, and, if we can, how does detection performance depend on the number of different reflecting surfaces in the two rooms? Clearly, in the limit where each room has only one reflecting surface, it is impossible to disambiguate a difference in illuminant from a difference in the spectral reflectance of the surfaces.

8.7.2.1 Method

Apparatus

Stimuli were generated by an RGB colour-graphics system with 8-bit resolution selected on each gun (Ramtek UK Ltd, Hampshire; 4660 series) under the control of a computer (Sun Microsystems Inc, CA, USA; type 3/160) and displayed on a 19-inch RGB monitor (Sony, Japan; Trinitron). Screen resolution was 1280 × 1024 pixels. Calibration procedures were as detailed in Craven and Foster (1992).

Stimuli

Stimuli were computer simulations of illuminated, square Mondrian patterns presented in a black surround. The individual colour patches physically comprising the Mondrian patterns were square and subtended on each side 0.86° visual angle. There were 49, 25 or 9 patches in each pattern; the pattern therefore subtended on each side 6.0°, 4.3°, or 2.6° respectively. Colour patches had the reflectance characteristics of Munsell surfaces. Surfaces were randomly selected from the 1976 *Munsell Book of Color*, and their spectral reflectances were taken from the basis-function decomposition by Parkkinen *et al.* (1989).

The illuminants were all formed from different combinations of the basis functions derived from the family of daylight illuminants (Judd *et al.*, 1964). Illuminant shifts were from a whitish origin, at $u' = 0.20$ on the daylight locus in the CIE (u', v') chromaticity diagram, to a range of points radiating out from this origin along four different colour directions: 'blue', 'orange', 'green', and 'pink' (in Figure 8.7.1, these directions are indicated by ' + ' signs). Specifically, blue and orange shifts were from the whitish origin at $u' = 0.20$ to various points along the tangent to the daylight locus at $u' = 0.20$, ranging to $u' = 0.18$ and to $u' = 0.22$ respectively; green and pink shifts were from the origin to equivalently spaced points along the normal to the daylight locus at $u' = 0.20$.

Procedure

On each trial one Mondrian was presented for 1 s and it was immediately replaced by a second Mondrian which was also presented for 1 s. Observers were asked to decide which change accounted for the change in the two Mondrians: either a change in the random sample of reflecting surfaces and their geometry, or a change in the illuminant *in addition to* a change in the random sample of reflecting surfaces and their geometry. The two alternatives were equally likely. When a change in illuminant occurred, it could have a magnitude represented by one of five different Euclidian distances in (u', v') space: 0.0066, 0.0131, 0.0197, 0.0263, 0.0438, in any one of the four different colour directions. An experimental session consisted of eight blocks of 80 trials. Within sessions, the number of colour surfaces from which each pattern was built remained the same, but across sessions it varied. In all, observers completed three sessions with each number of surfaces.

Observers

There were two naive observers, GP and TH; each had normal colour vision, as assessed with the Farnsworth-Munsell 100-hue test, and normal Snellen acuity.

Table 8.7.1. Values of the discrimination index d' as a function of increasing size of illuminant shift in the green direction. (SEMs of d' values ranged from 0.2, for d' values less than 1.0, to 1.0 or greater for the largest d' values between 2.0 and 3.0.)

Observer	**Number of surfaces**	**Illuminant shift**				
		0.0066	**0.0131**	**0.0197**	**0.0263**	**0.0438**
TH	49	0.3	1.4	1.9	2.9	2.9
	25	–0.1	0.6	0.9	1.6	2.7
	9	0.3	0.2	0.4	0.8	2.6
GP	49	–0.2	1.0	1.4	2.9	2.9
	25	0.0	0.5	1.3	1.4	2.7
	9	0.1	0.2	0.5	1.1	2.7

8.7.2.2 Results

Detection performance was quantified in terms of the discrimination index d' from signal-detection theory. Table 8.7.1 shows for each of the two observers d' values with the 49-surface, 25-surface and 9-surface patterns and illuminant shifts in the green direction. The d' values for illuminant shifts in the three other colour directions were similar.

For both observers, d' values were generally greater than zero, and increased with increasing size of illuminant shifts and number of colour surfaces.

8.7.2.3 Discussion

To determine whether these results are consistent with the assumption that observers relied upon estimates of space-average colour, a calculation was made of the performance of an ideal observer who was capable of perfectly extracting space-average colour. The ideal observer was presented with exactly the same stimuli as one of the real observers (TH). The first step to the calculation was to derive the space-average colour of each Mondrian pattern. If the Euclidian distance between the space-average colours of any two sequentially presented Mondrian patterns was greater than some criterion value, the ideal observer responded that an illuminant change had occurred, and if it was less than this criterion value the ideal observer responded that no illuminant change had occurred. Criterion values were chosen so that the false-alarm rates of the ideal observer were the same as those of the real observer; modest alterations to the criterion values did not substantially affect the outcome.

The d' values calculated from these hypothetical responses for 49-surface Mondrian patterns and illuminant shifts in the green direction through the five Euclidian distances in order of increasing magnitude were 0.6, 1.6, 2.9, 2.9 and 2.9 (2.9 was the maximum numerical value of d', given the false-alarm rate determined by the chosen criterion value). The d' values that observer TH (see Table 8.7.1) actually produced were of the same order of magnitude as these hypothetical ones, based on the assumption that space-average colour information and information about shifts in space-average colour were perfectly extracted. There was similar agreement between hypothetical and actual d' values for illuminant-shifts in the other directions and for the smaller patterns.

Even so, these results do not constitute evidence that space-average colour is the source of the observer's knowledge about the colour of the illuminant. There is another cue to illuminant colour that generally covaries with the space-average cue, namely, the colour of the highest-luminance patch in the pattern (McCann, 1992). Monte-Carlo simulations of Experiment 1 showed that the two observers' hit and false-alarm rates were predicted not only by shifts in space-average colour but also by shifts in the colours of the patches with the highest luminances.

To determine whether illuminant information can actually be derived from estimates of space-average colour alone, it was necessary to ensure that the information provided by the space-average cue was different from that provided by the highest-luminance patch. To this end, a second experiment (employing a different paradigm) was performed in which space-average colour was deliberately made a poor cue to illuminant colour while the cue provided by the colour of the highest-luminance patch remained good.

8.7.3 EXPERIMENT 2

The space-average colour cue was biased by building ('dual-cue') Mondrian patterns from a sample of 49 Munsell surfaces that, on average, reflected more light in one (orange) region of the visible spectrum than in any other region. One of the 49 surfaces was, however, a whitish surface that reflected light at all wavelengths approximately equally, and had the highest luminance under all the illuminants used to illuminate the dual-cue pattern. In the hypothetical case, in which observers used space-average colour to extract information about the colour of the illuminant, their information would differ substantially from that extracted in a second hypothetical case in which they used the colour of the highest-luminance patch.

To determine what colour observers actually assumed the illuminant to have, they were required to adjust the colour of the illuminant on a second ('comparison') Mondrian pattern until the illuminant was judged the same as that illuminating the dual-cue Mondrian pattern (see Beck, 1959, 1961, for description of a similar paradigm in which observers adjusted the luminance of a light illuminating various grey-level or essentially monochrome patterns). The two patterns were presented haploscopically; thus they could be compared simultaneously, but without interaction, in that the comparison pattern did not influence adaptation to the dual-cue pattern, and vice versa. An unbiased sample of 49 Munsell surfaces was used to build the comparison Mondrian pattern, for which both space-average colour and the colour of the highest-luminance patch were good predictors of illuminant colour. If observers adjusted the colour of the illuminant on the comparison pattern so that its space-average colour (which was effectively equivalent to the colour of the illuminant on the comparison pattern) approximated the space-average colour of the dual-cue pattern, then it could be concluded that they used space-average colour to extract illuminant information, even though space-average colour was not a good cue to illuminant colour. If, conversely, they set the colour of the illuminant on the comparison pattern so the colour of its highest-luminance patch (which was also effectively the colour of the illuminant on the comparison pattern) approximated the colour of the highest-luminance patch in the dual-cue pattern, then it could be concluded that they used the cue offered by the highest luminance patch to extract illuminant information.

At the limit, when individual colour patches are so small as to be unresolvable, only the space-average colour cue is available. In this experiment, although the dimensions of the square patches were always the same in the comparison and dual-cue patterns, they varied across experimental conditions from side 1 pixel (0.03°) to side 32 pixels (1.0°). (In Experiment 1, each patch had side slightly less than 1.0°.) To keep pattern size from varying with patch size, the number of individual patches generated from each of the 49 Munsell surfaces associated with each pattern was increased as patch size decreased. Mondrian patterns with 1-pixel patches appeared textured and it was impossible to identify individual colours; therefore, the only colour information available was from the space-average. In this case, observers were expected to set the colour of the comparison illuminant so that the space-average colour of the comparison pattern approximated the space-average colour of the dual-cue pattern. As patch size increased, however, individual colours, including the colour of the highest-luminance patch, were more and more easily identified. It then became an empirical issue whether space-average colour was still available and was still used to provide information about illuminant colour. The corresponding question for grey-level (or monochrome patterns) with only sparsely sampled grey-levels was addressed by Beck (1959, 1961), who considered patterns with and without clearly identifiable patches. Beck (1961) concluded that for patterns in which there are no clearly identifiable patches with high luminance, judgments about the illuminant are strongly influenced by space-average luminance, whereas for patterns in which there are clearly identifiable patches with high luminance, judgments about the illuminant are strongly influenced by the luminance of those patches.

8.7.3.1 Method

Apparatus

As in Experiment 1.

Stimuli

Stimuli were computer simulations of Mondrian patterns of illuminated Munsell surfaces. The simulations relied on the same basis-function descriptions of spectral reflectances and illuminants as used in Experiment 1.

Two square Mondrian patterns, each of side 7.0°, were presented side-by-side on a black background with their inside vertical edges separated by 3.0°. The comparison Mondrian pattern, presented on the left side of the display, and the dual-cue Mondrian pattern, presented on the right side, had equal numbers of colour patches. Across different conditions of the experiment, patch size varied from side 1 pixel (0.03°), through 2, 4, 8 and 16 pixels, to side 32 pixels (1.0°), but patch colour was always equally likely to be the colour of one of the 49 Munsell surfaces associated with each pattern. In Mondrian patterns with patches of side 32 pixels, each of the 49 Munsell surfaces associated with the pattern occurred only once, whereas in Mondrian patterns with 1-pixel patches, each of the 49 surfaces occurred 1024 times.

The 49 Munsell surfaces associated with the comparison pattern reflected on average the same amount of light in all regions of the visible spectrum, whereas those associated with the dual-cue pattern reflected on average more light in the orange region of the visible spectrum than in any other single region. One of the 49 surfaces associated with both comparison and dual-cue patterns was, however, a whitish sample which reflected light at all wavelengths approximately equally, and which was the highest-luminance surface under all the coloured illuminants used to illuminate the dual-cue pattern.

In total there were eight different illuminants used to illuminate the dual-cue Mondrian pattern across different trials: two blue, two orange, two green, and two pink. Like the illuminants in Experiment 1, they all fell along either the tangent or the normal to the daylight locus at the whitish point $u' = 0.20$ in (u', v') space. The two blue illuminants fell on the tangent to the daylight locus and had u' values of 0.18 (more saturated) and 0.19 (less saturated). The two orange illuminants also fell on the tangent to the daylight locus but had u' values of 0.22 (more saturated) and 0.21 (less saturated). The two green and the two pink illuminants fell at equivalently spaced points on the normal to the daylight locus, each side of the whitish point. The (u', v') co-ordinates of all of these illuminants except the more saturated orange one are plotted as the colours of the highest-luminance patch in the dual-cue pattern in each of the graphs in Figure 8.7.1.

Procedure

On each trial, comparison and dual-cue Mondrian patterns were presented simultaneously to the left and right eyes respectively. The dual-cue pattern was illuminated by one of the eight 'dual-cue' illuminants. Observers were required to adjust the colour of the illuminant on the comparison Mondrian pattern (the illuminant being set initially to be whitish) until it was judged the same as the illuminant on the dual-cue Mondrian pattern. While making these adjustments, they were encouraged to switch attention from one pattern to the other. In each experimental session the six patch-size conditions were undertaken in random order. For each condition, observers made eight matches, one for each of the eight dual-cue illuminants. There were six experimental sessions in all, and therefore observers repeated matches for each combination of illuminant and patch size six times. Different random arrangements of patches were used in different patch-size conditions and in different sessions.

Observers

There were two naive observers, GP and ME (one of whom participated in Experiment 1); each had normal colour vision, as assessed with the Farnsworth-Munsell 100-hue test, and normal Snellen acuity.

8.7.3.2 Results

Figure 8.7.1 shows the average matching performance (over six matches) of each of the two observers for each of seven of the eight dual-cue illuminants and three of the six patch sizes. Matching performance is represented by the space-average colour of the comparison pattern after matching. Each of the graphs in Figure 8.7.1 shows both the illuminant colour predicted by the space-average colour of the dual-cue pattern and the illuminant colour predicted by the colour of the highest-luminance patch in the dual-cue pattern for each illuminant. (Data are not presented for the more saturated of the two orange illuminants on the dual-cue patterns because when the illuminant on the comparison patterns was judged the same colour to observers, it was so saturated that the colour of the reflected light from some surfaces could not be reproduced by the monitor.)

8.7.3.3 Discussion

The colour of the illuminant on the comparison Mondrian pattern was set in such a way that the space-average colour of the comparison pattern was close to the space-average colour of the dual-cue pattern over most of the range of sizes of the patches making up the patterns. But with the largest patch size (1° side), the space-average colour of the comparison pattern was displaced part-way along a straight line joining the space-average colour of the dual-cue pattern and the colour of its highest-luminance patch (see Figure 8.7.1). It appears that

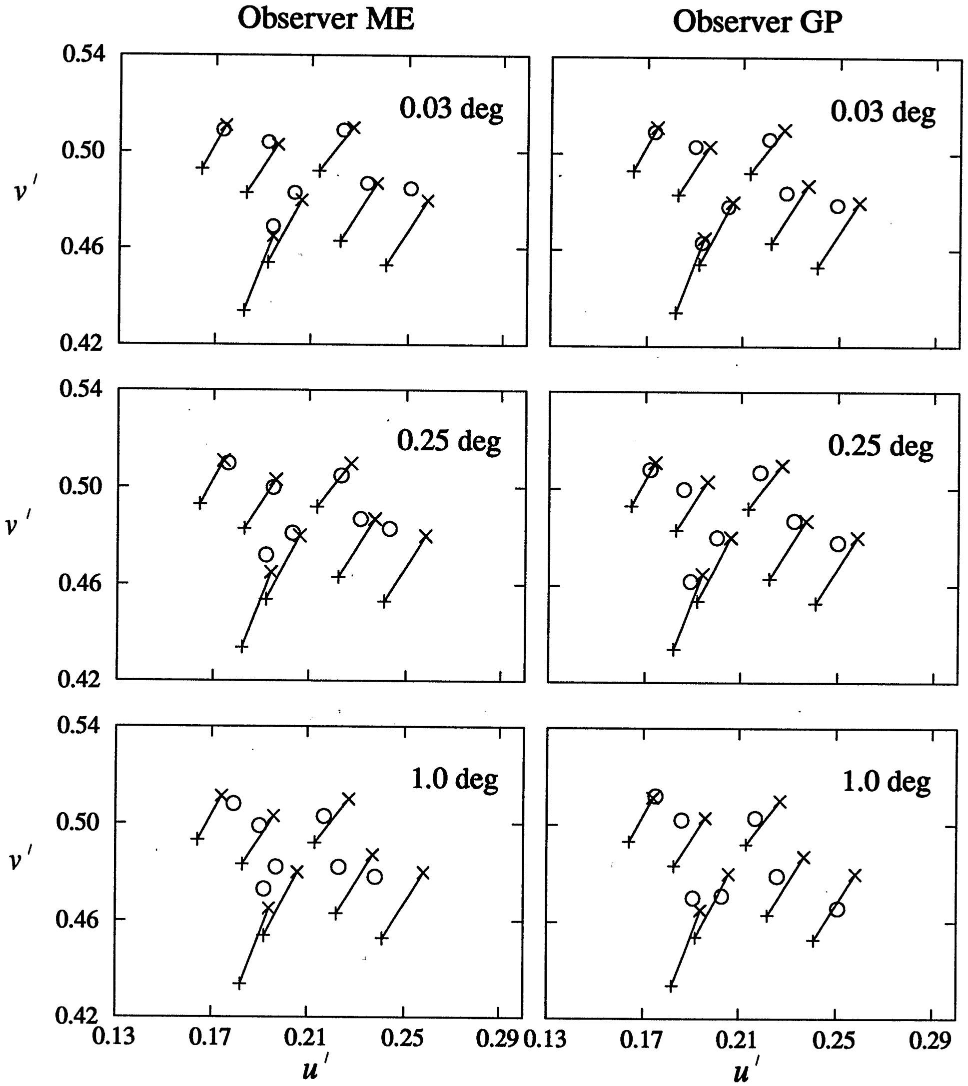

Figure 8.7.1. Illuminant matches for each of two observers for three different patch sizes. Each graph shows, for each of seven of the eight dual-cue illuminants, the average illuminant match made by the observer, defined as the average space-average colour of the comparison pattern after matching (O), and the illuminant matches predicted by both the space-average colour of the dual-cue pattern (×) and the colour of the highest-luminance patch in the dual-cue pattern (+). For each dual-cue illuminant, a straight line joins the matches predicted by the two cues. Patch size is indicated at the top right of each graph.

observers had access to information about space-average colour in patterns with the largest patches, even if they weighted this information less strongly than when the patches were smaller (compare Beck, 1959, 1961). Not only does space-average information appear to be extracted by the observer, it also appears to be extracted remarkably efficiently: performance altered little as patches increased in size from side 0.03° to side 0.25°.

8.7.4 SUMMARY AND CONCLUSIONS

The presence of different illuminants on different scenes can be reliably detected. Experiment 1 showed that observers' performance in this task increased with increasing size of illuminant change on the scenes and number of surfaces within them, and in a way that was similar to that shown by an ideal observer using space-average scene colour as the cue to the illuminant colour. But, in this experiment, as in naturally occurring scenes, space-average colour and the colour of the highest-luminance region covaried, and, in principle, human observers could have used either as the cue. In Experiment 2, these two cues were made to conflict with each other. In these circumstances, space-average colour was the preferred cue with patterns in which the constituent patches ranged in size from 0.03° to 0.5° across, and was a strong cue, if not the preferred one, with patterns in which the patches were 1.0° across (larger than the patches in Experiment 1).

If patch size had been increased further, and the gamut of colour surfaces decreased, the cue offered by the colour of the highest-luminance patch could have become more important in judging illuminant colour. In sparsely sampled grey-level patterns with clearly identifiable patches, the luminance of the highest-luminance patch has been shown to strongly influence illuminant matches (Beck, 1959, 1961). In a rather different task, the appearance of 'colour tautomi' arrays of just five colour patches has been shown to depend strongly on the presence of a white in the array (McCann, 1992). It is also possible that if the denser and more richly sampled Mondrian patterns used here had contained shape-from-shading information and the whitish patches were interpreted as highlights, observers might have attached greater significance to these local cues, and correspondingly less to space-average ones. For the two-dimensional Mondrian images used here, however, space-average colour seems to be the preferred cue to illuminant colour.

8.7.5 ACKNOWLEDGEMENTS

This work was supported by the Wellcome Trust (grant no. 034807). We thank J. J. McCann and S. M. C. Nascimento for helpful discussion; M. G. A. Thomson and S. Westland for critical review of the manuscript; and J. P. S. Parkkinen, J. Hallikainen and T. Jaaskelainen for providing their data for the spectral reflectances of the Munsell surfaces.

8.7.6 REFERENCES

BECK, J. (1959) Stimulus correlates for the judged illumination of a surface, *Journal of Experimental Psychology*, **58**, 267–274.

BECK, J. (1961) Judgments of surface illumination and lightness, *Journal of Experimental Psychology*, **61**, 368–375.

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

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

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

MCCANN, J. J. (1992) Rules for colour constancy, *Ophthalmic and Physiological Optics*, **12**, 175–177.

PARKKINEN, J. P. S., HALLIKAINEN, J. & JAASKELAINEN, T. (1989) Characteristic spectra of Munsell colors, *Journal of the Optical Society of America*, **A6**, 318–322.