Color Sensation, Color Perception and Mathematical Models of Color Vision

JOHN McCANN and KAREN L.HOUSTON

> By seeing objects of the same colour under these various illuminations, in spite of differences of illumination, we learn to form a correct idea of the colours of bodies, that is, to judge how such a body would look in white light; and since we are interested only in the colour that the body retains permanently, we are not conscious at all of the separate sensations which contribute to form our judgement.   H. von Helmholtz (1962 edition)

In "The Science of Color" the Committee on Colorimetry of the Optical Society of America (1953) makes a very important distinction between sensation and perception, following the definitions introduced by the Scottish philosopher Thomas Reid (1822). The Committee defines sensation as the "mode of mental functioning that is directly associated with stimulation of the organism". It defines perception as the "mode of mental functioning that includes the combination of different sensations and the utilization of past experience in recognizing the objects and facts from which the present stimulation arises". The distinction centers on the roles of cognition and recognition in the perception of objects in life-like complex images.

Helmholtz hypothesized that the colors we see are perceptions that involve a number of different sensations. If we consider the visual pathway in more detail, we note that the photopigments in the visual receptors respond to the spectral-energy distribution of the light entering the eye. That spectral-energy distribution is determined both by the spectral energy of the illumination falling on objects and the spectral reflectance of the objects themselves. The Helmholtz hypothesis in effect attributes to the visual system the ability to

discriminate between the physical entities of reflectance and illumination. This conclusion is assumed to follow from the fact that observers report constant colors from objects in very different illuminations.

The constant appearance of objects under widely varying conditions has been investigated in both chromatic and achromatic experiments. Examples of chromatic experiments are those of Katz(1935) and of Land (1974). Such experiments show that objects appear very nearly the same, despite large changes in the spectral distribution of the illumination. Achromatic experiments demonstrate asimilarconstancydespiteoverallchangesintheintensi-tyofillumination. Examples of such experiments are those of Gelb(1929) and Land and McCann (1971).

A shadow creates a situation in which the intensity of the illumination is considerably reduced. In fact, a shadow often creates changes in both the spectral distribution and the intensity of the light coming to the observers' eyes. Such a situation is exemplified by the photograph of a lake scene in New Hampshire shown in Plate Sa. The composition of the illumination on the two visible sides of t he swimming float in the center of the image is very different. The right side is lighted directly by the morning sun, while the left side, in the shadow, is illuminated by southern skylight and reflections off the water.
There are three totally different ways of measuring the two sides of the float.

Measure the Physical Stimuli

The physical measurements of the light from the two sides of the float are very different from each other. Measuring the spectra from the original photographic transparency, we calculate the correlated color temperature fortherightsidetobe3300K, and for the left side 7000K.

Conclusion: The two sides are very different.

Measure the Color Sensation

Psychophysical measurements of two sides of the float can determine their color and lightness *sensations*. These measurements could be made, for example, by asking observers to find matches for the left side and the right side from a catalogue of colored papers. An observer would select similar, but clearly distinguishable matches.

Conclusion: The two sides are slightly different.

Measure the Perception

In a third measurement, we might ask the same observer about the material properties of the float.   Is the float painted uniformly with the

same paint; that is, does it have the same reflectance on both sides? What are the illuminants; what are their spectral characteristics; where are they located with respect to the image? It is easy for the observer to report that the float is painted with white paint and that the sun is low and off to the right. The *perception* of the float is that it is composed of the same material both on the left and the right.

Conclusion: The two sides are the same.

Thus, identical sets of wavelength-energy distributions give rise to three entirely different conclusions. Depending on the level of the question we ask, the sides of the float can be described as very different, slightly different, or the same. In thinking about the visual mechanisms underlying these results, it becomes apparent that *perception* is a much more difficult concept to define and model than *sensation*. Perception requires many things, including the recognition of objects.

Reflectance and Illumination vs. Edges and Gradients

Assuming that the human visual mechanism actually establishes reflectance distinct from illumination implies that any computational model of the visual system must be able to make the same distinction. Calculating reflectance and illumination properties entails solving the equation

$$\text{Energy}(x,y) = \text{Reflectance } (x,y) \text{ x Illumination } (x,y)$$

given only the array of energies for all points in the image *(x,y)*. This approach is discussed, for example, by Marr (1982), Buchsbaum (1980), Horn (1975), and Rubin and Richards (in press).

We make the assumption that the visual system cannot distinguish sensations associated with reflectance from sensations associated with illumination, at each point *(x, y)*. Rather, we assume that the visual system generates sensations that correspond to patterns of luminance. We assume that perceptions of objects are then calculated from arrays of sensations. Such perceptions can recognize objects and illuminants.

A strategy for calculating sensations can be based on experimental observations showing that the visual system responds differently to a given change in luminance depending upon whether that change is abrupt or gradual. Abrupt changes in luminance are associated with large changes in sensation; gradual changes are associated with small changes in sensation. This point was made by Craik (1966) in his experiments with a spinning disk. He made a disk of whitepaperwitha45-degreeblacksector. On one edge of the black sector was a saw-tooth shaped projection such that along the radius there was an abrupt

increase, and then a gradual decrease of the black portion. When the disk was spun, the observer reported seeing a uniform, light center and a darker surround. Except for the region of the saw-tooth, the inner and outer regions of the spinning disk send the same amount of light to the eye. However, the visual system regards the stimulus as being composed of two nearly uniform, distinct sensations - the central region appears lighter than the outer ring because the abrupt change in luminance creates a larger change in sensation than does the gradual change in luminance. Other experiments by Cornsweet (1970), O'Brien (1958), Land and McCann (1971), and Land (1974) also illustrate this point.

In more complex situations the appearance of a segment of an image need not be simply related to the intensity of the light coming from that segment (Hering, 1964; Katz, 1935; Land and McCann, 1971; McKee, and Taylor, 1976).Plate5b is a photograph of a small part of the Black-and-White Mondrian used in an experiment described by Land and McCann (1971). In this experiment a complex array of black-and-white papers is lighted obliquely by a lamp placed below the array so that more light falls on the bottom of the display than on the top. The illumination decreases gradually from bottom to top. In this detail photograph of a part of the image there is a long, white, vertical strip. At point A the paper sends to the eye a certain amount of energy, L. Moving up the strip the amount of light falling on the paper gradually decreases. The intensity measured at B is 112 L. There is a small ,change in the sensation of lightness (range of sensations from white to black) associated with the change of luminance from point A to B. What is of particular interest is that the same change in luminance from L to *112* L causes a large change in lightness when presented as an abrupt change at an edge. The point indicated by B' in the lower part of Plate 5b sends to the eye the same amount of light *(1/2* L) as does the point B. The sensations associated with Band B' are not the same.

Gradients

The visual system neither responds vigorously to, nor completely ignores, gradients. Gradients are important in the recognition of three-dimensional objects, helping us to distinguish a spherical surface in three dimensions from a planar, circular disk. Gradients provide important information for calculating perceptions (Horn, 1975).

A wide variety of experiments have measured the threshold response to a luminance gradient. These experiments were designed to answer the question, "What is the smallest spatial rate of change on the retina that an observer can detect?" It was found that the gradient threshold is different for different situations, depending on both size and position on the retina, The responses to the two variables are interrelated in such a way as to endow the human visual system with the remarkable property that gradients appear the same over dramatic changes in viewing distance   (McCann, Savoy, Hall and Scarpetti,

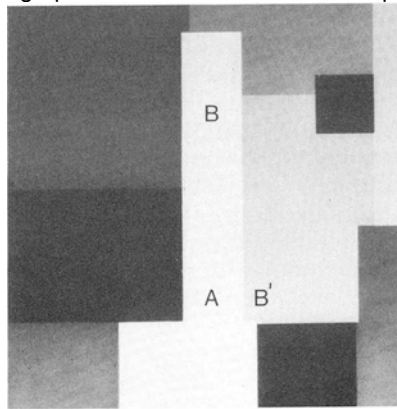Plate 5(a) Photograph of a lake scene in New Hampshire.

Plate 5(b). Photograph of a portion of Land and McCann's Black-and-white Mondrian experiment. The ratio of luminance between A and B is equal to that between A and B'. Gradual changes in luminance cause small changes in lightness, whereas abrupt changes in luminance produce large changes in lightness.
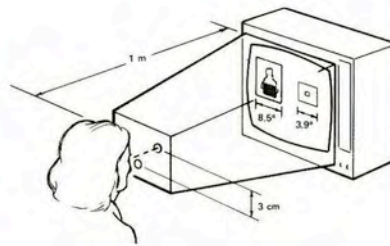
Plate 6. Images in the color-matching experiment and Reset Ratio Product calculations.

1974; Savoy and McCann, 1975; McCann, 1978). This invariance is of considerable advantage because it contributes substantially to the constant appearance of objects at various viewing distances.

Color Matching Experiments

Using the inferences gained from the study of lightness, we return to the general problem of computing color sensations. Land proposed that a particular color sensation is derived from the comparison of three lightness values. Lightnesses are formulated from intercomparisons of sets of independent long-, middle-, and short-wave responsive systems (Land and McCann, 1971; Land, 1964).

In a previous paper (McCann, McKee and Taylor, 1976) the Reset Ratio-Product model for computing lightness was described and used for predicting color sensations in a complex image. The image consisted of a collage of colored matte papers (a "Color Mondrian"). This collage was viewed by an observer under varying conditions of illumination, and color matches were selected for several areas in the image. Calculations were made with the model, using as input only the physical stimulus - the spectral energy distribution coming to the eye from the various patches. The calculation, described more fully later, consists of a sequence of simple arithmetic operations applied to pairs of picture elements in an iterative procedure. The predictions made with the model corresponded very well with the observer matches used to measure color sensations.

More recently, a series of experiments (McCann and Houston, unpublished) in our laboratory have extended the data array from 20 x 24 to 512 x 512 pixels and the sampling of the physical stimuli from 18measurements (one for each paper in the collage) to 262,144 (one for each pixel in the array). The experiments used a set of complex, real-life images (see Plate 6) made under widely differing illumination conditions: tungsten, daylight, and shade. The images were originally recorded on photographic film, then digitized and reproduced on a computer-driven color cathode-ray tube. These images were utilized in two ways: (1) as the stimuli for a series of psychophysical experiments in which observers made color matches to specific areas in the scenes, and (2) as' the basis for mathematical computation of the sensations represented by these areas.

In the color-matching experiments, the CRT screen was divided into two parts; on the left was the test image, and on the right, a standard matching display consisting of a central, color-adjustable square in a constant white surround. The observer viewed the two parts of the display through apertures in a black, pyramidal box arranged in such a way that the left eye saw only the test image and the right eye, only the adjustable square. In addition, the eyes were constrained to work alternately; the viewing apertures were separated by 3 cm vertically, so that by raising his head, the observer saw the adjustable

square with his right eye and nothing (black) with his left, and by lowering his head he saw the test image with his left eye and nothing with his right.

Each test image included a Macbeth ColorChecker Color Rendition Chart as part of the scene being photographed. The observer's task was to adjust the color appearance of the square in the white surround to obtain the visual match to particular segments in the color chart - these segments were the white square and the six primary hues (blue, green, red, yellow, magenta, and cyan). Because of the markedly different color balance of the illumination in the images, the same test square sends very different spectral energy distributions to the eye in the several cases.
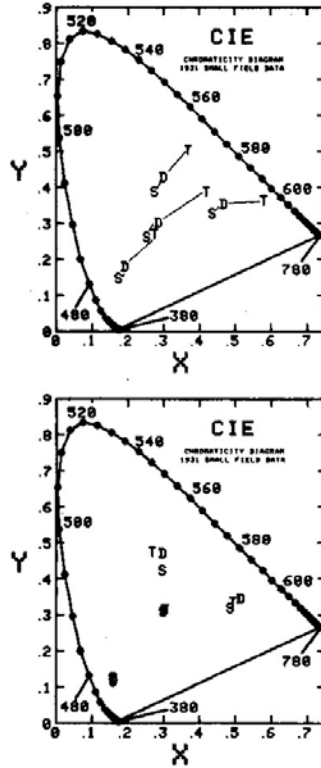

FIG. 1 The top graph shows the physical characteristics, plotted as chromaticities, of the red, green, blue, and white squares in the three test images in Plate 3. Data for tungsten, daylight, and shade are shown as "T", "D", and "S". Observer color matches of the same four squares are shown in the bottom graph. The sensations generated by a particular square in the test images were very similar despite large differences in chromaticities.

Figure 1(top) shows, as an example, the average chromaticities of the red, green, blue, and white test squares in the three images. The data, plotted on the 1931CIE diagram, show the large shifts in chromaticities associated with the different illuminants. Observer matches for these stimuli exhibit much smaller variation, as shown in Fig. 1(bottom), where the average matches made by observer KLH for the same test squares are plotted. These matches were selected under constant illumination conditions using a standard white surround, and so provide quantitative measures of the familiar phenomenon of color constancy.

Mathematical Predictions

Computed sensations for the various test squares are obtained from the Reset Ratio-Product model (McCann, McKee and Taylor, 1976; McCann and Houston, unpublished). The sensations are calculated in the form of lightness values for three overlapping wavelength bands approximating the long-, middle- and short-wave visual pigment absorption curves. In brief summary, the process involve:; five steps. First the radiant energy coming from the CRT phosphors for each target area is integrated under the pigment curves (Smith and Pokorny, 1975) to produce three input records, analogous to color separations. Each record is then processed separately in accordance with Retinex color mechanism independence (Land, 1974). The second calculation step compares local intensities by taking ratios, and the third step consists in propagating these ratios throughout the entire image by multiplication. Next the image is normalized with respect to white by resetting products greater than unity to 1.00. This normalized product at each image point is averaged with other products computed for the same point and also multiplied by other ratios to form new products at new locations. The final step then averages a large number of such products.

The average product thus computed for each point in the image contains information on the relationship of that point to all other points, and it is this average product that is taken as the predicted lightness for the point in a single waveband. Three such lightness predictions, one for each waveband, constitute the predicted color sensation.

The Reset Ratio-Product algorithm tends to normalize each record independently to its lightest point. This normalization has an effect similar to the application of von Kries coefficients to account for color constancy. The • algorithm includes, however, local contrast mechanisms that cause when appropriate the predictions to depart significantly from simple normalization. This is particularly evident in simple displays.

Next we compare the predicted sensations with observer color matches. The calculation for each pixel has a relative value between 0.0 (minimum-black)and 1.0(maximum-white) for each waveband. We have elected to plot the

independent long-, middle-, and short-wave data on three graphs. We now need to scale calculated values so that equal distances on the graphs represent equal increments in lightness. Therefore, the calculated values have been scaled with Glasser's lightness function (Glasser, McKinney, Reilly and Schnelle, 1958). Figure 2 shows a typical set of computer predictions plotted against corresponding experimental matches made by observer KLH. The *x-axis* values were obtained by averaging the matches for each test square, calculating the long-, middle-, and short-wave integrals of the spectral energy coming from this average matching color patch, normalizing to the corresponding integral for the white surround, and scaling the resulting fraction by Glasser's function.

Figure 2 shows the predicted lightness vs. observed lightness for all the test squares in the image taken in tungsten illumination, daylight illumination and shade. Examination of the graphs reveals good agreement between observed and computed lightness in that the points lie close to the line of unity slope. Similar agreement was found in the older experiments with Mondrian collages (McCann, McKee and Taylor, 1976).

Discussion

The strategy we have adopted for modeling the response of the human visual system to color is based on two observations. The first is philosophical: we consider the distinction between sensation and perception, and elect as a beginning to study sensation only, leaving for a later time the much more difficult task of calculating perception. For purposes of our model, color sensation is defined operationally as the quantity that is measured in observer matching experiments. The second distinction is psychophysical: we believe that the visual system evaluates color by comparing lightnesses in three independent fields (the long-, middle-, and short-wave records), and we observe that lightness is determined by the structure of a scene - its edges and gradients. At the level of sensation, the visual system need not separate reflectance and illumination.

The investigations reported here expand our techniques for studying complex color images. They reconfirm the idea that a normalization-like process is required for predicting color sensations in different illuminations. Various biological mechanisms for normalization have been proposed, such as photopigment bleaching, neural adaptation, and comparison of each image point to an average over the entire image. The computational model used in this paper involves ratios, products, resetting, and averaging of picture elements; there is no analog to photopigment bleaching or neural adaptation. Neither is there a direct comparison of each point to the average of the entire image. No *a priori* information about objects or illuminants is needed. The algorithm provides a procedure for calculating the three independent lightness records required for predicting color sensations. We conclude from the present .
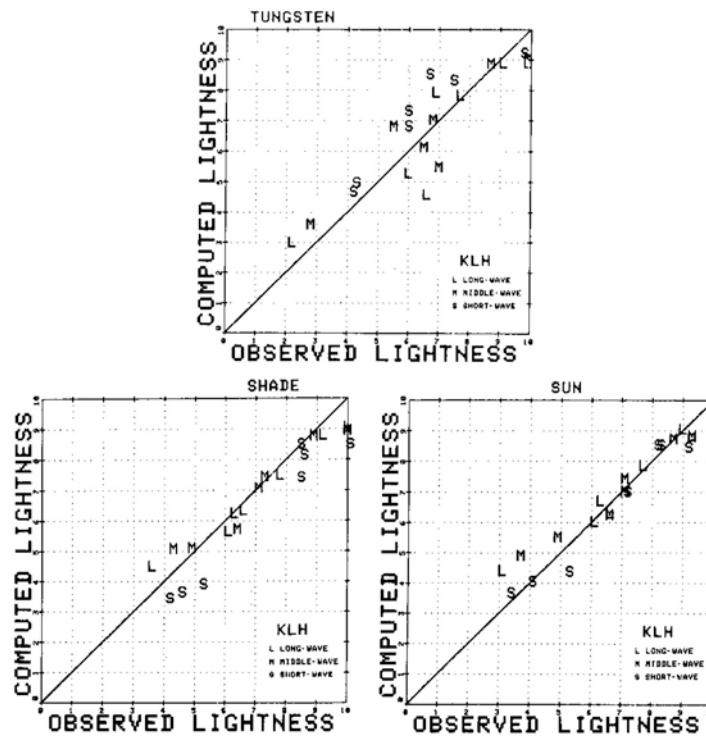
FIG. 2 Comparison of calculated lightness against observer results in tungsten, daylight and shade. The calculated values were obtained from the Reset Ratio-Product model, and the observer results come from binocular matching experiments. The good agreement shows that the model tends to produce accurate predictions of observed color sensation.

experiments that we can calculate a color sensation with reasonable accuracy, given only the array of spectral energy measurements characterizing the original scene.

Acknowledgements

References

Buchsbaum, G. (1980). *Journal 0/ the Franklin Institute* 310. 1-26.

Cornsweet, T. N. (1970). *Visual Perception,* pp. 270-279. New York, Academic Press.

Craik, K. J. W. (1966). *The Nature 0/ Psychology,* London, Cambridge University Press.

Gelb. A. (1929). Die "Farbenkonstanz" den Sehdinge. In *Handbuch der normalen und pathologischen Physiologie, Vol.* 12 (ed. Bethe, *A.).* Berlin, Springer.

Glasser, L. G., McKinney, A. H., Reilly, C. D., Schnelle, P. D. (1958). Cube-root color coordinate system. *J. Opt. Soc. Am.* 48, 736-740.

Helmholtz, H. von (1962). *Physiological Optics* (translated by Southall, J. P. C.). Optical Society of America (reprinted New York, Dover, 1962, Vol. 2, p. 287).

Hering, E. (1964). *Outlines 0/ a Theory 0/ the Light Sense* (translated by Hurvich. L. M. and Jameson. D.). Cambridge, Massachusetts, Harvard University Press.

Hom, B. P. K. (1975). Obtaining shape from shading information. In *The Psychology o/Computer Vision* (ed. Winston, P. H.). pp. 115-155. New York. McGraw-Hill.

Katz. D. (1935). *The World o/Colour.* London. Kegan Paul, Trench. Trubner and Co.

Land. E. H. (1964). The Retinex. *Am. Scientist* 52, 247-264.

Land, E. H. (1974). Smitty Stevens' Test of Retinex Theory. In *Sensation and Measurement* (eds Moskowitz. H. R., Scharf, B. and Stevens, J. C.), pp. 363-368. Dordrecht, Hoiland, Reidel Publishing Co.

Land, E. H. and McCann, J. J. (1971). Lightness and Retinex theory. *J. opt. Soc. Am.* 61,I-II.

Land, E. H. (1974). The Retinex theory of color vision. *Proe. roy. Inst. Gt. Britain* 47,23-58.

Marr, D. (1982). *Vision.* San Francisco, Freeman.

McCann, J. J. (1978). Visibility of gradients and low ~patial frequency sinusoids: evidence for a distance constancy mechanism. *Photogr. Sci. Eng.* 22,64-68.

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

McCann, J. J., Savoy, R. L., Hall, J. A., Jr and Scarpetti, J. J. (1974). Visibility of continuous luminance gradients. *Vision Res.* 14,917-927.

McCann, J. 1., Houston, K. (1983). Calculating Color Sensations from Arrays of Physical Stimulii. IEEE Transactions on Systems, Man, and Cybernetics. Special Issue on Neural and Sensory Information Processing, submitted for publication.

O'Brien, V. (1958). Contour perception, illusion and reality. *J. opt. Soc. Am. 48,* 112-119.

Optical Society of America, Committee on Colorimetry (1953). *The Science of Color,* pp. 58-59. New York, Crowell.

Reid, T. (1822). *Works 0/ Thomas Reid.* New York, Duyckinck.

Rubin, J. M. and Richards, W. A. *Biological Cybernetics,* in press.

Savoy, R. L. and McCann, J. J. (1975). Visibility of low-spatial-frequency sine-wave targets: Dependence on number of cycles. *J. opt. Soc. Am.* 65,343-350.

Smith, V. C., Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Res.* IS, 161-171.