Handbook of Human-Computer Interaction
Second, completely revised edition
M. Helander, T.K. Landauer, P. Prabhu (eds.)

Chapter 25

Color and Human-Computer Interaction

David L. Post
Armstrong Laboratory
Wright-Patterson Air Force Base
Ohio, USA

A decade or so ago, most of the displays used for human-computer interaction (HCI) were monochrome. Today, however, desktop computers come equipped routinely with color display systems offering 640 x 480 pixels or more, 256 or more colors, and non-interlaced refresh rates of 50 or 60 Hz or greater. Color printers and scanners are becoming increasingly common; they are available at prices many consumers can afford and offer levels of quality that a growing number of users find attractive. Furthermore, the price/performance ratios for these technologies continue to improve.

The software industry has taken advantage of the wide availability of color displays and increasing availability of color hardcopy equipment by using color in its products. To date, this use has been conservative for the most part, to assure compatibility with the large installed base of monochrome hardware. However, the general problem of using color in an effective and attractive manner to enhance HCI is no longer germane to only a few specialized applications—it is a problem faced by a substantial portion of the entire computing industry.

This chapter treats color-related topics that are relevant to HCI. The first section discusses the basics of color vision; the second and third sections introduce Commission Internationale de l'Éclairage (CIE) photometry and colorimetry, respectively, which are the bases for nearly all quantitative approaches to color; the fourth section covers alternatives to the CIE uniform color spaces that are especially relevant to computing; the fifth section describes equipment and procedures for measuring color; the sixth section treats a relatively new problem-area in human-computer interaction known as device-independent color transfer; the seventh section offers guidance for using color in computer systems; the eighth section discusses the use of computers to help solve color-related problems in HCI; the final section lists recommended sources for more detailed information.

25.1 Color Vision and Perception

A good way to approach the topic of color vision and perception is to start by defining *color:* from the psychophysical perspective, it is the aspect of visual perception by which an observer can distinguish among stimuli based on differences in the spectral composition of energy radiating from them. From the perceptual perspective, color is the attribute of vision consisting of chromatic and achromatic content in any combination, described by words such as red, white, etc.

25.1.1 Related versus Unrelated Color

An important dichotomy concerning color is whether it is *related* or *unrelated*. (*Surface* vs. *aperture* and *nonluminous* vs. *luminous* are closely related terms that are used instead, sometimes.) A related color is one that is perceived to belong to an area seen in relation to one or more other colors. An unrelated color is one that is perceived to belong to an area seen in isolation from other colors. Ordinarily, related colors are associated with reflecting and transmitting objects, whereas unrelated colors are associated with emissive sources, but in any event the visual system chooses an interpretation and perceives accordingly. For example, the greenness of grass appears to belong to the grass, whereas a green signal light at night appears to emit green light, even though the light entering the eye in these two cases may be the same. One of the interesting consequences of the distinction is that gray and brown can be perceived only as related colors; if they are viewed in isolation, gray will look white and brown will look like a dark orange or yellow.

25.1.2 Additive Versus Subtractive Color

Another important dichotomy concerning color is whether it is produced by an additive or subtractive process. Television is a familiar example of the additive-color process. A television screen consists of numerous tiny red, green, and blue dots (or stripes), each of which produces a variable amount of light. At normal viewing distances, the individual dots subtend very small visual angles at the eye, so the diffraction patterns they form on the retina overlap and mix. Thus, the tremendous range of colors we see on television is produced by adding only three primary colors together in varying proportions. Projection-CRT systems are another example: separate red, green, and blue images are projected and superimposed on a white screen, where they form a full-color image. The left side of Table 1 is a truth table that shows the eight colors that can be formed by mixing red, green, and blue primaries in an additive, all-or-nothing manner.

Most of the colors we encounter result from the subtractive-color process: light from the sun or some other source of radiant energy strikes objects; some of the wavelengths are absorbed to varying degrees by the objects, which subtracts them from the original light, and the remaining light is then reflected into our eyes. Thus, grass looks green because it tends to reflect wavelengths from the middle of the visible spectrum and absorb everything else. Similarly, the lens in a red traffic light transmits the longer visible wavelengths and absorbs others.

A color photograph works a bit differently because it would be impractical to use a different dye for every color in the picture. Instead, a photograph consists of three overlapping layers of transmissive cyan, magenta,

Table 1. Additive- and subtractive-color truth table.

Additive Primaries				Subtractive Primaries		
Red	Green	Blue	RESULT	Cyan	Magenta	Yellow
0	0	0	Black	1	1	1
1	0	0	Red	0	1	1
0	1	0	Green	1	0	1
0	0	1	Blue	1	1	0
0	1	1	Cyan	1	0	0
1	0	1	Magenta	0	1	0
1	1	0	Yellow	0	0	1
1	1	1	White	0	0	0
Note: 0 = "off" and 1 = "on." For subtractive primaries, "on" means filtering is active, so the associated wavelengths are being removed.						

and yellow dyes, applied in varying amounts on reflective white paper. The cyan layer absorbs red to varying degrees while leaving green and blue alone, the magenta layer controls green while leaving red and blue alone, and the yellow layer controls blue while leaving red and green alone. The right side of Table 1 is a truth table showing the eight colors that can be produced by mixing cyan, magenta, and yellow primaries in a subtractive all-or-none manner. It can be seen that the left and right sides of the table are logical "nots" of each other. This is because the cyan, magenta, and yellow subtractive primaries act basically as minus-red, minus-green, and minus-blue, respectively.

The ranges of colors that can be produced using the additive and subtractive processes differ and can lead to problems reproducing an image that was created using the additive process in a medium that uses the subtractive process (and vice versa). This issue is discussed in Section 25.6.2.

25.1.3 The Dimensions of Color Perception

Color perception can be decomposed into three fundamental attributes, or dimensions:

Hue. The main attribute of color stimuli by which observers distinguish among different portions of the spectrum, for example, blue versus green versus yellow, etc.

Brightness/Lightness. The former is associated with unrelated color and is the degree to which a stimulus appears to emit either more or less light, that is, appears "bright" or "dim"; the latter is associated with related color and is the degree to which a stimulus appears to reflect or transmit either more or less light, that is, appears "light" or "dark."

Saturation/Chroma. The former is the colorfulness of a stimulus, judged in proportion to its brightness. The latter is the colorfulness of a stimulus, judged as a proportion of the brightness of a similarly illuminated stimulus that appears white. Increasing (or decreasing) the brightness of a stimulus causes its chroma to increase (or decrease) but has no effect on its saturation, typically.

25.1.4 Photoreceptors

The three-dimensional character of color perception is reflected in visual physiology. For example, the normal human eye contains three types of color photoreceptors, called *cones* because of their shape, which have the spectral sensitivities illustrated in Figure 1. They are referred to as *long-*, *medium-*,and *short-wavelength sensitive (L, M,* and *S)* cones, according to their spectral sensitivities. The main physical difference among the three cone types is the photopigment each one contains; the sensitivity differences result from differences in the photopigments' absorption spectra. Notice that each cone type responds to a wide range of wavelengths and cannot discriminate among wavelengths within that range; therefore, the signal from a single cone carries no spectral information.

Figure 2 depicts photoreceptor densities in a normal right eye. The cones are concentrated in the central five degrees or so of the retina, in an area called the *fovea*, although they exist throughout the retina. Visual acuity is best in the central two degrees, where the cone concentrations are highest. Within the central eight degrees of the visual field, the ratio of L- to M-cones is roughly 2:1 (Nerger and Ciccrone, 1992). S-

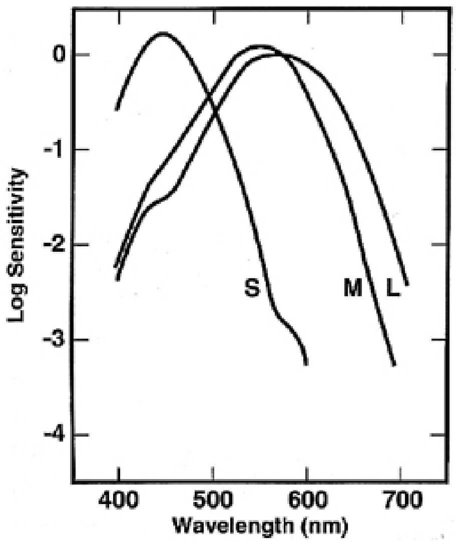

Figure 1. Spectral sensitivities of the cones, normalized for equal sensitivity to light having equal radiance at all visible wavelengths. *Figure provided by Jan Walraven, TNO Human Factors Institute, Soesterberg, The Netherlands*

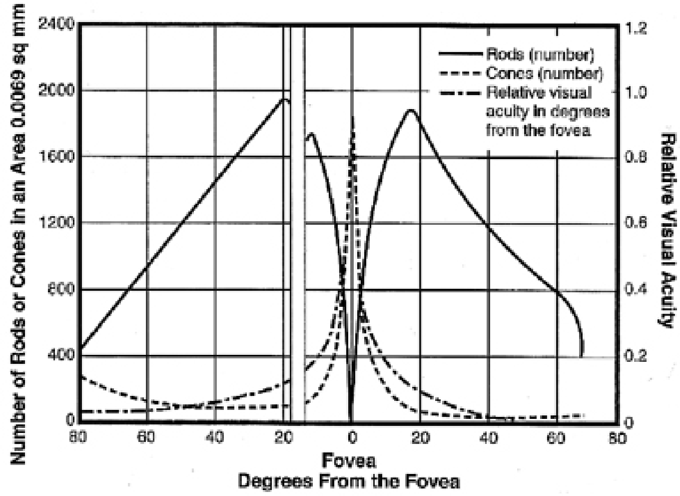

Figure 2. Distribution of cones and rods in the retina of a normal right eye. From Snyder, H.L. (1988). *Image quality. In M. Helander (Ed.), Handbook of human-computer interaction (pp. 437-474). Amsterdam: Elsevier*

cones are rare or absent altogether from the central 0.35 degrees of the retina; immediately outside this area, they constitute approximately 2% of the cones and become more common as eccentricity increases, rising to 7% at 3.5 degrees and beyond (Curcio et al., 1991).

Figure 2 also shows a second class of photoreceptors, called *rods* (again, because of their shape), which are responsible for night vision and are absent from the central fovea. Rods play only a minor role in color vision, so in introductory discussions such as this, they are treated usually as being inoperative at normal, daytime light levels (where color vision is fully operative) to avoid unnecessary complications. For present purposes, little is sacrificed by this simplification.

25.1.5 Color Channels

Figure 3 illustrates schematically how signals from the three types of cones are processed by the visual system: (1) signals from all three types are summed to produce an achromatic color channel that responds in proportion to the cone stimulation; (2) signals from L- and M-cones are differenced, yielding a red-green opponent color channel; and (3) signals from L- and M-cones are also summed to produce a signal that is differenced with S-cone output, yielding a yellow-blue opponent color channel. Thus, the three-dimensional character of color perception is evident at this level of the visual system also. The achromatic channel provides the basis for brightness and lightness perception,

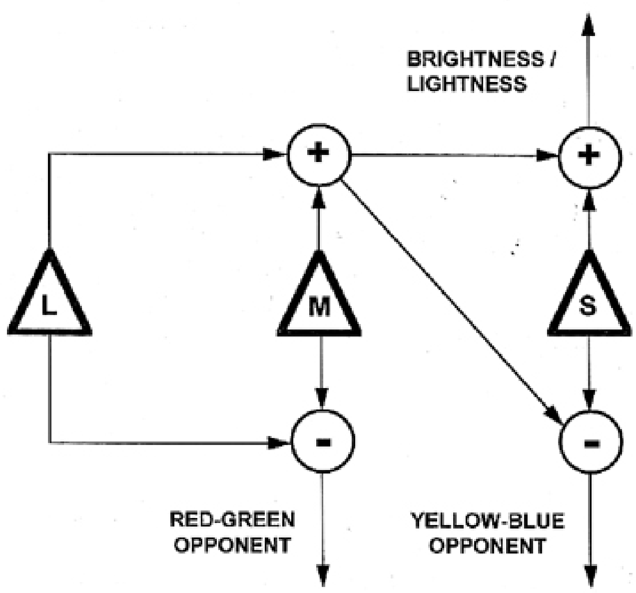

Figure 3. Color-channel schematic.

and the two opponent channels provide the basis for hue and chroma perception. Thus, the opponent-channel processing converts the spectrally ambiguous signals from the cones into ones that convey the chromatic aspects of light precisely.

25.1.6 Metamerism

Figure 4 illustrates two spectral distributions of light that produce identical stimulation of the three cone types. Consequently, the cone signals produced in re-

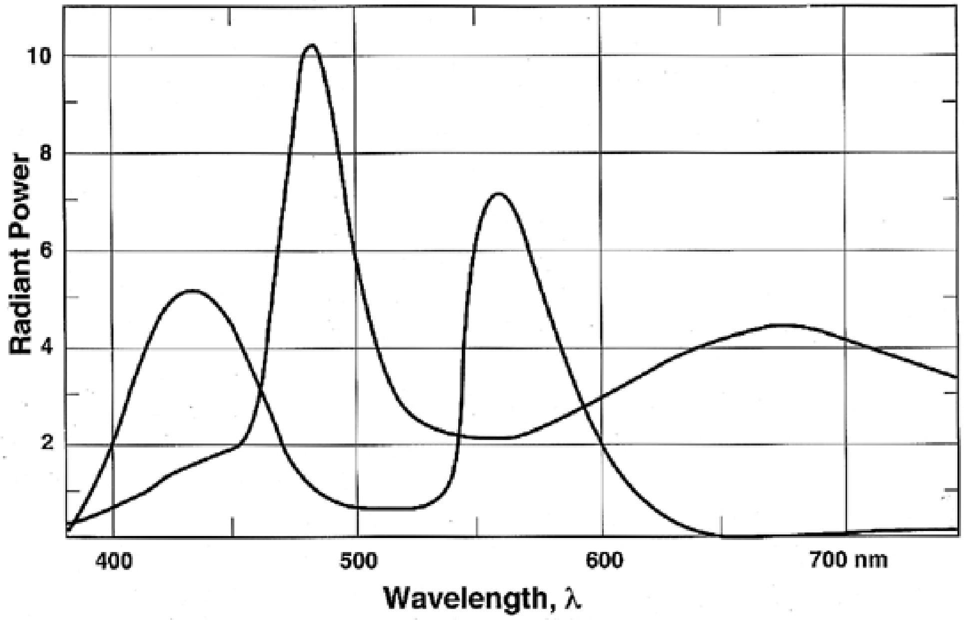

Figure 4. Metameric spectral distributions. From Wyszecki, G., and Stiles, W.S. (1982). Color science (2nd ed.). New York: Wiley. *Copyright © 1982 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.*

sponse to these lights are the same and the visual system cannot discriminate between them. Colors that have different spectral distributions but look the same are called *metamers*. Color television, which reproduces the colors of countless different spectral distributions by mixing only three color primaries in varying proportions, is a familiar example of metamerism. The fact that three primaries suffice to produce such a wide range of colors is a principle of color vision called *trichromacy* and is a direct consequence of the fact that there are only three types of cones.

25.1.7 Color Vision Deficiencies

Some people, called *dichromats,* can match all colors using only two primaries—at least, when they are viewing stimuli that subtend small visual angles (e.g., two degrees). They are classified as either *protanopes, deuteranopes*, or *tritanopes*, according to whether their L-, M-, or S-cones seem affected, respectively. Protanopes and deuteranopes are unable to discriminate among red, orange, yellow, and green. Tritanopes cannot distinguish blue from green, white from yellow, and red from purple. Interestingly, dichromats become trichromatic when larger stimuli are used, although their color vision is still abnormal in most cases.

A larger group of people has trichromatic vision even for small visual fields, but one of their cone types seems impaired to varying degrees. These people are called *anomalous trichromats* and are classified as either *protanomalous*, *deuteranomalous*, or *tritanomalous*. (Some people believe that tritanomalous observers are actually tritanopes whose rods partially replace the missing S-cone signals; see Pokorny, Smith, and Went, 1981.) Their color vision resembles that of the corresponding dichromats, in degrees that depend on the extent of their impairment.

A very small group of people can match all colors using only one primary. Some, called *cone monochromats*, behave as though they have only one cone type. Most commonly, they seem to have only S-cones and have poor (20/60 or worse) visual acuity. Another group, called *rod monochromats*, seem to have only rods. They have a blind spot in the center of their visual fields (i.e., at the fovea), especially poor (20/200) visual acuity, and are often unusually sensitive to light.

Many color-vision deficiencies are genetic in origin (they can result also from injury, illness, aging, and drugs) and their frequencies vary among the world's populations. Table 2, derived from Wyszecki and Stiles (1982, p. 464), shows the frequencies of occurrence for most types of deficiency in the US and Europe. In this population, the protan and deutan defects, which are sex-linked, are the most common. Ocular disorders and aging often preferentially reduce the amount of short-wavelength light reaching the retina, which tends to make blues look darker than they would otherwise. Other disorders cause a relatively greater loss of long-wavelength light, which tends to make reds look darker. The problem of designing displays for color-defective viewers is discussed in Section 25.7.8.

Table 2. Frequencies of Occurrence for Color-Vision Deficiencies in the US and Europe (Wyszecki and Stiles, 1982, p. 464)

Deficiency	Males (%)	Females (%)	Total (%)	Male/Female Ratio
Protanomalous	1.0	0.02	1.02	50:1
Deuteranomalous	4.9	0.38	5.28	13:1
Protanope	1.0	0.02	1.02	50:1
Deuteranope	1.1	0.01	1.11	91:1
Tritanope	0.002	0.001	0.003	2:1
Rod monochromat	0.003	0.002	0.005	1.5:1
TOTAL	8.005	0.433	8.438	18:1

25.1.8 Perceptual Phenomena

The processing that the visual system performs on the achromatic and opponent-color signals is complex and gives rise to numerous interesting perceptual phenomena. Space constraints prevent discussing them all in detail, but ones that are especially apt to be observed when designing HCI displays are summarized below.

Abney effect. If achromatic light is mixed progressively with monochromatic light (i.e., light consisting of only one wavelength), the hue of the resulting color changes gradually, in most cases. This means that, for most hues, lines of constant hue plot as curved lines on chromaticity diagrams (discussed below in Section 25.3). This, in turn, means it is difficult to write a computer program that allows a color's saturation to be changed without affecting its hue, or that generates colors having differing saturations but constant hue.

Assimilation. The color of a background may shift toward the color of a pattern placed on it, especially if the pattern is repetitive and consists largely of high spatial frequencies. This effect is the opposite of simultaneous contrast (described below). Figure 5 provides an example. When designing displays that may produce assimilation, one must be prepared to increase the colors' saturations or otherwise increase the color difference to restore the intended amount of color contrast.

Bezold-Brücke effect. For most hues, large changes in luminance (defined below in Section 25.2.1) cause the hue to shift. The main reason seems to be that the yellow-blue opponent signal changes with luminance more rapidly than the red-green signal. Like the Abney effect, the Bezold-Brücke effect is difficult to correct for automatically in software.

Chromostereopsis. If highly saturated colors having widely different hues are viewed simultaneously (e.g., red and blue characters on a black background), the colors may appear to lie in different depth planes. This effect is optical in origin, rather than a product of visual-system processing, and has two sources: (1) the optical and visual axes of the eyes are not aligned; and (2) the directional orientations of the cones vary. Figure 6 shows an example. The perception of chromostereopsis varies widely across observers, so it is better ordinarily to avoid color combinations and patterns that may create it, rather than trying to make use of it as a display feature.

***Color afterimages** (known also as successive contrast).* Staring at a color may produce an afterimage having the opposite hue, particularly if the color is highly saturated. Figure 7 gives a demonstration. Most viewers dislike HCI displays that produce afterimages, so a common practice is to avoid the use of saturated colors for characters in text-processing programs and screen backgrounds in general.

Color constancy. In many cases, the colors reported by a person for reflecting objects—particularly ones in spatially complex scenes—are largely unaffected by changes in the spectral distribution of the illumination, although the viewer is usually aware that their appearances have changed. The viewer seems able to "discount" the illumination and determine the "true" colors. After a few minutes, the visual system adapts to the illumination and the appearance changes diminish. This effect is beneficial in HCI applications because it tends to reduce complaints about illuminant-induced color errors in hardcopy and screen displays

Simultaneous contrast. The color of a stimulus tends to shift away from the color of its background and toward the background's complementary color, providing enhanced contrast. Figure 7 gives several demonstrations. This effect can be beneficial in HCI applications.

Small-field** (or **threshold**) **tritanopia. As the visual angle subtended by a color is reduced below two degrees or so, its saturation tends to diminish, making its hue harder to discern. The effect's name reflects the

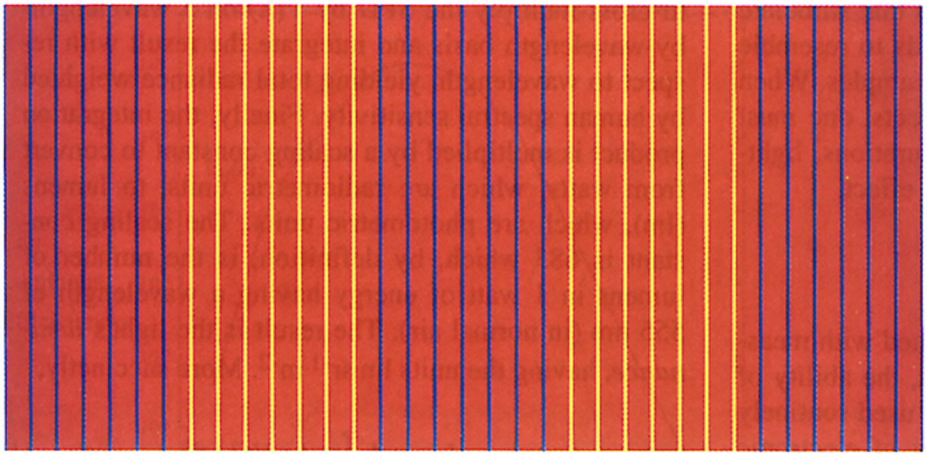

***Figure 5. Demonstration of assimiliation.** The red is invariant across the figure, but appears lighter behind the yellow stripes.*

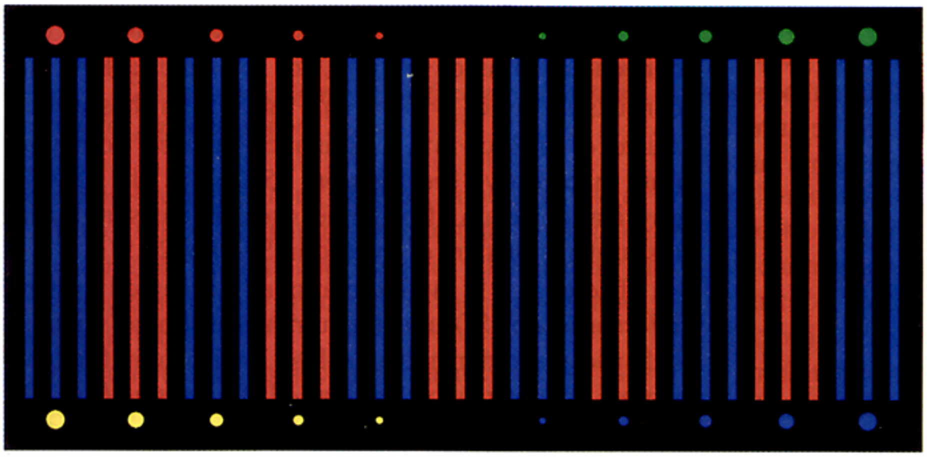

***Figure 6. Two demonstrations.** Chromostereopsis: most viewers will see the blue stripes as being farther in depth than the red stripes, some will see the opposite, and a few will see no difference. Small-field tritanopia: the colors of the dots become harder to discern as they become smaller; the effect is most pronounced for the blue dots.*

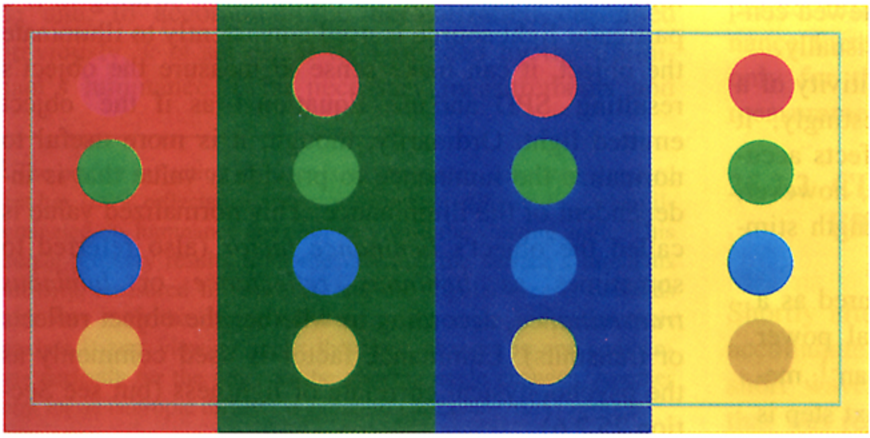

***Figure 7. Two demonstrations.** Color afterimages: stare at the figure under bright light for 30 seconds, then look at a white surface; an afterimage containing opponent hues should appear. Simultaneous contrast: the colors of the circles and the stripe are the same in all four quadrants, but change appearance on the different backgrounds.*

fact that it is more pronounced for colors that stimulate S-cones preferentially and therefore tends to resemble a tritan defect. Figure 6 shows several examples. When designing displays containing small objects, one must be prepared to increase the objects' saturations, lightnesses, or brightnesses to counteract this effect.

25.2 CIE Photometry

Photometry is the science that is concerned with measuring the visual efficacy of light, that is, the ability of light to produce a visual sensation. It is used routinely in HCI design to quantify the visibility of electronic displays and hardcopy. Photometry has its origins in the desire of scientists to measure brightness. It was recognized long ago that the human visual system does not respond at all to most of the electromagnetic spectrum, and even within the relatively narrow range of wavelengths that it does respond to, sensitivity varies with wavelength. Therefore, brightness is not simply a function of radiant energy—measuring it requires knowledge of how the various wavelengths are weighted by the visual system.

25.2.1 CIE Photopic Luminous Efficiency Function

Early in this century, an international standardizing body for the measurement of light, known as the *Commission Internationale de l'Éclairage (CIE)*, combined data from several researchers to produce a human spectral sensitivity function that was accepted as the standard by international agreement in 1924. This standard is the *CIE 1924 photopic luminous efficiency function, V(λ)*, and is illustrated in Figure 8. The CIE recommends $V(\lambda)$ for stimuli that are viewed centrally and subtend from one to four degrees visually.

$V(\lambda)$ represents the overall spectral sensitivity of a normal, trichromatic visual system. Interestingly, it represents people with deutan and tritan defects accurately, as well. Persons with protan defects, however, have much lower sensitivity to long-wavelength stimuli.

To use $V(\lambda)$, a light's radiance is measured as a function of wavelength, yielding its spectral power distribution (SPD) in units of watts·steradian^{-1}·meter^{-2}·nanometer^{-1} ($W \cdot sr^{-1} \cdot m^{-2} \cdot nm^{-1}$)[1]. The next step is to cross-multiply the SPD by $V(\lambda)$ on a wavelength-by-wavelength basis and integrate the result with respect to wavelength, yielding total radiance weighted by human spectral sensitivity. Finally, the integration product is multiplied by a scaling constant to convert from watts, which are radiometric units, to lumens (lm), which are photometric units. The scaling constant is 683 which, by definition, is the number of lumens in 1 watt of energy having a wavelength of 555 nm (in normal air). The result is the light's *luminance*, having the units $lm \cdot sr^{-1} \cdot m^{-2}$. More succinctly,

$$L = k \int_{\lambda} L_{\lambda} V(\lambda)\, d\lambda \ , \qquad (1)$$

where $k = 683\ lm \cdot W^{-1}$, L_{λ} is the SPD, and L is the resulting luminance. By definition, $1\ lm \cdot sr^{-1} = 1$ candela (cd), so the units are reported more commonly (and conveniently) as $cd \cdot m^{-2}$. So, for example, a light that consists solely of $1\ W \cdot sr^{-1} \cdot m^{-2}$ at 555 nm has a luminance of $683\ cd \cdot m^{-2}$. Luminance is used commonly (but incorrectly; see Section 25.2.6) as the psychophysical correlate of brightness and is measured using instruments that are discussed in Section 25.5.

The $cd \cdot m^{-2}$ is the internationally accepted Systéme Internationale (SI) unit for luminance. Some scientists still use the older and obsolete British unit for luminance, however, which is called the *footlambert (fL)*. The conversion is very simple: 1 fL = 3.43 $cd \cdot m^{-2}$. This conversion is explained below in Section 25.2.5.

Sometimes, an object that reflects or transmits light, rather than emitting it (e.g., a liquid-crystal display or hardcopy), must be characterized. In these cases, the object has no inherent luminance; instead, its luminance depends on the illumination striking it. If a particular light source is used consistently to illuminate the object, it can make sense to measure the object's resulting SPD and use Equation 1 as if the object emitted light. Ordinarily, though, it is more useful to normalize the luminance to provide a value that is independent of the illuminance. This normalized value is called the object's *luminance factor* (also referred to sometimes as *luminous reflectance* or *luminous transmittance*, according to whether the object reflects or transmits). Luminance factor is used commonly as the psychophysical correlate of lightness (but see Section 25.2.6).

The luminance factor of a reflecting object is the ratio of its luminance to that of the *perfect reflecting diffuser* under identical illuminating and measuring

[1] The steradian is a unit of solid angle and is the three-dimensional analogue of the radian. It is defined as the solid angle subtended at the center of a sphere by an area on the surface equal to the square of the radius. There are 4π steradians in a sphere.

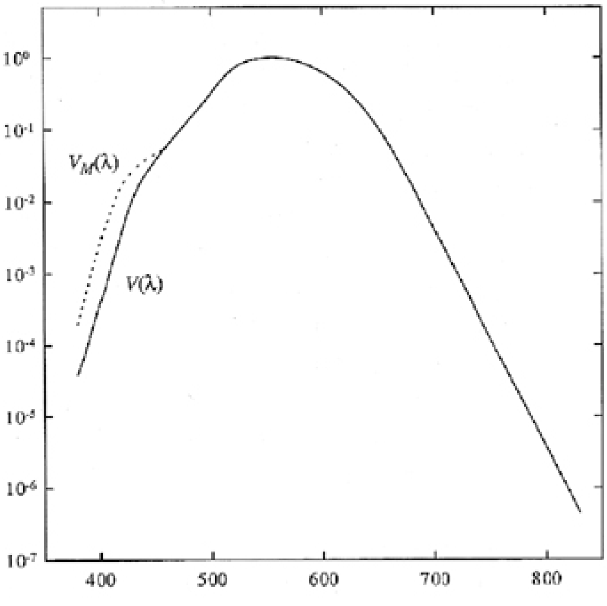

Figure 8. CIE photopic luminous efficiency function, $V(\lambda)$, and CIE 1988 modified 2° spectral luminous efficiency function for photopic vision, $V_M(\lambda)$.

conditions.[2] A transmitting object's luminance factor is the ratio of the object's luminance to that of the *perfect transmitting diffuser* under identical illuminating and measuring conditions.[3] Thus, luminance factors always range from 0 to 1.

To compute an object's luminance factor using Equation 1, the first step is to measure its spectral reflectance distribution or spectral transmittance distribution (as appropriate). The spectral distribution is multiplied by the SPD of an assumed illuminant (CIE standard illuminants C or D_{65} are typical choices[4]) on a wavelength-by-wavelength basis to determine the object's SPD under the illuminant. This SPD becomes L_λ and, to accomplish the normalization described previously, k is set equal to 1 divided by the illuminant's luminance. (The necessary measurements and calculations can be performed automatically by a spectrophotometer, as described in Section 25.5.2.) To avoid ambiguity, the illuminant used in the calculation must be specified whenever the resulting luminance factor is reported.

A simpler method is to measure the object's luminance under a given illuminant, substitute a white reflectance (or transmittance) standard for the object and measure the standard's luminance, and then divide the former by the latter. Ideally, the standard's reflectance (or transmittance) under the illuminant should be known, so the measured luminance can be corrected to yield the value that would have been obtained for the perfect reflecting (or transmitting) diffuser. The luminance factor that results from this procedure is valid only for the illuminant that was used to make the measurements.

25.2.2 CIE Modified Photopic Luminous Efficiency Function

Shortly after $V(\lambda)$ was introduced, evidence began to accumulate that it underestimates spectral sensitivity to short wavelengths. Judd (1951) derived a correction that was refined slightly by Vos (1978) and has since been accepted (CIE, 1990) as a supplement to $V(\lambda)$. This supplemental function is known as the *CIE 1988 modified 2° spectral luminous efficiency function for photopic vision*, $V_M(\lambda)$. Figure 8 compares $V(\lambda)$ with

---

[2] The perfect reflecting diffuser is an imaginary, idealized standard that has 100% reflectance and is an *isotropic diffuser*, i.e., when illuminated, its luminance does not vary with the viewing angle. This unique property results from the fact that the surface's luminous intensity, measured in candelas, and the angular area it subtends both vary with the cosine of the viewing angle and reach their maxima at zero viewing angle; therefore, changes in one compensate perfectly for the other as the viewing angle changes. Another term for an isotropic diffuser is *Lambert* (or *Lambertian*) *surface*.

[3] The perfect transmitting diffuser is an imaginary standard that has 100% transmittance and obeys the cosine law described in footnote 2.

[4] CIE standard illuminants represent phases of natural daylight and have SPDs that are defined officially in tables published by the CIE.

$V_M(\lambda)$, where it can be seen that the differences occur only at wavelengths below 460 nm. Therefore, $V(\lambda)$ and $V_M(\lambda)$ will yield significantly different results only for stimuli having a substantial proportion of their energy below 460 nm, in which case $V_M(\lambda)$ is arguably the appropriate choice. It remains to be seen whether $V_M(\lambda)$ will ultimately replace $V(\lambda)$ in general practice.

25.2.3 CIE Mesopic Photometry

CIE $V(\lambda)$ and $V_M(\lambda)$ are appropriate for luminances as low as 3 cd·m^{-2}. Between this level and 0.001 cd·m^{-2} lies the *mesopic* range, in which color vision operates but degrades progressively as luminance decreases. Because the changes are continuous, no single function can characterize mesopic vision; therefore, the CIE has no officially recommended mesopic luminous efficiency function. CIE (1989) offers several experimental approaches, but they are too complex for general use and are intended mainly for research on the problem.

25.2.4 CIE Scotopic Luminous Efficiency Function

At luminances below 0.001 cd·m^{-2} or so, only rods are operative and vision is possible only with the peripheral retina. For such cases, the CIE recommends a function that represents the sensitivity of the rods, called the *CIE 1951 relative scotopic luminous efficiency function (for young eyes), $V'(\lambda)$*. To compute luminance using this function, $V'(\lambda)$ is substituted for $V(\lambda)$ in Equation 1, k is set equal to 1700 lm·W^{-1}, and the result is referred to typically as *scotopic luminance* to distinguish it from the usual (photopic) case. Scotopic vision is necessarily colorless.

25.2.5 Illuminance

Luminance involves light leaving a surface, whereas illuminance involves light falling on a surface. The SI unit of illuminance is the lm·m^{-2} (called *lux* and abbreviated *lx*). The older and obsolete British unit, which is used sometimes, is the lm·ft^{-2} (called *footcandle* and abbreviated *fcd*). The conversion from one unit to the other is simple: since there are roughly 10.76 square feet in a square meter, 1 fcd = 10.76 lux.

An illuminance of 1 lm·m^{-2}, arriving at the surface of the perfect reflecting (or transmitting) diffuser, produces a luminance of $1/\pi$ cd·m^{-2}. It would be logical to suppose, therefore, that an illuminance of 1 fcd under the same circumstances produces a luminance of $1/\pi$ fL, but in fact the resulting luminance is 1 fL. This discrepancy results from the fact that the footlambert is defined as $1/\pi$ cd·ft^{-2}—a convention that was adopted to simplify converting from illuminance to luminance. Since the introduction of the SI units, however, this convention has generated confusion and errors because scientists who are accustomed to the 1 fcd ≡ 1 fL relation forget sometimes why it holds and assume a 1 lux ≡ 1 cd·m^{-2} relation.

The preceding discussion explains why 1 fL = 3.43 cd·m^{-2}. If not for the difference in the $1/\pi$ term, 1 fL would equal 10.76 cd·m^{-2} for the same reason that 1 fcd = 10.76 lux. Instead, however, 1 fL = $10.76/\pi$ cd·m^{-2} = 3.43 cd·m^{-2}.

The luminance of an object obviously does not depend on the distance from which it is measured. Illuminance, however, is related to the distance and angle between the illuminant and the measuring device (or illuminated surface). For a point source of illumination,

$$E = I \cos \varepsilon / r^2 , \tag{2}$$

where I is the source's luminous intensity in candelas, ε is the angle of incidence measured from the normal to the receiving surface, r is the distance in meters, and E is the resulting illuminance in lm·m^{-2}. (These units are correct, although the result appears to be cd·m^{-2}; see Wyszecki and Stiles, 1982, p. 266 for an explanation.) For extended sources, a commonly used rule is that Equation 2 produces an error less than 1% if the distance r is at least 10 times the largest maximum transverse dimensions of the source and receiving surface. Equation 2 is used frequently in such cases because the exact equations for extended sources are more complex and vary with the source's size and shape.

25.2.6 Luminance is not Brightness

Section 25.2 implied at the outset that $V(\lambda)$ provides a measure of brightness. The reader may have noticed, however, that the subsequent discussion has treated luminance and avoided further mention of brightness. The reason is that luminance and brightness are not the same, even when the more accurate $V_M(\lambda)$ function is used.

Brightness is a perception, which cannot be measured directly with instruments, whereas luminance is the psychophysical correlate, and the relationship be-

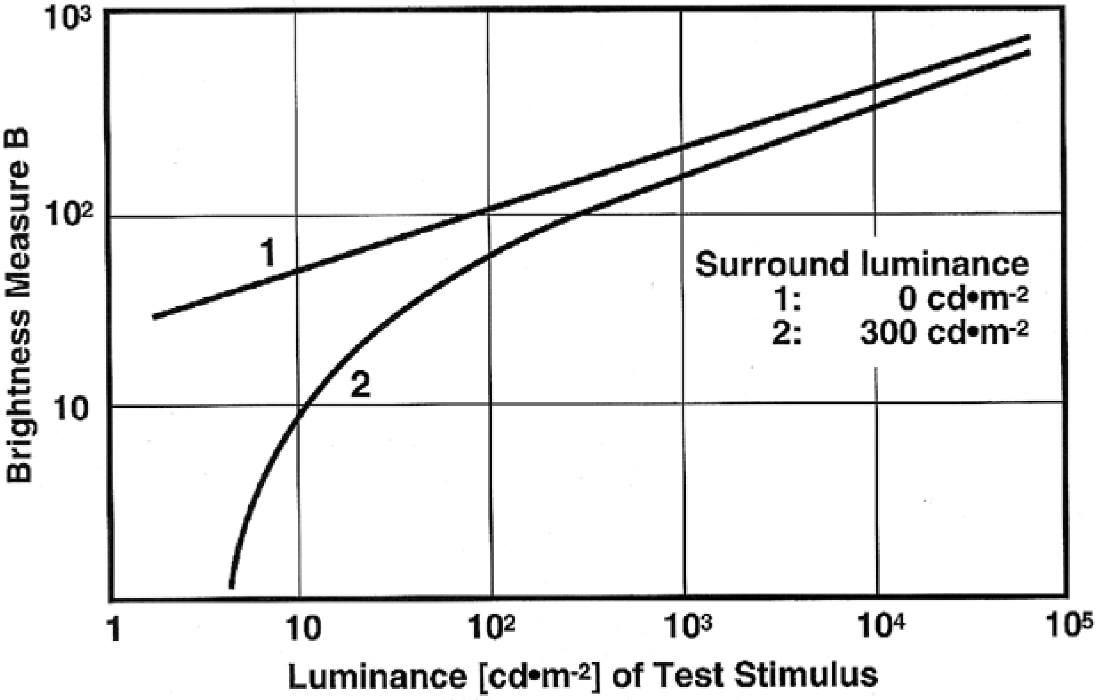

Figure 9. Brightness versus luminance of a stimulus viewed against two different surrounds. *From Wyszecki, G., and Stiles, W.S. (1982). Color science (2nd ed.). New York: Wiley. Copyright © 1982 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.*

tween the two tends to be nonlinear. Figure 9 shows brightnesses reported by observers as a function of luminance for a stimulus viewed against two different surrounds (note the use of log-log coordinates). It can be seen that the relation between the two is approximately a power function with an exponent of 1/3—at least, once the stimulus luminance surpasses the surround's. Therefore, doubling luminance, for example, produces less than a doubling of brightness, generally.

There is another way in which brightness and luminance differ. By definition, (photopic) luminance is a function of either $V(\lambda)$ or $V_M(\lambda)$, both of which were determined largely by a psychophysical method called *flicker photometry*. It is recognized now that this method eliminates contributions from the S-cones and, therefore, the resulting sensitivity functions reflect the L- and M-cones only. S-cones contribute significantly to the perception of brightness; hence, luminance cannot predict brightness accurately for stimuli having a substantial proportion of their energy at the shorter wavelengths of the visible spectrum (see Section 25.7.5 for a way to estimate luminances that will yield equal brightness for differing colors). On the other hand, S-cones contribute very little to visual acuity—probably because of their relative scarcity, as mentioned in Section 25.1.4—so luminance predicts visual acuity better than a more accurate psychophysical correlate of brightness would. Luminance predicts most other practical aspects of visual performance well also, so its use has continued and been widespread, even though it fails to meet its original purpose of correlating consistently with brightness.

There is one last point worth making: spectral sensitivity varies among observers, and $V(\lambda)$ and $V_M(\lambda)$ are averages; therefore, neither function predicts the spectral sensitivity of any specific person with complete accuracy. Ordinarily, though, it is impractical to determine and design for an individual's sensitivity, so it is necessary to rely on a reasonably accurate approximation. Experience has shown that $V(\lambda)$ and, particularly, $V_M(\lambda)$ serve this purpose adequately.

25.2.7 Practical Usage

Designers of display hardware often use the information that has been presented in Section 25.2 to make design predictions. For example, consider a hypothetical backlit transmissive display. It contains an illuminating system to which 5 watts of power can be delivered, has a luminous efficacy of 25 $\text{lm}\cdot\text{W}^{-1}$, measures 0.4 m^2, and delivers approximately uniform light to the display screen. The predicted illuminance on the screen is therefore (5 W · 25 $\text{lm}\cdot\text{W}^{-1}$ / 0.4 m^2 =) 313 $\text{lm}\cdot\text{m}^{-2}$. The screen measures 0.4 m^2, has a luminous transmittance of 0.5 relative to the illuminating system, and is

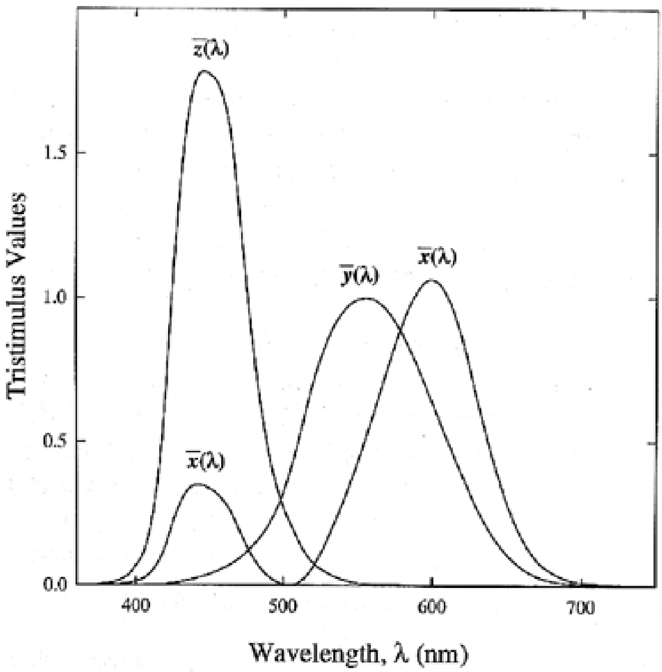

Figure 10. CIE 1931 $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ color-matching functions.

treated as a Lambert surface (discussed in footnote 1), so its predicted luminance is (313 $lm \cdot m^{-2} \cdot 0.5 / \pi =$) 50 $cd \cdot m^{-2}$. The same result is obtained for an emissive display having the same power, luminous efficacy, and area, if it is treated as an emitting Lambert surface with a non-diffusing faceplate that has 0.5 luminous transmittance.

As another example, consider a display under ambient illumination. The illuminant lies 2 m from the 20-cm display screen at an angle of 50 degrees, has a luminous intensity of 400 cd, and is treated as a point source. The screen is treated as a Lambert surface with a luminous reflectance of 0.1 relative to the illuminant. Using Equation 2, the predicted ambient illuminance on the screen (i.e., diffuse glare) is therefore (400 cd · cos 50^0 / $(2\ m)^2$ =) 64 $lm \cdot m^{-2}$ and the resulting screen luminance is therefore (64 $lm \cdot m^{-2} \cdot 0.1 / \pi =$) 2 $cd \cdot m^{-2}$. Section 25.3.7 shows how this result can be used to predict the illumination's impact on display color.

25.3 CIE Colorimetry and Color Spaces

Colorimetry is the science that is concerned with measuring the color-producing properties of light. Its development was motivated by the need in science and industry to have an objective, precise, and repeatable way to specify color. The most precise way of specifying a color is to show the spectral power distribution (SPD) of the light that produces it. This method is not very succinct, though, and it is inefficient because it overlooks metamerism; that is, it ignores the fact that different SPDs can produce identical colors.

25.3.1 CIE 1931 Standard Colorimetric Observer

In 1931, the CIE introduced a numerical method of color measurement and specification that takes advantage of the trichromacy of color vision and the metamerism that results. This method is based on color-matching experiments in which monochromatic lights were matched using mixtures of three monochromatic red, green, and blue primary lights. The CIE combined color-matching data from many different people and, for reasons that will be discussed momentarily, re-expressed them in terms of three imaginary primaries called ***X***, ***Y***, and ***Z***. The results, which are shown in Figure 10, are called the *CIE 1931 $\bar{x}(\lambda)$, $\bar{y}(\lambda)$*, and *$\bar{z}(\lambda)$ color-matching functions* (CMFs) and are referred to collectively as the *CIE 1931 standard colorimetric observer*. The CIE recommends its use for centrally fixated stimuli subtending one to four degrees visually.

The CIE 1931 $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ CMFs show how much of the ***X***, ***Y***, and ***Z*** primaries are needed to match any monochromatic light having unit radiance. (For example, referring to Figure 10, it can be seen that 1.06 units of ***X***, 0.63 units of ***Y***, and 0.00 units of ***Z*** are needed to match a monochromatic light having one unit of radiance at a wavelength of 600 nm.) Because the choice of radiance units is arbitrary, the CMFs allow the amounts of ***X***, ***Y***, and ***Z*** that are needed to match any monochromatic light to be computed, regardless of its radiance. Furthermore, any non-monochromatic light can be treated mathematically as a mixture of monochromatic lights. Consequently, the CMFs allow calculation of the amounts of ***X***, ***Y***, and ***Z*** that are needed to match any light at all. These three quantities, which are called *CIE 1931 X-, Y-*, and *Z-tristimulus values*, specify any color uniquely and precisely.

The ***X***, ***Y***, and ***Z*** imaginary primaries have two advantages over any set of real primaries that might be used instead. First, they can be mixed in positive amounts to match any real color. Second, the ***Y*** primary is defined so that it represents luminance only, so calculating an SPD's *Y*-tristimulus value yields the SPD's luminance. This definition was accomplished by making the $\overline{y}(\lambda)$ CMF the same as $V(\lambda)$. Because all luminance is contained in the Y primary, the X and Z primaries have no luminance at all. This observation underscores the imaginary nature of the X, Y, and Z primaries. Real light cannot have luminance only and zero X- and Z-tristimulus values, as Y does, nor can real light have zero luminance, as X and Z do. Furthermore, no finite set of real lights can be mixed in positive amounts to match all real colors. The X, Y, and Z primaries exist as mathematical concepts only and cannot be reproduced physically.

The CIE 1931 standard colorimetric observer represents the behavior of an imaginary, idealized person who has normal color vision that is representative of the average person and who performs the color-matching task with perfect consistency. Of course, no real person is perfectly consistent, and there are differences in color vision, even among persons whose color vision is classified as normal. Therefore, no real person will match colors in exactly the same way as the CIE standard observer. However, the standard observer provides a satisfactory approximation in most cases, as attested by the fact that it has survived intact for more than six decades.

25.3.2 Calculating and Using CIE 1931 Tristimulus Values

The calculation of X-, Y-, and Z-tristimulus values is analogous to the calculation of luminance, that is,

$$X = k \int_{\lambda} L_{\lambda}\, \overline{x}(\lambda)\, d\lambda\ , \quad (3)$$

$$Y = k \int_{\lambda} L_{\lambda}\, \overline{y}(\lambda)\, d\lambda\ , \text{ and} \quad (4)$$

$$Z = k \int_{\lambda} L_{\lambda}\, \overline{z}(\lambda)\, d\lambda\ , \quad (5)$$

where, for emitting objects, k and L_{λ} are as defined in Section 25.2.1; thus the Y-tristimulus value is equal to the object's luminance. For reflecting and transmitting objects, k is set equal to 100 divided by the illuminant's luminance; thus, the Y-tristimulus value is equal to 100 times the object's luminance factor and ranges from 0 to 100.

Knowledge of a color's tristimulus values allows anyone to produce a color that will be judged by most people to match the original reasonably well, given similar viewing conditions. This can be accomplished merely by assuring that the colors have the same tristimulus values, which is the same as assuring that they are metamers. It is not necessary to assure that the SPDs are the same, nor is it necessary to know the original color's SPD; only the color's tristimulus values are needed.

It is important to realize that, although the tristimulus values specify the requirements for a color match, they do not specify the resulting color perception. That is, there is no one-to-one correspondence between tristimulus values and colors. This is because color perception is subject to many influences besides the tristimulus values—as discussed and demonstrated in Section 25.1.8—so the same tristimulus values can produce different color perceptions under different viewing conditions. The CIE did not intend its colorimetric system to be used to predict color perceptions. It is only a method for specifying color by showing how to reproduce it.

25.3.3 CIE 1931 Chromaticity Diagram

It is often useful to transform X-, Y-, and Z-tristimulus values into numbers representing proportions among the tristimulus values. Let us define

$$x = X/(X + Y + Z), \quad (6)$$

$$y = Y/(X + Y + Z), \text{ and} \quad (7)$$

$$z = Z/(X + Y + Z). \quad (8)$$

The values of x, y, and z for a given color specify the proportions among X, Y, and Z that are needed to obtain a chromatic color match. That is, x, y, and z represent the purely chromatic aspects of color matching, independent of luminance. Notice that the values of x, y, and z always sum to 1. Therefore, in a three-dimensional space having x, y, and z as axes, the range of possible values for x, y, and z (i.e., the range where $x + y + z = 1$) defines a plane containing all possible chromaticities, both real and imaginary. The values of x, y, and z for a given color specify its location on this chromaticity plane and are referred to therefore as the color's *chromaticity coordinates*.

It is useful to have a diagram that shows the domain of real chromaticities and against which chromaticity coordinates can be referred. A diagram of this type can be produced easily by plotting the coordinates of the visible wavelengths. Because all chromaticities lie in the plane defined above, this diagram can be drawn in only two dimensions, that is, only two coordinates are needed. For example, if the coordinates x and y are plotted for the visible wavelengths and the endpoints (i.e., the points representing 360 and 830 nm) are joined by a straight line, the diagram shown in Figure 11 is produced. This diagram projects the real portion of the chromaticity plane onto the z plane.

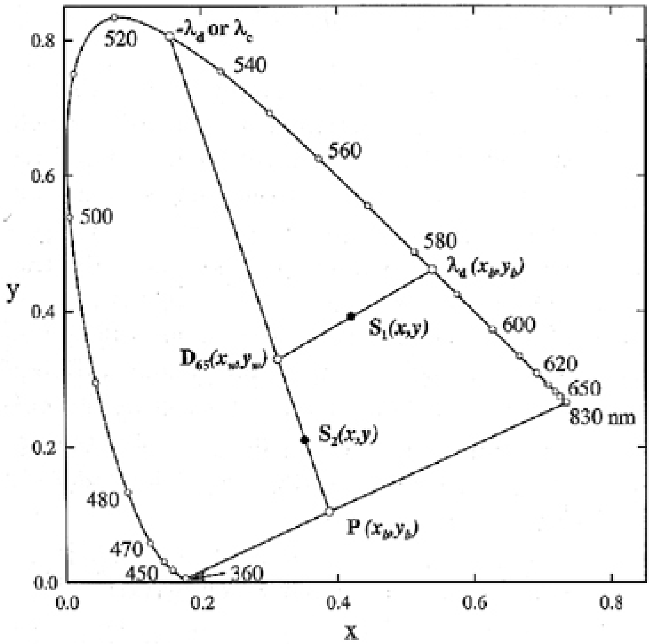

Figure 11. CIE 1931 (x,y)-chromaticity diagram.

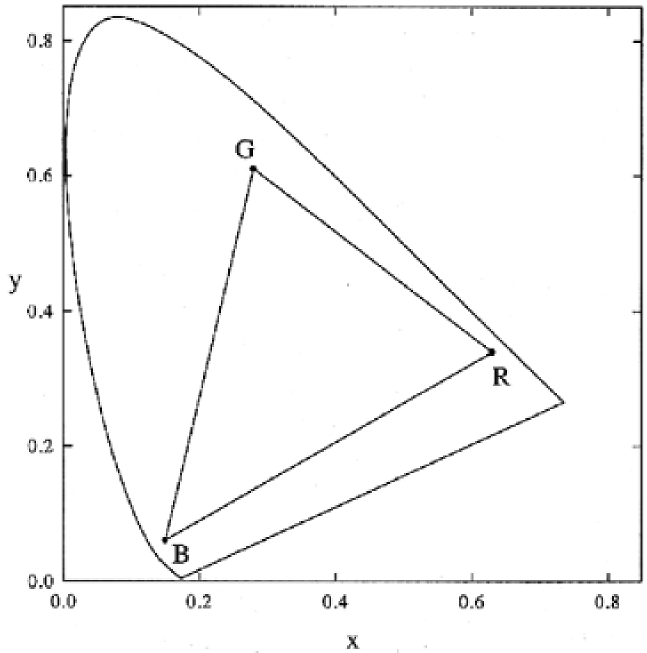

Figure 12. Typical color CRT monitor chromaticity gamut.

The diagram in Figure 11 is only one of an infinite number of projections that could be chosen, but it is convenient to choose one standard projection for universal use. The projection in Figure 11 has special significance because it is the standard that was chosen by the CIE. It is called the *CIE 1931 (x,y)-chromaticity diagram*. The curved line represents the visible wavelengths and is called the *spectrum locus* (because it is the locus of the spectrum). The straight line that closes the figure is called the *purple line*. All visible wavelengths lie on the spectrum locus, all pure purples (i.e., mixtures of 360 and 830 nm) lie on the purple line, and all other real colors lie somewhere in the interior.

A color can be specified completely by giving either its tristimulus values or its chromaticity coordinates and luminance. Since the chromaticity coordinates always sum to 1, only two need to be stated—by convention, x and y are used for this purpose. So, a color can be specified either in terms of X, Y, and Z or in terms of x, y, and Y.

The CIE 1931 (x,y)-chromaticity diagram has several useful properties. One is that all chromaticity coordinates that can be produced by mixing two primaries additively in positive amounts lie on a straight line between the coordinates of those two primaries. Similarly, all coordinates that can be produced by mixing three primaries lie on or within the triangle formed by the primaries on the diagram. (Figure 12, for example, shows the chromaticity gamut of a typical color CRT monitor. It was drawn by plotting the coordinates of the red, green, and blue channels and connecting them with straight lines.) This property generalizes to polygons formed by any number of primaries and is shared by all chromaticity diagrams.

The CIE 1931 (x,y)-chromaticity diagram can be used to define a quantity called *dominant wavelength* that correlates (imperfectly) with a color's hue. The dominant wavelength of a color is the wavelength of the monochromatic light that, when mixed in proper proportion with an achromatic light, matches the color. The achromatic light is defined typically as an equal energy source (this is usually the case if no explicit definition is stated) or CIE standard illuminant C or D_{65}.[5] Thus, in Figure 11, the dominant wavelength of S_1 is determined by drawing a straight line from the achromatic point (D_{65}, in this example) through S_1 to the spectrum locus at λ_d. The wavelength at λ_d is the dominant wavelength of S_1. Some colors, such as S_2 in Figure 11, have no dominant wavelength, but they can be mixed with a monochromatic light to match the achromatic light. The wavelength of this monochromatic light is the color's *complementary wavelength*. The complementary wavelength of S_2 in Figure 11 is the point at which a straight line, drawn from S_2 through the achromatic point, intersects the spectrum locus. This point can be denoted either $-\lambda_d$ or λ_c.

The chromaticity diagram can also be used to define a quantity called *excitation purity* that correlates

[5] An equal-energy source is one that has equal radiance at all wavelengths. Its (x,y)-chromaticity coordinates are both $=1/3$.

(imperfectly) with a color's saturation. The excitation purity of a color is the ratio of the color's distance from the achromatic point to the distance of the achromatic point from λ_d on the spectrum locus (or, for colors that have no dominant wavelength, the corresponding location on the purple line). This ratio is computed

$$p_e = (x - x_w) / (x_b - x_w) \text{ or, equivalently,} \quad (9)$$
$$= (y - y_w) / (y_b - y_w) \; , \quad (10)$$

where x and y are the color's chromaticity coordinates, x_w and y_w are the coordinates of the achromatic point, and x_b and y_b are the coordinates of the boundary point on the spectrum locus or purple line, as shown in Figure 11.

An alternative measure of saturation that is encountered sometimes is called *colorimetric purity*, which is defined

$$p_c = p_e \, y_b / y \; . \quad (11)$$

The preceding definition is the modern, officially sanctioned one (CIE, 1986). It can be useful to know that some of the literature on color vision and colorimetry uses an older definition for colors that have no dominant wavelength, however. In these cases,

$$p_c = (y_c / y)\,(x - x_w) / (x_c - x_w) \text{ or} \quad (12)$$
$$= (y_c / y)\,(y - y_w) / (y_c - y_w) \; , \quad (13)$$

where x_c and y_c are the chromaticity coordinates of the color's complementary wavelength.

25.3.4 CIE 1964 Supplementary Standard Colorimetric Observer and Chromaticity Diagram

For centrally fixated stimuli subtending four degrees visually or more, the CIE recommends the *CIE 1964 supplementary standard colorimetric observer*, often called either the *large-field* or *10-degree observer*. This observer consists of the $\bar{x}_{10}(\lambda)$, $\bar{y}_{10}(\lambda)$, and $\bar{z}_{10}(\lambda)$ CMFs, which are used to compute X_{10}-, Y_{10}-, and Z_{10}-tristimulus values just as X-, Y-, and Z-tristimulus values are computed using the 1931 observer. Figures 13 and 14 show the 1964 CMFs and associated *CIE 1964 (x_{10}, y_{10})-chromaticity diagram*. The definitions of dominant wavelength, excitation purity, and colorimetric purity for the 1964 chromaticity diagram are the same as the 1931 diagram, but with x_{10} and y_{10} chromaticity coordinates substituted, as appropriate.

It can be seen that the differences between Figures 13 and 14 and their 1931 counterparts (Figures 10 and 11) are not very large. The differences are due mainly to the fact that the 1931 observer was derived from two-degree stimuli, whereas the 1964 observer was derived from ten-degree stimuli.

The CIE has approved the use of the $\bar{y}_{10}(\lambda)$ color-matching function as a provisional substitute for $V(\lambda)$ for stimuli subtending more than four degrees (CIE, 1978a). Researchers use this substitution sometimes for peripherally viewed stimuli also, when the observers are light-adapted. In fact, in such cases, the entire 1964 observer is probably more appropriate than the 1931 observer. The CIE has not recommended any practices for peripheral stimuli, though, or, for that matter, any generally useful practices for stimuli subtending less than one degree.[6]

In most display-design applications, the stimuli subtend two degrees or less and either are or will be fixated centrally once the viewer attends to them, so the 1931 observer is the more appropriate choice. For characterizing larger or peripheral stimuli, though, the 1964 observer is appropriate.

25.3.5 CIE 1976 Uniform Chromaticity-Scale Diagram

Figure 15 illustrates color-matching data obtained by MacAdam (1942), plotted on the 1931 chromaticity diagram. The ellipses show the standard deviations of color matching at various locations on the diagram, multiplied by 10 to improve the figure's visibility. The fact that the standard deviations plot as ellipses of varying size, rather than as circles of constant size, shows that the 1931 diagram is not uniform perceptually. Therefore, the distance between two points on the diagram does not predict their perceived chromatic difference in any consistent way.

For cases where a perceptually uniform chromaticity diagram is desired, the CIE recommends a projective transformation of the 1931 diagram, called the *CIE 1976 (u',v')-uniform chromaticity-scale (UCS) diagram*. Because the transformation is projective, straight lines on the 1931 diagram remain straight on the UCS

6 The CIE has recommended luminous efficiency functions for brightness matching of monochromatic point sources and monochromatic fields subtending two and ten degrees (CIE, 1988).

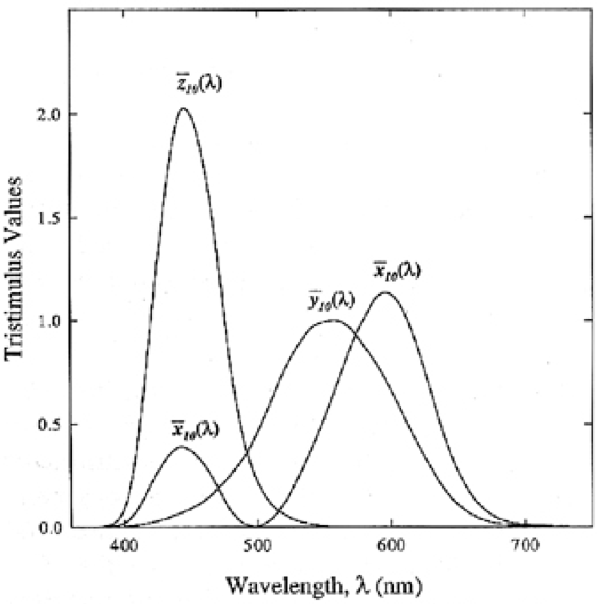

Figure 13. CIE $\bar{x}_{10}(\lambda)$, $\bar{y}_{10}(\lambda)$, and $\bar{z}_{10}(\lambda)$ color-matching functions.

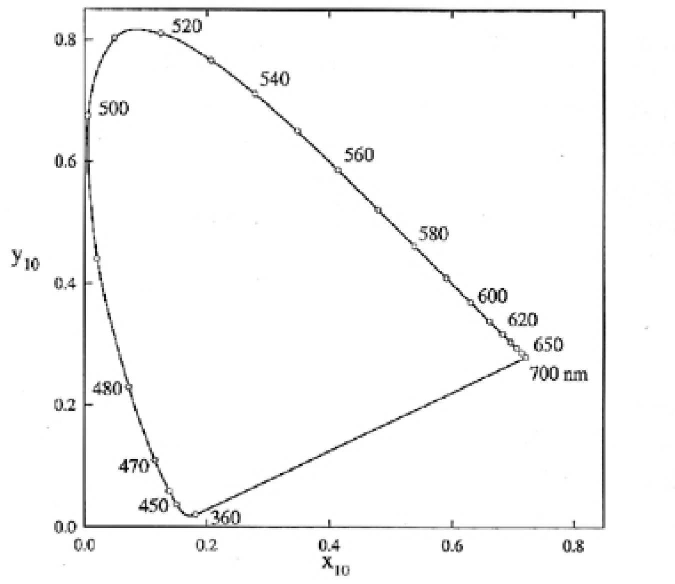

Figure 14. CIE 1964 (x_{10},y_{10})-chromaticity diagram.

diagram. This property simplifies the production of graphical representations of additive color mixtures, such as chromaticity gamuts (e.g., Figure 12).

The UCS diagram is illustrated in Figure 16 with MacAdam's (1942) ellipses replotted. Comparison with Figure 15 shows that the UCS diagram provides a small but useful improvement in perceptual uniformity, both by increasing the circularity of the ellipses and by reducing the variability of their sizes. The chromaticity coordinates of the UCS diagram are

$$u' = 4x / (-2x + 12y + 3) \quad (14)$$

or, using 1931 tristimulus values,

$$= 4X / (X + 15Y + 3Z)\,, \quad (15)$$

$$v' = 9y / (-2x + 12y + 3) \quad (16)$$

or, using 1931 tristimulus values,

$$= 9Y / X + 15Y + 3Z)\,,\ \text{and} \quad (17)$$

$$w' = 1 - u' - v' \quad (18)$$

Equations for performing the reverse transformations are

$$x = (27u'/4) / [(9u'/2) - 12v' + 9]\ ,\ \text{and} \quad (19)$$

$$y = 3v' / [(9u'/2) - 12v' + 9] \quad (20)$$

For stimuli subtending more than four degrees, the CIE recommends substituting X_{10}-, Y_{10}-, and Z_{10}-tristimulus values (or x_{10} and y_{10} chromaticity coordinates) for their counterparts in Equations 14-17, yielding coordinates denoted u'_{10}, v'_{10}, and w'_{10}.

The CIE recommends the UCS diagram for "comparisons of differences between object colors of the same size and shape, viewed in identical white to middle-gray surroundings, by an observer photopically adapted to a field of chromaticity not too different from that of average daylight" (CIE, 1986). An *object color* is one that is perceived as belonging to an object. The diagram is used routinely without regard to this restriction or the other guidance on observing conditions, though.

The 1976 UCS diagram replaces the older *CIE 1960 (u,v)-uniform chromaticity-scale (UCS) diagram.* The 1960 UCS diagram used the chromaticity coordinates $u = u'$ and $v = 2v' / 3$.

25.3.6 CIE Uniform Color Spaces

The CIE has adopted two systems that extend the notion of a perceptually uniform chromaticity diagram to the notion of a perceptually uniform three-dimensional color space; that is, one that includes an axis representing the luminance channel. The purpose is to provide color spaces based on the CIE 1931 system in which equal distance between colors produces equal color-difference perceptions.

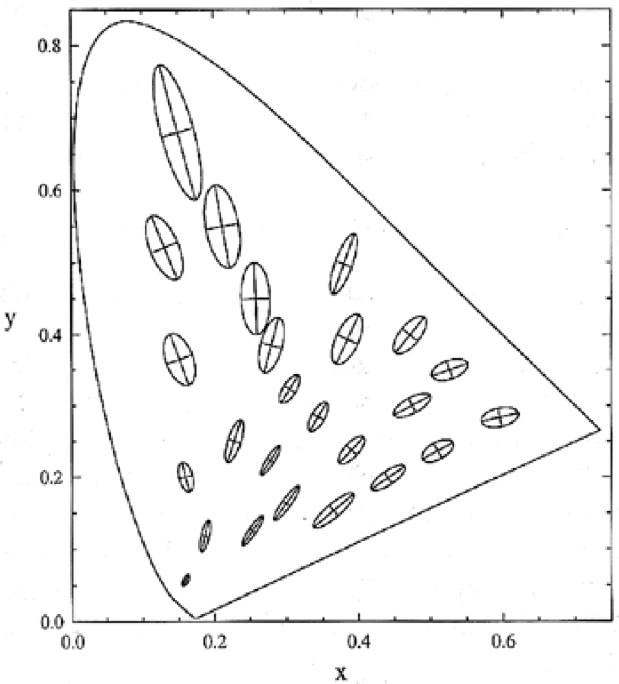

Figure 15. MacAdam's (1942) ellipses plotted on the CIE 1931 chromaticity diagram.

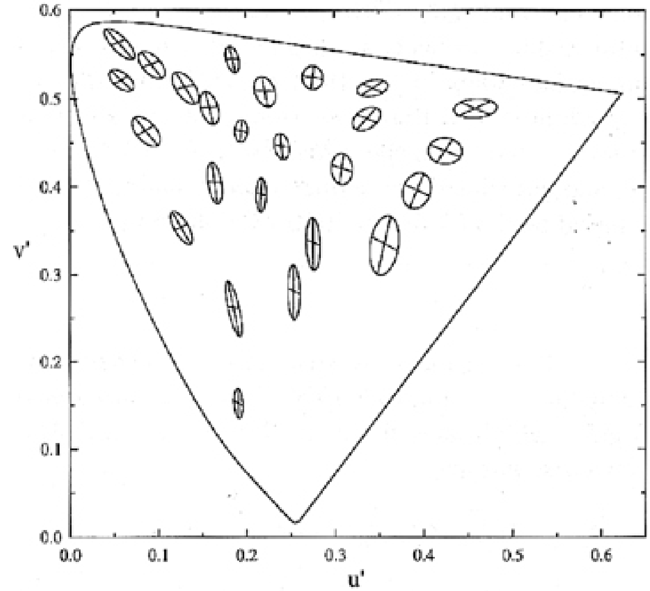

Figure 16. CIE 1976 (u',v')-uniform chromaticity-scale diagram with MacAdam's (1942) ellipses.

CIELUV

The first of the CIE uniform color spaces is a generalization of the 1976 UCS diagram, called the *CIE 1976 (L*u*v*) color space (CIELUV)*, having the axes

$$L^* = 116\,(Y / Y_n)^{1/3} - 16 \tag{21}$$

for $Y / Y_n \;\; > 0.008856$,

$$= 903.3\,(Y / Y_n) \tag{22}$$

for $Y / Y_n \;\; \leq 0.008856$,

$$u^* = 13L^*(u' - u'_n) \text{ , and} \tag{23}$$

$$v^* = 13L^*(v' - v'_n) \text{ ,} \tag{24}$$

where Y, u', and v' describe a given color, Y_n, u'_n, and v'_n describe a specified white object color (discussed at the end of this section), and the observing conditions are the same ones given above for the UCS diagram. The value of L^* is the color's *CIE 1976 lightness*. The measure of color difference between two colors is

$$\Delta E^*_{uv} \;=\; [(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2]^{1/2} \text{ ,} \tag{25}$$

where ΔL^* is the difference between the colors' L^* values, etc. Guidance concerning the use of ΔE^*_{uv} to ensure the adequacy of color differences is given in Section 25.7.1.

The CIELUV system includes psychophysical correlates of saturation, chroma, hue, and hue difference, which are called *CIE 1976 u,v saturation, CIE 1976 u,v chroma, CIE 1976 u,v hue-angle,* and *CIE 1976 u,v hue-difference,* respectively, and are defined, respectively,

$$s_{uv} \;=\; 13[(u' - u'_n)^2 + (v' - v'_n)^2]^{1/2} \text{ ,} \tag{26}$$

$$C^*_{uv} \;=\; (u^{*2} + v^{*2})^{1/2} \text{ ,} \tag{27}$$

$$= L^* s_{uv} \tag{28}$$

$$h_{uv} \;=\; \arctan\,[(v' - v'_n) / (u' - u'_n)] \text{ ,} \tag{29}$$

$$= \arctan\,(v^* / u^*) \text{ , and} \tag{30}$$

$$\Delta H^*_{uv} \;=\; [(\Delta E^*_{uv})^2 - (\Delta L^*)^2 - (\Delta C^*_{uv})^2]^{1/2} \text{ .} \tag{31}$$

By convention, h_{uv} lies between 0^0 and 90^0 if u^* and v^* are positive, between 90^0 and 180^0 if u^* is negative and v^* is positive, between 180^0 and 270^0 if u^* and v^* are negative, and between 270^0 and 360^0 if u^* is positive and v^* is negative. Notice that increasing (or decreasing) a color's Y-tristimulus value while holding its chromaticity coordinates constant leaves the color's s_{uv} unchanged but causes its C^*_{uv} to increase (or decrease). These features model the distinction between saturation and chroma that was described in Section 25.1.3.

For constant L^*, CIELUV provides a (u^*,v^*)-diagram (which is *not* a chromaticity diagram) in which straight lines in the 1931 or UCS diagrams remain straight. The CIELUV system replaces an older color space, known as the *1964 CIE (U*V*W*) color space*, and associated color-difference equation, which were based on the (defunct) 1960 UCS diagram.

CIELAB

The second uniform color space recommended currently by the CIE is the *CIE 1976 (L*a*b*) color space (CIELAB)*, which uses the same L^* axis as CIELUV but otherwise has axes

$$a^* = 500\,[f(X/X_n) - f(Y/Y_n)] \quad \text{and} \tag{32}$$

$$b^* = 200\,[f(Y/Y_n) - f(Z/Z_n)]\ , \tag{33}$$

where

$$f(X/X_n) = (X/X_n)^{1/3} \tag{34}$$

for $X/X_n > 0.008856$,

$$= 7.787\,(X/X_n) + 16/11 \tag{35}$$

for $X/X_n \leq 0.008856$,

$$f(Y/Y_n) = (Y/Y_n)^{1/3} \tag{36}$$

for $Y/Y_n > 0.008856$,

$$= 7.787\,(Y/Y_n) + 16/116 \tag{37}$$

for $Y/Y_n \leq 0.008856$,

$$f(Z/Z_n) = (Z/Z_n)^{1/3} \tag{38}$$

for $Z/Z_n > 0.008856$, and

$$= 7.787\,(Z/Z_n) + 16/116 \tag{39}$$

for $Z/Z_n \leq 0.008856$,

where X, Y, and Z describe a given color, X_n, Y_n, and Z_n describe a specified white object color (discussed below), and the observing conditions are the ones given previously for the UCS diagram. The CIELAB measure of color difference is

$$\Delta E^*_{ab} = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]^{1/2}\ , \tag{40}$$

where ΔL^* etc. have the same meanings as in CIELUV. *CIE 1976 a,b chroma*, *CIE 1976 a,b hue-angle*, and *CIE 1976 a,b hue-difference* are defined, respectively,

$$C^*_{ab} = (a^{*2} + b^{*2})\ , \tag{41}$$

$$h_{ab} = \arctan\,(b^*/a^*)\ , \text{and} \tag{42}$$

$$\Delta H^*_{ab} = [(\Delta E^*_{ab})^2 - (\Delta L^*)^2 - (\Delta C^*_{ab})^2]^{1/2}\ . \tag{43}$$

The conventions for h_{ab} are the same as for CIELUV's h_{uv}. CIELAB has no associated chromaticity diagram; therefore, CIELAB has no counterpart to CIELUV's s_{uv}. For constant L^*, CIELAB does provides an (a^*,b^*)-diagram, but straight lines on the (x,y)-, (u',v')-, and (u^*,v^*)-diagrams do not, in general, remain straight on (a^*,b^*)-diagrams.

CIE94

More recently, the CIE has recommended the *CIE 1994 (ΔL*ΔC*abΔH*ab) color-difference model (CIE94)* "when the size of the color difference can be considered small to moderate" (CIE, 1994). CIE94 uses

$$\Delta E^*_{94} = \left[\left(\frac{\Delta L^*}{k_L S_L}\right)^2 + \left(\frac{\Delta C^*_{ab}}{k_C S_C}\right)^2 + \left(\frac{\Delta H^*_{ab}}{k_H S_H}\right)^2\right]^{1/2} \tag{44}$$

as a replacement for Equations 25 and 40. For cases involving comparison against a color standard,

$$S_L = 1, \tag{45}$$

$$S_C = 1 + 0.045C^*_{ab}, \text{ and} \tag{46}$$

$$S_H = 1 + 0.015C^*_{ab}, \tag{47}$$

where C^*_{ab} is the standard's CIE 1976 *a,b* chroma. Otherwise,

$$S_C = 1 + 0.045(C^*_{ab,1}\,C^*_{ab,2})^{1/2}, \text{ and} \tag{48}$$

$$S_H = 1 + 0.015(C^*_{ab,1}\,C^*_{ab,2})^{1/2}, \tag{49}$$

(S_L is unchanged) where $C^*_{ab,1}$ is the first color's CIE 1976 *a,b* chroma, etc. The CIE assumes a specific set of viewing conditions, which includes 1000 lux of illumination from a source simulating standard illuminant D_{65}, a spatially uniform neutral background having $L^* = 50$, object-mode viewing, and spatially uniform colors that are immediately adjacent to each other, differ by 5 CIELAB units or less, and subtend a visual angle greater than four degrees. For these viewing conditions,

$$k_L = k_C = k_H = 1; \qquad (50)$$

otherwise, different values may be needed, in which case these values should be shown (for example) as $\Delta E^*_{94}(2:1:1)$.

The CIE (1994) says "the CIE 1976 $L^*a^*b^*$ and CIE 1976 $L^*u^*v^*$ recommendations as color spaces, and the recommended use of CIELUV for users who require a uniform chromaticity diagram, remain in effect." Thus, the CIELUV and CIELAB color spaces, along with their measures of chroma, saturation, etc., remain valid; only Equations 25 and 40 have been replaced. Although Equation 44 presumably models color-difference perception more accurately than Equations 25 and 40 for the assumed viewing conditions, most HCI design involves notable deviations from those conditions. The CIE (1994) has provided no guidance for treating deviations, so it is uncertain at the moment whether the CIE94 color-difference model will prove to be more accurate than its simpler CIELUV and CIELAB counterparts in such cases and be accepted widely among HCI practitioners.

Object Size

For objects that subtend more than four degrees visually, the CIE recommends substituting the 1964 supplementary standard colorimetric observer for the 1931 system, yielding quantities that are denoted by the subscript *10*, for example, L^*_{10}, $\Delta E^*_{uv,10}$, etc. Figures 17 and 18 show the CIELUV and CIELAB spaces, respectively, plotted with respect to the CIE 1964 observer and using CIE standard illuminant D_{65} as the specified white object color, that is, to define Y_n, etc. The closed shape at each figure's center represents the range of coordinates that can be produced by reflecting objects under D_{65}, whereas the outer edges represent the spectrum locus and purple line. (The figures thus also illustrate the fact that, for a fixed illuminant, reflecting objects can reproduce only a limited range of the visual system's color gamut.)

CIELUV vs. CIELAB

CIELUV has been more popular than CIELAB among workers concerned with electronic displays, whereas CIELAB has been more popular for applications involving reflecting and transmitting objects. CIELUV provides a chromaticity diagram, on which colors can be plotted independently of their L^* values and additive light mixtures (which are produced by most electronic displays) can be shown easily using straight lines. Otherwise, there is little practical basis for preferring one space over the other because most comparisons (e.g., Alman, Berns, Snyder, and Larsen, 1989; Ikeda, Nakayama, and Obara, 1979; Lippert, 1986; Lippert, Farley, Post, and Snyder, 1983; Mahy, Van Eycken, and Oosterlinck, 1994; Moroney and Fairchild, 1993; Pointer, 1981; Post, Costanza, and Lippert, 1982; Post, Lippert, and Snyder, 1983; Robertson, 1977) have failed to demonstrate substantial and consistent differences in their accuracies for predicting color-difference perception.[7] A main reason for the difference in modeling habits seems to be the mistaken idea—which appears in various articles and even some textbooks—that CIELUV is intended for modeling luminous sources and CIELAB is intended for modeling reflecting objects. This idea may have originated from recognition that CIELAB was derived from attempts to model a set of reflective color samples that represent the *Munsell color system* (see Section 25.4.2), whereas the UCS diagram (on which CIELUV is based) was derived from efforts to model the full range of chromatic vision (Wyszecki and Stiles, 1982, pp. 501-502). Wyszecki (1986, p. 9-47) has made it clear, however, that the CIE intended CIELUV and CIELAB for modeling reflecting objects exclusively and, as of 1986, had not addressed extensions to luminous sources. Regrettably, the CIE94 recommendations do not address luminous sources, either.

Application to Self-luminous Displays

The problem posed by CIELUV and CIELAB for self-luminous displays (e.g., color CRT monitors, backlit liquid-crystal displays, projection displays, etc.) concerns the definition of the "specified white object color" and, in particular, the definition of Y_n. For reflecting objects, the usual practice has been to equate the reference white with the perfect reflecting diffuser, illuminated by a CIE standard illuminant, as approved

---

[7] It is generally conceded, though, that CIELAB's modeling of chromatic adaptation is in better agreement with the visual system's true behavior—this is most relevant for modeling reflecting objects, which may be viewed under varying illuminants. See Kim, Berns, and Fairchild (1993) and Lo, Luo, and Rhodes (1996) for evidence on this point.

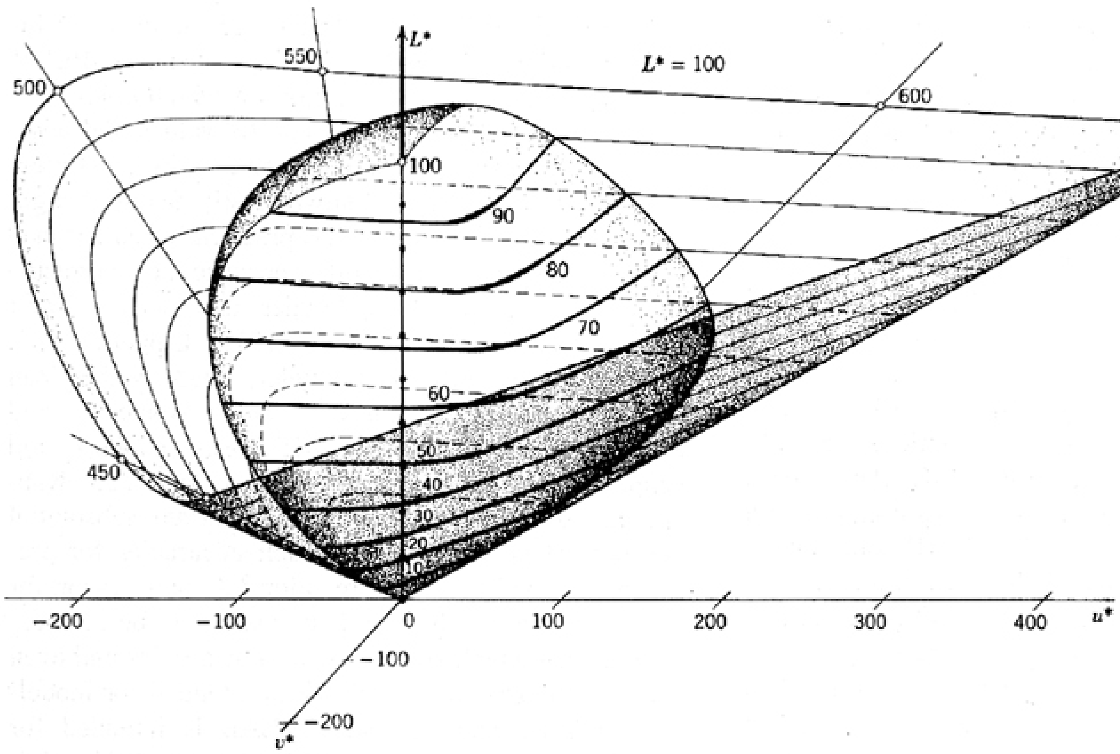

*Figure 17. CIE 1976 (L*u*v*) color space. From Judd, D.B., and Wyszecki, G. (1975). Color in business, science and industry (3rd ed.). New York: Wiley. Copyright © 1975 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.*

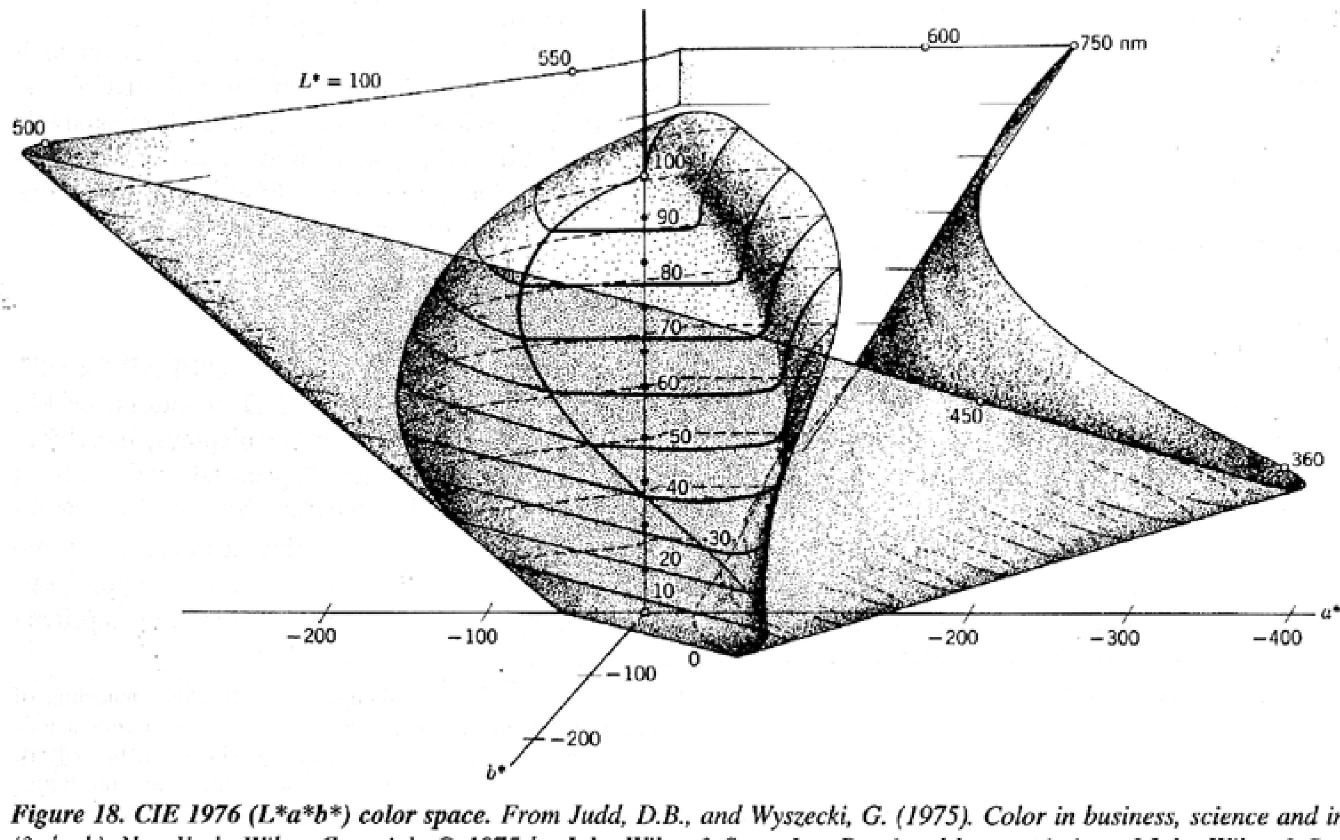

*Figure 18. CIE 1976 (L*a*b*) color space. From Judd, D.B., and Wyszecki, G. (1975). Color in business, science and industry (3rd ed.). New York: Wiley. Copyright © 1975 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.*

in CIE (1978b; superseded now by CIE, 1986). Some workers have used the actual illuminant or a white from the visual field, if a white was present, however. For self-luminous displays, the most popular convention has been the one suggested by Carter and Carter (1983), which equates Y_n with the luminance produced when the display's red, green, and blue channels are set to their maximum outputs and assigns the chromaticity coordinates of CIE standard illuminant D_{65} to the reference white. Some workers have used the coordinates produced by the display at its maximum output, which usually appear white, however.

Post (1984) noted, though, that the reference white is meant to represent the observer's state of adaptation, so the validity of equating it with a color that may not be visible is questionable. Furthermore, inconsistencies in displays' maximum luminances produce inconsistent CIELUV and CIELAB units under Carter and Carter's Y_n convention, which can lead to erroneous conclusions when attempts are made to generalize across displays. I argued that the issue is not whether the colors are truly luminous versus reflective, but whether they *appear* to be one or the other. Most vision researchers would agree, for example, that television images contain object colors typically. For modeling such cases, I recommended that a white be included in the visual field and used as the reference, regardless of whether the display is self-luminous or reflective. For modeling colors that appear luminous, I recommended dropping Y_n from the CIELUV and CIELAB equations and working in absolute units, as is conventional when dealing with colors that have luminance rather than lightness. More recently, CIE Technical Committee 1-27 has made recommendations for the object-color case that agree with mine (Alessi, 1994).

25.3.7 Practical Usage

The following relationships, which can be derived from Equations 6-8, are often useful when performing colorimetric calculations:

$$X = xY/y \,, \qquad (51)$$
$$Z = zY/y \,, \text{ and} \qquad (52)$$
$$X/x = Y/y = Z/z = X + Y + Z \qquad (53)$$

Two-color Mixtures

When computing additive color mixtures, the mixture's tristimulus values are equal to the sums of the constituent colors' tristimulus values. That is, if color A is an additive mixture of colors B and C, then

$$X_A = X_B + X_C \,, \qquad (54)$$
$$Y_A = Y_B + Y_C \,, \text{ and} \qquad (55)$$
$$Z_A = Z_B + Z_C \,, \qquad (56)$$

where X_A, Y_A, and Z_A are the tristimulus values of color A, etc. Equations 51-56 apply equally to computations involving the 1964 supplementary standard colorimetric observer.

Relationships equivalent with those in Equations 51-53 can be defined in terms of UCS chromaticity coordinates:

$$\overline{U} = u'Y/v' \,, \qquad (57)$$
$$\overline{W} = w'Y/v' \,, \text{ and} \qquad (58)$$
$$\overline{U}/u' = Y/v' = \overline{W}/w' = \overline{U} + Y + \overline{W} \,, \qquad (59)$$

where the quantities $\overline{U}$ and $\overline{W}$ are *not* UCS tristimulus values but are tristimulus values in a color space that uses Y to form one axis and has the UCS diagram as its chromaticity diagram. $\overline{U}$, Y, and $\overline{W}$ are a device that I find convenient when performing color-mixture calculations for colors given as u', v', and Y because they circumvent the need to transform into X, Y, and Z, do the sums, and then transform back into u', v', and Y. In the case of colors A, B, and C, for example, $\overline{U}_A = \overline{U}_B + \overline{U}_C$, etc. and $u'_A = \overline{U}_A/(\overline{U}_A + Y_A + \overline{W}_A)$, etc. Here too, the 1964 observer can be substituted.

Applying the preceding information to the example from Section 25.2.7 that involves a display under ambient illumination, assume that the UCS coordinates of the illumination are $u' = 0.2$ and $v' = 0.2$, the screen has uniform reflectance throughout the spectrum, and the display is producing $u' = 0.1$, $v' = 0.1$, and $Y = 10$. The u', v', and Y resulting on the screen from the mixture of the ambient illumination and display output are desired. Section 25.2.7 established that the illumination produces 2 $cd{\cdot}m^{-2}$ on the screen so, since the screen's reflectance is spectrally uniform (otherwise, the screen's tristimulus values relative to the illuminant would be needed), the values produced on the screen by the illumination are

$$\overline{U}_A = 0.2\,(2/0.2) = 2 \,, \qquad (60)$$
$$Y_A = 2\ cd{\cdot}m^{-2} \,, \text{ and} \qquad (61)$$
$$\overline{W}_A = 0.6\,(2/0.2) = 6 \qquad (62)$$

The values produced by the display are

$$\overline{U}_D = 0.1\,(10\,/\,0.1) = 10, \quad (63)$$
$$Y_D = 10\ cd{\cdot}m^{-2}\ \text{, and} \quad (64)$$
$$\overline{W}_D = 0.8\,(10\,/\,0.1) = 80 \quad (65)$$

The sums are

$$\overline{U}_T = 2 + 10 = 12\ , \quad (66)$$
$$Y_T = 2 + 10 = 12\ cd{\cdot}m^{-2}\text{, and} \quad (67)$$
$$\overline{W}_T = 6 + 80 = 86, \quad (68)$$

which yield the UCS chromaticity coordinates (the luminance is given in Equation 67)

$$u'_T = 12\,/\,(12 + 12 + 86) = 0.11\ , \quad (69)$$
$$v'_T = 12\,/\,(12 + 12 + 86) = 0.11\ . \quad (70)$$

Three-color Mixture

An example of color mixture that arises for self-luminous color displays is

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{vmatrix} x_R / y_R & x_G / y_G & x_B / y_B \\ 1 & 1 & 1 \\ z_R / y_R & z_G / y_G & z_B / y_B \end{vmatrix} \begin{bmatrix} Y_R \\ Y_G \\ Y_B \end{bmatrix} \quad (71)$$

where x_R, x_G, x_B, etc. are the 1931 chromaticity coordinates of the display's red, green, and blue channels, respectively, Y_R, Y_G, and Y_B are the channels' luminances, and X, Y, and Z are the tristimulus values that result on the screen. If the 3 x 1 vector of luminances is denoted $\boldsymbol{L}$, the 3 x 3 matrix of chromaticity coordinates is denoted $\boldsymbol{C}$, and the 3 x 1 vector of tristimulus values is denoted $\boldsymbol{T}$, then Equation 71 implies

$$\boldsymbol{L} = \boldsymbol{C}^{-1}\boldsymbol{T}\ . \quad (72)$$

That is, the luminances needed from the red, green, and blue channels to produce a desired set of tristimulus values can be calculated by multiplying the desired tristimulus values by the inverse of **C**. The same calculation suffices to decompose a known color (i.e., a set of displayed tristimulus values) into the display luminances that constitute it.

Equations 71 and 72 can be combined to determine the luminances needed on one display to duplicate the color produced by a given set of luminances on another display (i.e., to match colors across displays), for example,

$$\boldsymbol{L}_2 = \boldsymbol{C}_2^{-1}\boldsymbol{C}_1\boldsymbol{L}_1\ . \quad (73)$$

It is convenient sometimes to normalize the channel luminances so they range from 0 to 1. In this case,

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{vmatrix} X_{R,\,max} & X_{G,\,max} & X_{B,\,max} \\ Y_{R,\,max} & Y_{G,\,max} & Y_{B,\,max} \\ Z_{R,\,max} & Z_{G,\,max} & Z_{B,\,max} \end{vmatrix} \begin{bmatrix} \tilde{Y}_R \\ \tilde{Y}_G \\ \tilde{Y}_B \end{bmatrix} \quad (74)$$

where $X_{R,max}$ etc. are the tristimulus values of the red, green, and blue channels at their (individual) maximum outputs and $\tilde{Y}_R$ etc. are the normalized channel luminances, defined as $\tilde{Y}_R = Y / Y_{R,max}$, etc. Equations 72 and 73 hold for this case, given that the obvious substitutions are made. Equations 71-74 are valid also for computations made using the 1964 supplementary standard colorimetric observer and the $\overline{U}$, Y, $\overline{W}$ system.

25.4 Non-CIE Color Spaces and Systems

Alternatives to the CIE color spaces are so numerous that they cannot all be discussed in the space that is available for this chapter; therefore, attention is restricted to those that have the greatest current relevance to computer graphics. For this purpose, it is convenient to divide color spaces into two major categories: *device dependent* and *device independent*. The former use coordinates that relate to the specific hardware that implements them and have no consistent relationships with perception or coordinates in the CIE color spaces; the latter use coordinates that are intended to relate consistently with perception and often relate consistently with CIE coordinates.

25.4.1 Device-Dependent Color Spaces

RGB

The red, green, and blue luminances that appear on a display screen are determined by voltages that are applied to the display. On a computer-driven display, these voltages are related linearly to numbers that are stored in the graphics hardware. For convenience in general discussions, the numbers are often treated as

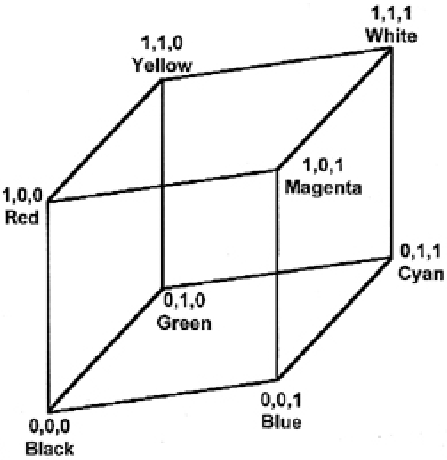

Figure 19. RGB additive-color space.

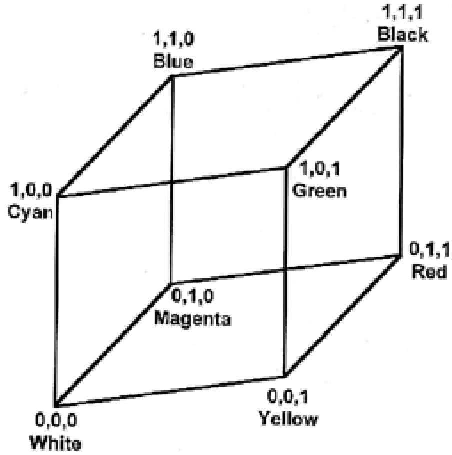

Figure 20. CMY subtractive-color space.

normalized values ranging from 0 to 1 and will be referred to here as *R, G,* and *B*. Color film recorders (which are used to produce photographic slides) also accept *RGB* inputs; scanners and high-end video cameras produce *RGB* outputs.

Figure 19, which depicts the left side of Table 1, illustrates the display color space defined by *R, G,* and *B*. If one views the colors associated with the corners of the figure on a display, they usually will match the appearances suggested by their labels, but the outcome depends on the display's state of adjustment and is not guaranteed. Further, if the colors are compared across different displays, their appearances are not apt to match because the displays will often respond differently to the voltages sent to them and the chromaticity coordinates of the displays' red, green, and blue primaries may differ.

CMY

Most color printing technologies use subtractive cyan, magenta, and yellow (*C, M,* and *Y*) primaries. If the numbers defining *C, M,* and *Y* are normalized to range from 0 to 1, the color space they form can be depicted as shown in Figure 20, which represents the right side of Table 1. Normalized *CMY* coordinates can be converted to normalized *RGB* coordinates and vice versa by the vector expressions

$$[R \;\; G \;\; B] = [1 \;\; 1 \;\; 1] - [C \;\; M \;\; Y] \text{ and} \tag{75}$$

$$[C \;\; M \;\; Y] = [1 \;\; 1 \;\; 1] - [R \;\; G \;\; B] . \tag{76}$$

The same sorts of limitations that were noted above concerning the appearance of colors defined in terms of *RGB* apply also to colors defined in terms of *CMY:* different printers may respond differently to the signals sent to them and may use different colorants and papers. Further, although Equations 75 and 76 convert between *CMY* and *RGB* coordinates, it is unlikely that the resulting colors will match.

CMYK

It can be difficult to achieve a good black by mixing *CMY* colorants, which tend to be more expensive than black ones. Therefore, some printing technologies use black as a fourth primary and conserve *CMY* by replacing them with black wherever possible by means of the expressions

$$K = min\,(C, M, Y), \tag{77}$$
$$C' = C - K, \tag{78}$$
$$M' = M - K \text{, and} \tag{79}$$
$$Y' = Y - K, \tag{80}$$

where *K* represents the normalized black coordinate and *C', M',* and *Y'* are the black-adjusted *CMY* coordinates.

HSV and HSL

The coordinates of the preceding color spaces are inconvenient for the adjustments people want to make ordinarily. For example, if a user wants to increase the saturation of a displayed yellow without affecting its color otherwise, examination of its *RGB* coordinates is not apt to make the required adjustments obvious. Two device-dependent color spaces are used commonly to address this problem: Hue, Saturation, and Value *(HSV)* and Hue, Saturation, and Lightness *(HSL),*

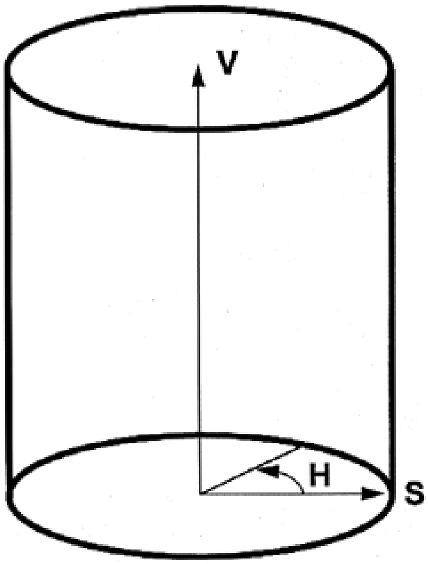

Figure 21. HSV color space.[8]

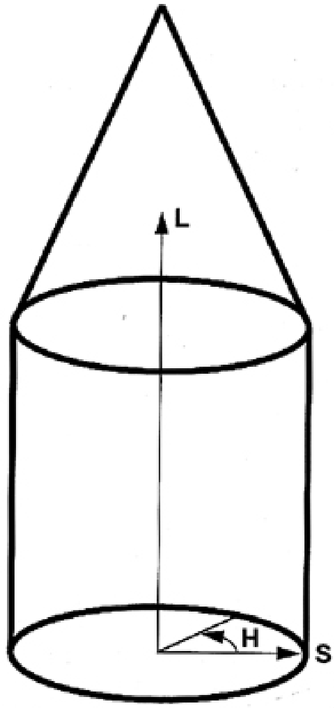

Figure 22. HSL color space.[8]

which are illustrated in Figures 21 and 22.[8] The *RGB* <—> *HSV* and *RGB* <—> *HSL* conversions are too lengthy to show here, but are given in Smith (1978) and Metrick (1979), respectively, as well as various texts (e.g., Foley, Van Dam, Feiner, and Hughes, 1990, pp. 590-595).

Both spaces use polar coordinates in which Hue is specified in degrees while Saturation and Value (or Lightness) range from 0 to 1. Grays lie along the central ($S = 0$) axis, with black at V or $L = 0$ and maximum white at V or $L = 1$. The main difference concerns the representation of Value versus Lightness: in *HSV*, $V = 1$ whenever R, G, or $B = 1$; in *HSL*, $L = 1$ only when $R = G = B = 1$. On most displays, $Y_G > Y_R > Y_B$ when $R = G = B$. Thus, each V plane in *HSV* contains the display's full chromatic gamut, but consists of colors having widely varying luminances. In *HSL*, luminance is inconsistent within each L plane, but the variance is smaller and, over the range $0.5 \geq L \geq 1$, the planes shrink to the display's maximum white.

Some people believe that *HSV* represents conventional thinking about color better than *HSL* and is therefore easier to use; others prefer *HSL* because it yields more consistent luminances and shows how the display's chromatic gamut shrinks at higher luminances. In neither case, however, do the axes correspond with constant hue, saturation, or lightness/brightness perception, so the objective of allowing one dimension to be adjusted while leaving the others unchanged is only approximated.

25.4.2 Device-Independent Color Spaces

Munsell

The Munsell color system is one of the oldest and most familiar device-independent color spaces. More specifically, it is a *color appearance system,* which means it was derived from perceptual scaling experiments and is meant to represent color in a perceptually uniform way. It uses a cylindrical arrangement, like *HSV*, with coordinates labeled Hue, Value, and Chroma and is exemplified by a physical standard called the *Munsell Book of Color*. Each page in the book is a constant Hue chart, with square color-sample chips arranged in rows and columns representing constant Value and Chroma, respectively. Each chip on a given row or column represents an equal perceptual step along its associated dimension and is identified by three numbers that identify its coordinates. For example, 5R 2/6 signifies a red having Hue 5R and lying 6 Chroma steps away from the central, neutral axis at Value 2. The perceptual spacing holds only for comparisons of chips against a middle-gray to white background under day-

[8] *HSV* is referred to and illustrated usually as a hexcone, or sometimes a cone, but it is really a cylinder because $S = 1$ whenever R, G, or $B = 0$. Similarly, *HSL* is shown usually as a double hexcone or double cone, but is actually a cylinder at $L < 0.5$ because, in this case also, $S = 1$ whenever R, G, or $B = 0$. Both spaces have a discontinuity at $R = G = B = 0$, where they collapse to a single point representing black.

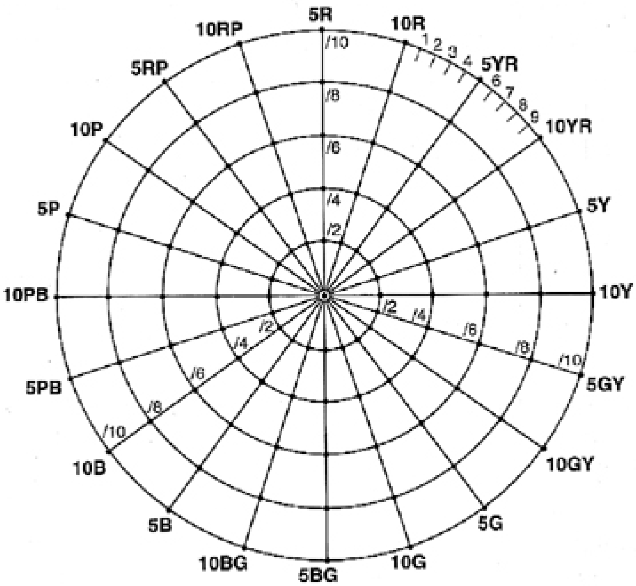

Figure 23. Munsell color space. From Wyszecki, G., and Stiles, W.S. (1982). Color science (2nd ed.). New York: Wiley. Copyright © 1982 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

time illumination. Figure 23 illustrates a constant Value plane in the Munsell space.

Examination of Figure 23 shows that, although the Hue spacing is constant perceptually at a given Chroma, the spacing increases with Chroma. Thus, the perceived difference between 10R 2/6 and 5R 2/6 is much greater than the difference between 10R 2/2 and 5R 2/2. Inconsistent Hue spacing is an unavoidable consequence of the system's cylindrical arrangement. Another noteworthy point is that the Hue charts do not include chips representing all 100 possible combinations of Value and Chroma because they cannot all be realized physically. For example, there can be only one chip representing Value 10 (i.e., a white having Chroma = 0) because addition of pigment to produce Hue or other Chromas would cause absorption of light that would reduce the Value.

Newhall, Nickerson, and Judd (1943) published a revised spacing for the original Munsell colors, which is known as the Munsell Renotation System and has since become the standard Munsell system. Their paper provides CIE 1931 x, y, and Y relative to CIE standard illuminant C for all Munsell colors with Values ranging from 1 to 9. Most lines of constant Hue plot as curves on the 1931 chromaticity diagram, while loci of constant Chroma have varying degrees of curvature. Value is related to Y by a somewhat complicated polynomial (Wyszecki and Stiles, 1982, p. 508) that is approximated usually by using Equation 21 (i.e., L^*). A computer program is available that converts among the Munsell, CIE 1931, CIELUV, and CIELAB systems, as well as DIN and NCS (described below) and others (Smith, Whitfield, and Wiltshire, 1990).

DIN

The Deutsches Institut für Normung (DIN) system (Richter, 1955) is another color appearance system, which was developed by the German Standards Association and is the official German standard color space. It is arranged in a conical configuration having coordinates labeled DIN-Farbton (hue), DIN-Sättingung (saturation), and DIN-Dunkelstufe (relative lightness). The colors constituting this system are represented in the DIN Color Chart, which shows their CIE 1931 x, y, and Y relative to CIE standard illuminant C. Lines of constant DIN-Farbton have constant dominant or complementary wavelength relative to CIE standard illuminant C and therefore plot as straight lines on the 1931 chromaticity diagram. Loci of constant DIN-Sättingung plot as ovals, however. DIN-Dunkelstufe for a given color is a logarithmic function of Y/Y_0, where Y_0 is the greatest luminance factor that can be realized physically for that chromaticity.

NCS

The Swedish Natural Color System (NCS; Hård, Sivik, and Tonnquist, 1996a, 1996b) is a color appearance system that was produced by the Swedish Standards Institution and has been gaining popularity internationally since it was introduced in 1979. It is arranged in a double cone (i.e., two cones, joined at their bases) with the coordinates hue (Φ), blackness (s), and chromaticness (c). Black and white appear at the apices of the cones; red, blue, green, and yellow appear along the perimeter of the joined bases at 45-degree increments with red opposite green and blue opposite yellow. Planes of constant s resemble Figure 23 but shrink as they move away from the central plane, toward black or white. The NCS and Munsell color spaces seem closely related and transformations between the two are fairly simple (Nickerson and Judd, 1975).

CNS

The Color Naming System (CNS) (Berk, Brownstone, and Kaufmann, 1982) uses standardized English words to denote colors according to their hue, saturation, and lightness or brightness. It was developed for computer graphics use and was derived from a more complex naming system introduced by the U.S. National Bureau of Standards (known now as the National Institute of Standards and Technology) and the Inter-Society Color Council (Kelly and Judd, 1955). The CNS provides six basic hue names (red, orange, yellow, green, blue, and purple), plus brown, which can be substituted for orange. These names can be paired and used with "ish" as a suffix to denote 31 hues. Four levels of saturation (grayish, moderate, strong, and vivid) and five levels of lightness or brightness (very dark, dark, medium, light, and very light) are provided as modifiers, and there are seven achromatic hue names (black, very dark gray, dark gray, gray or medium gray, light gray, very light gray, and white), yielding (31 x 4 x 5 + 7 =) 627 possible color names. Of these, 480 are associated with specific Munsell chips and therefore defined colorimetrically, but the orange and brown names denote identical chips, so the number of unique chip correspondences is 340.

Pantone

The Pantone Matching System is used widely in commercial art for specifying colors by referencing physical standards. The main reference is the Pantone Color Selector: a book that shows 1012 colors that can be produced using the Pantone-licensed inks. Many more Pantone colors are available, however, and are documented in additional, special-purpose references that focus, for example, on metallic colors, pastels, and tints. All of the colors can be reproduced accurately by commercial printers, providing they use Pantone-licensed inks. Pantone has licensed a color management system (see Section 25.6) to several applications-software vendors, printer vendors, and a CRT monitor vendor, to allow the use of its system in integrated, color-calibrated computer graphics environments.

YIQ* and *YUV

Color television broadcasting in the US converts *RGB* signals from video cameras into a color space defined by the National Television Standards Committee (NTSC) standard, using the transformation

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{vmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.212 & -0.523 & 0.311 \end{vmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (81)$$

where Y is a luminance signal, I is a red-cyan opponent signal, and Q is a green-magenta opponent signal. Europe uses the Phase Alternate Line (PAL) and Sequential Couleur á Memoire (SECAM) standards, both of which use the transformation

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{vmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.100 \end{vmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (82)$$

where U is a blue-yellow opponent signal and V is a red-cyan opponent signal (note that Y is the same as NTSC).

The strategy of separating the luminance and chrominance information has two advantages. First, monochrome displays need decode only the Y signal (this was important when color broadcasting was introduced because, otherwise, it would have been incompatible with the existing base of monochrome receivers). Second, transmission bandwidth can be conserved by allocating more to the luminance signal and less to the chrominance signals. Experience shows that, providing the bandwidths are adequate, most viewers do not notice the resulting relative loss of spatial chrominance modulation. One reason for this tolerance is the fact that chrominance modulation is accompanied by luminance modulation in most natural images.

One can debate whether *YIQ* and *YUV* are truly device-independent because they meet this criterion only if the displays and cameras have identical *RGB* primaries and white points.[9] This requirement implies matching the standards shown in Table 3, but many cameras and most displays deviate from these standards. (Contemporary television-phosphor standards for the US, Europe, and high-definition television are shown in Table 4; note the differences from the broadcast standards shown in Table 3. Many displays deviate from the Table 4 standards, too.) In such cases, conversions of the form shown in Equation 73 must be used to convert between the camera and display *RGB* primaries (after rescaling to account for the change in white point, if necessary) if accurate transformation is desired. Conversion from *YIQ* and *YUV* to *RGB* can be accomplished by inverting Equations 81 and 82, in the manner illustrated in Equation 72. Transformations to and from CIE *XYZ* can be made also, using the information in Tables 3 and 4.

---

[9] The white point of a display is the chromaticity produced when it receives equal *RGB* input voltages.

Table 3. NTSC, PAL, and SECAM Chromaticity Coordinates.

System	x	y	u'	v'
NTSC				
R	0.67	0.33	0.48	0.53
G	0.21	0.71	0.08	0.58
B	0.14	0.08	0.15	0.20
White (C)	0.3101	0.3162	0.2009	0.4609
PAL				
R	0.64	0.33	0.45	0.52
G	0.29	0.60	0.12	0.56
B	0.15	0.06	0.18	0.16
White (D_{65})	0.3127	0.3290	0.1978	0.4683
SECAM	Same as PAL with White = CIE std. illuminant C			

YCbCr

The Joint Photographic Experts Group (JPEG) and Motion Picture Experts Group (MPEG) digital image-compression standards use a color space based on *YUV*. It uses the NTSC definition of *Y*, the PAL chromaticities, and defines the opponent-color signals as

$$C_b = (U/2) + 0.5 \text{ , and} \tag{83}$$
$$C_r = (V/1.6) + 0.5 \text{ .} \tag{84}$$

Here again, separation of luminance and chrominance information and emphasis on preserving the former is used to advantage. Due partly to this strategy, JPEG achieves compression ratios better than 2:1 (20:1 is not unusual) for color images in its lossy modes (i.e., modes in which the reproduction is allowed to degrade relative to the original), the exact ratio depending on the user's willingness to sacrifice fidelity. MPEG achieves roughly three times greater compression by taking advantage of spatial correlations among successive frames of typical moving images (Pennebaker and Mitchell, 1993, pp. 21 and 253-258).

25.5 Color Measurement Devices

Many commercial instruments that perform the measurements needed for CIE colorimetry are available.These instruments form three categories: spectro-

Table 4. Contemporary Standards for Television-Phosphor Chromaticity Coordinates

System	x	y	u'	v'
SMPTE C (US)				
R	0.63	0.34	0.43	0.53
G	0.31	0.60	0.13	0.56
B	0.16	0.07	0.18	0.18
EBU (Europe)				
Same as PAL—see Table 3				
CCIR 709 (High-Definition Television)				
R	0.64	0.33	0.45	0.52
G	0.30	0.60	0.13	0.56
B	0.15	0.06	0.18	0.16

radiometers, spectrophotometers, and filter colorimeters. Colorimetric instruments also provide photometric measurements, of course. Instruments that are designed solely for photometry (e.g., photometers and illumination meters) are simplified versions of colorimetric instruments, basically. Color scanners resemble filter colorimeters but incorporate cost-saving design compromises that complicate efforts to obtain accurate CIE values.

25.5.1 Spectroradiometers

Spectroradiometers are used to measure spectral power distributions (SPDs). Light from the target to be measured is gathered by optics and dispersed into a spectrum by a prism or diffraction grating. In a *scanning spectroradiometer*, the spectrum is sampled by a slit, which allows only a narrow range of wavelengths (e.g., 1 or 5 nm) to pass through and illuminate a photosensor. An SPD is obtained by alternately recording the photosensor's signal and moving to sample the adjacent portion of the spectrum, starting at one end of the spectrum and ending at the other. Afterwards, the SPD is corrected by multiplying it by the instrument's spectral calibration function on a wavelength-by-wavelength basis, thereby compensating for imperfections in its spectral sensitivity and yielding the final, calibrated SPD.

A newer type of spectroradiometer is illustrated in Figure 24. The slit and photosensor are replaced by a multi-element photosensor, the elements of which are arranged in a row that is aligned with the spectrum. Thus, all the elements are illuminated simultaneously

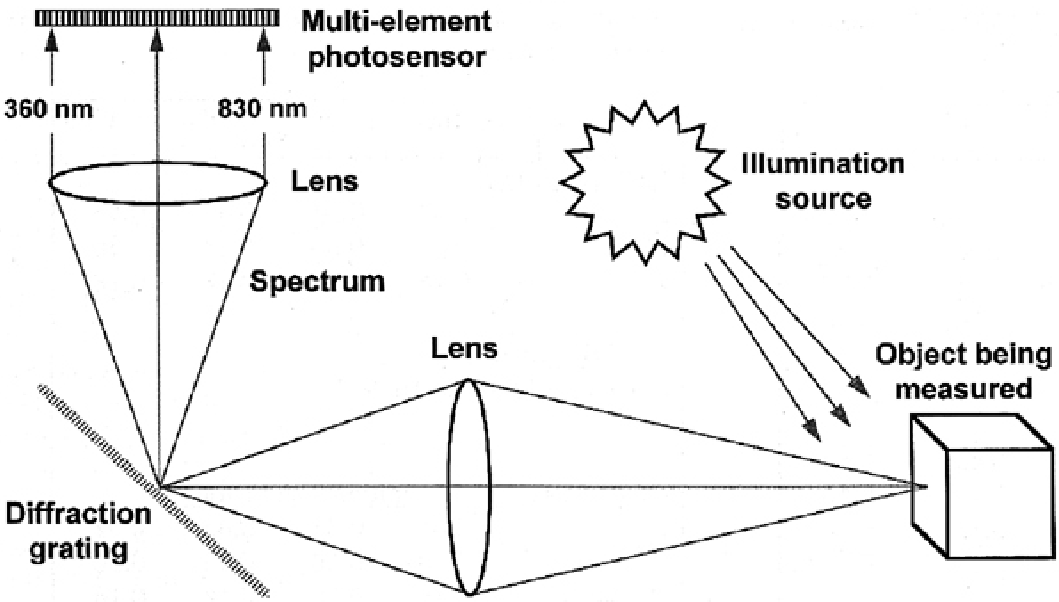

Figure 24. Simple schematic of a spectroradiometer that uses a multi-element photosensor.

and each element samples a narrow range of wavelengths, so the entire spectrum is measured at once without any moving parts. This arrangement tends to be faster than the scanning approach and eliminates wavelength errors caused by random variations in the positioning system.

Most modern spectroradiometers are automated and either include a computer and software or are designed to operate under the control of manufacturer provided software that executes on a user-provided computer. Thus, to obtain a measurement, the user needs only to assure that the light to be measured is being sampled correctly by the optics and issue the appropriate commands to the software. The instrument then performs the measurement, computes tristimulus values, etc., and displays the results. All commercial systems of which I am aware provide only the CIE 1931 color-matching functions (CMFs), though, and do not include the 1964 CMFs, $V_M(\lambda)$, or $V'(\lambda)$ as options; therefore, users who desire these alternatives must write custom software that reads the SPDs produced by the instrument and performs the necessary calculations.

The instrument's spectral calibration function is obtained by measuring a *standard light source*, which has a known SPD, and dividing the true SPD by the measured SPD on a wavelength-by-wavelength basis. For instruments that use a photomultiplier tube as the photosensor, the spectral calibration function must be redetermined every hour or so because the tube's spectral response tends to drift; instruments that use a solid-state photosensor are stable for months usually, but tend to be less sensitive.

Most users purchase a standard light source from a commercial vendor and rely on the vendor to provide the true SPD and perform occasional recalibration and maintenance. The true SPD is determined by comparison against a standard that is maintained by or traceable to a national standardizing body, such as the U.S. National Institute of Standards and Technology (NIST, known formerly as the National Bureau of Standards).

25.5.2 Spectrophotometers

Spectrophotometers are used to measure spectral reflectance distributions and spectral transmittance distributions; they can use either the sequential-scanning approach to measurement or the multi-element photosensor approach. The main differences between a spectroradiometer and a spectrophotometer are that the latter includes a chamber in which the target is placed, plus a lamp to illuminate the chamber. Light from the lamp illuminates the target and the target's SPD is measured with a photosensor. Next, the process is repeated while measuring a white comparison standard. Finally, the target's SPD is divided by the standard's SPD and multiplied by the standard's calibrated spectral reflectance (or transmittance) distribution on a wavelength-by-wavelength basis, yielding the target's spectral reflectance (or transmittance) distribution.

The lamp's SPD must be continuous, but need not be known to obtain a calibrated measurement because it cancels out when the target's SPD is divided by the standard's. The standard's spectral reflectance or transmittance distribution is needed, however. This in-

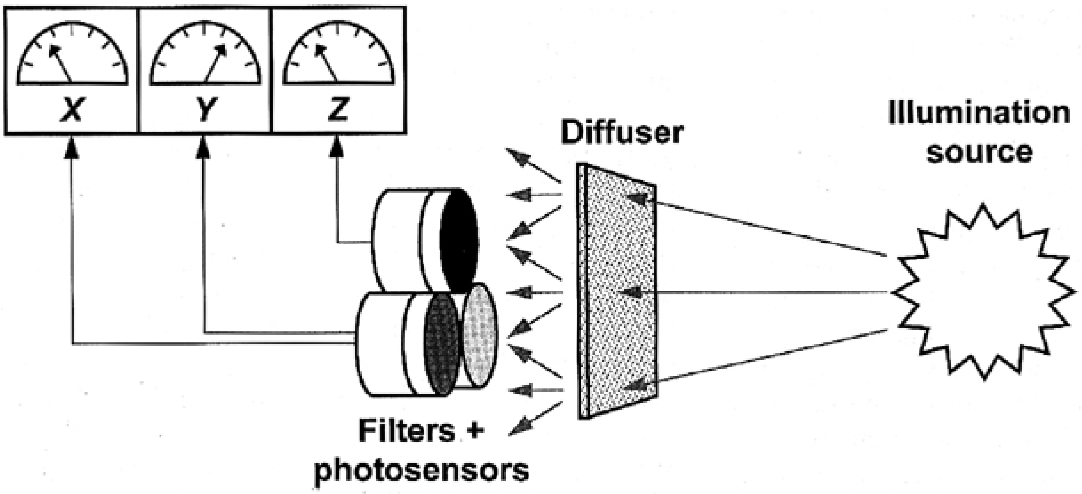

Figure 25. Simple schematic of a filter colorimeter that incorporates a diffuser.

formation is obtained by purchasing a standard that has a known spectral distribution and replacing it occasionally, to compensate for shifts in its spectral characteristics that occur unavoidably over time. Most modern spectrophotometers use solid-state multi-element photosensors, rather than photomultiplier tubes. The resulting loss of sensitivity is unimportant because the lamp's radiance can be increased to maintain a high measurement signal-to-noise ratio, when necessary.

25.5.3 Filter Colorimeters

Filter colorimeters yield CIE tristimulus values directly, without determining the target's SPD. The design illustrated in Figure 25 uses a set of three colored filters, each of which is paired with its own photosensor. Each filter/photosensor pair (FPP) has a spectral sensitivity that matches one of the CIE CMFs. Thus, the FPPs perform the spectral weightings and integrations called for in Equations 3-5 and, when the outputs are scaled properly, yield tristimulus values. A cheaper, alternative design uses only one photosensor. The filters are placed in a rotating wheel or some other moving system and the tristimulus values are obtained sequentially.

In most colorimeters, the FPPs are placed behind a diffuser, but other light-gathering methods can be used. (Photometers, for example—which are merely filter colorimeters containing only one FPP—typically include a lens that focuses an image of the target onto the FPP.) In colorimeters that are intended specifically for measuring self-luminous displays, the FPPs and diffuser are contained in a housing that is designed to be placed directly against the display's faceplate and positions the FPPs at a known, (fairly) consistent distance from the emitting surface. Thus, the luminance can be estimated from the measured illuminance.

Colorimeters that are designed for measuring reflecting or transmitting objects are similar to spectrophotometers. They include a lamp in the housing that contains the FPPs, to provide a constant set of illuminating tristimulus values that can be determined by measuring a white comparison standard. The housing assures a consistent measuring geometry and prevents external light from reaching the target. Thus, dividing the FPP outputs obtained for the target by the outputs obtained for the standard yields the target's tristimulus values under the built-in lamp. Some units attempt to convert these values to their equivalents under another illuminant (e.g., CIE standard illuminant C or D_{65}), but this yields inaccurate results because no mathematically valid conversion exists. A simple example proves this claim: the instrument will yield identical tristimulus values for two colors that are metamers under the colorimeter's built-in illuminant, both before and after conversion, even if they are not metamers under the target illuminant.

One important design issue for filter colorimeters concerns the CIE $\bar{x}(\lambda)$ CMF, which is difficult to duplicate with FPPs because it has two peaks: one at 442 and another at 599 nm (see Figure 10). Two different techniques are used to deal with this problem. The cheaper, but less desirable, technique takes advantage of the fact that the $\bar{z}(\lambda)$ CMF resembles the shorter-wavelength portion of the $\bar{x}(\lambda)$ CMF; therefore, if the output from the Z photosensor is scaled properly and added to the output from an FPP having spectral sensitivity that matches only the longer-wavelength portion of the $\bar{x}(\lambda)$ CMF, an approximation of X can be obtained. Thus, the Z FPP does double duty. The better design technique uses one FPP that matches the shorter-wavelength portion of the $\bar{x}(\lambda)$ CMF, plus another FPP that matches the longer-wavelength portion, and sums their outputs to obtain *X*. This approach is more expensive because four FPPs are needed.

Another design issue concerns the filters. One approach is to use single- or multi-layer filters; another is to build a mosaic filter from differently colored pieces of material that have the desired composite spectral transmittance. Accuracy can be increased by customizing each filter according to the spectral sensitivity of the specific photosensor it is paired with; this is easier to do with mosaic filters. Mosaic filters tend to yield the greatest accuracy and sensitivity but are more expensive to make, so single- and multi-layer filters are more common.

Each photosensor in a filter colorimeter samples a broad range of wavelengths at once, whereas the photosensors in spectroradiometers and spectrophotometers sample narrow ranges and therefore receive weaker signals for a given target. Hence, filter colorimeters often have greater sensitivity and yield measurements more quickly. Most filter colorimeters also have good repeatability, that is, yield the same measurement results for the same input. These qualities, plus their compactness and ruggedness, make filter colorimeters well-suited for purposes such as adjusting displays coming off a production line to the same white point. With the possible exception of well-built mosaic-filter designs, though, colorimeters tend to be significantly less accurate than spectroradiometers and spectrophotometers. Furthermore, it is impractical to adjust the spectral sensitivity of the FPPs after they are built, to compensate for errors in the original construction or changes in their spectral response that develop over time. At best, the processing of the FPP signals can be rebalanced to yield accurate measurements for a particular SPD.

25.6 Device-Independent Color Transfer

Readers who have scanned a color image into a computer and examined the result on a color display, or designed a color image on a display and printed a hardcopy, have probably found that the colors are inconsistent. The inconsistencies reflect a failure to achieve *device-independent color transfer* (DICT). Sometimes, this failure has no practical importance, but it is often displeasing aesthetically, and a growing number of people are using their computers for professional art and publishing purposes that demand accurate color transfer.

One obstacle to achieving DICT is the fact that every color scanner, display, film recorder, and printer has its own, unique color space and produces or interprets color coordinates in terms of that space, as discussed in Section 25.4.1. Several vendors have introduced proprietary *color management systems* (CMSs) that approach this problem by mapping color from one device-dependent space to another. Often, the mapping involves trilinear interpolation within a *rendering table* that describes color transfer across a given pair of input and output devices, although simpler transformations of the sort illustrated in Equation 71 are used instead, sometimes, to conserve computational resources.

The proprietary approach fails when files must be shared between application programs that use different, incompatible CMSs (or none at all) or when a needed rendering table is lacking. Platform vendors have begun addressing this problem at the system level: Apple has added a CMS extension called *ColorSync* to its Macintosh operating system, Microsoft has included a CMS module called *Image Color Matching* in its Windows '95 operating system, SUN Microsystems has announced plans to add a CMS to its Solaris operating system, and Silicon Graphics International plans apparently to do the same with its IRIS operating system; furthermore, Adobe Systems has included DICT features in the PostScript Level 2 page description language. Moves like these toward open-architecture system-level DICT support should help to eliminate the compatibility problems eventually.

Another important development has been the growing adoption of a standard profile format, created by a large group of computer hardware and software vendors that make up the International Color Consortium (ICC) to support cross-platform color transfer. A key feature of the ICC format is standardization on the use of CIE *X*-, *Y*-, and *Z*-tristimulus values (or, optionally, $L^*a^*b^*$ values) as base units for representing color, along with a standard output medium and set of viewing conditons to reference those values against. Scanner outputs are converted to *X*, *Y*, and *Z* (or L^*, a^*, and b^*) using a *device profile*. These values can then be converted to the data needed by an output device using that device's profile and knowledge of the actual viewing conditions. This two-step mapping technique eliminates the need for a rendering table for every possible pairing of input and output devices, thereby simplifying the addition of devices to a system. At the moment, the ICC profile format specification is being considered as a draft international standard by the International Standards Organization.

System-level support and the use of platform-independent device profiles promise to remove unnecessary barriers to achieving DICT, but more fundamental problems remain. These problems include device-profile error, gamut mismatch, inconsistencies in the viewing and scanning conditions, contrast limitations, and quantization error.

25.6.1 Device-Profile Error

Even if the manufacturer calibrates a unit accurately before it leaves the factory, there will be unit-to-unit variations that are not reflected in the device profile unless a custom profile is created for each unit. Furthermore, the user may adjust the device's color transfer functions afterwards, or those functions may shift as components age, optics become dirty, etc. Printers can exhibit sudden shifts due to a change in their inks, ribbons, or paper, and are sensitive to changes in temperature and humidity. Scanners are sensitive to temperature changes, too.

In principle, users can correct device-profile errors. For example, some CRT monitors come equipped with measuring devices that, in conjunction with built-in self-test functions, enable them to calibrate themselves. A custom profile can be produced for a scanner by scanning a test pattern having known tristimulus values under the scanner's built-in illuminant. A custom profile can be produced for a printer by printing a test pattern and scanning the result with an accurately profiled scanner. Users who are serious about DICT may be willing to exercise these remedies, but most people probably are not.

25.6.2 Gamut Mismatch

One problem that cannot be solved completely by a CMS concerns mismatches among the color gamuts of differing output devices. For example, a user working at a display may create an image containing a saturated blue that cannot be reproduced by the printer, or a user may scan a color photograph containing a saturated red that cannot be reproduced on either the display or the printer. Problems can occur in the lightness/brightness dimension, too. Photographic slides have contrast ratios (i.e., the ratio between the maximum and minimum luminances or luminance factors that can be achieved) as high as 1000:1. Under normal viewing conditions, typical color displays provide 100:1 or less (20:1 is a more common maximum, although a carefully adjusted CRT monitor, viewed in a dark room, can deliver 1000:1). Photographic prints and high quality offset and digital printers rarely achieve contrast ratios greater than 100:1, and many inkjet printers produce less. The CMS can detect problems like these when it consults the output device's profile, but the question of what to do about them remains.

The simplest compromise, which is used by some application programs, is to alert the user and offer a chance to correct the problems manually. Two automated fixes have been devised, also: *clipping,* which involves moving the colors inside the output device's gamut by desaturating and/or darkening them while trying to retain their original hues, and *compression,* which involves desaturating and/or darkening all colors in the image until the outliers are brought within the gamut. The optimality of these transformations is limited, of course, by the fact that no perceptually uniform color space having known lines of constant hue, saturation, and lightness/brightness exists to perform them in. Furthermore, both methods have weaknesses, and it can be difficult to decide which yields the better result. Clipping retains accurate rendering of most colors in the image, but it can eliminate differences among colors that are meant to differ. Compression is intended to retain the relative differences among the image's colors, but it reduces saturation and contrast and may spoil colors that need accurate rendering, such as flesh tones. Some people use methods (referred to sometimes as *soft clipping*) that combine clipping and compression in varying amounts in an attempt to obtain a satisfactory compromise.

25.6.3 Viewing and Scanning Inconsistencies

Suppose that a photograph is scanned and the resulting image file is displayed. If the scanner's accuracy is perfect and the display reproduces the scanner tristimulus values exactly, one might expect the displayed image's color appearance to match the photograph's. A match is unlikely, though, because the photograph and display reflect the room illumination differently and the images appear against different backgrounds. Color-appearance models have been developed to correct for such influences (Fairchild, 1994; Guth, 1995; Hunt, 1994, 1995; Luo, 1996; Nayatani, 1995; Nayatani, Sobagaki, Hashimoto, and Yano, 1995; Seim and Valberg, 1986), but they are complex and imperfect.

One of the problems described in Section 25.5.3 concerning filter colorimeters arises, too: the tristimulus values are obtained with the scanner's built-in illuminant, which probably does not match the illumination in the room. Therefore, the scanner values will not match those under the room illumination and a perfect match will result only if the original image is copied using the same colorants and paper, or if the scanner and room illumination match; exact reproduction of the scanner tristimulus values is appropriate in these cases. One potential solution to this problem would be the development of scanners that yield spectral reflectance distributions. Coupled with knowledge of the room illumination and an accurate color appearance model, this information would permit appropriate corrections to be computed. For the moment, however, the problems of developing such scanners and the likely cost of producing them prohibit this solution. A more feasible approach is to use assumptions about the spectral characteristics of the colorants used in the scanned materials to estimate the spectral reflectance distributions. Berns and Shyu (1995) have shown that fairly accurate estimates can be obtained, if the assumptions, predictive model, and scanner profile are accurate.

25.6.4 Quantization Error

A typical photocopy of a black and white photograph provides a familiar example of quantization error. The numerous shades of gray that are present in the photograph are reproduced with a smaller number of shades, causing contrast to be reduced in some areas and exaggerated in others. Some degree of contrast distortion would occur even if the photocopier had the same number of shades available, or even many more, because it is unlikely that a perfect match would exist for every shade in the photograph. Color images contain hue and saturation, for which additional distortions can occur. Readers who have examined color photographs on computer systems having a small color repertoire have seen examples of these distortions.

Quantization error is inherent in digitization and, hence, unavoidable in computer imaging. Errors that occur during the analog-to-digital sampling process can be corrected partially using image reconstruction methods. Errors that occur during the digital-to-analog output process can be compensated in several ways. One common method is *dithering*, which involves distributing the information from each pixel over adjacent pixels. For example, orange can be created on a printer that lacks it by printing a few red and yellow pixels close together and, thus, using the additive-color principle described in Section 25.1.2 in connection with television. Another output compensation method is *palette optimization,* which is used routinely on displays when the graphics card cannot produce the required number of colors simultaneously. Algorithms for palette optimization vary widely in sophistication, but their objective is to reduce the number of unique colors in a digital image while maintaining a color-appearance match.

The simplest, albeit expensive, way to reduce the visibility of the errors is to increase the number of bits per pixel and the number of simultaneously available colors. Most high-end color-graphics hardware today provides 24 bits/pixel (i.e., 8 bits each for R, G, and B—although special-purpose graphics cards providing up to 15 bits each are available for displays) and three (or more) 8-bit-deep image planes in the graphics card so that all ($2^{24} \approx$) 16.7 million colors are available at once. This level of performance appears adequate for most purposes if the system's color transfer functions are optimized. Theoretically, the optimum transfer functions are linear, but in practice nonlinear functions that are customized for each image often yield more pleasing results. That is, color transfer that is inaccurate in carefully chosen ways can look better than accurate transfer. This paradox has been known for years by artists, television engineers, printers, and photographers, but has been recognized only recently by the computer DICT community.

25.7 Color Usage

The main practical uses of color in computer images involve coding information and aiding visual search by making items in the image more discriminable from one another. For example, color can be used to indicate groupings or shared characteristics, to indicate state, to draw attention, or to communicate qualitative or quantitative differences or changes. The following subsections summarize information useful for these purposes.

25.7.1 Color Discrimination

If color is to assist visual search, the target's color must differ sufficiently from the color(s) of other, distracting objects on the display. One way to assess the magnitude of a color difference is to use the ΔE^*_{uv} formula given in Equation 25. This method is recommended by ANSI (1988, p. 21), along with a suggested minimum difference of 40 units that was derived by Carter and Carter (1982). However, on computer displays, the target and distracters are often smaller than the CIE 1931 standard observer's one-degree minimum (see Section

25.3.1). As explained in Section 25.1.8, small stimuli have reduced saturation so, for computer displays, Equation 25 tends to overestimate their perceived color differences. Silverstein and Merrifield (1985, pp. 149-154) therefore introduced the modification

$$\Delta E^*_{uv\text{-}sf} = [(K_L \Delta L^*)^2 + (K_u \Delta u^*)^2 + (K_v \Delta v^*)^2]^{.5}, \quad (85)$$

where $\Delta E^*_{uv\text{-}sf}$ is a small-field-corrected version of ΔE^*_{uv}; ΔL^*, Δu^*, and Δv^* are the target-distracter color differences along the corresponding CIELUV axes; and K_L, K_u, and K_v are correction factors that vary with the size of the stimuli.[10]

Carter (1989) subsequently developed equations for estimating the correction factors:

$$K_L = 1.0366 - e^{0.15263 - 0.05766A} \quad (86)$$

for $0 < A < 60$,

$$K_u = -0.0065 + 0.008991A \quad (87)$$

for $0 < A \leq 32$,

$$= -0.5403 + 0.0257A \quad (88)$$

for $32 < A < 60$,

$$K_v = -0.0420 + 0.005446A \quad (89)$$

for $0 < A \leq 32$, and

$$= -0.8594 + 0.0310A \quad (90)$$

for $32 < A < 60$,

where A is the visual angle subtended by the stimuli in arc-minutes. Negative solutions (which result, for example, from Equation 89 when A is less than 7.71 arc-minutes) should be rounded to zero. Carter (1989) also analyzed data on visual search for color-coded stimuli and revised the Carter and Carter (1982) finding by concluding that asymptotic performance occurs when the CIELUV color difference (corrected for size, if necessary) reaches 20 units. Therefore, the use of Equations 85-90 and a 20-unit minimum seems preferable to the ANSI (1988) recommendation.

[10] Nagy (1994) has presented evidence that applying stimulus-size corrections in a color space based on cone excitations would yield more accurate results, but a practical procedure for doing this has not been developed yet.

25.7.2 Symbol Legibility

To help assure the legibility of colored symbols, ANSI (1988, p. 21) recommends a minimum symbol-background color difference of 100 units, as assessed by an equation developed by Lippert (1986):

$$\Delta E_{Cu'v'} = [(155C)^2 + (367\Delta u')^2 + (167\Delta v')^2]^{.5}, \quad (91)$$

where C is the symbol-background luminance contrast (i.e., $[L_{max} - L_{min}]/L_{max}$, where L_{max} is the greater of the two luminances and L_{min} is the lesser).

More generally, ANSI (1988, p. 20) requires a symbol-background luminance contrast of 0.67 (which yields $\Delta E_{Cu'v'} = 103$ all by itself) and recommends a value of 0.86. ISO (1992) also requires a minimum of 0.67. Therefore, satisfying the ANSI- and ISO-recommended minimum for luminance contrast guarantees that the ANSI color-difference recommendation will be met and assures legibility on monochrome displays, also.

25.7.3 Color-Code Size

An important question that arises when designing color codes concerns the number of colors that can be used. The answer depends in part on the colors because, if they are not readily discriminable, user performance will be poor, even for a code containing only two colors. The answer depends also on the application. For pseudo- and false-colored images (i.e., images in which a variable, such as X-ray transmission or infrared reflectance, is mapped to differing colors), for example, the number can be very large because discrimination among adjacent, simultaneously presented colors is all that is required and most viewers can discriminate among millions of colors under these circumstances (Nickerson and Newhall, 1943). For situations where the colors must be recognized, the number is much smaller because recognition is more difficult than discrimination. The maximum number of colors that can be recognized on an absolute basis varies, depending on whether they differ in hue only or in brightness/lightness and saturation/chroma as well. In the former case, 11 is a reasonable upper limit (Smallman and Boynton, 1993). In the latter case, as many as 50 can be used if the viewers are given training (Hanes and Rhoades, 1959); otherwise, a limit of 24 to 30 colors is more realistic, (Derefeldt and Swartling, 1995). In all cases, though, it is best to minimize the number, because this tends to reduce errors and search time. See Post (1992a) for further discussion of these last points.

25.7.4 Color Selection

Given that the size of a color code has been decided, the problem of choosing the colors remains. Three approaches have been developed to assist in this task. One involves spacing colors in a perceptually uniform representation of the display's color gamut, another requires testing numerous colors to find a discriminable set, and the third makes use of population stereotypes.

Color-spacing Algorithms

Carter and Carter (1982) pioneered the development of algorithms that, given a desired color-code size N and a description of a display's color gamut, choose N colors within the gamut such that the perceptual difference between the two nearest colors is maximized. Thus, a set of N displayable and maximally discriminable colors is obtained. Others have improved on this idea; the most recent and mathematically sophisticated work is DeCorte's (1990). This approach to color selection has two weaknesses, though. First, a perceptually uniform color space is needed to gauge the color distances accurately, and no truly uniform space is available. Second, the algorithms have difficulty considering the colors' appearances; this poses problems if, for example, three maximally discriminable colors that look like red, yellow, and green are desired. Therefore, color sets produced by these algorithms should always be verified and, if necessary, modified empirically.

Discrimination Testing

McFadden (1992) performed a series of experiments to identify a set of colors that would remain discriminable and identifiable, even when members of the set were used as backgrounds for other members. The intended applications for the set include radar displays, medical images, and electronic maps and charts, where colors are often juxtaposed or superimposed and may therefore influence one another's color appearance. The test stimuli in the experiments consisted of solid circles subtending 0.5 degrees visually and having a luminance of 9 $cd \cdot m^{-2}$, centered within solid circles subtending 2 degrees and having a luminance of 10 $cd \cdot m^{-2}$. The resulting set of 11 colors is shown in Table 5; however, it must be noted that the set has never been tested in an actual color coding application, to the best of my knowledge.

Population Stereotypes

The third approach to color selection uses population stereotypes. Berlin and Kay (1969) and Boynton and Olson (1987) have presented evidence that there are only 11 (or so) color names for which stereotypical representatives can be established (the names are black, gray, white, red, pink, orange, yellow, green, blue, purple, and brown), so this technique works best for relatively small color codes. For such cases, though, it is preferable to the color-spacing technique described above, because Travis and Johns (1994) have found that visual search is less error-prone on displays coded using stereotypical colors than with non-stereotypical colors, even if the color differences (assessed using ΔE^*_{uv}) for the two cases are equal.

The first step in this method is to choose, by name, colors that are likely to convey the intended meanings without explanation. Table 6 (Bergum and Bergum, 1981), which provides data on typical color-meaning associations, can assist this process. (For cartography, there are special conventions, of course—see Olson, 1987 and Grossman, 1992 for useful reviews.) The second step is to choose specific chromaticity coordinates to represent the color names, also in accordance with population stereotypes. Post and Greene (1986), Post and Calhoun (1988, 1989), Kaufmann (1990), and Kaufmann and O'Neill (1993) have published chromaticity diagrams, showing the probabilities of obtaining common color names as a function of location on the diagram for various typical viewing conditions. These diagrams can be used to determine both the optimal chromaticity coordinates for representing colors and their colorimetric tolerances. Alternatively, if one's application allows for customization, the user can be asked to select optimal representatives for each color name, using a "color picker" interface. Smallman and Boynton (1993) have shown that this procedure yields good visual search performance on color-coded displays.

25.7.5 Heterochromatic Brightness Matching

It is desirable sometimes to equalize the brightnesses of two or more colors. One might suppose that an acceptable solution is to equalize their luminances, but this often fails to produce the expected result. The most obvious discrepancies involve blue stimuli frequently—for example, for a typical color CRT monitor, roughly twice as much luminance is needed from the green channel to produce a green having the same brightness as the blue produced by the blue channel—

Table 5. Eleven colors that remain discriminable and recognizable when juxtaposed or superimposed on one another (McFadden, 1992)

Name	*x*	*y*	*u'*	*v'*
Blue	.225	.216	.175	.378
Green	.298	.453	.152	.520
Red Purple	.317	.192	.271	.370
Orange Red	.520	.332	.350	.503
Yellow	.376	.398	.214	.510
Purple	.275	.213	.220	.383
Yellow Green	.349	.465	.177	.531
Red Orange	.437	.328	.288	.487
Red	.484	.283	.357	.469
Orange	.391	.354	.242	.493
Gray	.313	.328	.198	.468

Note: Color names are shown only to help distinguish the colors and suggest their appearances—you might not agree with the choices.

Table 6. Color-meaning associations for U.S. college students (Bergum and Bergum, 1981)

Meaning	Red	Orange	Yellow	Green	Blue	Purple
Stop	100.0	0.0	0.0	0.0	0.0	0.0
Go	0.0	0.0	0.0	99.2	0.8	0.0
Hot	94.5	2.4	0.8	0.0	2.4	0.0
Cold	0.8	0.8	0.8	0.8	96.1	0.8
Danger	89.8	5.5	4.7	0.0	0.0	0.0
Caution	11.0	7.1	81.1	0.0	0.8	0.0
Safe	0.0	2.4	16.5	61.4	18.1	1.6
Radiation	59.1	13.4	15.7	0.0	3.9	7.9
On	50.4	3.1	4.7	37.8	3.1	0.8
Off	29.9	6.3	4.7	15.0	31.5	12.6
Near	0.2	19.7	38.6	9.4	15.0	7.1
Far	2.4	15.0	11.0	6.2	30.7	34.6

Note: Tabled values are percentages.

but differences occur for many other pairings, too. Ware and Cowan (1983) analyzed most of the data relating to brightness-luminance relations as a function of color and produced the corrective equation

$$K = [alog\,(0.256 - 0.184y - 2.527xy + 4.656x^3y + 4.657xy^4]^{-1} \quad ,(92)$$

where x and y are a color's CIE 1931 chromaticity coordinates and K is a luminance weighting factor. For example, for a blue having the coordinates $x = 0.2$ and $y = 0.2$, Equation 92 yields $K = 0.746$. The equation is designed to yield approximately unity for CIE standard illuminant D_{65}; therefore, to make the blue's brightness match a D_{65} white having a luminance of 10 cd/m^2, the blue should be 7.46 cd/m^2. Calculations like this predict the luminances needed to equate the brightness of any color set, or to assure that some colors are brighter than others.

Ware and Cowan (1983) specified several restrictions on the equation's use: (1) the stimuli should subtend 30 arc-minutes or more; (2) their luminances should be 2 cd/m^2 or more; (3) their y-chromaticity coordinates must not be less than 0.02; and (4) they should have similar backgrounds. Calhoun and Post (1990) tested Equation 92 under conditions that met the preceding requirements and found that, in general, it provided better brightness matches than did equalizing luminance. A simpler alternative is to let users adjust the colors to produce equal brightness for their own eyes, if one's application can permit this.

25.7.6 Background Color

For a color stimulus, viewed against an achromatic background, increasing the background luminance increases the stimulus' saturation—at least, until the background luminance exceeds the stimulus luminance. Beyond this point, increases in background luminance reduce stimulus saturation until, finally, the stimulus appears black. Thus, an achromatic background tends to increase the perceived differences among stimulus colors and, for stimuli having equal luminance, the differences are maximized when the background and stimulus luminances are the same.[11]

[11] One should not conclude, however, that it is a good idea to equalize the stimulus and background luminances; this can tend to make the edges of the stimulus indistinct (Boynton, 1978) and may pose problems for color-defective viewers, as discussed in Section 28.7.8. Furthermore, motion detection for a stimulus that differs only in chromaticity from its background is poor (Anstis, 1986, p. 16-19).

For stimuli that differ in luminance, the maximizing background luminance will be the highest level that does not desaturate the lower-luminance stimuli too seriously. Carter and Carter (1988), for example, found that visual search on a color-coded display was enhanced by an achromatic background having a luminance that was intermediate to the stimulus luminances. Similarly, Jacobsen (1986) found that recognition of colors was better on an achromatic background than on a black background.

If the stimulus' background is colored, the stimulus' hue will shift away from the background hue and toward the background's complement ordinarily, as discussed in Section 25.1.8. Thus, colored backgrounds complicate efforts to produce specific color perceptions and control perceived color differences. There is another reason to prefer an achromatic background or, at least, a desaturated one if it must be colored for some reason: Pastoor (1990) assessed many symbol/background color pairs and found that the less saturated backgrounds were always preferred. Furthermore, for dark-on-light text (i.e., text luminance < background luminance), his subjects were largely indifferent to the text color when desaturated backgrounds were used. (For light-on-dark text, however, subjects preferred desaturated text colors, also.)

25.7.7 Peripheral Color Vision

Color vision can change in several important ways for stimuli that are not imaged on the fovea: their relative brightnesses may change, they tend to be less saturated, and their hues may change. For light-adapted viewers, the changes in brightness are due mainly to the *macular pigment:* a yellow pigment that produces significant filtering within the central ± 4 degrees or so of the retina (Stabell and Stabell, 1980), the effects of which are built into $V(\lambda)$. Luminance adjustments for peripheral{ XE "Color:peripheral" } stimuli can be computed by correcting $V(\lambda)$ for the pigment's spectral transmittance function. For most HCI practitioners, though, it is probably more useful to note simply that the effects are most apt to be noticeable for stimuli that have substantial radiance at wavelengths below 535 nm (and near 460 nm, particularly), such as typical blue and cyan stimuli on a CRT monitor.

Red/green discrimination tends to fail at eccentricities beyond 20 to 30 degrees (causing red and green stimuli to appear yellow) and complete color blindness sets in beyond 40 to 50 degrees (Kinney, 1979). To some extent, these effects can be offset by increasing the size or luminance of the stimulus. For

Table 7. Good and bad color pairings for color-defective viewers (Arditi and Knoblauch, 1994).

	Color 1	Color 2
Good	Any light color	Black
	Any dark color	White
	Light yellow	Dark blue
	Light green	Dark red
Bad	Light red	Dark green
	White	Yellow
	Light gray	Yellow
	Turquoise	Green
	Lavender	Pink

light-adapted viewers, minimum size recommendations for reliable color identification can be derived from Kuyk (1982): at 20, 30, and 45 degrees eccentricity, the angle subtended at the eye by the stimulus should be at least 2.2, 4.2, and 5.5 degrees, respectively.

Clearly, it is best to avoid requiring accurate color perception in the periphery. This is not a problem in most cases because, ordinarily, the viewer will fixate the stimulus after detecting it and peripheral detection can be triggered by making the stimulus bright or causing it to flash or move. If peripheral color perception is required, though, one should make the stimulus as large, bright, and saturated a possible, to maximize the chances that it will be perceived correctly.

25.7.8 Color-Defective Viewers

When designing color images for general use, it is important to avoid creating problems for color-defective viewers (see Section 25.1.7). Therefore, do not rely on color alone to identify or distinguish image content; instead, assure that color differences are redundant with shape or brightness/lightness differences. This design strategy is often mandatory anyway because applications must usually be compatible with monochrome displays.

One simple test of whether an image is problematic for color-defective viewers is to look at it, first with the display in monochrome mode and then with each of the display's color channels disabled (one at a time—alternatively, one can view the display through cyan, magenta, and yellow filters) and see whether the desired discriminations remain possible.[12] Here is a general guideline that helps assure success on these checks: when using colors from the spectral extremes, set their brightness/lightness low and pair them with colors from the spectral midrange, set at high brightness/lightness. Table 7, adapted from Arditi and Knoblauch (1994), may also be helpful. It gives some examples of color pairings that tend to yield good and bad contrast for the color-defective population. Another strategy that helps color-defective viewers is to design the application so the user can adjust the colors.

25.7.9 Psychological Effects of Color

Some of the published guidance on color usage presumes the validity of the common belief that "warm" colors such as red and orange have an arousing effect on humans, whereas "cool" colors such as green and blue are calming. This color-induced state of arousal is often supposed to influence such things as mood and productivity. Efforts to test these ideas using physiological measures have been reviewed by Kaiser (1985). He concluded that color can influence galvanic skin response and electroencephalograms, but the effects are inconsistent and may have more to do with color preferences and learned associations than with direct impact on physiology. Other physiologically based studies have been performed since Kaiser's (1985) review, but have not produced results that contradict the preceding summary.

Many other measures, such as muscle strength, motor performance, questionnaire-based assessments of mood, intelligence tests, and behavioral observation have also been used to assess effects of color on arousal. The most dramatic claims have concerned "Baker-Miller pink": an ill-defined color that is occasionally the subject of anecdotal reports, asserting that placing violent persons in a room painted with this color calms them. To date, however, no large or consistent effects of color on these alternative measures have been demonstrated in the scientific literature. With specific regard to Baker-Miller pink, perhaps the most telling observation is this: recently, I phoned the commanding officer at the U.S. Navy prison that reportedly obtained the original, near-miraculous results (Schauss, 1979). He was familiar with the claims concerning pink cells, but was unaware of his facility's early role in those claims and stated that, presently, pink is used only to help guards distinguish the female prisoners' areas from the male prisoners' (which are blue).

12 Do not suppose that exercises like these allow one to experience the perception of color-defective viewers. Usually, it is impossible to determine what these perceptions are, and in at least one case where determination was possible, the results were surprising. See Kaiser and Boynton (1996, pp. 452-455) for an interesting discussion of these points.

Most people have color preferences and preconceived notions about their significance (many of which are culturally dependent). It is plausible, therefore, that these psychological factors can cause color to influence behavior and mood. So far, however, no effects that would provide a sound basis for design recommendations have been demonstrated scientifically.

25.8 Computer Assistance

Perhaps the most interesting recent development in color and human-computer interaction is a trend toward using the computer as a color design aid. One of the simpler examples is an enhanced "color picker" that depicts the display screen's color gamut using a perceptually uniform, three-dimensional representation (Bauersfeld and Price, 1990). The user can select a color swatch from this space and move it about to see how it looks in the context of other colors that are on the screen already. Other improved color pickers have also been designed. For example, Beretta (1990, 1993, 1994) has developed one for working with colors of familiar objects and another for more general design. The former contains preset, colorimetrically calibrated palettes, such as skin tones or vegetation, and allows the user to mix them as an artist would mix paints. The latter allows the user to specify a base color and then automatically generates a palette of "harmonious" partner colors, based on geometric relationships in a perceptually uniform color space.

A logical extension of color pickers is the provision of expert assistance that applies color-design rules, derived from the sort of knowledge and guidelines contained in Section 25.7. Meier (1988) describes an early example: a program that recommends colors for objects on a desktop environment after the screen designer provides information about the objects and their inter-relationships. More recently, Hedin and Derefeldt (1990) have produced an aid that generates palettes according to different perceptual rules, depending on the user's intent. For example, given a background color specification, palettes that yield legible text per Equation 91 can be created automatically. Rogowitz, Gerth, and Rabenhorst (1995) have developed a similar but more complex interactive tool that gradually constrains the designer's choices of font and color as a screen design progresses to assure legibility. That is, the program "understands" spatial vision and the spatial characteristics of the fonts, as well as color vision, and decisions made early in the design process are analyzed in terms of perceptual rules to determine what remaining choices should be permitted. Bergman, Rogowitz, and Treinish (1995) have created a program that generates colormaps for data visualization, based on data type (nominal, ordinal, interval, or ratio), spatial frequency content, and the representation task (isomorphic, segmentation, or highlighting).

Automated assistance has also been created for checking screens after they have been designed. Jiang, Murphy, Carter, Bailin, and Truszkowski (1993) describe a rule-based program that analyzes screen designs and determines whether they violate human factors guidance for color usage. If problems are found, the program suggests changes.

Sophisticated "color adjusters" have been designed for enhancing digital images. Kanamori and Kotera (1991) have developed an interactive color-correction system that uses fuzzy logic to help identify image areas that the user wishes to correct. In their demonstration, skin tone was adjusted globally without affecting other colors in the image, even though the pixels representing skin had varying *RGB* values. Sanger, Asada, Haneishi, and Miyake (1994) have created a program that analyzes images and identifies areas that correspond to human faces, to support color correction of skin tones. Katajamäki, Laihanen, and Saarelma (1996) have gone even further by developing an algorithm that analyzes images and corrects skin tones, overall color balance, and contrast automatically, without human intervention.

25.9 Additional Reading

Durrett (1987), Travis (1991), Widdel and Post (1992), and Jackson, MacDonald, and Freeman (1994) are fairly recent, overview-type texts that expand usefully on the material presented in this chapter, as well as covering additional, relevant topics. Readers desiring greater detail on specific sections of the present chapter may wish to consult the following:

Sections 25.1-25.3. For a good introductory text on color vision and CIE colorimetry, see Hurvich (1981). For more advanced treatments, see Judd and Wyszecki (1975), Wyszecki and Stiles (1982), MacAdam (1985), and Kaiser and Boynton (1996). Birch (1993) provides detailed coverage of defective color vision. Derefeldt, Menu, and Swartling (1995) offer a quick introduction to cognitive and physiological aspects of color and include many useful references.

Section 25.4. For information about color appearance systems, see Judd and Wyszecki (1975), Wyszecki and Stiles (1982), and Derefeldt (1991). Sproson's (1983) book is devoted entirely to television colorimetry; Hunt (1995) also provides good coverage of this subject. Pennebaker and Mitchell (1993) cover the JPEG standard in detail.

Section 25.5. Judd and Wyszecki (1975) and Wyszecki and Stiles (1982) discuss colorimetric equipment in general; Zwinkels (1996) provides more detail concerning use and calibration. Berns, Gorzynski, and Motta (1993) have presented a thorough study of several commercial devices.

Sections 25.6 and 25.8. Burger (1993) and MacDonald (1996) provide good introductions to color management systems. Post (1992b), Berns (1996), and Howard (1996) discuss issues relevant to color management on displays; Johnson (1996a, 1996b) does the same for scanners, digital cameras, and printers. Rhodes and Luo (1996) describe a state-of-the-art system for achieving device-independent color. Information about the ICC, including a copy of the latest profile format specification, can be obtained via the World-Wide Web at http://www.color.org/. The Society for Imaging Science and Technology (IS&T) has sponsored an annual *Non-Impact Printing Congress* since 1985. IS&T and the Society of Photo-Optical Instrumentation Engineers (SPIE) have jointly sponsored an annual *Color Hard Copy and the Graphic Arts* conference since 1992. These same two societies have also sponsored an annual *Device-Independent Color Imaging* conference since 1993. IS&T and the Society for Information Display (SID) have sponsored an annual *Color Imaging Conference* since 1993. There are published proceedings for all these conferences. It is best to obtain the most recent volumes of these proceedings and work backwards, because work on the topics discussed in Sections 25.6 and 25.8 is very active; any effort to capture the state of the art in a textbook is apt to be outdated quickly.

Section 25.7. Travis (1991), Post (1992a), and Jackson, MacDonald, and Freeman (1994) provide guidance on the use of color on displays. Also, IS&T and SPIE have jointly sponsored a *Human Vision, Visual Processing, and Digital Display* conference almost every year since 1989, plus an occasional *Human Vision and Electronic Imaging* conference since 1990, where relevant work is often presented.

25.10 Acknowledgments

I offer my sincere thanks to Alan Nagy (Wright State University) and Louis D. Silverstein (VCD Sciences, Inc.), who provided many helpful comments on an early draft of this chapter, and to Roy S. Berns (Munsell Color Science Laboratory), Gunilla Derefeldt (National Defense Research Establishment), and David S. Travis (System Concepts Ltd.), who did the same for a later version. I also thank Gary Beckler, Chris Calhoun and Gary Rankin (Logicon Technical Services, Inc.), who produced most of the figures.

25.11 References

Alessi, P. (1994). CIE guidelines for coordinated research on evaluation of colour appearance models for reflection print and self-luminous display image comparisons. *Color Research and Application, 19,* 48-58.

Alman, D. H., Berns, R. S., Snyder, G. D., and Larsen, W. A. (1989). Performance testing of color-difference metrics using a color tolerance dataset. *Color Research and Application, 14,* 139-151.

ANSI (1988). *American national standard for human factors engineering of visual display terminal workstations* (ANSI/HFS Standard No. 100-1988). Santa Monica, CA: Human Factors Society.

Anstis, S. (1986). Motion perception in the frontal plane. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of perception and human performance. Volume I: Sensory processes and perception* (pp. 16-1 thru 16-16-27). New York: Wiley.

Arditi, A., and Knoblauch, K. (1994). Choosing effective display colors for the partially sighted. *1994 Society for Information Display International Symposium Digest of Technical Papers, 25,* 32-35.

Bauersfeld, P. F., and Price, B. A. (1990). The 3D perceptual picker: Color selection in 3D. *1990 Society for Information Display International Symposium Digest of Technical Papers, 21,* 180-183.

Beretta, G. B. (1990). Color palette selection tools. *Conference Summaries of SPSE's 43rd Annual Conference,* 94-96.

Beretta, G. B. (1993). *Reference color selection system.* U.S. Patent #5,254,978.

Beretta, G. B. (1994). *Functional color selection system.* U.S. Patent #5,311,212.

Bergman, L. D., Rogowitz, B. E., and Treinish, L. A.

(1995). A rule-based tool for assisting colormap selection. *Proceedings of IEEE Visualization '95*, 118-125.

Bergum, B. O., and Bergum, J. E. (1981). Population stereotypes: An attempt to measure and define. *Proceedings of the Human Factors Society 25th Annual Meeting*, 662-665.

Berk, T., Brownstone, L., and Kaufmann, A. (1982). A new color naming system for computer graphics. *IEEE Computer Graphics and Applications, 3*, 37-44.

Berlin, B., and Kay, P. (1969). *Basic color terms*. Berkeley, CA: University of California Press.

Berns, R. S. (1996). Methods for characterizing CRT displays. *Displays, 16*, 173-182.

Berns, R. S., Gorzynski, M. E., and Motta, R. J. (1993). CRT colorimetry. Part II: Metrology. *Color Research and Application, 18*, 315-325.

Berns, R. S., and Shyu, M. J. (1995). Colorimetric characterization of a desktop drum scanner using a spectral model. *Journal of Electronic Imaging, 4*, 360-372.

Birch, J. (1993). *Diagnosis of defective colour vision*. Oxford: Oxford University Press.

Boynton, R. M. (1978). Ten years of research with the minimally distinct border. In J. C. Armington, J. Krauskopf, and B. R. Wooten (Eds.), *Visual psychophysics and physiology* (193-207). New York: Academic Press.

Boynton, R. M., and Olson, C. X. (1987). Locating basic colors in the OSA space. *Color Research and Application, 12*, 94-105.

Burger, R. E. (1993). *Color management systems*. San Francisco: The Color Resource.

Calhoun, C. S., and Post, D. L. (1990). Heterochromatic brightness matches via Ware and Cowan's luminance correction equation. *1990 Society for Information Display International Symposium Digest of Technical Papers, 21*, 261-264.

Carter, R. C. (1989). Calculate (don't guess) the effect of symbol size on usefulness of color. *Proceedings of the Human Factors Society 33rd Annual Meeting*, 1368-1372.

Carter, R. C., and Carter, E. C. (1982). High contrast sets of colors. *Applied Optics, 21*, 2936-2939.

Carter, R. C., and Carter, E. C. (1983). CIE L*u*v* color difference equations for self-luminous displays. *Color Research and Application, 8*, 252-253.

Carter, R. C., and Carter, E. C. (1988). Color coding for rapid location of small symbols. *Color Research and Application, 13*, 226-234.

CIE (1978a). *Light as a true visual quantity: Principles of measurement* (CIE Publication No. 41). Paris: Author.

CIE (1978b). *Recommendations on uniform color spaces - color-difference equations - psychometric color terms* (Supplement No. 2 to CIE Publication No. 15). Paris: Author.

CIE (1986). *Colorimetry* (2nd ed., CIE Publication No. 15.2). Vienna: Author.

CIE (1988). *Spectral luminous efficiency functions based upon brightness matching for monochromatic point sources 2^0 and 10^0 fields* (CIE Publication No. 75). Vienna: Author.

CIE (1989). *Mesopic Photometry: History, Special Problems, and Practical solutions* (CIE Publication No. 81). Vienna: Author.

CIE (1990). *CIE 1988 2^0 spectral luminous efficiency function for photopic vision.* (CIE Publication No. 86). Vienna: Author.

CIE (1994). *Industrial colour-difference evaluation.* (CIE Publication No. 116). Vienna: Author.

Curcio, C. A., Allen, K. A., Sloan, K. R., Lerea, C. L., Hurley, J. B., Klock, I. B., and Milam, A. H. (1991). Distribution and morphology of human cone photoreceptors stained with anti-blue opsin. *The Journal of Comparative Neurology, 312*, 610-624.

De Corte, W. (1990). Recent developments in the computation of ergonomically optimal contrast sets of CRT colours. *Displays, 11*, 123-128.

Derefeldt, G. (1991). Colour appearance systems. In P. Gouras (Ed.), *The Perception of Colour* (pp. 218-261). Volume 6 in J. R. Cronly-Dillon (Series Ed.), *Vision and visual dysfunction*. London: MacMillan.

Derefeldt, G., Menu, J.-P., and Swartling, T. (1995). Cognitive aspects of color. *Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE): Human Vision, Visual Processing, and Digital Display VI, 2411*, 16-24.

Derefeldt, G., and Swartling, T. (1995). Colour concept retrieval by free colour naming. Identification of up to 30 colours without training. *Displays, 16*, 69-77.

Durrett, H. J. (Ed.). (1987). *Color and the Computer.* San Diego: Academic Press.

Fairchild, M. D. (1994). Visual evaluation and evolution of the RLAB color space. *The Second IS&T/SID*

Color Imaging Conference: Color Science, Systems and Applications, 9-13.

Foley, J. D., Van Dam, A., Feiner, S. K., and Hughes, J. F. (1990). *Computer Graphics Principles and Practice* (2nd ed.). Reading, MA: Addison-Wesley.

Grossman, J. D. (1992). Color conventions and application standards. In H. Widdel and D. L. Post (Eds.), *Color in electronic displays* (pp. 209-218). New York: Plenum.

Guth, S. L. (1995). Further applications of the ATD model for color vision. *Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE): Device-Independent Color Imaging II, 2414,* 12-26.

Hård, A., Sivik, L., and Tonnquist, G. (1996a). NCS, natural color system—from concept to research and applications. Part II. *Color Research and Application, 21,* 206-220.

Hård, A., Sivik, L., and Tonnquist, G. (1996b). NCS, natural color system—from concept to research and applications. Part I. *Color Research and Application, 21,* 180-205.

Hanes, R. M., and Rhoades, M. V. (1959). Color identification as a function of extended practice. *Journal of the Optical Society of America, 49,* 1060-1064.

Hedin, C. E., and Derefeldt, G. (1990). Palette - A color selection aid for VDU images. *Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE): Perceiving, Measuring, and Using Color, 1250,* 165-176.

Howard, C. M. (1996). Managing color appearance in self-luminous displays. *Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE): Human Vision and Electronic Imaging, 2657,* 2-9.

Hunt, R. W. G. (1995). *The Reproduction of Colour* (5th ed.). Kingston-upon-Thames, England: Fountain Press.

Hunt, R. W. G. (1994). An improved predictor of colourfulness in a model of colour vision. *Color Research and Application, 19,* 23-26.

Hurvich, L. M. (1981). *Color Vision.* Sunderland, MA: Sinauer Associates, Inc.

Ikeda, K., Nakayama, M., and Obara, K. (1979). Comparison of perceived colour-differences of colour chips with their colorimetric ones in the CIE 1976 L*u*v* and the CIE L*a*b* uniform colour spaces. *CIE Proceedings, 19,* 83-89.

ISO (1992). *Ergonomic requirements for office work with visual display terminals (VDTs). Part 3. Visual display requirements* (ISO standard 9241-3). Geneva, Switzerland: Author.

Jackson, R., MacDonald, L., and Freeman, K. (1994). *Computer generated colour: A Practical Guide to Presentation and Display.* Chichester, West Sussex, England: Wiley.

Jacobsen, A. R. (1986). The effect of background luminance on color recognition. *Color Research and Application, 11,* 263-269.

Jiang, J., Murphy, E. D., Carter, L. E., Bailin, S. C., and Truszkowski, W. F. (1993). Knowledge-based evaluation of GUI color usage. *MOTIF '93 and COSE International User Conference Proceedings,* 78-87.

Johnson, T. (1996a). Methods for characterizing colour scanners and digital cameras. *Displays, 16,* 183-191.

Johnson, T. (1996b). Methods for characterizing colour printers. *Displays, 16,* 193-202.

Judd, D. B. (1951). Report of the U.S. secretariat committee on colorimetry and artificial daylight. In *CIE Proceedings, 1* (Part 7, p. 11). Paris: CIE.

Judd, D. B., and Nickerson, D. (1975). Relation between Munsell and Swedish Natural Color System scales. *Journal of the Optical Society of America, 65,* 85-90.

Judd, D. B., and Wyszecki, G. (1975). *Color in Business, Science and Industry* (3rd ed.). New York: Wiley.

Kaiser, P. K. (1985). Physiological response to color: A critical review. *Color Research and Application, 9,* 29-36.

Kaiser, P. K., and Boynton, R. M. (1996). *Human Color Vision* (2nd ed.). Washington, DC: Optical Society of America.

Kanamori, K., and Kotera, H. (1991). A method for selective color control in perceptual color space. *Journal of Imaging Science, 35,* 307-316.

Katajamäki, J. Laihanen, P., and Saarelma, H. (1996). Classification of images for automatic colour correction. *IS&T and SID's 3rd Color Imaging Conference: Color Science, Systems, and Applications,* 109-111.

Kaufmann, R. (1990). Effects of low-level red ambient lighting and stimulus size on classification of colours on displays. *Displays, 11,* 146-156.

Kaufmann, R., and O'Neill, M. C. (1993). Colour names and focal colours on displays. *Ergonomics, 36,* 881-890.

Kelly, K. L., and Judd, D. B. (1955). *The ISCC-NBS method of designating colors and a dictionary of color names* (NBS Circular 553). Washington, DC: U.S. Government Printing Office.

Kim, T. G., Berns, R. S., and Fairchild, M. D. (1993). Comparing appearance models using pictorial images. *The First IS&T and SID Color Imaging Conference: Transforms and Transportability of Color,* 72-77.

Kinney, J. A. S. (1979). The use of color in wide-angle displays. *Proceedings of the Society for Information Display, 20,* 33-40.

Kuyk, T. K. (1982). Spectral sensitivity of the peripheral retina to large and small stimuli. *Vision Research, 22,* 1293-1297.

Lippert, T. M. (1986). Color difference prediction of legibility performance for raster CRT imagery. *1986 Society for Information Display International Symposium Digest of Technical Papers, 17,* 86-89.

Lippert, T. M., Farley, W. W., Post, D. L., and Snyder, H. L. (1983). Color contrast effects on human performance. *1983 Society for Information Display International Symposium Digest of Technical Papers, 14,* 170-171.

Lo, M. C., Luo, M. R., and Rhodes, P. A. (1996). Evaluating colour models' performance between monitor and print images. *Color Research and Application, 21,* 277-291.

Luo, M. R. (1996). The LLAB model for colour appearance and colour difference evaluation. *Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE): Color Imaging: Device-Independent Color Hard Copy and Graphic Arts, 2658,* 261-269.

MacAdam, D. L. (1942). Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America, 32,* 247-274.

MacAdam, D. L. (1985). *Color measurement: Theme and variations* (2nd ed.). New York: Springer-Verlag.

MacDonald, L. W. (1996). Developments in colour management systems. *Displays, 16,* 203-211.

Mahy, M., Van Eycken, L., and Oosterlinck, A. (1994). Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV. *Color Research and Application, 19,* 105-121.

McFadden, S. (1992). Discrimination of colours presented against different coloured backgrounds. *Color Research and Application, 17,* 339-351.

Meier, B. (1988). ACE: A color expert system for user interface design. *Proceedings of the SIGGRAPH Symposium on User Interface Software,* 117-128.

Metrick, L. (1978). Status report of the Graphic Standards Planning Committee. *Computer Graphics, 13(3),* III-6 & III-7; III-37 & III-38.

Moroney, N. M., and Fairchild, M. D. (1993). Color space selection for JPEG image compression. *The First IS&T and SID Color Imaging Conference: Transforms & Transportability of Color,* 157-159.

Nagy, A. (1994). Red/green color discrimination and stimulus size. *Color Research and Application, 19,* 99-104.

Nayatani, Y. (1995). Revision of the chroma and hue scales of a nonlinear color-appearance model. *Color Research and Application, 20,* 143-155.

Nayatani, Y., Sobagaki, H., Hashimoto, K., and Yano, T. (1995). Lightness dependency of chroma scales of a nonlinear color-appearance model and its latest formulation. *Color Research and Application, 20,* 156-167.

Nerger, J. L., and Cicerone, C. M. (1992). The ratio of L cones to M cones in the human parafoveal retina. *Vision Research, 32,* 879-888.

Newhall, S. M., Nickerson, D. M., and Judd, D. B. (1943). Final report of the O.S.A. subcommittee on the spacing of the Munsell colors. *Journal of the Optical Society of America, 33,* 385-418.

Nickerson, D., and Newhall, S. M. (1943). A psychological color solid. *Journal of the Optical Society of America, 33,* 419-422.

Olson, J. (1987). Color and the computer in cartography. In H. J. Durrett (Ed.), *Color and the computer* (pp. 205-219). San Diego, CA: Academic Press.

Pastoor, S. (1990). Legibility and subjective preference for color combinations in text. *Human Factors, 32,* 157-171.

Pennebaker, W. B., and Mitchell, J. L. (1993). *JPEG still image compression standard.* New York: Van Nostrand Reinhold.

Pointer, M. R. (1981). A comparison of the CIE 1976 colour spaces. *Color Research and Application, 6,* 108-118.

Pokorny, J., Smith, V. C., and Went, L. N. (1981). Color matching in autosomal dominant tritan defect. *Journal of the Optical Society of America, 71,* 1327-1334.

Post, D. L. (1984). CIELUV/CIELAB and self-luminous displays: Another perspective. *Color Research and Application, 9,* 244-245.

Post D. L. (1992a). Applied color vision research. In H. Widdel and D. L. Post (Eds.),*Color in electronic displays* (pp. 137-173). New York: Plenum.

Post, D. L. (1992b). Colorimetric measurement, calibration, and characterization of self-luminous displays.

In H. Widdel and D. L. Post (Eds.), *Color in Electronic Displays* (pp. 299-312). New York: Plenum.

Post, D. L., and Calhoun, C. S. (1988). Color-name boundaries for equally bright stimuli on a CRT: Phase II. *1988 Society for Information Display International Symposium Digest of Technical Papers, 19,* 65-68.

Post, D. L., and Calhoun, C. S. (1989). Color-name boundaries for equally bright stimuli on a CRT: Phase III. *1989 Society for Information Display International Symposium Digest of Technical Papers, 20,* 284-287.

Post, D. L., Costanza, E. B., and Lippert, T. M. (1982). Expressions of color contrast as equivalent achromatic contrast. *Proceedings of the Human Factors Society 26th Annual Meeting,* 581-585.

Post, D. L., and Greene, F. A. (1986). Color-name boundaries for equally bright stimuli on a CRT: Phase I. *1986 Society for Information Display International Symposium Digest of Technical Papers, 17,* 70-73.

Post, D. L., Lippert, T. M., and Snyder, H. L. (1983). Color contrast metrics for head-up displays. *Proceedings of the Human Factors Society 27th Annual Meeting, 2,* 933-937.

Rhodes, P., and Luo, M. R. (1996). A system for WYSIWIG colour communication. *Displays, 16,* 213-221.

Richter, M. (1955). The official German Standard Color Chart. *Journal of the Optical Society of America, 45,* 223-226.

Robertson, A. R. (1977). The CIE 1976 colour spaces. *Color Research and Application, 2,* 7-11.

Rogowitz, B. E., Gerth, J. A., and Rabenhorst, D. A. (1995). An intelligent tool for selecting colors and fonts for graphical user interfaces. *Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE): Human Vision, Visual Processing, and Digital Display VI, 2411,* 35-43.

Sanger, D., Asada, T., Haneishi, H., and Miyake, Y. (1994). Facial pattern recognition and its preferred color reproduction. *IS&T and SID's 2nd Color Imaging Conference: Color Science, Systems, and Applications,* 149-153.

Schauss, A. G. (1979). Tranquilizing effect of color reduces aggressive behavior and potential violence. *Orthomolecular Psychiatry, 8,* 218-221.

Seim, T., and Valberg, A. (1986). Towards a uniform color space: A better formula to describe the Munsell and OSA color scales. *Color Research and Application, 11,* 11-24.

Silverstein, L. D., and Merrifield, R. M. (1985, July). *The Development and Evaluation of Color Systems for Airborne Applications* (Report No. DOT/FAA/PM-85-19). Washington, DC: U.S. Department of Transportation.

Smallman, H. S., and Boynton, R. M. (1993). On the usefulness of basic colour coding in an information display. *Displays, 14,* 158-165.

Smith, A. R. (1978). Color gamut transform pairs. *Computer Graphics, 12(3),* 12-19.

Smith, H. S., Whitfield, T. W. A., and Wiltshire, T. J. (1990). A colour notation conversion program. *Color Research and Application, 15,* 338-343.

Sproson, W. N. (1983). *Colour science in television and display systems.* Bristol, England: Adam Hilger.

Stabell, U., and Stabell, B. (1980). Variation in density of macular pigmentation and in short-wave cone sensitivity with eccentricity. *Journal of the Optical Society of America, 70,* 706-711.

Travis, D. S. (1991). *Effective color displays: Theory and practice.* San Diego: Academic Press.

Travis, D. S., and Johns, A. M. (1994). Back to basics. *1994 Society for Information Display International Symposium Digest of Technical Papers, 25,* 877-880.

Vos, J. J. (1978). Colorimetric and photometric properties of a 2 deg fundamental observer. *Color Research and Application, 3,* 125-128.

Ware, C., and Cowan, W. B. (1983). *Specification of heterochromatic brightness matches: A conversion factor for calculating luminances of stimuli that are equal in brightness* (Tech. Report 26055). Ottawa, Canada: National Research Council of Canada.

Widdel, H., and Post, D. L. (Eds.). (1992). *Color in Electronic Displays.* New York: Plenum.

Williams, D. R., MacLeod, D. I. A., and Hayhoe, M. M. (1980). Foveal tritanopia. *Vision Research, 21,* 1341-1356.

Wyszecki, G., and Stiles, W. S. (1982). *Color science* (2nd ed.). New York: Wiley.

Wyszecki, G. (1986). Color appearance. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of perception and human performance. Volume I: Sensory processes and perception* (pp. 9-1 thru 9-57). New York: Wiley.

Zwinkels, J. C. (1996). Colour-measuring instruments and their calibration. *Displays, 16,* 163-171.