Light and Color

02

The emerging field of HDR imaging is directly linked to diverse existing disciplines such as radiometry, photometry, colorimetry, and color appearance — each dealing with specific aspects of light and its perception by humans. In this chapter we discuss all aspects of color that are relevant to HDR imaging. This chapter is intended to provide background information that will form the basis of later chapters.

2.1 RADIOMETRY

The term *scene* indicates either an artificial or real environment that may become the topic of an image. Such environments contain objects that reflect light. The ability of materials to reflect light is called "reflectance."

Radiometry is the science concerned with measuring light. This section first briefly summarizes some of the quantities that may be measured, as well as their units. Then, properties of light and how they relate to digital imaging are discussed.

Light is radiant energy, measured in joules. Because light propagates through media such as space, air, and water, we are interested in derived quantities that measure how light propagates. These include radiant energy measured over time, space, or angle. The definitions of these quantities and their units are outlined in Table 2.1 and should be interpreted as follows.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 2

Quantity	**Unit**	**Definition**
Radiant energy (Q_e)	J (joule)	Q_e
Radiant power (P_e)	J s^{-1} = W (watt)	$P_e = \frac{dQ_e}{dt}$
Radiant exitance (M_e)	W m^{-2}	$M_e = \frac{dP_e}{dA_e}$
Irradiance (E_e)	W m^{-2}	$E_e = \frac{dP_e}{dA_e}$
Radiant intensity (I_e)	W sr^{-1}	$I_e = \frac{dP_e}{d\omega}$
Radiance (L_e)	W m^{-2} sr^{-1}	$L_e = \frac{d^2 P_e}{dA \cos\theta \, d\omega}$

TABLE 2.1 *Radiometric quantities. The cosine term in the definition of L_e is the angle between the surface normal and the angle of incidence, as shown in Figure 2.4. Other quantities are shown in Figures 2.1 through 2.3.*

Because light travels through space, the flow of radiant energy may be measured. It is indicated with radiant power or radiant flux and is measured in joules per second, or watts. It is thus a measure of energy per unit of time.

Radiant flux density is the radiant flux per unit area, known as *irradiance* if we are interested in flux arriving from all possible directions at a point on a surface (Figure 2.1) and as *radiant exitance* for flux leaving a point on a surface in all possible directions (Figure 2.2). Both irradiance and radiant exitance are measured in watts per square meter. These are therefore measures of energy per unit of time as well as per unit of area.

If we consider an infinitesimally small point light source, the light emitted into a particular direction is called radiant intensity measured in watts per steradian (Figure 2.3). A steradian is a measure of solid angle corresponding to area on the unit sphere. Radiant intensity thus measures energy per unit of time per unit of direction.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 3

FIGURE 2.1 Irradiance: power incident upon unit area dA.

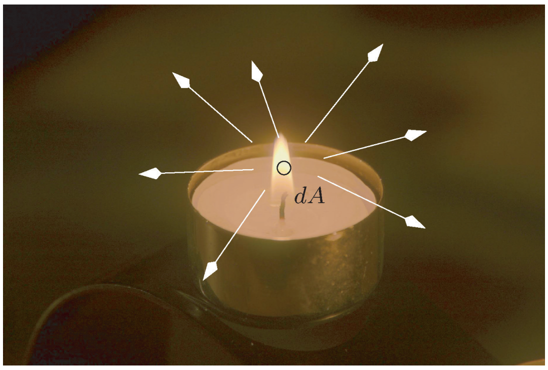


FIGURE 2.2 Radiant exitance: power emitted per unit area.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 4

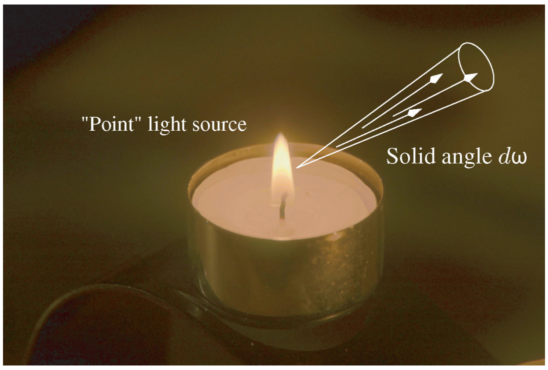


FIGURE 2.3 Radiant intensity: power per solid angle $d\omega$.

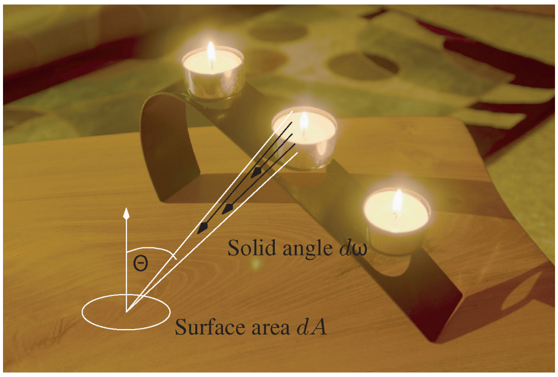


FIGURE 2.4 Radiance: power incident on a unit surface area dA from a unit set of directions $d\omega$.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 5

Flux passing through, leaving, or arriving at a point in a particular direction is known as radiance measured in watts per square meter per steradian (Figure 2.4). It is a measure of energy per unit of time as well as per unit of area and per unit of direction. Light that hits a point on a surface from a particular direction is at the heart of image formation. For instance, the combination of shutter, lens, and sensor in a (digital) camera restricts incoming light in this fashion.

When a picture is taken, the shutter is open for a small amount of time. During that time, light is focused through a lens that limits the number of directions from which light is received. The image sensor is partitioned into small pixels, so that each pixel records light over a small area. The light recorded by a pixel may be modeled by the "measurement equation" (see, for example, [66] for details). Because a camera records radiance, it is therefore possible to relate the voltages extracted from the camera sensor to radiance, provided pixels are neither under- nor overexposed [104,105].

Each of the quantities given in Table 2.1 may also be defined per unit wavelength interval, which are then referred to as spectral radiance $L_{e,\lambda}$, spectral flux $P_{e,\lambda}$, and so on. The subscript e indicates radiometric quantities and differentiates them from photometric quantities (discussed in the following section). In the remainder of this book, these subscripts are dropped unless this leads to confusion.

Light may be considered to consist of photons that can be emitted, reflected, transmitted, and absorbed. Photons normally travel in straight lines until they hit a surface. The interaction between photons and surfaces is twofold. Photons may be absorbed by the surface, where they are converted into thermal energy, or they may be reflected in some direction. The distribution of reflected directions, given an angle of incidence, gives rise to a surface's appearance. Matte surfaces distribute light almost evenly in all directions (Figure 2.5), whereas glossy and shiny surfaces reflect light in a preferred direction. Mirrors are the opposite of matte surfaces and emit light specularly in almost a single direction. This causes highlights that may be nearly as strong as light sources (Figure 2.6). The depiction of specular surfaces may therefore require HDR techniques for accuracy.

For the purpose of lighting simulations, the exact distribution of light reflected from surfaces as a function of angle of incidence is important (compare Figures 2.5 and 2.6). It may be modeled with bidirectional reflection distribution functions (BRDFs), which then become part of the surface material description. Advanced

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 6

FIGURE 2.5 *This object only reflects light diffusely. Because of the bright lighting conditions under which this photograph was taken, this image should look bright overall and without a large variation in tone.*

rendering algorithms use this information to compute how light is distributed in a scene, from which an HDR of the scene may be generated [24,58].

2.2 PHOTOMETRY

Surfaces reflect light and by doing so may alter the spectral composition of it. Thus, reflected light conveys spectral information of both the light source illuminating a surface point and the reflectance of the surface at that point.

There are many wavelengths that are not detectable by the human eye, which is sensitive to wavelengths between approximately 380 to 830 nanometers (nm).

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 7

FIGURE 2.6 *The metal surface of this clock causes highlights that are nearly as strong as the light sources they reflect. The environment in which this image was taken is much darker than the one depicted in Figure 2.5. Even so, the highlights are much brighter.*

Within this range, the human eye is not equally sensitive to all wavelengths. In addition, there are differences in sensitivity to the spectral composition of light among individuals. However, this range of sensitivity is small enough that the spectral sensitivity of any human observer with normal vision may be approximated with a single curve. Such a curve is standardized by the Commission Internationale de l'Eclairage (CIE) and is known as the $V(\lambda)$ curve (pronounced vee-lambda), or CIE photopic luminous efficiency curve. This curve is plotted in Figure 2.7.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 8

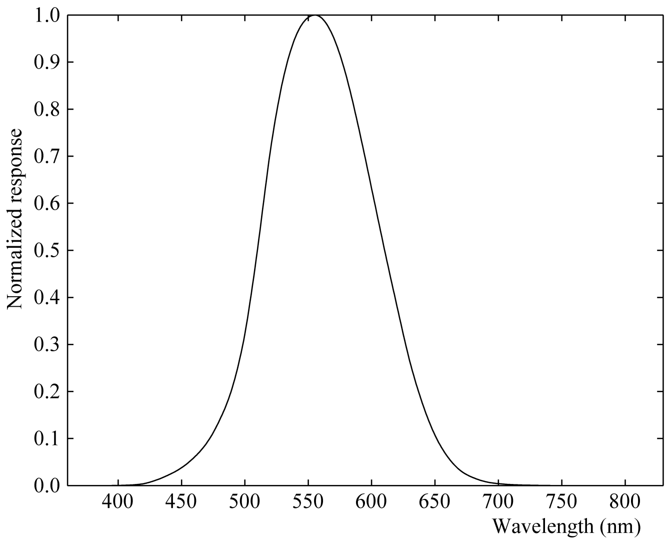


FIGURE 2.7 *CIE standard observer photopic luminous efficiency curve.*

In that we are typically interested in how humans perceive light, its spectral composition may be weighted according to $V(\lambda)$. The science of measuring light in units that are weighted in this fashion is called photometry. All radiometric terms introduced in the previous section have photometric counterparts, which are outlined in Table 2.2. By spectrally weighting radiometric quantities with $V(\lambda)$, they are converted into photometric quantities.

Luminous flux (or luminous power) is photometrically weighted radiant flux. It is measured in lumens, which is defined as $1/683$ watt of radiant power at a frequency of 540×10^{12} Hz. This frequency corresponds to the wavelength for which humans are maximally sensitive (about 555 nm). If luminous flux is measured over a differential solid angle, the quantity obtained is luminous intensity, measured in

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 9

Quantity	Unit
Luminous power (P_v)	*lm* (lumen)
Luminous energy (Q_v)	*lm* s
Luminous exitance (M_v)	*lm m*$^{-2}$
Illuminance (E_v)	*lm m*$^{-2}$
Luminous intensity (I_v)	*lm sr*$^{-1}$ = *cd* (candela)
Luminance (L_v)	*cd m*$^{-2}$ = *nit*

TABLE 2.2 *Photometric quantities.*

lumens per steradian. One lumen per steradian is equivalent to one candela. Luminous exitance and illuminance are both given in lumens per square meter, whereas luminance is specified in candela per square meter (a.k.a. "nits").

Luminance is a perceived quantity. It is a photometrically weighted radiance and constitutes an approximate measure of how bright a surface appears. Luminance is the most relevant photometric unit to HDR imaging. Spectrally weighting radiance amounts to multiplying each spectral component with the corresponding value given by the weight function and then integrating all results, as follows.

$$L_v = \int_{380}^{830} L_{e,\lambda} V(\lambda)\, d\lambda$$

The consequence of this equation is that there are many different spectral compositions of radiance L_e possible that would cause the same luminance value L_v. It is therefore not possible to apply this formula and expect the resulting luminance value to be a unique representation of the associated radiance value.

The importance of luminance in HDR imaging lies in the fact that it provides a natural boundary of visible wavelengths. Any wavelength outside the visible range

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 10

does not need to be recorded, stored, or manipulated, in that human vision is not capable of detecting those wavelengths. Many tone-reproduction operators first extract a luminance value from the red, green, and blue components of each pixel prior to reducing the dynamic range, in that large variations in luminance over orders of magnitude have a greater bearing on perception than extremes of color (see also Section 7.1.2).

2.3 COLORIMETRY

The field of colorimetry is concerned with assigning numbers to physically defined stimuli such that stimuli with the same specification look alike (i.e., match). One of the main results from color-matching experiments is that over a wide range of conditions almost all colors may be visually matched by adding light from three suitably pure stimuli. These three fixed stimuli are called primary stimuli. Color-matching experiments take three light sources and project them to one side of a white screen. A fourth light source, the target color, is projected to the other side of the screen. Participants in the experiments are given control over the intensity of each of the three primary light sources and are asked to match the target color.

For each spectral target, the intensity of the three primaries may be adjusted to create a match. By recording the intensities of the three primaries for each target wavelength, three functions $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, and $\bar{b}(\lambda)$ may be created. These are called color-matching functions. The color-matching functions obtained by Stiles and Burch are plotted in Figure 2.8. They used primary light sources that were nearly monochromatic with peaks centered on $\lambda_{\mathrm{R}} = 645.2$ nm, $\lambda_{\mathrm{G}} = 525.3$ nm, and $\lambda_{\mathrm{B}} = 444.4$ nm [122]. The stimuli presented to the observers in these experiments span 10 degrees of visual angle, and hence these functions are called 10-degree color-matching functions. Because the recorded responses vary only a small amount between observers, these color-matching functions are representative of normal human vision. As a result, they were adopted by the CIE to describe the "CIE 1964 standard observer." Thus, a linear combination of three spectral functions will yield a fourth, Q_λ, which may be visually matched to a linear combination of primary stimuli as follows.

$$Q_\lambda = \bar{r}(\lambda)R + \bar{g}(\lambda)G + \bar{b}(\lambda)B$$

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 11

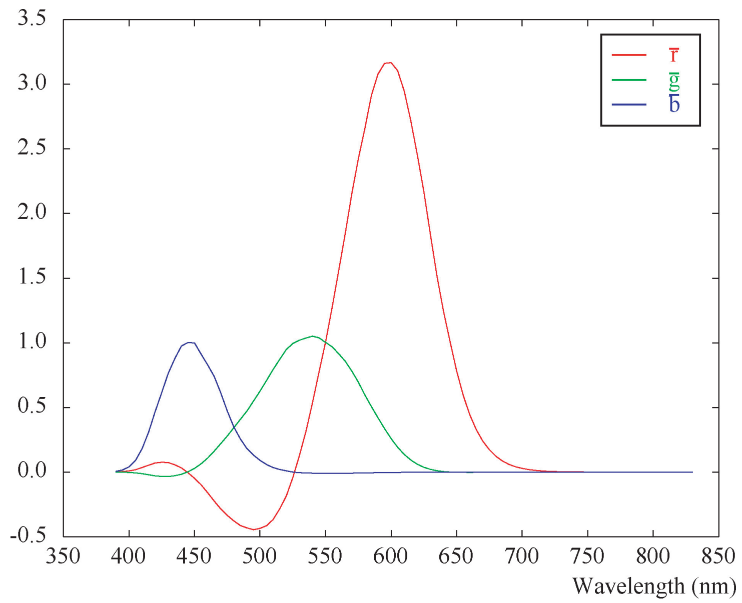


FIGURE 2.8 *Stiles and Burch (1959) 10-degree color-matching functions.*

Here, R, G, and B are scalar multipliers. Because the primaries are fixed, the stimulus Q_λ may be represented as a triplet by listing R, G, and B. This (R, G, B) triplet is then called the tristimulus value of Q.

For any three real primaries, it is sometimes necessary to supply a negative amount to reach some colors (i.e., there may be one or more negative components of a tristimulus value). In that it is simpler to deal with a color space whose tristimulus values are always positive, the CIE has defined alternative color-matching functions chosen such that any color may be matched with positive primary coef-

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 12

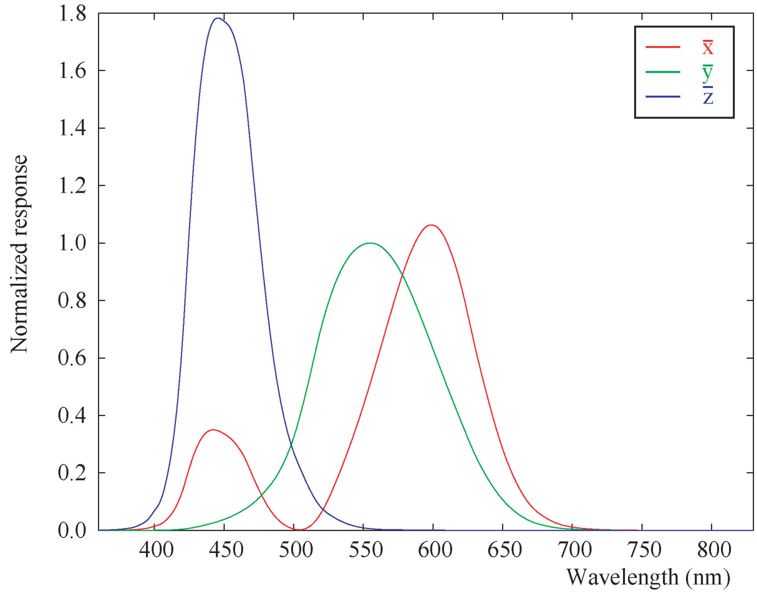


FIGURE 2.9 CIE 1931 2-degree XYZ color-matching functions.

ficients.[1] These color-matching functions are named $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ (plotted in Figure 2.9). These functions are the result of experiments in which the stimulus spanned 2 degrees of visual angle and are therefore known as the "CIE 1931 standard observer" [149]. A spectral stimulus may now be matched in terms of these

1 Real or realizable primaries are those that can be obtained by physical devices. For such primaries it is not possible to supply negative amounts because light cannot be subtracted from a scene. However, although less desirable in practice there is no mathematical reason a tristimulus value could not be converted such that it would be represented by a different set of primaries. Some of the values might then become negative. Such conversion issues are discussed further in Section 2.4.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 13

color-matching functions, as follows.

$$Q_\lambda = \bar{x}(\lambda)X + \bar{y}(\lambda)Y + \bar{z}(\lambda)Z$$

For a given stimulus Q_λ, the tristimulus values (X, Y, Z) are obtained by integration, as follows.

$$X = \int_{380}^{830} Q_\lambda\, \bar{x}(\lambda)\, d\lambda$$
$$Y = \int_{380}^{830} Q_\lambda\, \bar{y}(\lambda)\, d\lambda$$
$$Z = \int_{380}^{830} Q_\lambda\, \bar{z}(\lambda)\, d\lambda$$

The CIE XYZ matching functions are defined such that a theoretical equal-energy stimulus, which would have unit radiant power at all wavelengths, maps to tristimulus value $(1, 1, 1)$. Further, note that $\bar{y}(\lambda)$ is equal to $V(\lambda)$ — another intentional choice by the CIE. Thus, Y represents photometrically weighted quantities.

For any visible color, the tristimulus values in XYZ space are all positive. However, as a result the CIE primaries are not realizable by any physical device. Such primaries are called "imaginary," as opposed to realizable, primaries which are called "real."[2] Associated with tristimulus values are chromaticity coordinates, which may be computed from tristimulus values as follows.

$$x = \frac{X}{X+Y+Z}$$
$$y = \frac{Y}{X+Y+Z}$$
$$z = \frac{Z}{X+Y+Z} = 1 - x - y$$

Because z is known if x and y are known, only the latter two chromaticity coordinates need to be kept. Chromaticity coordinates are relative, which means that

2 This has nothing to do with the mathematical formulation of "real" and "imaginary" numbers.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 14

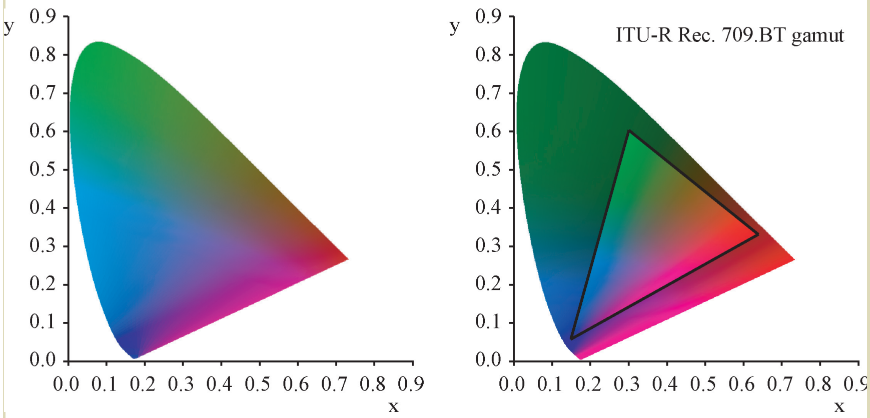


FIGURE 2.10 *CIE xy chromaticity diagram showing the range of colors humans can distinguish (left). On the right, the triangular gamut spanned by the primaries defined by ITU Recommendation (ITU-R) BT.709 color space [57] is shown.*

within a given system of primary stimuli two colors with the same relative spectral power distribution will map to the same chromaticity coordinates. An equal-energy stimulus will map to coordinates ($x = 1/3$, $y = 1/3$).

Chromaticity coordinates may be plotted in a chromaticity diagram with two axes. A CIE xy chromaticity diagram is shown in Figure 2.10. All monochromatic wavelengths map to a position along the curved boundary, called the spectral locus, which is of horseshoe shape. The line between red and blue is called the "purple line," which represents the locus of additive mixtures of short- and long-wave stimuli.

The three primaries used for any given color space will map to three points in a chromaticity diagram and thus span a triangle. This triangle contains the range of colors that may be represented by these primaries (assuming nonnegative tristimulus values). The range of realizable colors given a set of primaries is called

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 15

the *color gamut*. Colors that are not representable in a given color space are called out-of-gamut colors.

The gamut for the primaries defined by ITU-R (International Telecommunication Union Recommendations) BT.709 is shown on the right in Figure 2.10. These primaries are a reasonable approximation of most CRT computer monitors and officially define the boundaries of the sRGB color space [124] (see Section 2.11). The triangular region shown in this figure marks the range of colors that may be displayed on a standard monitor. The colors outside this triangle cannot be represented on most displays. They also cannot be stored in an sRGB file, such as the one used for this figure. We are therefore forced to show incorrect colors outside the sRGB gamut in all chromaticity diagrams in this book.

The diagrams in Figure 2.10 show two dimensions of what is a 3D space. The third dimension (luminance) goes out of the page, and the color gamut is really a volume of which a slice is depicted. In the case of the sRGB color space, the gamut is shaped as a six-sided polyhedron, often referred to as the "RGB color cube." This is misleading, however, in that the sides are only equal in the encoding (0–255 thrice) and are not very equal perceptually.

It may be possible for two stimuli with different spectral radiant power distributions to match against the same linear combination of primaries, and thus are represented by the same set of tristimulus values. This phenomenon is called metamerism. Whereas metameric stimuli will map to the same location in a chromaticity diagram, stimuli that appear different will map to different locations. The magnitude of the perceived difference between two stimuli may be expressed as the Cartesian distance between the two points in a chromaticity diagram. However, in the 1931 CIE primary system the chromaticity diagram is not uniform (i.e., the distance between two points located in one part of the diagram corresponds to a different perceived color difference than two points located elsewhere in the diagram). Although CIE XYZ is still the basis for all color theory, this nonuniformity has given rise to alternative color spaces (discussed in the following sections).

2.4 COLOR SPACES

Color spaces encompass two different concepts. First, they are represented by a set of formulas that define a relationship between a color vector (or triplet) and the

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 16

standard CIE XYZ color space. This is most often given in the form of a 3-by-3 color transformation matrix, although there are additional formulas if the space is nonlinear. Second, a color space is a 2D boundary on the volume defined by this vector, usually determined by the minimum and maximum value of each primary — the color gamut. Optionally, the color space may have an associated quantization if it has an explicit binary representation. In this section, linear transformations are discussed, whereas subsequent sections introduce nonlinear encodings and quantization.

We can convert from one tristimulus color space to any other tristimulus space using a 3-by-3 matrix transformation. Usually the primaries are known by their xy chromaticity coordinates. In addition, the white point needs to be specified, which is given as an xy chromaticity pair (x_W, y_W) plus maximum luminance Y_W. The white point is the color associated with equal contributions of each primary (discussed further in the following section).

Given the chromaticity coordinates of the primaries, first the z chromaticity coordinate for each primary is computed to yield chromaticity triplets for each primary; namely, (x_R, y_R, z_R), (x_G, y_G, z_G), and (x_B, y_B, z_B). From the white point's chromaticities and its maximum luminance, the tristimulus values (X_W, Y_W, Z_W) are calculated. Then, the following set of linear equations is solved for S_R, S_G, and S_B.

$$
\begin{aligned}
X_W &= x_R S_R + x_G S_G + x_B S_B \\
Y_W &= y_R S_R + y_G S_G + y_B S_B \\
Z_W &= z_R S_R + z_G S_G + z_B S_B
\end{aligned}
$$

The conversion matrix to convert from RGB to XYZ is then given by

$$
\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x_R S_R & x_G S_G & x_B S_B \\ y_R S_R & y_G S_G & y_B S_B \\ z_R S_R & z_G S_G & z_B S_B \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}
$$

The conversion from XYZ to RGB may be computed by inverting this matrix. If the primaries are unknown, or if the white point is unknown, a second best solution is

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 17

	R	G	B	White
x	0.6400	0.3000	0.1500	0.3127
y	0.3300	0.6000	0.0600	0.3290

TABLE 2.3 *Primaries and white point specified by ITU-Recommendation BT.709.*

to use a standard matrix such as that specified by ITU-R BT.709 [57]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.2405 & -1.5371 & -0.4985 \\ -0.9693 & 1.8760 & 0.0416 \\ 0.0556 & -0.2040 & 1.0572 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

The primaries and white point used to create this conversion matrix are outlined in Table 2.3.

There are several standard color spaces, each used in a particular field of science and engineering. Each is reached by constructing a conversion matrix, the previous matrix being an example. Several of these color spaces include a nonlinear transform akin to gamma correction, which is explained in Section 2.9. We therefore defer a discussion of other standard color spaces until Section 2.11.

In addition to standard color spaces, most cameras, scanners, monitors, and TVs use their own primaries (called spectral responsivities in the case of capturing devices). Thus, each device may use a different color space. Conversion between these color spaces is thus essential for the faithful reproduction of an image on any given display device.

If a color is specified in a device-dependent RGB color space, its luminance may be computed because the Y component in the XYZ color space represents luminance (recall that $V(\lambda)$ equals $\bar{y}(\lambda)$). Thus, a representation of luminance is obtained by computing a linear combination of the red, green, and blue components according to the middle row of the RGB-to-XYZ conversion matrix. For instance, luminance may be computed from ITU-R BT.709 RGB as follows.

$$Y = 0.2126R + 0.7152G + 0.0722B$$

Finally, an important consequence of color metamerism is that if the spectral responsivities (primaries) associated with a camera are known, as well as the emissive spectra of the three phosphors of a CRT display, we may be able to specify a transformation between the tristimulus values captured with the camera and the tristimulus values of the display and thus reproduce the captured image on the display. This would, of course, only be possible if the camera and display technologies did not impose restrictions on the dynamic range of captured and displayed data.

2.5 WHITE POINT AND ILLUMINANTS

For the conversion of tristimulus values between XYZ and a specific RGB color space, the primaries of the RGB color space must be specified. In addition, the white point needs to be known. For a display device, the white point is the color emitted if all three color channels are contributing equally.

Similarly, within a given scene the dominant light source will produce a color cast that will affect the appearance of the objects in the scene. The color of a light source (illuminant) may be determined by measuring a diffusely reflecting white patch. The color of the illuminant therefore determines the color of a scene the human visual system normally associates with white.

An often-used reference light source is CIE illuminant D_{65}. This light source may be chosen if no further information is available regarding the white point of a device, or regarding the illuminant of a scene. Its spectral power distribution is shown in Figure 2.11, along with two related standard illuminants, D_{55} (commonly used in photography) and D_{75}.

Cameras often operate under the assumption that the scene is lit by a specific light source, such as a D_{65}. If the lighting in a scene has a substantially different color, an adjustment to the gain of the red, green, and blue sensors in the camera

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 19

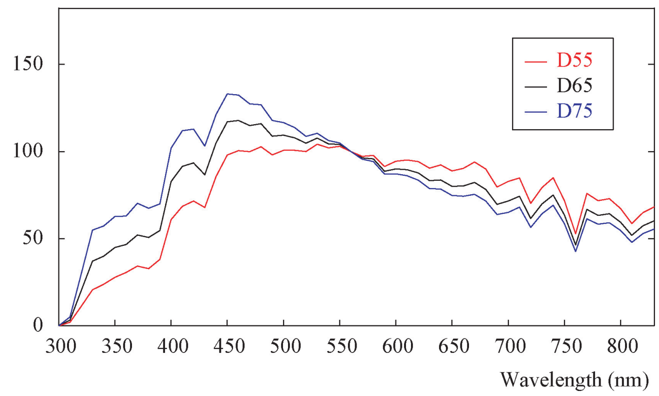


FIGURE 2.11 *Spectral power distribution of CIE illuminants D_{55}, D_{65}, and D_{75}.*

may be made. This is known as white balancing [100]. If the white balance chosen for a particular scene were incorrect, white balancing might be attempted as an image-processing step.

The difference between illuminants may be expressed in terms of chromaticity coordinates, but a more commonly used measure is correlated color temperature. Consider a blackbody radiator, a cavity in a block of material heated to a certain temperature. The spectral power distribution emitted by the walls of this cavity is a function of the temperature of the material only. The color of a blackbody radiator may thus be characterized by its temperature, which is measured in degrees Kelvin (K).

The term *color temperature* refers to the temperature of a selective radiator that has chromaticity coordinates very close to that of a blackbody. The lower the temperature the redder the appearance of the radiator. For instance, tungsten illumination (about 3,200° K) appears somewhat yellow. Higher color temperatures have a more bluish appearance.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 20

Scene	*T* (in °K)	*x*	*y*
Candle flame	1850	0.543	0.410
Sunrise/sunset	2000	0.527	0.413
Tungsten (TV/film)	3200	0.427	0.398
Summer sunlight at noon	5400	0.326	0.343
CIE A (incandescent)	2854	0.448	0.408
CIE B (direct sunlight)	4874	0.384	0.352
CIE C (overcast sky)	6774	0.310	0.316
CIE D50 (noon skylight)	5000	0.346	0.359
CIE D65 (average daylight)	6504	0.313	0.329
CIE E (equal energy)	5500	0.333	0.333
CIE F2 (office fluorescent)	4150	0.372	0.375

TABLE 2.4 *Correlated color temperature T and chromaticity coordinates (xy) for common scene types and a selection of CIE luminaires.*

The term *correlated color temperature* is more generally used for illuminants that do not have chromaticity coordinates close to those generated by blackbody radiators. It refers to the blackbody's temperature that most closely resembles the perceived color of the given selective radiator under the same brightness and specified viewing conditions. Table 2.4 outlines the correlated color temperature of several common scene types and CIE luminaires, as well as their associated chromaticity coordinates.

The CIE standard illuminant D_{65}, shown in Figure 2.11, is defined as natural daylight with a correlated color temperature of 6,504 K. The D_{55} and D_{75} illuminants have correlated color temperatures of 5,503 and 7,504 K, respectively. Many color spaces are defined with a D_{65} white point. In photography, D_{55} is often used. Display devices often use a white point of 9,300 K, which tends toward blue. The reason for this is that blue phosphors are relatively efficient and allow the overall display brightness to be somewhat higher, at the cost of color accuracy [100].

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 21

Humans are very capable of adapting to the color of the light source in a scene. The impression of color given by a surface depends on its reflectance as well as the light source illuminating it. If the light source is gradually changed in color, humans will adapt and still perceive the color of the surface the same, although light measurements of the surface would indicate a different spectral composition and CIE XYZ tristimulus value [125]. This phenomenon is called chromatic adaptation. The ability to perceive the color of a surface independently of the light source illuminating it is called color constancy.

Typically, when viewing a real scene an observer would be chromatically adapted to that scene. If an image of the same scene were displayed on a display device, the observer would be adapted to the display device and the scene in which the observer viewed the image. It is reasonable to assume that these two states of adaptation will generally be different. As such, the image shown is likely to be perceived differently than the real scene. Accounting for such differences should be an important aspect of HDR imaging, and in particular tone reproduction. Unfortunately, too many tone-reproduction operators ignore these issues, although the photoreceptor-based operator, iCAM, and the Multiscale Observer Model include a model of chromatic adaptation (see Sections 7.2.7, 7.3.3, and 7.3.4), and Akyuz et al. have shown that tone reproduction and color appearance modeling may be separated into two steps [4].

In 1902, von Kries speculated that chromatic adaptation is mediated by the three cone types in the retina [90]. Chromatic adaptation occurs as the red, green, and blue cones each independently adapts to the illuminant.

A model of chromatic adaptation may thus be implemented by transforming tristimulus values into a cone response domain and then individually scaling the red, green, and blue components according to the current and desired illuminants. There exist different definitions of cone response domains leading to different transforms. The first cone response domain is given by the LMS color space, with L, M, and S standing respectively for long, medium, and short wavelengths. The matrix that converts between XYZ and LMS lies at the heart of the von Kries transform and is denoted M_{vonKries}, as in the following.

$$M_{\mathrm{vonKries}} = \begin{bmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 22

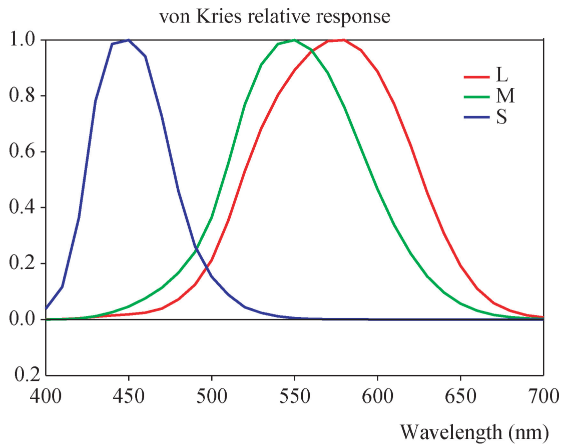


FIGURE 2.12 *Relative response functions for the von Kries chromatic adaptation transform. (Reprinted from [36]).*

$$M_{\text{vonKries}}^{-1} = \begin{bmatrix} 1.9102 & -1.1121 & 0.2019 \\ 0.3710 & 0.6291 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

As the LMS cone space represents the response of the cones in the human visual system, it is a useful starting place for computational models of human vision. It is also a component in the iCAM color appearance model (see Section 7.3.3). The relative response as a function of wavelength is plotted in Figure 2.12.

A newer cone response domain is given by the Bradford chromatic adaptation transform [64,76] (see Figure 2.13), as follows.

$$M_{\text{Bradford}} = \begin{bmatrix} 0.8951 & 0.2664 & -0.1614 \\ -0.7502 & 1.7135 & 0.0367 \\ 0.0389 & -0.0685 & 1.0296 \end{bmatrix}$$

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 23

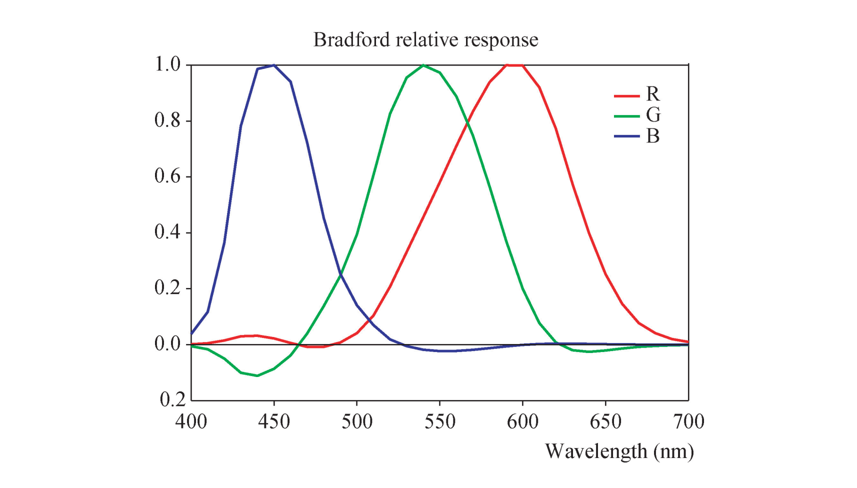


FIGURE 2.13 *Relative response functions for the Bradford chromatic adaptation transform. (Reprinted from [36].)*

$$M_{\text{Bradford}}^{-1} = \begin{bmatrix} 0.9870 & -0.1471 & 0.1600 \\ 0.4323 & 0.5184 & 0.0493 \\ -0.0085 & 0.0400 & 0.9685 \end{bmatrix}$$

A third chromatic adaptation transform (see Figure 2.14) is used in the CIECAM02 color appearance model (described in Section 2.8), as follows

$$M_{\text{CAT02}} = \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{bmatrix}$$

$$M_{\text{CAT02}}^{-1} = \begin{bmatrix} 1.0961 & -0.2789 & 0.1827 \\ 0.4544 & 0.4735 & 0.0721 \\ -0.0096 & -0.0057 & 1.0153 \end{bmatrix}$$

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 24

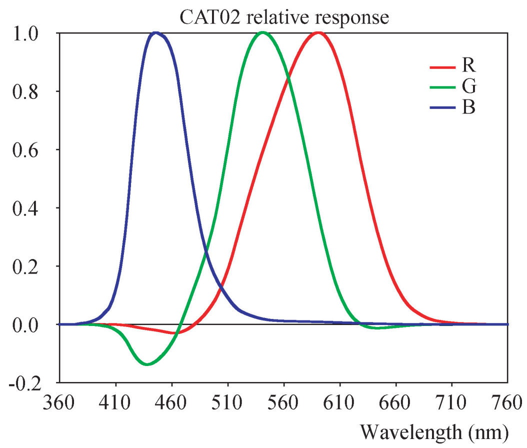


FIGURE 2.14 *Relative response functions for the CAT02 chromatic adaptation transform. (Reprinted from Mark Fairchild's slides.)*

These three chromatic adaptation transforms may be used to construct a matrix that will transform XYZ tristimulus values for a given white point to a new white point [126]. If the source white point is given as (X_S, Y_S, Z_S) and the destination white point as (X_D, Y_D, Z_D), their transformed values are

$$\begin{bmatrix} \rho_S \\ \gamma_S \\ \beta_S \end{bmatrix} = M_{\mathrm{cat}} \begin{bmatrix} X_S \\ Y_S \\ Z_S \end{bmatrix}$$

$$\begin{bmatrix} \rho_D \\ \gamma_D \\ \beta_D \end{bmatrix} = M_{\mathrm{cat}} \begin{bmatrix} X_D \\ Y_D \\ Z_D \end{bmatrix},$$

where M_{cat} is one of the three chromatic adaptation matrices M_{vonKries}, M_{Bradford}, or M_{CAT02}. A chromatic adaptation matrix for these specific white points may be constructed by concatenating the previously cited von Kries or Bradford matrices with a diagonal matrix that independently scales the three cone responses, as follows.

$$M = M_{\text{cat}}^{-1} \begin{bmatrix} \rho_{\text{D}}/\rho_{\text{S}} & 0 & 0 \\ 0 & \gamma_{\text{D}}/\gamma_{\text{S}} & 0 \\ 0 & 0 & \beta_{\text{D}}/\beta_{\text{S}} \end{bmatrix} M_{\text{cat}}$$

Chromatically adapting an XYZ tristimulus value is now a matter of transforming it with matrix M, as follows.

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = M \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Here, (X', Y', Z') is the CIE tristimulus value whose appearance under the target illuminant most closely matches the original XYZ tristimulus under the source illuminant.

Chromatic adaptation transforms are useful for preparing an image for display under different lighting conditions. Thus, if the scene were lit by daylight and an image of that scene viewed under tungsten lighting, a chromatic adaptation transform might be used to account for this difference. After applying the chromatic adaptation transform, the (X', Y', Z') tristimulus values need to be converted to an RGB color space with a matrix that takes into account the white point of the display environment. Thus, if the image is to be viewed under tungsten lighting, the XYZ-to-RGB transformation matrix should be constructed using the white point of a tungsten light source.

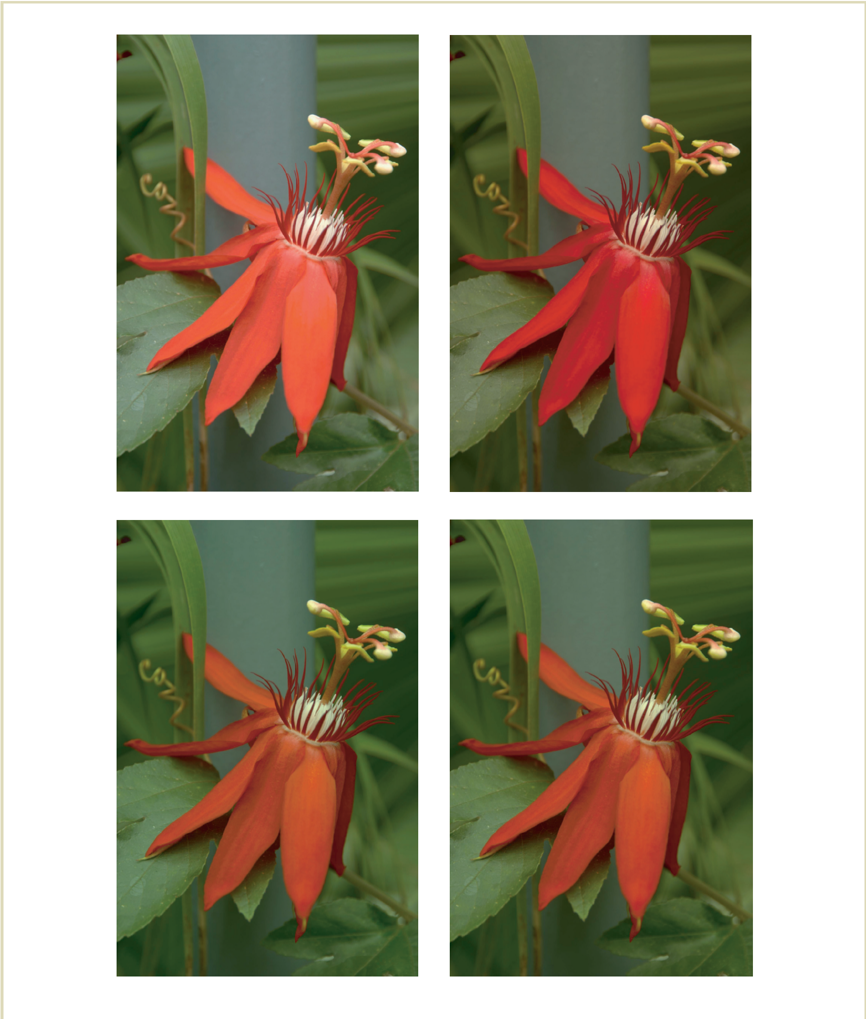
As an example, Figure 2.15 shows an image lit with daylight approximating D_{65}.[3] This figure shows the image prepared for several different viewing environments. In each case, the CAT02 chromatic adaptation transform was used, and the conversion to RGB color space was achieved by constructing a conversion matrix with the appropriate white point.

3 This image was taken in a conservatory in Rochester, New York, under cloud cover. The CIE D_{65} standard light source was derived from measurements originally taken from similar daylight conditions in Rochester.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 26

FIGURE 2.15 CAT02 chromatic adaptation. In reading order: original image, followed by five images chromatically adapted from D_{65} to incandescent, tungsten, D_{50}, E, and F2.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 27



reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 28

FIGURE 2.16 *(continued) Comparison of different chromatic adaptation transforms. In reading order: original image, followed by von Kries, Bradford, and CAT02 transforms. The final image is the chromatic adaptation transform applied directly in XYZ space. The transform is from D_{65} to tungsten.*

The difference among the three different chromatic adaptation transforms is illustrated in Figure 2.16. Also shown in this figure is a chromatic adaptation performed directly in XYZ space, here termed *XYZ scaling*. The scene depicted here was created with only the outdoor lighting available and was taken in the same conservatory as the images in Figure 2.15. Thus, the lighting in this scene would be reasonably well approximated with a D_{65} luminant. Figure 2.16 shows transforms from D_{65} to tungsten.

The spectral sensitivities of the cones in the human visual system are broadband; that is, each of the red, green, and blue cone types (as well as the rods) are sensitive to a wide range of wavelengths, as indicated by their absorbance spectra (shown

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 29

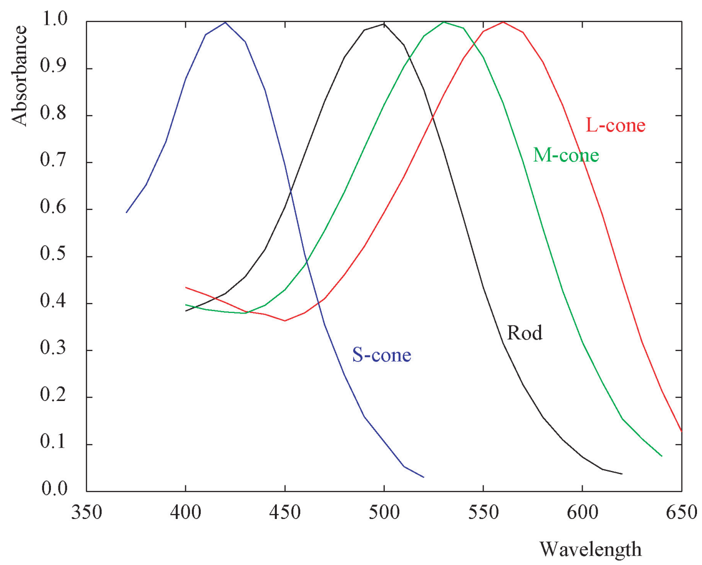


FIGURE 2.17 *Spectral absorbance spectra for the L, M, and S cones, as well as the rods (after [16]).*

in Figure 2.17) [16]. As a result, there is significant overlap between the different cone types, although their peak sensitivities lie at different wavelengths.

It is possible to construct new spectral response functions that are more narrow-band by computing a linear combination of the original response functions. The graphs of the resulting response functions look sharper, and the method is therefore called "spectral sharpening." Within a chromaticity diagram, the three corners of the color gamut lie closer to the spectral locus, or even outside, and therefore the gamut is "wider" so that a greater range of visible colors can be represented.

A second advantage of applying such a transform is that the resulting tristimulus values become more decorrelated. This has advantages in color constancy algo-

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 30

rithms; that is, algorithms that aim to recover surface reflectance from an image that has recorded the combined effect of surface reflectance and illuminance [7]. It also helps to reduce visible errors in color-rendering algorithms [208].

2.6 COLOR CORRECTION

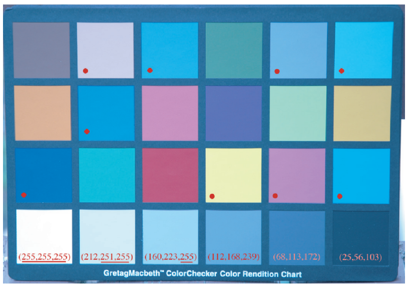
Without the camera response function (Section 4.6), one cannot linearize the input as needed for color correction. Thus, a color-correction value will not apply equally to all parts of the tone scale. For instance, darker colors may end up too blue compared to lighter colors. Furthermore, colors with primary values clamped to the upper limit (255 in an 8-bit image) have effectively been desaturated by the camera. Although users are accustomed to this effect in highlights, after color correction such desaturated colors may end up somewhere in the midtones, where desaturation is unexpected. In a naïve method, whites may even be moved to some nonneutral value, which can be very disturbing.

Figure 2.18 demonstrates the problem of color correction from an LDR original. If the user chooses one of the lighter patches for color balancing, the result may be incorrect due to clamping in its value. (The captured RGB values for the gray patches are shown in red.) Choosing a gray patch without clamping avoids this problem, but it is impossible to recover colors for the clamped patches. In particular, the lighter neutral patches end up turning pink in this example. The final image shows how these problems are avoided when an HDR original is available. Because the camera response curve has been eliminated along with clamping, the simple

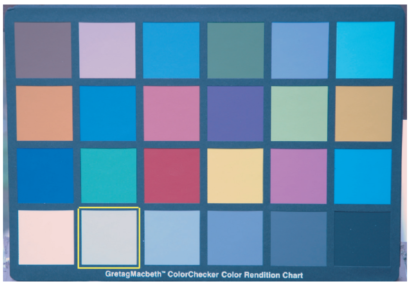
FIGURE 2.18 *(a) A Macbeth ColorChecker chart captured with the appropriate white balance setting under an overcast sky; (b) the same scene captured using the "incandescent" white balance setting, resulting in a bluish color cast (red dots mark patches that cannot be corrected because one or more primaries are clamped to 255); (c) an attempt to balance white using the second gray patch, which was out of range in the original; (d) the best attempt at correction using the fourth gray patch, which was at least in range in the original; and (e) range issues disappear in an HDR original, allowing for proper post-correction.*

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 31

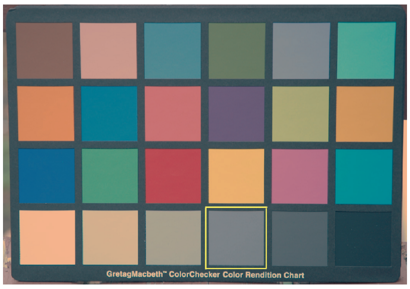
(a) *White balanced capture*



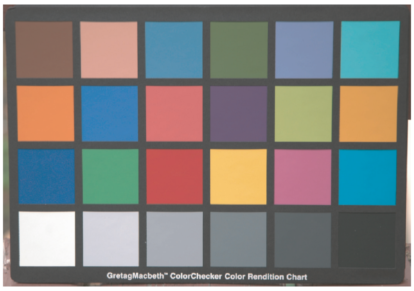
(b) *Off-balance capture*



(c) *LDR white balancing using second patch*



(d) *LDR white balancing using fourth patch*



(e) *HDR white balancing*

approach of balancing colors by choosing a neutral patch and multiplying the image by its inverse works quite well.

2.7 COLOR OPPONENT SPACES

With a 3-by-3 matrix, pixel data may be rotated into different variants of RGB color spaces to account for different primaries. A feature shared by all RGB color spaces is that for natural images correlations exist between the values in each color channel. In other words, if a pixel of a natural image has a large value for the red component, the probability of also finding a large value for the green and blue components is high. Thus, the three channels are highly correlated.

An example image is shown in Figure 2.19. A set of randomly selected pixels is plotted three times in the same figure, where the axes of the plot are R-G, R-B, and G-B. This plot shows a point cloud of pixel data at an angle of about 45 degrees, no matter which channel is plotted against which. Thus, for this natural image strong correlations exist between the channels in RGB color space.

This means that the amount of information carried by the three values comprising a pixel is less than three times the amount of information carried by each of the values. Thus, each color pixel carries some unquantified amount of redundant information.

The human visual system deals with a similar problem. The information captured by the photoreceptors needs to be transmitted to the brain through the optic nerve. The amount of information that can pass through the optic nerve is limited and constitutes a bottleneck. In particular, the number of photoreceptors in the retina is far larger than the number of nerve endings that connect the eye to the brain.

After light is absorbed by the photoreceptors, a significant amount of processing occurs in the next several layers of cells before the signal leaves the eye. One type of processing is a color space transformation to a *color opponent space*. Such a color space is characterized by three channels; a luminance channel, a red-green channel, and a yellow-blue channel.

The luminance channel ranges from dark to light and bears resemblance to the Y channel in CIE XYZ color space. The red-green channel ranges from red to green via neutral gray. The yellow-blue channel encodes the amount of blue versus the amount of yellow in a similar way to the red-green channel (Figure 2.20). This

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 33

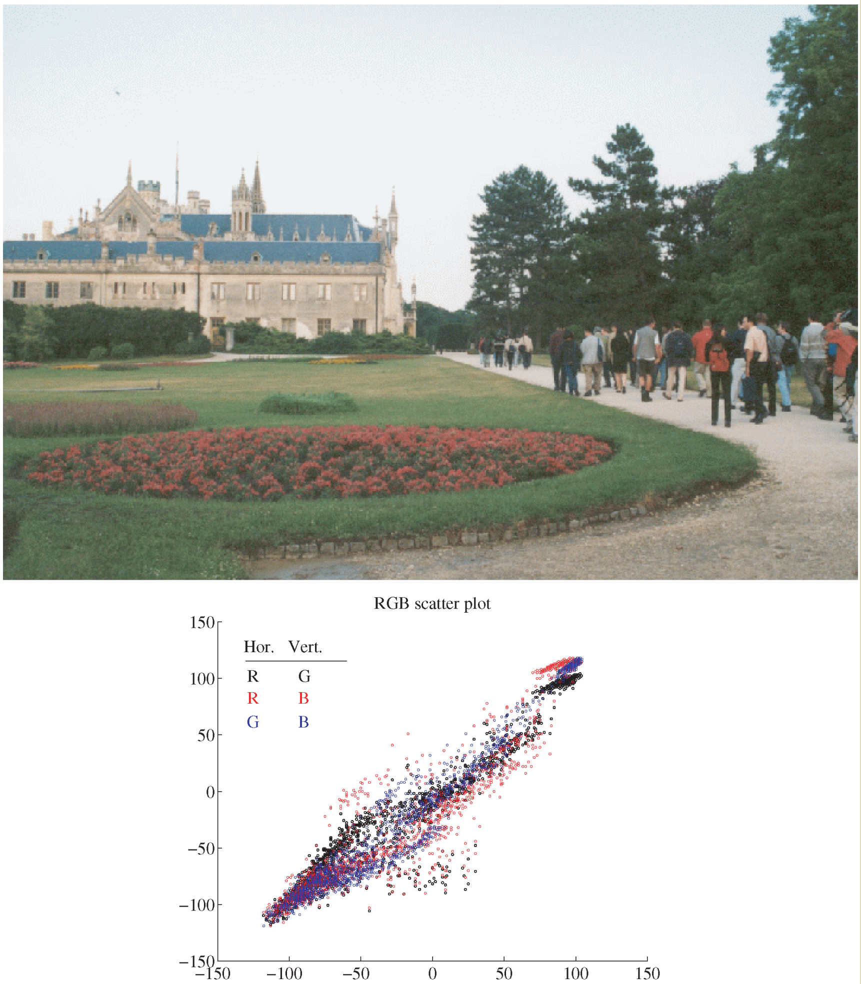


FIGURE 2.19 Scatter plot of RGB pixels randomly selected from the image at the top.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 34

FIGURE 2.20 *Original image (top left) split into a luminance channel (top right), a yellow-blue channel (bottom left), and a red-green channel (bottom right). For the purpose of visualization, the images depicting the yellow-blue and red-green channels are shown with the luminance component present.*

encoding of chromatic information is the reason humans are able to describe colors as reddish yellow (orange) or greenish blue. However, colors such as bluish yellow and reddish green are never described because of this encoding (see [92]).

It is possible to analyze sets of natural images by means of principal components analysis (PCA) [110]. This technique rotates multidimensional data such that the

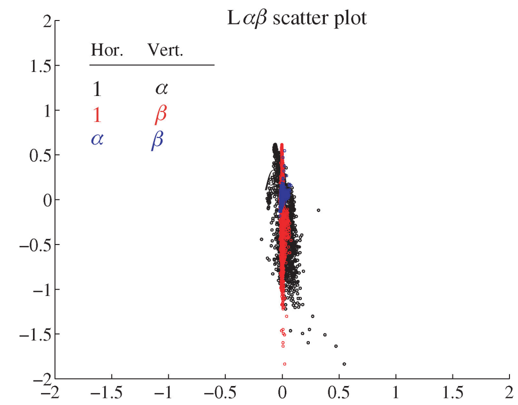


FIGURE 2.21 Scatter plot of $L\alpha\beta$ pixels randomly selected from the image of Figure 2.19.

axes align with the data as well as possible. Thus, the most important axis aligns with the direction in space that shows the largest variation of data points. This is the first principal component. The second principal component describes the direction accounting for the second greatest variation in the data. This rotation therefore decorrelates the data.

If the technique is applied to images encoded in LMS color space (i.e., images represented in a format as thought to be output by the photoreceptors), a new set of decorrelated axes is produced. The surprising result is that the application of PCA to a set of natural images produces a color space that is closely matched to the color opponent space the human visual system employs [110].

A scatter plot of the image of Figure 2.19 in a color opponent space ($L\alpha\beta$, discussed later in this section) is shown in Figure 2.21. Here, the point clouds are reasonably well aligned with one of the axes, indicating that the data is now decorrelated. The elongated shape of the point clouds indicates the ordering of the principal axes, luminance being most important and therefore most elongated.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 36

The decorrelation of data may be important, for instance, for color-correction algorithms. What would otherwise be a complicated 3D problem may be cast into three simpler 1D problems by solving the problem in a color opponent space [107].

At the same time, the first principal component (the luminance channel) accounts for the greatest amount of variation, whereas the two chromatic color opponent channels carry less information. Converting an image into a color space with a luminance channel and two chromatic channels thus presents an opportunity to compress data because the latter channels would not require the same number of bits as the luminance channel to accurately represent the image. The color opponent space $L\alpha\beta$ that results from applying PCA to natural images may be approximated by the following matrix transform, which converts between $L\alpha\beta$ and LMS (see Section 2.5).

$$\begin{bmatrix} L \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix}$$

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{3} & 0 & 0 \\ 0 & \frac{\sqrt{6}}{6} & 0 \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} L \\ \alpha \\ \beta \end{bmatrix}$$

This color space has proved useful in algorithms such as the transfer of color between images, where the colors are borrowed from one image and applied to a second image [107]. This algorithm computes means and standard deviations for each channel separately in both source and target images. Then, the pixel data in the target image are shifted and scaled such that the same mean and standard deviation as the source image are obtained. Applications of color transfer include the work of colorists, compositing, and matching rendered imagery with live video footage in mixed-reality applications.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 37

In addition, human sensitivity to chromatic variations is lower than to changes in luminance. Chromatic channels may therefore be represented at a lower spatial resolution than the luminance channel. This feature may be exploited in image encodings by sampling the image at a lower resolution for the color opponent channels than for the luminance channel. This is demonstrated in Figure 2.22, where the

FIGURE 2.22 *The red-green and yellow-blue channels are reduced in spatial resolution by a factor of 1, 2, 4, 8, 16, and 32.*

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 38

full resolution image is shown on the left. The spatial resolution of the red-green and yellow-blue channels is reduced by a factor of two for each subsequent image. In Figure 2.23, the luminance channel was also reduced by a factor of two. The artifacts in Figure 2.22 are much more benign than those in Figure 2.23.

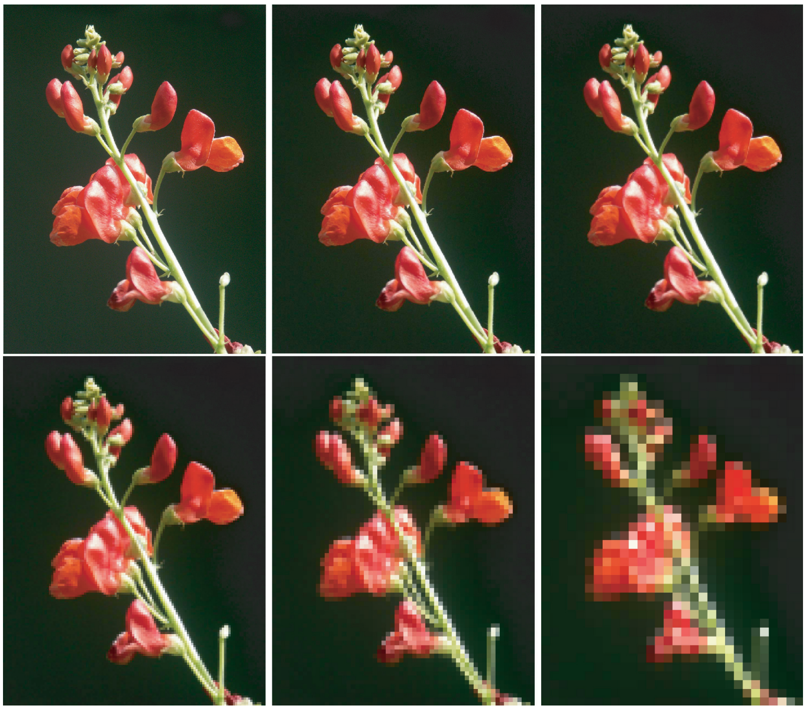


FIGURE 2.23 *All three channels in $L\alpha\beta$ space are reduced in spatial resolution by a factor of 1, 2, 4, 8, 16, and 32.*

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 39

Subsampling of chromatic channels is used, for instance, in the YC_BC_R encoding that is part of the JPEG file format and part of various broadcast standards, including HDTV [100]. Conversion from RGB to YC_BC_R and back as used for JPEG is given by

$$\begin{bmatrix} Y \\ C_B \\ C_R \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.168 & -0.333 & 0.498 \\ 0.498 & -0.417 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.000 & 1.397 \\ 1.000 & -0.343 & -0.711 \\ 1.000 & 1.765 & 0.000 \end{bmatrix} \begin{bmatrix} Y \\ C_B \\ C_R \end{bmatrix}.$$

This conversion is based on ITU-R BT.601 [100]. Other color spaces which have one luminance channel and two chromatic channels, such as CIELUV and CIELAB, are discussed in the following section.

2.8 COLOR APPEARANCE

The human visual system adapts to the environment it is viewing (see Chapter 6 for more information). Observing a scene directly therefore generally creates a different visual sensation than observing an image of that scene on a (LDR) display. In the case of viewing a scene directly, the observer will be adapted to the scene. When looking at an image of a display, the observer will be adapted to the light emitted from the display, as well as to the environment in which the observer is located.

There may therefore be a significant mismatch between the state of adaptation of the observer in these two cases. This mismatch may cause the displayed image to be perceived differently from the actual scene. The higher the dynamic range of the scene the larger this difference may be. In HDR imaging, and in particular tone reproduction, it is therefore important to understand how human vision adapts to various lighting conditions and to develop models that predict how colors will be perceived under such different lighting conditions. This is the domain of color appearance modeling [27].

A color's appearance is influenced by various aspects of the viewing environment, such as the illuminant under which the stimulus is viewed. The chromatic adaptation transforms discussed in Section 2.5 are an important component of most color appearance models.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 40

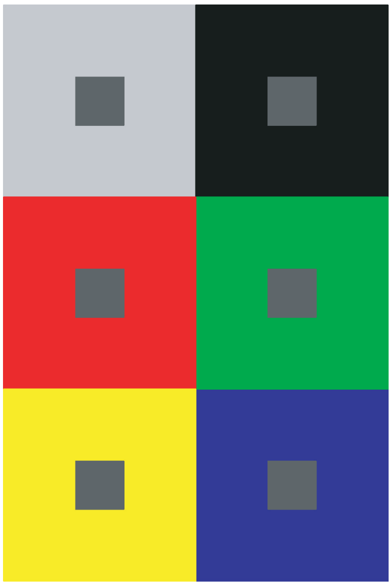


FIGURE 2.24 *Simultaneous color contrast shown for an identical gray patch displayed on differently colored backgrounds.*

The color of the area surrounding the stimulus also plays an important role, as demonstrated in Figure 2.24, where the same gray patch is shown on different backgrounds. The color of the patch will appear different in each case — an effect known as simultaneous color contrast.

To characterize a stimulus within a specific environment, first its tristimulus value is specified in CIE XYZ color space. Second, attributes of the environment in which the stimulus is viewed need to be provided. If the stimulus is a homoge-

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 41

neous reflecting patch of color on a neutral (gray) background, this characterization of the environment may be as simple as the specification of an illuminant.

The appearance of a color is then described by "appearance correlates" that may be computed from the color's tristimulus values as well as the description of the environment. Useful appearance correlates include lightness, chroma, hue, and saturation, which are defined later in this section.

Appearance correlates are not computed directly in XYZ color space, but require an intermediate color space such as the CIE 1976 $L^*u^*v^*$ or CIE 1976 $L^*a^*b^*$ color spaces. The names of these color spaces may be abbreviated as CIELUV and CIELAB, respectively.

For both of these color spaces it is assumed that a stimulus (X, Y, Z) is formed by a white reflecting surface that is lit by a known illuminant with tristimulus values $(X_\mathrm{n}, Y_\mathrm{n}, Z_\mathrm{n})$. The conversion from CIE 1931 tristimulus values to CIELUV is then given by the following.

$$
\begin{aligned}
L^* &= 116\left(\frac{Y}{Y_\mathrm{n}}\right)^{1/3} - 16 \\
u^* &= 13L^*(u' - u'_\mathrm{n}) \\
v^* &= 13L^*(v' - v'_\mathrm{n})
\end{aligned}
$$

This conversion is under the constraint that $Y/Y_\mathrm{n} > 0.008856$. For ratios smaller than 0.008856, L^*_m is applied as follows.

$$L^*_\mathrm{m} = 903.3\frac{Y}{Y_\mathrm{n}}$$

The primed quantities in these equations are computed from (X, Y, Z) as follows.

$$
\begin{aligned}
u' &= \frac{4X}{X + 15Y + 3Z} & \qquad u'_\mathrm{n} &= \frac{4X_\mathrm{n}}{X_\mathrm{n} + 15Y_\mathrm{n} + 3Z_\mathrm{n}} \\
v' &= \frac{9Y}{X + 15Y + 3Z} & \qquad v'_\mathrm{n} &= \frac{9Y_\mathrm{n}}{X_\mathrm{n} + 15Y_\mathrm{n} + 3Z_\mathrm{n}}
\end{aligned}
$$

This transformation creates a more or less uniform color space, such that equal distances anywhere within this space encode equal perceived color differences. It is

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 42

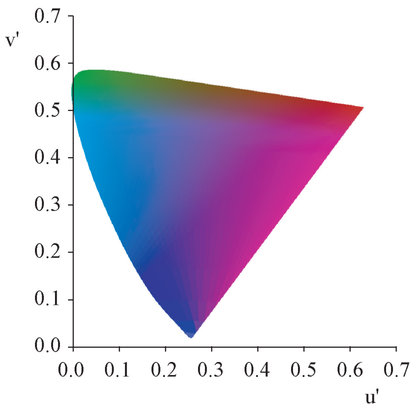


FIGURE 2.25 CIE (u', v') *chromaticity diagram showing the range of colors humans can distinguish.*

therefore possible to measure the difference between two stimuli (L_1^*, u_1^*, v_1^*) and (L_2^*, u_2^*, v_2^*) by encoding them in CIELUV space, and applying the color difference formula

$$\Delta E_{\mathrm{uv}}^* = \left[(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2\right]^{1/2},$$

where $\Delta L^* = L_1^* - L_2^*$, etc.

In addition, u' and v' may be plotted on separate axes to form a chromaticity diagram, as shown in Figure 2.25. Equal distances in this diagram represent approximately equal perceptual differences. For this reason, in the remainder of this book CIE (u', v') chromaticity diagrams are shown rather than perceptually nonuniform CIE (x, y) chromaticity diagrams. The CIELAB color space follows a similar approach. For the ratios X/X_{n}, Y/Y_{n} and Z/Z_{n}, each being larger than 0.008856,

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 43

the color space is defined by

$$L^* = 116\left(\frac{Y}{Y_\mathrm{n}}\right)^{1/3} - 16$$

$$a^* = 500\left[\left(\frac{X}{X_\mathrm{n}}\right)^{1/3} - \left(\frac{Y}{Y_\mathrm{n}}\right)^{1/3}\right]$$

$$b^* = 200\left[\left(\frac{Y}{Y_\mathrm{n}}\right)^{1/3} - \left(\frac{Z}{Z_\mathrm{n}}\right)^{1/3}\right].$$

If any ratio is smaller than 0.008856, the modified quantities L^*_m, a^*_m, and b^*_m may be computed as follows.

$$L^*_\mathrm{m} = 903.3\frac{Y}{Y_\mathrm{n}} \qquad \text{for } \frac{Y}{Y_\mathrm{n}} \le 0.008856$$

$$a^*_\mathrm{m} = 500\left[f\left(\frac{X}{X_\mathrm{n}}\right) - f\left(\frac{Y}{Y_\mathrm{n}}\right)\right]$$

$$b^*_\mathrm{m} = 200\left[f\left(\frac{Y}{Y_\mathrm{n}}\right) - f\left(d\frac{Z}{Z_\mathrm{n}}\right)\right]$$

The function $f(.)$ takes a ratio as argument in the previous equations. If either of these ratios is denoted as r, $f(r)$ is defined as

$$f(r) = \begin{cases} (r)^{1/3} & \text{for } r > 0.008856 \\ 7.787r + \dfrac{16}{116} & \text{for } r \le 0.008856. \end{cases}$$

Within this color space, which is also approximately perceptually linear, the difference between two stimuli may be quantified with the following color difference formula.

$$\Delta E^*_\mathrm{ab} = \left[(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2\right]^{1/2}$$

The reason for the existence of both of these color spaces is largely historical. Both color spaces are in use today, with CIELUV more common in the television and display industries and CIELAB in the printing and materials industries [125].

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 44

Although CIELUV and CIELAB by themselves are perceptually uniform color spaces, they may also form the basis for color appearance models. The perception of a set of tristimulus values may be characterized by computing appearance correlates [27]. Our definitions are based on Wyszecki and Stiles' book *Color Science* [149].

> *Brightness:* The attribute of visual sensation according to which a visual stimulus appears to emit more or less light is called brightness, which ranges from bright to dim.
>
> *Lightness:* The area in which a visual stimulus is presented may appear to emit more or less light in proportion to a similarly illuminated area that is perceived as a white stimulus. Lightness is therefore a relative measure and may be seen as relative brightness. Lightness ranges from light to dark. In both CIELUV and CIELAB color spaces, L^* is the correlate for lightness. Note that if the luminance value of the stimulus is about 18% of Y_{n} (i.e., $Y/Y_{\mathrm{n}} = 0.18$), the correlate for lightness becomes about 50, which is halfway on the scale between light and dark. In other words, surfaces with 18% reflectance appear as middle gray. In photography, 18% gray cards are often used as calibration targets for this reason.[4]
>
> *Hue:* The attribute of color perception denoted by red, green, blue, yellow, purple, and so on is called hue. A chromatic color is perceived as possessing hue. An achromatic color is not perceived as possessing hue. Hue angles h_{uv} and h_{ab} may be computed as follows.
>
> $$h_{\mathrm{uv}} = \arctan \frac{v^*}{u^*}$$
> $$h_{\mathrm{ab}} = \arctan \frac{a^*}{b^*}$$
>
> *Chroma:* A visual stimulus may be judged in terms of its difference with an achromatic stimulus with the same brightness. This attribute of visual sen-

4 Although tradition is maintained and 18% gray cards continue to be used, the average scene reflectance is often closer to 13%.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 45

sation is called chroma. Correlates of chroma may be computed in both CIELUV (C^*_{uv}) and CIELAB (C^*_{uv}) as follows.

$$C^*_{\mathrm{uv}} = \left[(u^*)^2 + (v^*)^2\right]^{1/2}$$
$$C^*_{\mathrm{ab}} = \left[(a^*)^2 + (b^*)^2\right]^{1/2}$$

Saturation: Whereas chroma pertains to stimuli of equal brightness, saturation is an attribute of visual sensation which allows the difference of a visual stimulus and an achromatic stimulus to be judged regardless of any differences in brightness. In CIELUV, a correlate for saturation s^*_{uv} may be computed as follows.

$$s^*_{\mathrm{uv}} = \frac{C^*_{\mathrm{uv}}}{L^*}$$

A similar correlate for saturation is not available in CIELAB.

Several more color appearance models have recently appeared. The most notable among these are CIECAM97 [12,28,54,85], which exists in both full and simplified versions, and CIECAM02 [74,84]. As with the color spaces mentioned previously, their use is in predicting the appearance of stimuli placed in a simplified environment. They also allow conversion of stimuli between different display media, such as different computer displays that may be located in different lighting environments. These recent color appearance models are generally more complicated than the procedures described in this section, but are also deemed more accurate.

The CIECAM97 and CIECAM02 color appearance models, as well as several of their predecessors, follow a general structure but differ in their details. We outline this structure using the CIECAM02 model as an example [74,84].

This model works under the assumption that a target patch with given relative tristimulus value XYZ is viewed on a neutral background and in the presence of a white reflective patch, which acts as the reference white (i.e., it is the brightest part of the environment under consideration). The background is again a field of limited size. The remainder of the visual field is taken up by the surround. This simple environment is lit by an illuminant with given relative tristimulus values $X_{\mathrm{W}}Y_{\mathrm{W}}Z_{\mathrm{W}}$. Both of these relative tristimulus values are specified as input and are normalized between 0 and 100.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 46

Surround	F	c	N_c
Average	1.0	0.69	1.0
Dim	0.9	0.59	0.95
Dark	0.8	0.525	0.8

TABLE 2.5 *Values for intermediary parameters in the CIECAM02 model as a function of the surround description.*

The luminance measured from the reference white patch is then assumed to be the adapting field luminance L_a—the only absolute input parameter, measured in cd/m^2. The neutral gray background has a luminance less than or equal to the adapting field luminance. It is denoted Y_b and is specified as a fraction of L_a, also normalized between 0 and 100.

The final input to the CIECAM02 color appearance model is a classifier describing the surround as average, dim, or dark. This viewing condition parameter is used to select values for the intermediary parameters F, c, and N_c according to Table 2.5. Further intermediary parameters n, N_{bb}, N_{cb}, and z are computed from the input as follows.

$$n = \frac{Y_b}{Y_W}$$

$$N_{cb} = 0.725\left(\frac{1}{n}\right)^{0.2}$$

$$N_{bb} = N_{cb}$$

$$z = 1.48 + \sqrt{n}$$

Next, a factor F_{L} is computed from the adapting field luminance, which accounts for the partial adaptation to overall light levels. This takes the following form.

$$k = \frac{1}{5L_{\mathrm{a}} + 1}$$
$$F_{\mathrm{L}} = 0.2k^4(5L_{\mathrm{a}}) + 0.1(1 - k^4)^2(5L_{\mathrm{a}})^{1/3} \tag{2.1}$$

The CIECAM02 color appearance model, and related models, proceed with the following three main steps.

- Chromatic adaptation
- Nonlinear response compression
- Computation of perceptual appearance correlates

The chromatic adaptation transform is performed in the CAT02 space, outlined in Section 2.5. The XYZ and $X_{\mathrm{W}}Y_{\mathrm{W}}Z_{\mathrm{W}}$ tristimulus values are first converted to this space, as follows.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M_{\mathrm{CAT02}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Then a degree of adaptation D is computed, which determines how complete the adaptation is. It is a function of the adapting field luminance as well as the surround (through the parameters L_{a} and F). This takes the following form.

$$D = F\left[1 - \frac{1}{3.6}\exp\left(\frac{-L_{\mathrm{a}} - 42}{92}\right)\right]$$

The chromatically adapted signals are then computed, as follows.

$$R_{\mathrm{c}} = R\left[\left(D\,\frac{Y_{\mathrm{W}}}{R_{\mathrm{W}}}\right) + (1 - D)\right]$$

$$G_{\mathrm{c}} = G\left[\left(D\,\frac{Y_{\mathrm{W}}}{G_{\mathrm{W}}}\right) + (1 - D)\right]$$

$$B_{\mathrm{c}} = B\left[\left(D\,\frac{Y_{\mathrm{W}}}{B_{\mathrm{W}}}\right) + (1 - D)\right]$$

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 48

After applying this chromatic adaptation transform, the result is converted back to XYZ space.

The second step of the CIECAM02 model is the nonlinear response compression, which is carried out in the Hunt–Pointer–Estevez color space, which is close to a cone fundamental space such as LMS (see Section 2.5). Conversion from XYZ to this color space is governed by the following matrix.

$$M_{\mathrm{H}} = \begin{bmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

The chromatically adapted signal after conversion to the Hunt–Pointer–Estevez color space is indicated with the $(R'G'B')$ triplet. The nonlinear response compression yields a compressed signal $(R'_{\mathrm{a}}G'_{\mathrm{a}}B'_{\mathrm{a}})$, as follows.

$$R'_{\mathrm{a}} = \frac{400(F_{\mathrm{L}}R'/100)^{0.42}}{27.13 + (F_{\mathrm{L}}R'/100)^{0.42}} + 0.1$$

$$G'_{\mathrm{a}} = \frac{400(F_{\mathrm{L}}G'/100)^{0.42}}{27.13 + (F_{\mathrm{L}}B'/100)^{0.42}} + 0.1$$

$$B'_{\mathrm{a}} = \frac{400(F_{\mathrm{L}}B'/100)^{0.42}}{27.13 + (F_{\mathrm{L}}B'/100)^{0.42}} + 0.1$$

This response compression function follows an S shape on a log-log plot, as shown in Figure 2.26.

The final step consists of computing perceptual appearance correlates. These describe the perception of the patch in its environment, and include lightness, brightness, hue, chroma, colorfulness, and saturation. First a set of intermediary parameters is computed, as follows, which includes a set of color opponent signals a and b, a magnitude parameter t, an achromatic response A, hue angle h, and eccentricity factor e.

$$a = R'_{\mathrm{a}} - 12G'_{\mathrm{a}}/11 + B'_{\mathrm{a}}/11$$

$$b = (R'_{\mathrm{a}} + G'_{\mathrm{a}} - 2B'_{\mathrm{a}})/9$$

$$h = \tan^{-1}(b/a)$$

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 49

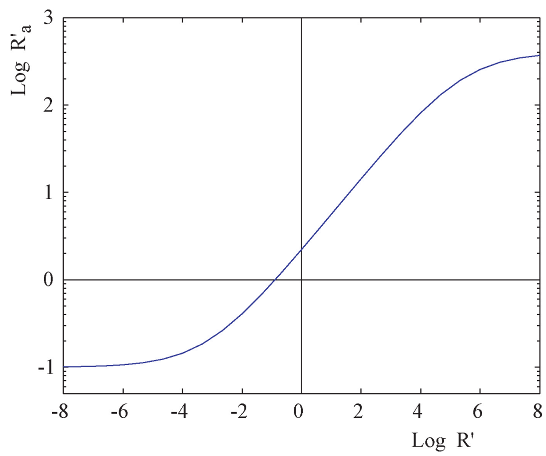


FIGURE 2.26 CIECAM02 nonlinear response compression on log-log axes.

$$
\begin{aligned}
e &= \left(\frac{12{,}500}{13} N_\mathrm{c} N_\mathrm{cb}\right)\left[\cos\left(\frac{h\pi}{180}+2\right)+3.8\right] \\
t &= \frac{e\sqrt{a^2+b^2}}{R'_\mathrm{a}+G'_\mathrm{a}+21B'_\mathrm{a}/20} \\
A &= [2R'_\mathrm{a}+G'_\mathrm{a}+B'_\mathrm{a}/20-0.305]N_\mathrm{bb}
\end{aligned}
$$

The unique hues red, yellow, green, and blue have values for h as given in Table 2.6. The hue angles h_1 and h_2 for the two nearest unique hues are determined from the value of h and Table 2.6. Similarly, eccentricity factors e_1 and e_2 are derived from this table and the value of e. The hue composition term H_i of the next lower unique hue is also read from this table. The appearance correlates may then be computed with the following equations, which are estimates for hue H, lightness

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 50

Unique Hue	Hue Angle	Eccentricity Factor	Hue Composition
Red	20.14	0.8	0
Yellow	90.00	0.7	100
Green	164.25	1.0	200
Blue	237.53	1.2	300

TABLE 2.6 *Hue angles h, eccentricity factors e, and hue composition H_i for the unique hues red, yellow, green, and blue.*

J, brightness Q, chroma C, colorfulness M, and saturation s.

$$H = H_i + \frac{100(h - h_1)/e_1}{(h - h_1)/e_1 + (h2 - h)/e_2}$$

$$J = 100\left(\frac{A}{A_W}\right)^{cz}$$

$$Q = \left(\frac{4}{c}\right)\sqrt{\frac{J}{100}}(A_W + 4)F_L^{0.25}$$

$$C = t^{0.9}\sqrt{\frac{J}{100}}(1.64 - 0.29^n)^{0.73}$$

$$M = C F_L^{0.25}$$

$$s = 100\sqrt{\frac{M}{Q}}$$

These appearance correlates thus describe the tristimulus value XYZ in the context of its environment. Thus, by changing the environment only the perception of this patch will change and this will be reflected in the values found for these appearance correlates. In practice, this would occur, for instance, when an image displayed on a monitor and printed on a printer needs to appear the same. Although colorimetry may account for the different primaries of the two devices, color appearance modeling additionally predicts differences in color perception due to the state of adaptation of the human observer in both viewing conditions.

If source and target viewing conditions are known, color appearance models may be used to convert a tristimulus value from one viewing condition to the other. The first two steps of the model (chromatic adaptation and nonlinear response compression) would then be applied, followed by the inverse of these two steps. During execution of the inverse model, the parameters describing the target environment (adapting field luminance, tristimulus value of the reference white, and so on) would be substituted into the model.

The field of color appearance modeling is currently dominated by two trends. The first is that there is a realization that the visual environment in which a stimulus is observed is in practice much more complicated than a uniform field with a given luminance. In particular, recent models are aimed at modeling the appearance of a pixel's tristimulus values in the presence of neighboring pixels in an image. Examples of models that begin to address these spatial configurations are the S-CIELAB and iCAM models [29,30,61,86,151].

A second trend in color appearance modeling constitutes a novel interest in applying color appearance models to HDR data. In particular, there is a mismatch in adaptation of the human visual system in a typical scene involving high contrast ratios and a human observer in front of a typical display device. Thus, if an accurate HDR capture of a scene is tone mapped and displayed on a computer monitor, the state of adaptation of the human observer in the latter case may cause the scene to appear different from the original scene.

The iCAM "image appearance model," derived from CIECAM02, is specifically aimed at addressing these issues [29,61], and in fact may be seen as a tone-reproduction operator. This model is presented in detail in Chapter 8.

2.9 DISPLAY GAMMA

Cathode ray tubes have a nonlinear relationship between input voltage V and light output L_v. This relationship is well approximated with the following power

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 52

law function.

$$L_v = kV^{\gamma}$$

The exponent γ models the nonlinearity introduced by the specific operation of the CRT, and is different for different monitors. If V is normalized between 0 and 1, the constant k simply becomes the maximum output luminance $L_{\max}$.

In practice, typical monitors have a gamma value between 2.4 and 2.8. However, further nonlinearities may be introduced by the lookup tables used to convert values into voltages. For instance, Macintosh computers have a default gamma of about 1.8, which is achieved by the interaction of a system lookup table with the attached display device. Whereas the Macintosh display system may have a gamma of 1.8, the monitor attached to a Macintosh will still have a gamma closer to 2.5 [100].

Thus, starting with a linear set of values that are sent to a CRT display, the result is a nonlinear set of luminance values. For the luminances produced by the monitor to be linear, the gamma of the display system needs to be taken into account. To undo the effect of gamma, the image data needs to be gamma corrected before sending it to the display, as explained in material following.

Before the gamma value of the display can be measured, the black level needs to be set appropriately [100]. To set the black point on a monitor, you first display a predominantly black image and adjust the brightness control on the monitor to its minimum. You then increase its value until the black image just starts to deviate from black. The contrast control may then be used to maximize the amount of contrast.

The gamma value of a display device may then be estimated, as in the image shown in Figure 2.27. Based on an original idea by Paul Haeberli, this figure consists of alternating black and white lines on one side and solid gray patches on the other. By viewing this chart from a reasonable distance and matching the solid gray that comes closest to the gray formed by fusing the alternating black and white lines, the gamma value for the display device may be read from the chart. Note that this gamma estimation chart should only be used for displays that follow a power-law transfer function, such as CRT monitors. This gamma estimation technique may not work for LCD displays, which do not follow a simple power law.

Once the gamma value for the display is known, images may be pre-corrected before sending them to the display device. This is achieved by applying the follow-

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 53

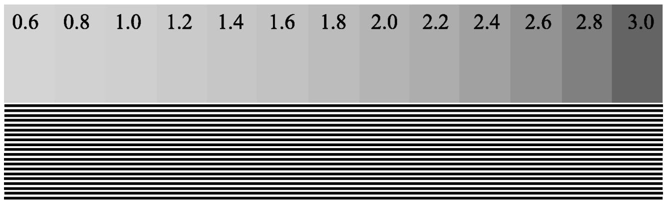


FIGURE 2.27 *Gamma estimation for CRT displays. The alternating black and white lines should be matched to the solid grays to determine the gamma of a display device.*

ing correction to the values in the image, which should contain normalized values between 0 and 1.

$$R' = R^{1/\gamma}$$
$$G' = G^{1/\gamma}$$
$$B' = B^{1/\gamma}$$

An image corrected with different gamma values is shown in Figure 2.28.

The technology employed in LCD display devices is fundamentally different from CRT displays, and the transfer function for such devices is often very different. However, many LCD display devices incorporate circuitry to mimic the transfer function of a CRT display device. This provides some backward compatibility. Thus, although gamma encoding is specifically aimed at correcting for the nonlinear transfer function of CRT devices, often (but not always) gamma correction may be applied to images prior to display on LCD.

Many display programs perform incomplete gamma correction (i.e., the image is corrected such that the displayed material is intentionally left nonlinear). Often, a gamma value of 2.2 is used. The effect of incomplete gamma correction is that contrast is boosted, which viewers tend to prefer [29]. In addition, display devices reflect some of their environment, which reduces contrast. Partial gamma correction may help regain some of this loss of contrast [145].

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 54

FIGURE 2.28 *An image corrected with different gamma values. In reading order:* $\gamma = 1.0$, 1.5, 2.0, *and* 2.5. *(Image courtesy of the Albin Polasek museum, Winter Park, Florida.)*

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 55

One of the main advantages of using gamma encoding is that it reduces visible noise and quantization artifacts by mimicking the human contrast sensitivity curve. However, gamma correction and gamma encoding are separate issues, as explained next.

2.10 BRIGHTNESS ENCODING

Digital color encoding requires quantization, and errors are inevitable during this process. In the case of a quantized color space, it is preferable for reasons of perceptual uniformity to establish a nonlinear relationship between color values and the intensity or luminance. The goal is to keep errors below the visible threshold as much as possible.

The eye has a nonlinear response to brightness. That is, at most adaptation levels, brightness is perceived roughly as the cube root of intensity (see, for instance, the encoding of $L*$ of the CIELAB and CIELUV color spaces in Section 2.8). Applying a linear quantization of color values would yield more visible steps in darker regions than in the brighter regions, as shown in Figure 2.29.[5] A power-law encoding with a γ value of 2.2 produces a much more even distribution of quantization steps, although the behavior near black is still not ideal. For this reason and others, some encodings (such as sRGB) add a short linear range of values near zero (see Section 2.11).

However, such encodings may not be efficient when luminance values range over several thousand or even a million to one. Simply adding bits to a gamma encoding does not result in a good distribution of steps, because it can no longer be assumed that the viewer is adapted to a particular luminance level, and the relative quantization error continues to increase as the luminance gets smaller. A gamma encoding does not hold enough information at the low end to allow exposure readjustment without introducing visible quantization artifacts.

To encompass a large range of values when the adaptation luminance is unknown, an encoding with a constant or nearly constant relative error is required. A log encoding quantizes values using the following formula rather than the power

5 We have chosen a quantization to 6 bits to emphasize the visible steps.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 56

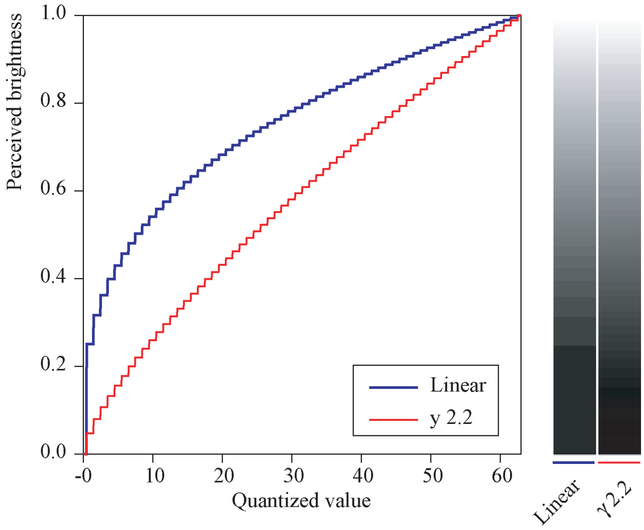


FIGURE 2.29 *Perception of quantization steps using a linear and a gamma encoding. Only 6 bits are used in this example encoding to make the banding more apparent, but the same effect takes place in smaller steps using 8 bits per primary.*

law cited earlier.

$$I_{\text{out}} = I_{\text{min}} \left[\frac{I_{\text{max}}}{I_{\text{min}}} \right]^{v}$$

This formula assumes that the encoded value v is normalized between 0 and 1, and is quantized in uniform steps over this range. Adjacent values in this encoding thus differ by a constant factor equal to

$$\left[\frac{I_{\text{max}}}{I_{\text{min}}} \right]^{1/N},$$

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 57

where N is the number of steps in the quantization. This is in contrast to a gamma encoding, whose relative step size varies over its range, tending toward infinity at zero. The advantage of constant steps is offset by a minimum representable value, $I_{\min}$, in addition to the maximum intensity we had before.

Another alternative closely related to the log encoding is a separate exponent and mantissa representation, better known as floating point. Floating-point representations do not have perfectly equal step sizes but follow a slight sawtooth pattern in their error envelope, as shown in Figure 2.30. To illustrate the quantization dif-

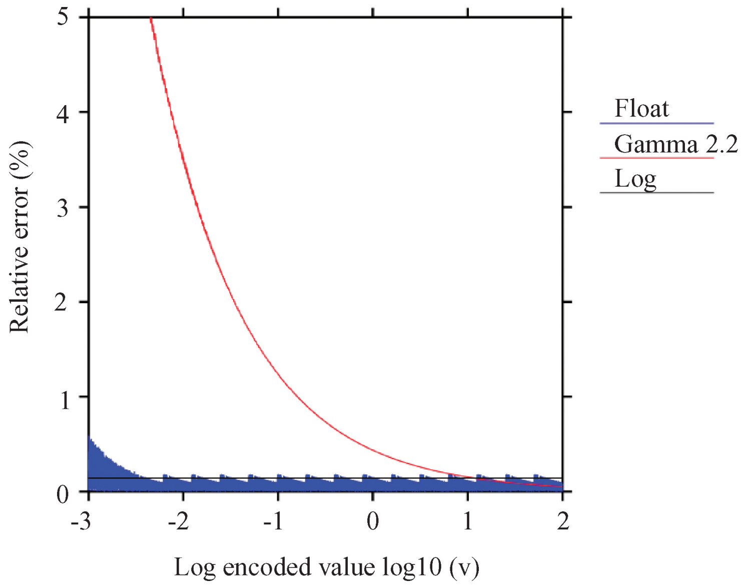


FIGURE 2.30 *Relative error percentage plotted against* $\log_{10}$ *of image value for three encoding methods.*

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 58

ferences between gamma, log, and floating-point encodings, a bit size (12) and range (0.001 to 100) are chosen that can be reasonably covered by all three types. A floating-point representation with 4 bits in the exponent, 8 bits in the mantissa, and no sign bit is chosen, because only positive values are required to represent light.

By denormalizing the mantissa at the bottom end of the range, values between $I_{\min}$ and zero may also be represented in a linear fashion, as shown in this figure.[6] By comparison, the error envelope of the log encoding is constant over the full range, whereas the gamma encoding error increases dramatically after just two orders of magnitude. Using a larger constant for γ helps this situation somewhat, but ultimately gamma encodings are not well suited to full HDR imagery where the input and/or output ranges are unknown.

2.11 STANDARD RGB COLOR SPACES

Most capture and display devices have their own native color space, generically referred to as device-dependent RGB. Although it is entirely possible to convert an image between two device-dependent color spaces, it is more convenient to define a single standard color space that can serve as an intermediary between device-dependent color spaces.

On the positive side, such standards are now available. On the negative side, there is not one single standard but several competing standards. Most image encodings fall into a class called *output-referred standards*, meaning that they employ a color space corresponding to a particular output device rather than to the original scene they are meant to represent. The advantage of such a standard is that it does not require any manipulation prior to display on a targeted device, and it does not "waste" resources on colors that are out of this device gamut. Conversely, the disadvantage of such a standard is that it cannot represent colors that may be displayable on other output devices or that may be useful in image processing operations along the way.

A *scene-referred* standard follows a different philosophy, which is to represent the original captured scene values as closely as possible. Display on a particular output

6 Floating-point denormalization refers to the linear representation of values whose exponent is at the minimum. The mantissa is allowed to have a zero leading bit, which is otherwise assumed to be 1 for normalized values, and this leads to a steady increase in relative error at the very bottom end, rather than an abrupt cutoff.

device then requires some method of mapping the pixels to the device's gamut. This operation is referred to as tone mapping, which may be as simple as clamping RGB values to a 0-to-1 range or something more sophisticated, such as compressing the dynamic range or simulating human visual abilities and disabilities (see Chapters 6 through 8). The chief advantage gained by moving tone mapping to the image decoding and display stage is that correct output can be produced for any display device, now and in the future. In addition, there is the freedom to apply complex image operations without suffering losses due to a presumed range of values.

The challenge of encoding a scene-referred standard is finding an efficient representation that covers the full range of color values. This is precisely where HDR image encodings come into play, as discussed in Chapter 3.

For reference, we discuss several current output referenced standards. In Section 2.4, we already introduced the ITU-R RGB color space. In the remainder of this section conversions to several other color spaces are introduced. Such conversions all follow a matrix multiplication followed by a nonlinear encoding. The sRGB color space is introduced as an example, before generalizing the concept to other color spaces.

The nonlinear sRGB color space is based on a virtual display. It is a standard specified by the International Electrotechnical Commission (IEC 61966-2-1). The primaries as well as the white point are specified in terms of xy chromaticities according to Table 2.7 (this table also shows information for other color spaces, discussed in material following). The maximum luminance for white is specified as 80 cd/m^2.

Because the specification of sRGB is with respect to a virtual monitor, it includes a nonlinearity similar to gamma correction. This makes sRGB suitable for Internet applications as well as scanner-to-printer applications. Many digital cameras now produce images in sRGB space. Because this color space already includes a nonlinear transfer function, images produced by such cameras may be displayed directly on typical monitors. There is generally no further need for gamma correction, except perhaps in critical viewing applications.

The conversion of CIE XYZ tristimulus values to sRGB consists of a 3-by-3 matrix multiplication followed by a nonlinear transfer function. The linear part of the transform is identical to the matrix specified in ITU-R BT.709, introduced in Section 2.4. The resulting RGB values are converted into sRGB using the following

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 60

Color Space		R	G	B	White Point (Illuminant)	
Adobe RGB (1998)	x	0.6400	0.2100	0.1500	D65	0.3127
	y	0.3300	0.7100	0.0600		0.3290
sRGB	x	0.6400	0.3000	0.1500	D65	0.3127
	y	0.3300	0.6000	0.0600		0.3290
HDTV (HD-CIF)	x	0.6400	0.3000	0.1500	D65	0.3127
	y	0.3300	0.6000	0.0600		0.3290
NTSC (1953)	x	0.6700	0.2100	0.1400	C	0.3101
	y	0.3300	0.7100	0.0800		0.3161
SMPTE-C	x	0.6300	0.3100	0.1550	D65	0.3127
	y	0.3400	0.5950	0.0700		0.3290
PAL/SECAM	x	0.6400	0.2900	0.1500	D65	0.3127
	y	0.3300	0.6000	0.0600		0.3290
Wide gamut	x	0.7347	0.1152	0.1566	D50	0.3457
	y	0.2653	0.8264	0.0177		0.3584

TABLE 2.7 *Chromaticity coordinates for primaries and white points defining several RGB color spaces.*

transfer function (for R, G, and $B > 0.0031308$).

$$R_{\mathrm{sRGB}} = 1.055R^{1/2.4} - 0.055$$
$$G_{\mathrm{sRGB}} = 1.055G^{1/2.4} - 0.055$$
$$B_{\mathrm{sRGB}} = 1.055B^{1/2.4} - 0.055$$

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 61

For values smaller than 0.0031308, a linear function is specified, as follows.

$$R_{\mathrm{sRGB}} = 12.92R$$
$$G_{\mathrm{sRGB}} = 12.92G$$
$$B_{\mathrm{sRGB}} = 12.92B$$

This conversion follows a general pattern that is found in other standards. First, a 3-by-3 matrix is defined, which transforms from XYZ to a color space with different primaries. Then a nonlinear transform is applied to the tristimulus values. This transform takes the following general form [93].

$$R' = \begin{cases} (1+f)R^{\gamma} - f & \text{for } t \leqslant R \leqslant 1 \\ sR & \text{for } \quad 0 < R < t \end{cases}$$

$$G' = \begin{cases} (1+f)G^{\gamma} - f & \text{for } t \leqslant G \leqslant 1 \\ sG & \text{for } 0 < G < t \end{cases}$$

$$B' = \begin{cases} (1+f)B^{\gamma} - f & \text{for } t \leqslant B \leqslant 1 \\ sB & \text{for } 0 < B < t \end{cases}$$

Note that the conversion is linear in a small dark region, and follows a gamma curve for the remainder of the range. The value of s determines the slope of the linear segment, and f is a small offset. Table 2.8 lists several RGB standards, which are defined by their conversion matrices as well as their nonlinear transform specified by the γ, f, s, and t parameters [93]. The primaries and white points for each color space are outlined in Table 2.7. The gamuts spanned by each color space are shown in Figure 2.31. The gamut for the HDTV color space is identical to the sRGB standard and is therefore not shown again.

The Adobe RGB color space was formerly known as SMPTE-240M, but was renamed after SMPTE's gamut was reduced. It has a larger gamut than sRGB, as shown in the chromaticity diagrams of Figure 2.31. This color space was developed with the printing industry in mind. Many digital cameras provide an option to output images in Adobe RGB color, as well as sRGB.

The HDTV and sRGB standards specify identical primaries, but differ in their definition of viewing conditions. As such, the difference lies in the nonlinear transform.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 62

Color Space	XYZ to RGB Matrix	RGB to XYZ Matrix	Nonlinear Transform
Adobe RGB (1998)	$\begin{bmatrix} 2.0414 & -0.5649 & -0.3447 \\ -0.9693 & 1.8760 & 0.0416 \\ 0.0134 & -0.1184 & 1.0154 \end{bmatrix}$	$\begin{bmatrix} 0.5767 & 0.1856 & 0.1882 \\ 0.2974 & 0.6273 & 0.0753 \\ 0.0270 & 0.0707 & 0.9911 \end{bmatrix}$	$\gamma = \text{N/A}$ $f = \text{N/A}$ $s = \text{N/A}$ $t = \text{N/A}$
sRGB	$\begin{bmatrix} 3.2405 & -1.5371 & -0.4985 \\ -0.9693 & 1.8760 & 0.0416 \\ 0.0556 & -0.2040 & 1.0572 \end{bmatrix}$	$\begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix}$	$\gamma = 0.42$ $f = 0.055$ $s = 12.92$ $t = 0.003$
HDTV (HD-CIF)	$\begin{bmatrix} 3.2405 & -1.5371 & -0.4985 \\ -0.9693 & 1.8760 & 0.0416 \\ 0.0556 & -0.2040 & 1.0572 \end{bmatrix}$	$\begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix}$	$\gamma = 0.45$ $f = 0.099$ $s = 4.5$ $t = 0.018$
NTSC (1953)	$\begin{bmatrix} 1.9100 & -0.5325 & -0.2882 \\ -0.9847 & 1.9992 & -0.0283 \\ 0.0583 & -0.1184 & 0.8976 \end{bmatrix}$	$\begin{bmatrix} 0.6069 & 0.1735 & 0.2003 \\ 0.2989 & 0.5866 & 0.1145 \\ 0.0000 & 0.0661 & 1.1162 \end{bmatrix}$	$\gamma = 0.45$ $f = 0.099$ $s = 4.5$ $t = 0.018$

TABLE 2.8 *Transformations for standard RGB color spaces (after [93]).*

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 63

Color Space	XYZ to RGB Matrix	RGB to XYZ Matrix	Nonlinear Transform
SMPTE-C	$\begin{bmatrix} 3.5054 & -1.7395 & -0.5440 \\ -1.0691 & 1.9778 & 0.0352 \\ 0.0563 & -0.1970 & 1.0502 \end{bmatrix}$	$\begin{bmatrix} 0.3936 & 0.3652 & 0.1916 \\ 0.2124 & 0.7010 & 0.0865 \\ 0.0187 & 0.1119 & 0.9582 \end{bmatrix}$	$\gamma = 0.45$ $f = 0.099$ $s = 4.5$ $t = 0.018$
PAL/SECAM	$\begin{bmatrix} 3.0629 & -1.3932 & -0.4758 \\ -0.9693 & 1.8760 & 0.0416 \\ 0.0679 & -0.2289 & 1.0694 \end{bmatrix}$	$\begin{bmatrix} 0.4306 & 0.3415 & 0.1783 \\ 0.2220 & 0.7066 & 0.0713 \\ 0.0202 & 0.1296 & 0.9391 \end{bmatrix}$	$\gamma = 0.45$ $f = 0.099$ $s = 4.5$ $t = 0.018$
Wide gamut	$\begin{bmatrix} 1.4625 & -0.1845 & -0.2734 \\ -0.5228 & 1.4479 & 0.0681 \\ 0.0346 & -0.0958 & 1.2875 \end{bmatrix}$	$\begin{bmatrix} 0.7164 & 0.1010 & 0.1468 \\ 0.2587 & 0.7247 & 0.0166 \\ 0.0000 & 0.0512 & 0.7740 \end{bmatrix}$	$\gamma =$ N/A $f =$ N/A $s =$ N/A $t =$ N/A

TABLE 2.8 *(continued)*

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 64

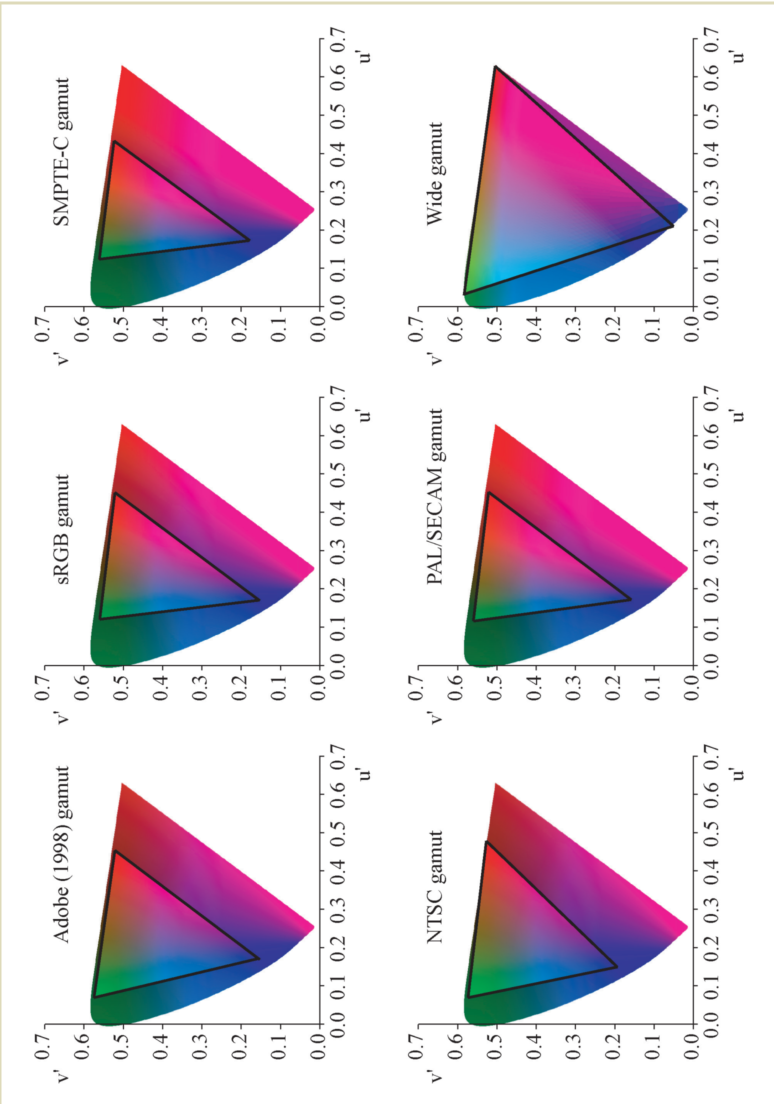


FIGURE 2.31 CIE (u', v') chromaticity diagrams showing the color gamuts for various color spaces.

reinhard v.2005/03/22 Prn:13/06/2005; 15:50 F:reinhard02.tex; VTEX/JOL p. 65

The National Television System Committee (NTSC) standard was used as the color space for TV in North America. It has now been replaced with SMPTE-C to match phosphors in current display devices, which are more efficient and brighter. Phase Alternating Line (PAL) and Systeme Electronique Couleur Avec Memoire (SECAM) are the standards used for television in Europe.

Finally, the Wide gamut color space is shown for comparison [93]. Its primaries are monochromatic light sources with wavelengths of 450, 525, and 700 nm. This color space is much closer to the spectrally sharpened chromatic adaptation transforms discussed in Section 2.5.