Color constancy: A case for multiple levels and paradigms

Michael H. Brill

Sarnoff Corporation, CN 5300, Princeton, NJ 08543-5300.
mbrill@sarnoff.com

Abstract: Shepard claims that color constancy needs linear basis-function spectra, and infers the illuminant before removing its dependency. However, of the models of color constancy that have exact (and reasonable) spectral regimes, some do not need linear basis-function expansions of reflectance and illuminant spectra, some do not solve for the illuminant, and some estimate only partial object-reflectance information for single or multiple objects.
[SHEPARD]

To discuss color constancy, I must first (by assumption) exclude metamerism: If two reflectances match under illuminant "I" but not under illuminant "J," there is no visual transformation that can compensate the change from J to I. In SHEPARD's world of more than three reflectance basis functions, my assumption would mean the eye is blind to all but a three-dimensional (3D) subspace of reflectances under all allowed illuminants.

Having said this, I think SHEPARD aptly describes a natural illuminant spectrum as "a terrestrial transformation of the invariant solar source." However, I do not agree with him that, to extract useful reflectance information from a scene, a visual system must find the illuminant transformation, invert it, and retrieve three reflectance-dependent quantities. Furthermore, even these tasks do not require (as he implies) that the illuminant and reflectance spectra are linear expansions of limited sets of basis functions.

Some investigators *define* color constancy using SHEPARD's constraints, but this choice denies possibilities for strong invariance – under very general conditions. Furthermore, there is evidence that cognitive universals need not be represented one-to-one on perceptual space (one color for each reflectance): people report scene colors differently if asked "what is the color of the light?" as opposed to "what color is that surface?" (Arend & Reeves 1986). Hence we see illuminant biases in a scene, even though we also see illuminant-invariant attributes. To extrapolate, I think we might answer still differently if asked for particular aspects of a colored surface (such as chromaticness) or for color relationships among parts of an object.

1. Older models outside Shepard's framework. Here are some examples of other color-constancy models that have been available for some time (see Brill & West 1986 for a review). A nonlinear model that first solves for the illuminant (Nikolaev 1985) assumes the illuminant and reflectance spectra are Gaussians in a monotonic function of wavelength. The parameters of the Gaussians are analogous to the basis-function coefficients of the linear models. A model that assesses only reflectance *relationships* (Brill & Hemmendinger 1985) makes no illuminant assumptions, but depends on the fact that the spectrum locus in chromaticity space is convex. In this model, the only invariant quantity is the right- or left-handed ordering of the chromaticities of three reflectances. Finally, there is a model that assesses reflectance relationships but does not assume the reflecting objects are coplanar (Petrov 1992). This model begins to come to terms with shading and shadows. However, like all the others discussed so far, it compares points in space, assuming that illuminants vary more smoothly in space than reflectances. This is a problem, for cast shadows and material boundaries have the same sizes and shapes (Arend 1994).

2. A new model. The above problem can be avoided by posing an illuminant-invariant map based on the analysis of a single point in space. Let x be a monotonic function of visible wavelength. Suppose, at one point on the retina, the sensor values are R, G, B, and the sensors have peak x-values r, g, b. Define at this point the following function of R, G, B:

$$P = (g - b)log(R) + (b - r)log(G) + (r - g)\log(B). \tag{1}$$

Although difficult at first to believe, there are several alternative sets of spectral conditions under which *P* depends only on reflectance and not on illuminant spectrum:

Regime 1. Equation 1 was originally applied to von Kries ratios rather than to R, G, B (Brill & West 1981). Later (Finlayson et al. 2000), it was applied to R, G, B. In both cases the following assumptions were used. Let the visual spectral sensitivities be equal-spread Gaussians in x, with standard deviation equal to 1 (chosen, without loss of generality, as the natural unit for x). Let the reflectance spectrum be Gaussian in x:

$$S(x) = a\ exp[-(x - p)^2/(2\,s^2)] \tag{2}$$

Finally, let the illuminant spectrum be exponential in x:

$$E(x) = c\ exp(f\,x). \tag{3}$$

Here, a, p, s, c, and f are coefficients.

Then, *P* is invariant to illuminant change (change of c and f), but depends on the reflectance parameter *s (i.e., gives incomplete information about the reflectance):*

$$P = -0.5[(g - b)r^2 + (b - r)g^2 + (r - g)\,b^2]\ /(s^2 + 1). \tag{4}$$

Regime 2. Equation 1 is also illuminant-invariant for completely general reflectance spectra, provided the following conditions are satisfied (Finlayson et al. 2000; Marchant & Onyango 2000): The sensor spectral sensitivities are narrow-band, approaching delta functions in wavelength, and the illuminant spectrum have the following form:

$$E(x) = e\ F(x)\ exp(-fx). \tag{5}$$

[A noteworthy special case: for the Wien approximation to a blackbody radiator (Wyszecki & Stiles 1982, p. 13), x = 1/wavelength, e and f depend on the black-body temperature T, and $F(x) = x^3$].

3. Outlook. If we generalize SHEPARD's definition of color constancy to include invariants that do not span the three dimensions of color, metamerism might yet be allowed in a color-constant system, so long as the additional freedom in the reflectance does not affect the values of the invariants. More generally, perceptual incompleteness of cognitive universals is compatible with vision models that make multiple incomplete representations of a scene (e.g., Lubin 1995). Such representations may be needed in our variegated world, allowing several hypotheses for visual truth to compete as reality unfolds.

Colour perception may optimize biologically relevant surface discriminations – rather than type-I constancy

Nicola Bruno[a] and Stephen Westland[b]

[a]*Dipartimento di Psicologia, Università di Trieste, 34143, Trieste, Italy;*
[b]*Colour and Imaging Institute, University of Derby, Derby, DEE2 3HL, United Kingdom.* **nicola.bruno@univ.trieste.itd**
http://www.psico.univ.trieste.it/users/nick/
s.westland@colour.derby.ac.uk
http://www.colour.derby.ac.uk/colour/people/westland

Abstract: Trichromacy may result from an adaptation to the regularities in terrestrial illumination. However, we suggest that a complete characterization of the challenges faced by colour perception must include changes in surface surround and illuminant changes due to inter-reflections between surfaces in cluttered scenes. Furthermore, our trichromatic system may have evolved to allow the detection of brownish-reddish edibles against greenish backgrounds.
[SHEPARD]

Introduction. Human colour perception has evolved as a trichromatic system with specific receptoral sensitivities and post-receptoral transformations. SHEPARD is almost certainly right in proposing that ecological forces have played a crucial role in shaping such a system. However, he may be wrong in characterizing these forces in terms of the inherent three-dimensionality of variations in the power spectra of natural illumination. Two arguments are raised against SHEPARD's view. First, a complete characterization of the challenges faced by colour perception must include not only illuminant changes, but also changes in surface surround, and illuminant changes due to inter-reflections between surfaces in cluttered environments. We claim that the sole ability to compensate for variations in illumination power spectra is probably inadequate to produce adaptive surface colours. Second, a number of recent results on the statistics of natural reflectance spectra and their relationship to human spectral sensitivities suggest that cone sensitivities optimize surface discriminations that were biologically important to our progenitors, most notably, those involving red-green discriminations. Under this view, the approximate colour constancy of the human visual system derives from the need to guarantee that such discriminations can be performed, rather than being a major evolutionary goal that required the internalization of the global statistics of reflectance and illumination variations.

Two types of constancy in the light of mutual illumination. The environment in which our progenitors evolved was likely to be cluttered with natural formations of various kinds and was subjected to circadian variations in the spectral composition of daylight. To detect edible materials, such as fruit or roots, our species evolved the ability to use information in colour signals (Mollon 1989; Osorio & Vorobyev 1996) and the spatial relationships between colour signals (Foster & Nascimento 1994; Nascimento & Foster 1997). This ability amounts to solving three related challenges: achieving colour descriptors for surface materials despite changes in phases of illuminaton (Type-1 constancy); achieving constant descriptors despite changes in the surrounds (Type-2 constancy); and properly treating changes in intensity and spectral composition of the illumination due to mutual inter-reflections, shadowing, and transparency effects. The first challenge could conceivably be solved by exploiting statistical constraints on the variability of the phases of daylight (Shepard 1994). However, it is doubtful that the other two could. In fact, there is some consensus that solving the Type-2 constancy problem entails exploiting regularities in the distribution of surface reflectances, possibly using maxima in the distribution of colour signals (e.g., McCann 1992) or their variability (Brown & MacLeod 1997). In addition, there is a growing consensus that colour constancy will eventually require taking into account spatial structure (Schirillo 1999). In this respect, a standing problem for the field of colour vision is to connect a number of important facts that have emerged from the study of such effects of spatial structure in the perception of achromatic colours (Agostini & Galmonte 1999; Bruno et al. 1997; Cataliotti & Gilchrist 1995).

How is the sampling of colour signals "optimal"? Since the pioneering contributions of Cohen (1964) and Maloney (1986), attempts at measuring the statistics of natural reflectance spectra have been performed in several laboratories (Parkkinen et al. 1989; Westland et al. 2000). Two crucial questions have been raised: how many basis functions are practically necessary to fully capture the variability of natural reflectances, and how does the abstract space defined by such bases relate to the coding of colour signals by the cones and the chromatically-opponent channels in the visual system? Answers to the first question have varied from three (Cohen 1964) to as many as twelve (Westland et al. 2000), depending on the intepretation one gives to the word "practically." Early answers to the second question (e.g., Buchsbaum & Gottschalk 1984) suggested that the first three basis functions are closely related to a luminance channel, red-green opponency, and yellow-blue opponency. Underlying this characterization of the mutuality of bases and opponent coding is the implicit assumption that chromatic coding is optimized to recover surface reflectance from the image intensity equation. However, whether this early answer is true in general is presently not clear. In a recent set of measurements (Castellarin 2000) on a large sample of natural reflectance spectra collected in Italy and the UK, we consistently found the first basis function to be an approximately increasing monotonic function of wavelength, which closely mirrors the sample average of our measured reflectances, not luminance. On the other hand, we find the second base to be highly similar to a red-green opponent signal; whereas the third and the fourth base show a much weaker relation to luminance signals and to yellow-blue opponent code. Our findings seem consistent with the notion that chromatic coding is optimized to capture a single dimension of variation in natural spectra, the red-green dimension, rather than fully reconstructing spectra reflectances. A similar proposal has been advanced by Nagle and Osorio (1993) using different statistical techniques and a different sample. The ability to perform most accurate discriminations along the red-green dimension may reflect pressures from the terrestrial environment of our progenitors, who hunted and gathered to detect brownish-reddish edibles against greenish backgrounds.

Universal generalization and universal inter-item confusability

Nick Chater,[a] Paul M. B. Vitányi,[b] and Neil Stewart[c]

[a]*Institute for Applied Cognitive Science, Department of Psychology, University of Warwick, Coventry, CV4 7AL, United Kingdom;* [b]*Centrum voor Wiskunde en Informatica, 1098 SJ, Amsterdam, The Netherlands;* [c]*Department of Psychology, University of Warwick, Coventry, CV4 7AL, United Kingdom.* **nick.chater@warwick.ac.uk** **paulv@cwi.nl** **neil.stewart@warwick.ac.uk**
http://www.warwick.ac.uk/fac/sci/Psychology/staff/academic.html#NC/
http://www.cwi.nl/~paulv/
http://www.warwick.ac.uk/fac/sci/Psychology/staff/postgraduates.html#NS/

Abstract: We argue that confusability between items should be distinguished from generalization between items. Shepard's data concern confusability, but the theories proposed by Shepard and by Tenenbaum & Griffiths concern generalization, indicating a gap between theory and data. We consider the empirical and theoretical work involved in bridging this gap. [SHEPARD; TENENBAUM & GRIFFITHS]

SHEPARD shows a robust psychological law that relates the distance between a pair of items in psychological space and the probability that they will be confused with each other. Specifically, the probability of confusion is a negative exponential function of the distance between the pair of items. In experimental contexts, items are assumed to be mentally represented as points in a multidimensional Euclidean space, and confusability is assumed to be determined according to the distance between items in that underlying mental space. The array of data that SHEPARD amasses for the universal law has impressive range and scope.

Although intended to have broader application, the law is primarily associated with a specific experimental paradigm – the identification paradigm. In this paradigm, human or animal agents are repeatedly presented with stimuli concerning a (typically small) number of items. We denote the items themselves as a, b, . . . , corresponding stimuli as S(A), S(B), . . . , and the corresponding responses as R(A), R(B), ... People have to learn to associate a specific, and distinct, response with each item – a response that can be viewed as "identifying" the item concerned.

How does a law concerning confusability in the identification paradigm relate to the question of generalization? We suggest that there is no direct relationship. Generalization from item A to item B in the sense discussed by **SHEPARD**, involves deciding that an item b has property f, because item a has property f. This is an inductive inference: f(A), therefore f(B). By contrast, confusing item A with B means misidentifying item A as being item B. Generalization typically does not involve any such misidentification: on learning that a person has a spleen, I may suspect that a goldfish has a spleen – but there is no need to misidentify or mix up people and goldfish.

These observations suggest that there may be a gap between