3

A System of Photometry and Colorimetry Based on Cone Excitations

Color Research and Application 11, 244-252 (1986)

I became involved with the CIE in the Spring of 1959 when Dick Blackwell and Glenn Fry cornered me in a bar somewhere in Washington. We then drove out to NBS (now NIST) in Gaithersburg, where Lou Barbrow, Secretary of the U.S. National Committee, signed me up as a last-minute junior replacement for the ailing Harry Helson, and I became a U.S. delegate to the 14th quadrennial convention in Brussels in June of 1959. After attending that meeting, and those in Vienna and Washington during the 60s, I dropped out of CIE activities for the first two meetings in the 70s but returned in 1979 for the 19th convention in Kyoto.

There, on August 28th, with Gunter Wyszecki's enthusiastic backing, a new subcommittee was formed at my instigation and with me as chairman. We had two tasks, one of which has already been mentioned in the comments preceding Paper No. 1. The part that particularly interested me forms the subject of this paper, which is based on an invited talk I gave at the 20th quadrennial meeting in Amsterdam on 1 September, 1983, one of only five presentations to the full convention.

The final paragraph in the section titled Summary and Prospect, at the top of page 251, outlines my hope that the CIE would one day adopt this system of photometry and colorimetry as an alternative (not a replacement) for the existing ones. I worked toward this goal some time, without much success, after which Allen Nagy chaired the committee for a while. More recently,, under the direction of Glenn Fry, it seems to be going off in an entirely different direction.

The system is based on the Smith-Pokorny fundamentals and the crucial assumption that blue cones contribute nothing to luminance, an idea that is anathema to some, particularly the Dutch contingent. Nobody, not even a Dutchman, doubts that the blue-cone contribution is other than miniscule. The chromaticity diagram that Don MacLeod and I developed (see Paper No. 11), and which reappears here, absolutely depends on the assumption that the blue-cone contribution to luminance is zero. The system then allows the number of trolands of retinal illuminance to be conveniently divided into red and green components (here called L and M) with the number of S trolands being separately calculated. The system has several definite advantages compared to the hoary CIE system; these are clearly spelled out.

The equations that appear in the Appendix, which were derived by Naotake Kambe, allow one to calculate chromaticity coefficients in the new chromaticity diagram, given those of the old one. As printed, Equation (2) contained a stray symbol which has been removed in this edition.

Robert M. Boynton

Department of Psychology
University of California at San Diego
La Jolla, California 92093

A System of Photometry and Colorimetry Based on Cone Excitations*

A system of photometry and colorimetry is proposed that is based upon cone action spectra. Instead of X, Y, Z *tristimulus values, the new system divides the visual stimulus into* L, M, *and* S *components, which are related to the relative excitation levels of the three classes of human cone photoreceptors (long-wavelength-sensitive* L, *middle-wavelength-sensitive* M, *and short-wavelength-sensitive* S). *On the assumption that luminance is proportional to* L + M, *with* S *cones making no contribution to it, a chromaticity diagram results in which the relation between chromaticity coordinates and cone excitations is transparent, rather than inadvertently obscured as in the CIE system.*

Introduction

At the Kyoto meeting of the CIE in August, 1979, I suggested the formation of a subcommittee to work toward the development of a system of colorimetric specification based upon a set of physiological reference stimuli.[1] The subcommittee that resulted has been replaced by Technical Committee 1–04 of the reorganized CIE, with Allen L. Nagy as the current Chairman.

The essential content of this article has been presented by invitation on three occasions. At the Amsterdam meeting of the CIE in September, 1983, I presented a paper[2] in which I described the version of the new system that I favor. Similar material was subsequently presented at the inaugural of the Munsell Color Science Laboratory at the Rochester Institute of Technology in February, 1984, and at the Inter-Society Color Council annual meeting in Pittsburgh in April, 1985. The reaction to these presentations has led me to believe that, without waiting for the CIE to finish its deliberations, the color community may be ready to try what is proposed.

The need for an alternative system of photometry and colorimetry arises because of certain properties of the CIE system that inadvertently disguise essential relationships. This problem could not have been appreciated in 1931 when the CIE adopted its familiar triad of functions as the basis for a system of colorimetry and photometry for 2° fields. In that system, two stimuli are predicted to match if they have the same tristimulus values. This implies that their spectral radiance distributions, when integrated in turn with the functions $\bar{x}$, $\bar{y}$, and $\bar{z}$, yield the same tristimulus values X, Y, Z. Regardless of differences in X or Z, two stimuli are defined to be equal in luminance if their values of Y are equal. These computations also provide the basis for the familiar CIE chromaticity diagram on which all stimuli that match a given color (whether physically the same or metameric) plot at a given point and where all colors plot within the region staked out by the spectrum locus and the purple line connecting the ends of the visible spectrum. The CIE system predicts color matches for an ideal observer who is more or less representative of an average human having normal color vision.

It is because the $\bar{x}$, $\bar{y}$, $\bar{z}$ functions approximate linear transformations of the action spectra of an average person's three kinds of cones that they can be used to predict color matches. However, because $\bar{x}$, $\bar{y}$, and $\bar{z}$ are not the action spectra themselves, they provide an unnecessarily difficult conceptual foundation for those interested in teaching the subject of color vision, as well as for others who, in their technical work, would prefer a system wherein colorimetric specification relates more directly to physiology. The system proposed in this article relates to current concepts about color vision, with special attention paid to the excitation properties of the cone photoreceptors.

Physiological Concepts

The idea that there are three different kinds of cone photoreceptors that are responsible for the first stage of color

*Presented at the 20th CIE Congress, Amsterdam, August 31–September 8, 1983.

 CCC 0361-2317/86/040244-09$04.00

processing was suggested nearly 200 years ago by Thomas Young,[3] long before photoreceptors had been seen under the microscope. His notion was that the perception of color depends in the first instance upon the relative excitation of three different classes of elements, now known to be cones, whose absorption spectra overlap. Helmholtz[4] incorporated this idea into his second volume on physiological optics.

Following a revival promoted by Hurvich and Jameson[5] starting in the 1950s, it is now generally agreed, as the German physiologist Ewald Hering[6] originally proposed, that the signals generated at the initial trichromatic stage are soon transformed into two opponent-color signals and one luminance signal, as illustrated in Fig. 1. Defined with respect to the wavelength region of the visible spectrum in which each exhibits maximum sensitivity, the three kinds of cones are designated as longwave-sensitive (L), middle-wave-sensitive (M), and shortwave-sensitive (S). The initial cone signals are all positive. Signals from M cones are both added to and subtracted from L-cone signals. The summed signals relate to luminance, which is discussed later in this article. The difference signals relate to one dimension of color. Relative to their contribution to luminance, L cones contribute only half as much to red–green color balance (see Appendix I). If this attenuated L-cone output exceeds that of M, the result is positive and the difference signal is interpreted at a higher stage of processing to mean red. If negative, it means green. For this reason, the difference signal is said to be transmitted through a red–green (r–g) chromatic channel. The opponent-color signal varies in strength depending upon the degree to which the input from L and M cones is unbalanced, and this depends in turn upon the spectral distribution of the incoming light.

Physiological and behavioral evidence, some of it from macaque monkeys whose vision is remarkably similar to ours, supports the conception, originally based on introspection, that such opponent processing does occur.[7–10] At the retinal level, the code is literally in terms of signals that can be either positive or negative relative to a resting voltage

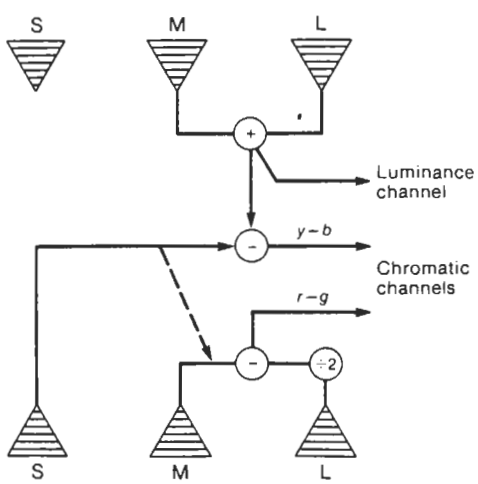

FIG. 1. Schematic diagram of an opponent-color model.

established when there is either zero or balanced excitation of L and M cones. At later stages of visual processing, for example in the optic nerve fibers where all-or-none impulses prevail, the signals are encoded instead as variations of a resting frequency of firing rate.

A second type of opponent-color channel, called yellow–blue (y–b), is also depicted in Fig. 1. It should be imagined that the L and M cones at the top are the same ones shown at the bottom; they have been drawn twice to simplify the connections depicted in the diagram. The yellow-blue opponent channel receives an amplified input with a negative sign from the relatively small number of S cones and a longwave input from the sum of excitations of L and M cones. The S cones also influence the balance of the red–green channel as suggested by the dotted line in the figure. Ultimately, the hue that is perceived depends upon a comparison of the information supplied by the two kinds of opponent channels.

The luminance channel, shown at the top of Fig. 1, receives additive input from the L and M cones, and thus it signals the total rate of photon absorption by these cones without regard to how the photons are distributed between the two. The signal in the luminance channel is always positive. The luminance channel receives no input from S cones. Some $L + M$ signals are transmitted through different nerve fibers than those which carry $L - M$ messages. It is possible, however, that other fibers carry both kinds of information,[11,12] depending upon the spatial and temporal parameters of the stimulus.

The Appearance of Spectral Lights

Although seldom seen in nature, spectral lights are useful for colorimetry and for the analysis of visual function. Suppose that we look at a patch of monochromatic light and vary its wavelength from long to short. To understand what happens, we need to consider the cone action spectra (Fig. 2). At the long wavelengths, $L/2$ cone excitation exceeds that of M and red is seen. As wavelength is shortened, the red–green difference becomes less until, at around 575 nm, the $L/2$ and M cone signals are in balance. Because S cones are insensitive to long wavelengths, the result is a sensation that is neither red nor green (it turns out to be yellow). As wavelength is further shortened, M exceeds $L/2$, and green is seen, with diminishing yellowness, until, near 500 nm, $S = L + M$, the yellow–blue channel is balanced, and unique green is seen. At still shorter wavelengths, blue increasingly replaces green. At the shortest wavelengths, the high level of S-cone excitation unbalances the red–green opponent channel in the red direction (suggested by the dotted line in Fig. 1), producing violet.

Before concluding discussion of the model, I wish to make it clear that it is by no means intended to account for everything about color perception. In the first place, it is a gross oversimplification even of the domain it intends to cover. There are many important facts of color vision with which the model does not attempt to deal. For example, it is not concerned with lateral neural interactions, whereas in

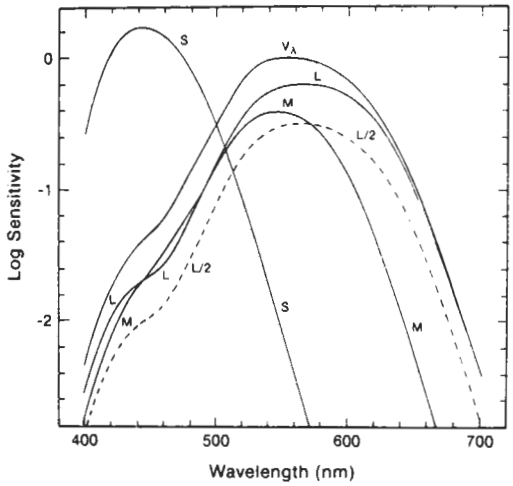

FIG. 2. The cone action spectra of Smith and Pokorny.[19] These have been fit with equations discussed in ref. 1. See Appendix I for a discussion of the significance of the *L*/2 function. The curve labeled V_λ is the spectral luminous efficiency $V(\lambda)$ as modified by Judd.[26]

fact these are strongly implied because the perceived color of a given patch of visual space depends to a considerable extent upon every other color in the visual field. Nor does it deal with the important problems of chromatic adaptation and system nonlinearities. These are not major considerations here because we are attempting to develop a system of photometry and colorimetry that relates to cone excitations and not to complex color perception.

Cone Action Spectra and the Initial Stages of Visual Processing

The proposed system begins with a specification of the action spectra of the three kinds of cones found in the human retina. These are not yet known for certain. Moreover, they differ somewhat from one person to another, so that whatever set we adopt is somewhat arbitrary. On the other hand, we do know quite a bit about the action spectra of cones—certainly much more than was known when the 1931 CIE colorimetric system was being developed.

Most convincing to many people is the direct evidence of microspectrophotometry, initially developed by MacNichol and his students[13] at the Johns Hopkins University in Baltimore in the mid 1960s, and in Wald's laboratory at Harvard.[14] Absorption spectra were measured, initially in goldfish retinas, by transmitting tiny beams of light through cones that had been isolated on a microscope slide. An initial spate of activity in these laboratories produced data of limited accuracy for *L* and *M* cones, but nothing certain about the *S* cones of primates.[15] Within the last few years, the method has been revived and refined by a group of British workers.[16,17] This new work, some of which has been carried out using human material, has confirmed that there are three types of cones, each of which contains its own type of photosensitive pigment.

A good deal of physiological investigation (much of it involving recordings of the electrical activity in single neural units) has led to the following conception of the initial visual process. Vision begins with the absorption of single photons of light by single molecules of the three types of photopigments. The action spectra of cones is a chemical matter, depending upon the fact that the probability of photon absorption is wavelength dependent, with a different wavelength of peak absorption for each of the three kinds of pigments. Each absorptive event elicits a tiny physiological response which shows itself as a change in the membrane potential of the receptor and is correlated with the liberation of a small amount of neurotransmitter at the base of the cone, capable of exciting the post-synaptic cells that enable the messages initiated in the cones to be transmitted toward the brain.[18]

Psychophysical Evidence

In addition to the optical, chemical, and electrophysiological research that has given rise to the foregoing concepts about visual function, we must not overlook the evidence of psychophysics. A psychophysical experiment involves the use of intact subjects, usually human, who are asked to make various kinds of discriminations or judgments of stimuli that are very carefully chosen, controlled, and calibrated. The most-accurate available estimates of cone action spectra are based upon psychophysical data. A fundamental requirement is that, in order to account for color matching, cone action spectra inferred from psychophysics must be linear transformations of color-matching data. Another way of saying this is that the cone action spectra define a unique set of three imaginary reference lights, or primaries, such that each primary is related to the unique excitation of only one type of cone. Because cone action spectra overlap, no real lights can achieve this.

Of special interest are the two most common forms of color blindness (actually, color deficiency), those of dichromats. Dichromats come almost exclusively in two varieties, the protanopes and deuteranopes, who under appropriate conditions behave as if they are lacking the *L*- or *M*-cone photopigments.[19,20] This leaves each type of dichromat with a two-dimensional form of color vision based upon the relative excitations of the *S* cones and only one of the others. Moreover, it is believed that the remaining pigment (*M* cones for protanopes, *L* cones for deuteranopes) is the same as one of those found in the retinas of normal trichromatic observers. If the responses of *S* cones can be kept out of the way, the spectral sensitivities of protanopes and deuteranopes should be those of the *M* and *L* cones, respectively.

Proposed Cone Action Spectra

Modern sets of estimated cone action spectra differ only in detail. The set shown in Fig. 2 was derived a decade ago by Smith and Pokorny at the University of Chicago.[21,22] They are very similar to functions derived by Vos and Wal-

raven.[23] All such sets of functions agree with respect to the following major features:

- The L and M cones are sensitive across the entire visible spectrum, with their peak sensitivities lying not too far apart at about 565 and 540 nm, respectively.
- In the long wavelengths, L-cone sensitivity far exceeds that of the M cones.
- The ratio of M- to L-cone excitation is maximal at about 465 nm; for every wavelength shorter than 465 nm, there is another longer than 465 nm that, if suitably adjusted for luminance, can produce stimuli that would match for a person lacking S cones.
- The S-cone action spectrum, which peaks at around 440 nm, descends rapidly toward the long wavelengths so that, for all practical purposes, the S cones are insensitive beyond about 520 nm.

There are many technical difficulties and arguments that one can raise about such action spectra. For example, one must be concerned about the effects of inert pigments, such as those in the lens of the eye, which selectively absorb light before it can reach the receptors; these vary considerably among observers. Other concerns include individual differences in cone action spectra, and the validity of the assumption that dichromacy is a reduced form of trichromacy. These problems are important and of considerable theoretical interest to the basic vision scientist. But in the present context, they are of second-order importance. What is proposed here ignores these problems in order to focus, just as was done 50 years ago, upon a hypothetical standard observer. From this point of view, the choice of a particular set of cone action spectra, from various ones that might be proposed, is not so very critical.

I have proposed[1] that the set of cone action spectra derived by Smith and Pokorny be used, for which equations are available. These functions are heavily based on information derived from dichromats. Like all such estimates, these must be transformations of the color-matching functions of normal observers if they are to predict their color matches. This raises a question concerning which color-matching data should be used. A decision must also be made concerning whether the proposed system should be compatible with the current CIE system so that, for example, chromaticity coordinates in one system can be transferred to the other.

Choice of Color-Matching Data

The color-matching data used by the CIE in 1931 were those of Wright[24] and Guild.[25] With a certain amount of smoothing, appropriate transformations were developed so that luminance values would be proportional to tristimulus values related to the $\bar{y}$ function alone, compatible with the $V(\lambda)$function which had been standardized in the previous decade. Over the years it became evident that the $V(\lambda)$function seriously underestimates the stimulating power of the shortest visible wavelengths. In 1951, Judd[26] proposed a "corrected" set of data to remedy this problem. Since that time, these have been widely used for scientific purposes, but without official CIE sanction. I mention this because the Smith–Pokorny and Vos–Walraven functions are transformations of the Judd data. An alternative is to adopt functions that are transformations of what many regard as the best set of color matching data available,[27] those of Stiles and Burch,[28] which were developed under CIE auspices in the 1950s. The Stiles–Burch data are not linear transformations of the 1931 CIE color mixture functions, or of Judd's modified values. Stiles–Burch data for the ten individual subjects, which had been lost for many years, were recently recovered and made available by Trezona.[29]

Specification of Luminance

For the purposes of the remainder of this article, I refer to the Smith–Pokorny data. It has already been noted that these functions attempt to account for the color-matching behavior of normal trichromatic human observers, as well as that of protanopes and deuteranopes on the assumption that either the L- or M-cone function is missing. In addition, these functions attempt to account for luminance, the subject of this section, where the emphasis now shifts from colorimetry to photometry.

We are concerned here with how to compare the visual effectiveness of two lights of different spectral distributions. There is no way to do this that is valid for all circumstances. The early photometrists searched for and discovered a method that yielded data having some highly desirable properties.[30] The method is flicker photometry, in which the stimuli to be compared are alternated at about 10 to 15 Hz, with their relative radiances adjusted to minimize or eliminate the sensation of flicker. In addition to the high precision of the method, another desirable property is the additivity and transitivity of the quantities that it specifies, making it possible to build electronic photometers with a spectral sensitivity approximating $V(\lambda)$, which thereby optically scale and integrate the effects of incident light at each wavelength to yield a single luminance value. (Eventually the CIE redefined luminance in terms of an ideal detector having such properties. The same kind of evolution has taken place in the domain of colorimetry.)

Given some additional well-documented facts concerning the temporal properties of chromatic and luminance channels, flicker photometry can easily be understood in terms of the model presented earlier. It has been shown that the chromatic channels respond best to steady lights and to low temporal frequencies, whereas the luminance channels respond best to intermediate and relatively high frequencies to which the chromatic channels are insensitive.[31] Consequently, when lights are alternated at 10 to 15 Hz, as they are in flicker photometry, only the luminance channels can follow this alternation. The chromatic channels register the integrated hue that is appropriate to the alternating stimuli, but so far as flicker is concerned, they are silent. Referring again to the model of Fig. 1, recall that the S cones have access to the visual brain only through the blue–yellow opponent channels. For this reason, S cones do not contribute to the sensation of flicker that one is asked to elim-

inate or minimize in flicker photometry. Therefore, if flicker photometry is said to define luminance, the *S* cones do not contribute to it, and the *L* and *M* cones contribute to the perception of flicker only by way of the luminance channels.

It has sometimes been argued that luminance, defined in this way, is not a meaningful expression of light quantity. For example, there is no doubt that the *S* cones contribute to brightness, which is one of the reasons why stimuli of equal luminance usually do not appear equally bright. There is a body of research to show that chromatic channels are also relatively insensitive to high *spatial* frequencies.[32] As a result, if we had only the chromatic channels to see with, we would not be able to perceive fine details and our visual acuity would be very poor. *S* cones, having access to the brain only by way of chromatic channels, are not able to participate in the perception of fine details. Moreover, there are far fewer *S* than *L* or *M* cones in the retina, especially in the critical central fovea.[33] Even if there were as many, detailed spatial information would be lost anyway in the yellow–blue chromatic channels.

The foregoing ideas are well supported by anatomical and physiological evidence.[34,35] The *S* cones, and the blue–yellow opponent pathways into which they feed, seem to exist for the sole purpose of providing the additional yellow–blue dimension of color perception, without sacrificing a significant amount of spatial visual acuity.

Simplification of Chromaticity Specification at Constant Luminance

From these facts, it may be concluded that there is a meaningful physiological basis for the idea that luminance depends upon the combined output of *L* and *M* cones only. The *L* and *M* functions of Fig. 2, if added together, predict the Judd-corrected $V(\lambda)$ function, also shown in Fig. 2, to within about 3% error. As we now see, this leads to conceptual simplifications and a new kind of chromaticity diagram.

The conceptual simplification seems first to have been appreciated by LeGrand[36] in 1949, when he re-analyzed MacAdam's chromatic-discrimination data. Le Grand noted that if luminance depends only upon the sum of *L*- and *M*-cone excitations, then, for the case of equal luminance, the initial trichromacy of color vision reduces to two special dimensions. These are (1) an exchange of *L*- for *M*-cone excitation, and (2) the level of *S*-cone excitation. The first can easily be visualized in Fig. 3, where the Smith–Pokorny cone action spectra are plotted for an equal-luminance spectrum. When this is done, the relative sensitivities of the *L* and *M* cones become mirror images of each other because they must sum to the same unit value at all wavelengths. As a consequence, only one of these need be considered; we focus on the *L* cones arbitrarily. Reading from right to left, the graph shows a high level of *L*-cone excitation at the longest visible wavelengths, and a monotonic but wiggly descent to a minimum at 465 nm, with a subsequent rise at the shortest visible wavelengths. This plot reveals at once, if one draws horizontal lines that intersect the $L/(L + M)$ function twice, the pairs of wavelengths that would be confused by a person lacking blue cones.

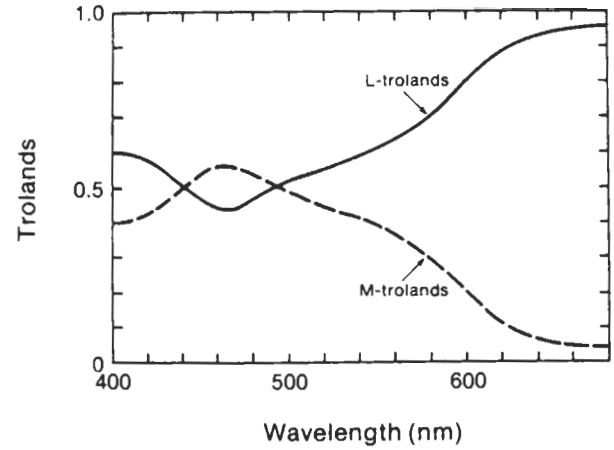

FIG. 3. The Smith–Pokorny *L* and *M* functions plotted on an equal-luminance basis.

L, M, *and* S *Trolands*

Consider a stimulus that produces a retinal illuminance of 1.0 troland (td). From the foregoing, we conceptualize that this implies a summed output of the *L* and *M* cones sufficient to produce a level of visual response appropriate to a 1-td input. But without further information, we do not know how this unit-troland value is divided between the *L* and *M* cones. This information is provided if we know the spectral radiance distribution of the stimulus and the action spectra of the *L* and *M* cones, with the latter properly normalized so that they sum to the Judd-modified $V(\lambda)$ function. Figure 3 gives this information for spectral lights. These components may be called *L* and *M* trolands.

This leaves the excitation level of *S* cones unspecified. This can be as little as zero (for long-wavelength spectral stimuli) or very great (for very-short wavelengths). To calculate a value proportional to *S*-cone excitation level, the spectral radiance distribution of the stimulus must now be integrated with the action spectrum of the *S* cones. What then remains is a specification of the units in which the result will be given.

In accord with an earlier suggestion of Boynton and Kambe,[37] I propose that we define a new unit, the *S* troland (we called it a "blue troland" in that article). By definition, specify, as 1 *S* troland, the level of *S*-cone excitation aroused by 1 td of an equal-energy white. The monochromatic light which also produces 1 S-td per td can be determined by drawing a straight line on the CIE chromaticity diagram from the point (1,0) through the equal-energy-white point (0.333, 0.333) until it intersects the spectrum locus. This occurs at about 498 nm. Figure 4 shows the number of *S* trolands associated with 1 td of monochromatic light for each of the specified wavelengths.

What is proposed here is essentially a substitution, for tristimulus values *X*, *Y*, and *Z*, of a transformed set of

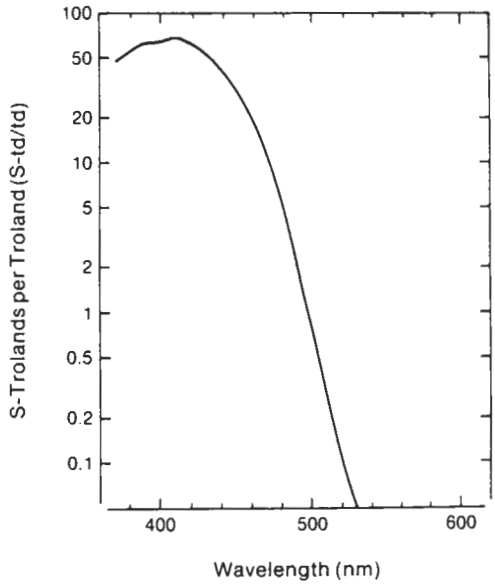

FIG. 4. The number of *S*-trolands per troland for monochromatic lights.

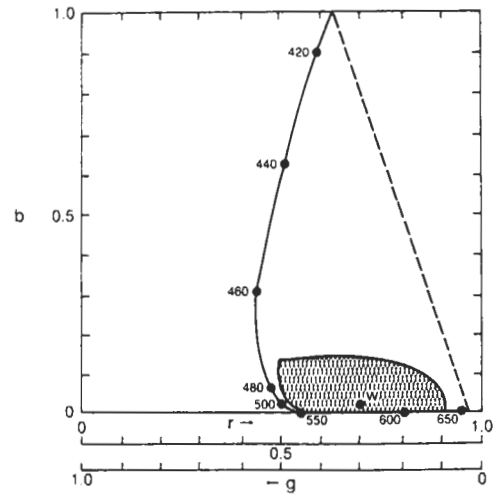

FIG. 5. An equal-luminance cone-excitation chromaticity diagram. Red–green ($L - M$) exchange is represented horizontally; varying levels of *S*-cone excitation, vertically. The domain of optimal surface colors, assuming illuminant D_{65} and 10% reflectance, is shown as the shaded area. The dashed line connects the spectral extremes.

tristimulus values, *L*, *M*, and *S*, which relate directly to cone excitations.

The Luther Diagram

The final step is to develop a chromaticity diagram based upon the foregoing concepts. This is shown in Fig. 5. A diagram somewhat similar to this was first proposed in the 1920s by Luther[38] in Germany and developed independently by MacLeod and Boynton.[39] (A diagram of similar appearance was published by Fry,[40] but it lacks the convenient properties of this one because of his assumption that *S* cones make a small contribution to luminance.) This "Luther diagram" is a unit-luminance chromaticity plot. On the abscissa, reading from left to right, we have the fractional amount of *L*-cone excitation, *r*, which in principle could vary from zero to one. Also shown is a scale of *M*-cone excitation, *g*, running from right to left. The sum of *L*- and *M*-cone excitations is everywhere equal to one.

On the ordinate, we plot the level of *S*-cone excitation, *b*. Remember that this does not affect luminance. Relative to the fractional parts of unit luminance attributable to *L*- and *M*-cone excitation, the units in which *S*-cone excitation is specified must be independently defined. The scaling of the plotted figure is such that the entire spectrum locus fits into a square.

This diagram retains most of the familiar properties of the CIE chromaticity diagram:

- Mixtures of colors plot along connecting straight lines.
- The spectrum locus is everywhere linear or convex, so that the domain of non-spectral colors falls within the curved spectrum locus and the purple line that connects the ends of the visible spectrum.
- The long-wavelength section of the spectrum locus is still straight, the familiar bend at about 520 nm can be seen, and the convex portion of the spectrum locus in the blue-green region is evident.
- The center-of-gravity rule applies for determining where, along the mixture line, a color plots.

In addition, however, the diagram has the following special advantages not found in the CIE chromaticity diagram:

- This is a cone-excitation diagram. At every location in the chart, a given horizontal distance represents the same amount of change of *L*-cone excitation, accompanied by an equal and opposite change of *M*-cone excitation. Similarly, a given vertical distance anywhere in the diagram represents a given change of *S*-cone excitation.
- The center-of-gravity principle, which for the CIE diagram requires special trichromatic units that are not otherwise useful, is based instead on luminance measurements. As examples, a mixture of equal luminances of any two stimuli plots halfway between the chromaticities of those stimuli, or if one color has four times the luminance of another, the mixture plots one-fifth of the distance from the stronger toward the weaker color.

Luther Diagram vs. CIE Diagram

The new diagram represents chromaticities in a way that is directly linked to cone excitations. Why are these relations so hard to visualize in the CIE chromaticity diagram? The answer lies in the oblique relation of the coordinate system of the new diagram, if plotted on the old one. This is shown in Fig. 6. On the CIE diagram, vertical lines in the Luther

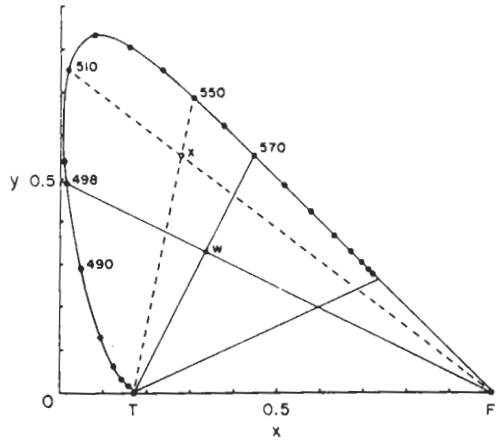

FIG. 6. Horizontal lines in the Luther diagram, representing constant levels of *S*-cone excitation, plot in the CIE diagram as straight lines radiating from the point *F* (1,0). Vertical lines in the Luther diagram, representing constant ratios of *L/M* cone excitation, plot here as straight lines radiating from the tritanopic copunctal point *T*. To estimate the relative *S*- and *M*-troland values for a stimulus, for example at *X*, draw a straight line from *T* through *X* until it intersects the spectrum locus at 550 nm. Refer to Fig. 3 for that wavelength to find the fractional values of *L* and *M* trolands per troland (approximately 0.6 and 0.4). Draw a second line from *F* through *X* until it intersects the spectrum locus at 510 nm. Refer to Fig. 4 to find the number of *S* trolands per troland at that wavelength (about 0.2). (This procedure ignores the small differences between the CIE chromaticity diagram and this one, which is based on the Judd-modified color-matching functions.)

plot, which represent loci of equal *L/M* excitation ratio, translate into lines of various slopes, all radiating from the tritanopic copunctal point (*T*). Horizontal lines in the Luther plot, which represent loci of constant *S*-cone excitation, translate in the CIE diagram into lines radiating from the point *F* (1,0). Along each of these lines, the level of *S*-cone excitation is constant and the ratio of *L*- to *M*-cone excitation varies. Angular relations in the CIE diagram are meaningless for understanding the underlying cone excitations unless they are referred to the superimposed oblique coordinate system, which is intuitively difficult. Furthermore, there is no simple relationship between distances on the CIE diagram and the changes in cone excitations to which they relate.

The legend of Fig. 6 shows how to estimate relative *L*-, *M*-, and *S*-troland values for a given chromaticity on the CIE diagram. The equations describing the relationships discussed in this section are given in Appendix II.

Limitations of the Luther Diagram

A few words are in order regarding the limitations of the proposed diagram.

- It is not a "uniform" chromaticity diagram. For example, equally discriminable steps in the vertical direction are represented by increasing distances as one moves up the chart. This reflects a basic fact of psychophysiology: The greater the level of prevailing *S*-cone excitation, the larger the amount of additional excitation required to produce a discriminable step.[18,36]
- If the spectrum colors are to be represented as in Fig. 5, the white point plots very close to the red–green spectrum locus. This surely would be a defect if one wanted to make graphical determinations using this diagram. On the other hand, the nearness of white to the spectrum locus dramatically represents a fact of physiology that is not evident on the CIE plot: Relative to the level of excitation of *S* cones that can be elicited with very-short-wavelength light, only about two percent as much *S*-cone excitation is needed to turn a spectrum yellow into a neutral white.

In the world of reflecting surfaces, the levels of *S*-cone excitation produced by monochromatic lights of the shortest visible wavelengths are not even approached. With this in mind, it is of interest to represent the gamut of surface colors on the Luther plot. The theoretical limits can be specified in terms of the concept of optimal colors, as first developed by Schrödinger[41] and later specified for the CIE system by MacAdam.[42] My calculations relate to the optimal colors that can be achieved at 10% reflectance assuming illuminant D_{65}, in which the short wavelengths are richly represented. The shaded area of Fig. 5 shows that there is no need to use the upper part of the chart because surface colors are not capable of producing these high levels of *S*-cone excitation at reasonable reflectances. By rescaling the ordinate, one could move the white point farther away from the spectrum locus to make the chart more useful for visualizing mixtures of colors within the domain that is theoretically realizable by surfaces that reflect 10% or more of the incident light.

Summary and Prospect

Committee TC 1–04 of the CIE is working toward the final development of a system of colorimetry and photometry, following the general outlines of what is described here, for possible adoption as an alternative to, but not as a replacement for, the current CIE procedures. The principal features of the new system bear repeating.

- Like the current CIE system, it is based upon the concept of an ideal observer.
- The new system will relate to an agreed-upon set of cone action spectra that are representative of those found in human eyes.
- In the new system, luminance is proportional to the sum of *L*- and *M*-cone excitations, having a real basis in physiology, one that relates to cone action spectra, and luminance channels that are distinct from chromatic ones.
- Chromatic differences are defined as those that take place at equal luminance.
- A new chromaticity diagram is developed that, while retaining the desirable features of all linear chromaticity diagrams, including that of the current CIE sys-

tem, has a number of useful properties which derive from the direct relation of the coordinate axes to cone excitation.

Following the preparation of a detailed technical report, it is anticipated that the CIE will be asked to adopt the new system on a trial basis with the expectation that, with further modifications as experience might dictate, the alternative system of colorimetry and photometry, designed to supplement but not to replace the present one, will be officially sanctioned. Finally, I wish to acknowledge the encouragement and support of the late Gunter Wyszecki, who strongly endorsed these objectives.

Appendix I. Relative Contributions of *L* and *M* Cones to Luminance and Color

The absolute sensitivities of the *L*, *M*, and *S* cones are not known. Perhaps the most fundamental way to define them would be in terms of a set of three probabilities, summing to one, that a photon absorbed in a cone at some reference wavelength, say 500 nm, will be absorbed by each of them. Relative to this ideal, the *S*-cone function of Fig. 2 is probably about a hundred times too high. Given that it contributes nothing to luminance, its position in relation to $V(\lambda)$ accounts for yellow–blue color balance instead. Here it is positioned vertically to predict a unique green at a wavelength where the *S* and $V(\lambda)$ (in the figure, V_λ) curves cross, near 500 nm.

The Judd-modified $V(\lambda)$ function can be fit very closely by the sum of the *L* and *M* curves vertically positioned as they are shown in Fig. 2. But this does not predict color balance at all well. For example, near 570 nm the *L* curve is 0.3 log unit higher than the *M* curve. If the red–green opponent-color signal were dependent upon $L - M$ so defined, the prediction would be for a reddish orange at this wavelength, which instead appears nearly unique yellow. Moreover, the *M* curve rises only slightly above the *L* curve in the short wavelengths, suggesting that little green would be seen there, which again is contrary to observation.

These problems are solved by the assumption that, for whatever reason, the *L* cones contribute only half as much to color balance as they do to luminance. The *L*/2 curve of Fig. 2 has the same shape as the *L* curve; it has merely been shifted down by a factor of 2, or 0.3 log unit. This puts the red–green color balance about where it should be, and allows the *M* curve to rise well above the *L*/2 curve at all shorter wavelengths.

Why should *L* cones contribute only half as much to color balance as to luminance? The physiological answer to this question is not known, nor does there seem to be any teleological significance, although it may help to make a linear photometry possible (see below). One possibility is that, although all *L* cones feed the luminance channel, only every second one feeds the longwave input to the red–green opponent-color channel. A second possibility is that the *M*-cone signals are somehow amplified, relative to those from *L* cones, before feeding the shortwave input to the red–green opponent-color channel. (Such amplification must take place for the *S* cones, which exist in relatively tiny numbers in the retina, yet contribute mightily to color balance.)

Whatever the exact mechanism, this arrangement seems to serve a useful purpose relative to the linear behavior of the luminance channels. From physiology, we know that cone signals are related in a nonlinear fashion to photon absorption, even at intensity levels for which the effects of bleaching are negligible. Yet the luminance channel, defined in terms of the experimental operation of flicker photometry, exhibits linear properties of a sort that would be expected only from signals derived from a single type of receptor. Much of this may simply be due to the dominance of *L*-cone input to the luminance channels, the spectral sensitivity of which would not be very different (compare the shape of $V(\lambda)$ in Fig. 2 with that of *L*) even if the *M* cones made no contribution at all.

Appendix II. Constant-Luminance Relations on the CIE Chromaticity Diagram

On the Judd-modified CIE chromaticity diagram, lines of constant *S*-cone excitation radiate upward and to the left from the point 1,0. The equation that expresses this relation is

$$y = \frac{1 - x}{1 + S} \tag{1}$$

where *S* is expressed in *S* trolands. As examples, if $S = 0$, then $y = 1 - x$, which is the line of zero *S*-cone excitation that includes the long-wavelength spectrum locus. As *S* approaches infinity, *y* approaches zero for all values of *x*, defining the alychne. For $S = 1, y = -0.5x + 0.5$. This is the line shown in Fig. 6 that is drawn from *F*, passing through *W*, which then intersects the spectrum locus at 498 nm and ends at the ordinate axis at +0.5.

Lines of constant *L*/*M* radiate upward from the point *T* ($x = 0.175$, $y = 0$) and have a slope that depends upon the value of *L*/*M*. The equation that expresses this relation is

$$y = \left\{\frac{1.21\ [1 + (L/M\times)]}{2.74(L/M) - 3.71}\right\}(x - 0.175). \tag{2}$$

As a first example of how this equation works, consider the case where $L/M = 2$, the line of chromatic balance. In eq. (2), the bracketed term represents the slope of the line running upward from *T*. For this example, it is 2.05, and the line intersects the long-wavelength spectrum locus, where $x + y = 1$, at $x = 0.446$, $y = 0.554$, very close to 570 nm. On the way, the line passes through *W*, as shown in Fig. 6.

As a second example, consider the hypothetical case for which only *L* cones are stimulated. As *L*/*M* approaches infinity, the bracketed slope term approaches 0.4416, and a line from *T* would intersect an extension of the spectrum locus at $x = 0.747$, $y = 0.253$, just beyond the domain of real colors. Smaller positive slopes than this are impossible,

even for imaginary stimuli. As L/M approaches zero, representing pure M-cone excitation, the slope term approaches -0.326 and a line drawn from T would in this case intersect an extension of the spectrum locus at $x = 1.640$, $y = -0.640$, well off the diagram, downward and to the right. (This curious representation of the M-cone fundamental illustrates what is perhaps the outstanding counter-intuitive feature of the CIE chromaticity diagram.)

As another example, consider the value of L/M that yields an infinite slope—a vertical line drawn upward from T. Setting the value of L/M in the bracketed term to progressively larger numbers reveals that L/M approaches $371/274 = 1.354$.

Finally, where is $L/M = 1$? In this case, the slope term works out to -2.494; a line from T with this slope intersects the spectrum locus twice, at the two wavelengths where the L- and M-troland curves of Fig. 3 cross. (Because of the compression in the "blue corner" of the CIE diagram, the shorter of these two wavelengths is almost impossible to visualize).

Equations (1) and (2) are based on derivations of my former colleague, Naotake Kambe, to whom I give thanks.

1. *Proceedings, 19th Session CIE (Kyoto 1979)*, Publication CIE No. 50 (1980), Bureau Control de la CIE, Paris, 1980, p. 484.
2. R. M. Boynton, "A System of Photometry and Colorimetry Based on Cone Excitations," in *CIE Proceedings, 20th Session, Amsterdam*, Vol. 1, Papers, Bureau Central de la CIE, Paris, 1983, pp. B1/1–B1/6.
3. T. Young, On the theory of light and colours, *Phil. Trans.* **1802,** 12–48.
4. H. von Helmholtz, *Physiological Optics*, Ed. by J. P. C. Southall, Optical Society of America, Rochester, NY, 1924.
5. L. M. Hurvich, *Color Vision*, Sinauer, Sunderland, MA 1981.
6. E. Hering, *Outlines of a Theory of the Light Sense*, translated by L. M. Hurvich and D. Jameson, Harvard University Press, Cambridge, MA, 1964.
7. R. M. Boynton, *Human Color Vision*, Holt, Rinehart, and Winston, New York, 1979.
8. R. M. Boynton, "Spatial and Temporal Approaches for Studying Color Vision," in *Colour Vision Deficiencies VI (Doc. Ophthalmol. Proc. Ser.* **33.** Junk, The Hague, 1982, pp. 1–14.
9. J. E. Thornton and E. N. Pugh, Jr., "Relationship of Opponent-Colours Cancellation Measures to Cone-Antagonistic Signals deduced from Increment Threshold Data," in J. D. Mollon and L. T. Sharpe, Eds., *Colour Vision*, Academic Press, N.Y., 1983.
10. J. D. Mollon, Color vision, *Ann. Rev. Psychol.* **33,** 41–85 (1982).
11. E. Zrenner, *Neurophysiological Aspects of Color Vision in Primates*, Springer, New York, 1983.
12. C. R. Ingling and E. Martinez, "The Spatiochromatic Signal of the *r–g* Channel," in J. D. Mollon and L. T. Sharpe, Eds., *Colour Vision*, Academic Press, New York, 1983.
13. W. B. Marks, W. H. Dobelle, and E. F. MacNichol, Visual pigments of single primate cones, *Science* **143,** 1181–1183 (1964).
14. P. K. Brown and G. Wald, Visual pigments in human and monkey retinas, *Nature* **200,** 37–43 (1963).
15. P. Liebman, "Microspectrophotometry of Photoreceptors," in H. J. A. Dartnall, Ed., *Photochemistry of Vision; Handbook of Sensory Physiology, VII/1*, Springer, New York, 1972.
16. J. K. Bowmaker, H. J. A. Dartnall, J. N. Lythgoe, and J. D. Mollon, The visual pigments of rods and cones in the rhesus monkey, *Macaca mulatta*, *J. Physiol.* **274,** 329–348 (1978).
17. J. K. Bowmaker and H. J. A. Dartnall, Visual pigments of rods and cones in a human retina, *J. Physiol.* **293,** 501–511 (1980).
18. R. W. Rodieck, *The Vertebrate Retina*, Freeman, San Francisco, 1973.
19. W. A. H. Rushton, A cone pigment in the protanope, *J. Physiol.* **168,** 345–359 (1963).
20. W. A. H. Rushton, A foveal pigment in the deuteranope, *J. Physiol.* **176,** 24–37 (1965).
21. V. C. Smith and J. Pokorny, Spectral sensitivity of color-blind observers and the cone photopigments, *Vision Res.* **12,** 2059–2071 (1972).
22. V. C. Smith and J. Pokorny, Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm, *Vision Res.* **15,** 161–171 (1975).
23. J. J. Vos and P. L. Walraven, On the derivation of the foveal receptor primaries, *Vision Res.* **11,** 799–818 (1971).
24. W. D. Wright, A re-determination of the trichromatic coefficients of the spectral colours, *Trans. Opt. Soc. London* **30,** 141–164 (1928–1929).
25. J. Guild, The colorimetric properties of the spectrum, *Phil. Trans. R. Soc. London Ser. A.* **230,** 149–187 (1931).
26. D. B. Judd, Report of U.S. Secretariat Committee on Colorimetry and Artificial Daylight, *CIE Proceedings, 12th Session, Stockholm*, Bureau Central de la CIE, Paris, 1951, Vol. 1, Sec. 7, p. 11.
27. O. Estévez, *On the Fundamental Data-Base of Normal and Dichromatic Vision*, Doctoral dissertation, University of Amsterdam, 1979.
28. W. S. Stiles and J. M. Burch, N.P.L. colour-matching investigation. Mean results from a pilot group of ten subjects, *Opt. Acta* **2,** 176–181 (1955).
29. P. W. Trezona, *Individual Observer Data for the 1955 Stiles–Burch 2° Pilot Investigation*, (NPL Report QU68), National Physical Laboratory, Teddington, England, 1984.
30. Y. Le Grand, *Light, Colour and Vision*, 2nd ed., translated by R. W. G. Hunt, J. W. T. Walsh and F. R. W. Hunt, Chapman and Hall, London, 1968, pp. 67–72.
31. D. H. Kelly and D. van Norren, Two-band model of heterochromatic flicker, *J. Opt. Soc. Am.* **67,** 1081–1091 (1976).
32. D. H. Kelly, Spatiotemporal variation of chromatic and chromatic contrast thresholds, *J. Opt. Soc. Am.* **73,** 742–750 (1983).
33. J. D. Mollon, "A Taxonomy of Tritanopias," in G. Verriest, Ed. *Colour Vision Deficiencies VI*, Junk, The Hague, 1983.
34. E. P. McCrane, F. M. de Monasterio, S. J. Schein, and R. C. Caruso, Non-fluorescent dye staining of blue cones, *Inv. Ophthalmol. Visual Sci.* **24,** 1449–1455 (1983).
35. D. R. Williams, D. I. A. MacLeod, and M. M. Hayhoe, Punctate sensitivity of the blue mechanism, *Vision Res.* **21,** 1357–1376 (1981).
36. Y. Le Grand, Les seuils différentiels de couleurs dans la théorie de Young, *Rev. d'Opt.* **28,** 261–278 (1949).
37. R. M. Boynton and N. Kambe, Chromatic difference steps of moderate size measured along theoretically critical axes, *Color Res. Appl.* **5,** 13–23 (1980).
38. R. Luther, Aus dem Gebiet der Farbreizmetrik, *Z. Tech. Phys.* **8,** 540–558 (1927).
39. D. I. A. MacLeod and R. M. Boynton, A chromaticity diagram showing cone excitation by stimuli of equal luminance. *J. Opt. Soc. Am.* **69,** 1183–1186 (1979).
40. G. A. Fry, A photo-receptor mechanism for the modulation theory of color vision, *J. Opt. Soc. Am.* **25,** 361–367 (1935).
41. E. Schrödinger, Grundlinien einer Theorie der Farbenmetrik im Tagessehen, *Ann. Physik* **63,** 481–520 (1920).
42. D. L. MacAdam, Maximum visual efficiency of colored materials, *J. Opt. Soc. Am.* **25,** 361–367 (1935).

Received August 10, 1985; accepted December 28, 1985