PROCESSING DIGITAL COLOR IMAGES: FROM CAPTURE TO DISPLAY

Because the devices in an electronic imaging system represent color in different ways, getting them to communicate in a manner that preserves color fidelity and is transparent to the user is a challenging task.

Jan P. Allebach

Color is a vital part of our everyday experience. It provides essential cues about our environment, adds aesthetic value to the world around us and even has a strong effect on our mood. Yet the role of color in this age of electronic information processing is surprisingly incomplete. We use color in separate and isolated systems and have no way to connect them. We photograph our families and look at the prints in albums; we view time-varying color images on our television sets; we sit in front of computer monitors that display multicolor windows and icons and perhaps some color pictures. How do we take the family photographs and display them on our televisions, or grab snapshots from television broadcasts and put them into our personal computers, or combine those images with those in the photo albums and print them as color images in a newsletter? These scenarios and others that involve capture, manipulation and display or printing of color images in heretofore unimagined ways are becoming realized in the workplace and home thanks to digital color. (Figure 1 shows a monitor display of one of the window-based products for processing digital color images, in this case Adobe Systems' Adobe-Photoshop.)

Several key features of digital color make handling color images possible:

▷ Digital color provides a common language with which devices that "see" colors differently may communicate.

▷ Digital color allows maximum flexibility in manipulating color information for calibration, color correction, enhancement, feature extraction and conversion between different "languages" for representing color that are native to different color devices, such as scanners and printers.

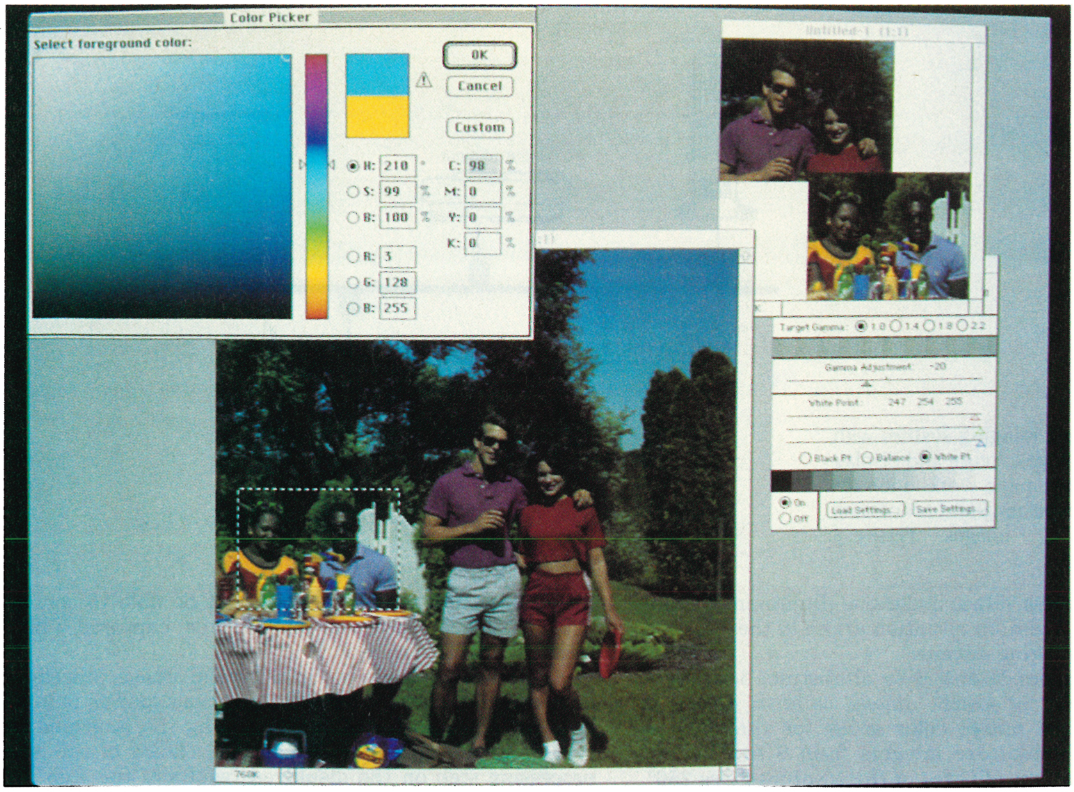
▷ Digital color allows us to exploit the pervasive and rapidly growing technologies for digital processing, transmission and storage of information.

To understand these issues more clearly, it is helpful to consider a representative electronic imaging system, such as the one depicted in figure 2. At the input to the system a digital camera acquires an image of the scene before it. Or one might digitally capture an image by scanning it from an analog source, such as a photograph or print. In either case, the resulting digital image consists of a two-dimensional array of sample values taken from the scene or print. Each sample value specifies the color information at one spatial point in the image. To completely characterize the light incident on the detectors in the camera or scanner it is necessary to specify the radiance at all wavelengths. However, we are interested only in the *minimum* specification that allows us to recreate the sensation of a particular color in the original image when that color is reproduced at the output of our system, a display or printer device.

Trichromatic representation of input color

It is indeed fortunate for designers of imaging systems that one can *largely* specify a color using just three numbers, called the tristimulus values.[1] That three values can, to a reasonable level of approximation, represent a color is a consequence of the fact that human perception of color is mediated by the responses of three different types of photoreceptors in the retina, called cones. Colorimetry, the mathematical specification of color, provides the basis for specifying color in terms of tristimulus values. The word "largely" is important here, because colorimetry

Jan Allebach is a professor of electrical engineering at Purdue University, in West Lafayette, Indiana.



Window-based environment on the monitor of a personal computer or workstation used for processing and displaying digital color images. **Figure 1**

based on a simple tristimulus model cannot specify color appearance completely. Nonetheless, the tristimulus model is a powerful one that lends itself well to analysis by the standard tools of linear algebra.[2,3] Color perception is discussed in more depth in the article by Alan Robertson on page 24.

Much of the complexity of digital color derives from the fact that the specific meaning of the set of three numbers, or "3-tuple," used to represent color depends on the context in which it is used. At the output of a digital camera or scanner, this vector refers to the responses of three different sensors, or photoreceptors, each of which may be modeled as the integral over wavelength of the product of a spectral sensitivity function and the spectral power distribution of the light incident on the photoreceptor. Although the three response curves are fairly broadband, they typically peak in different parts of the visible spectrum—one each in the long-, medium- and short-wavelength regions. Therefore it is convenient to refer to the three responses as red (R), green (G) and blue (B), even though we may not perceive the pure spectral color at the peak wavelength of any of the three response curves as red, green or blue.

The output 3-tuple corresponding to a particular input stimulus with spectral power distribution $S(\lambda)$ would thus be given by

$$R_S = \int S(\lambda)\, S_R(\lambda)\, d\lambda \tag{1a}$$

$$G_S = \int S(\lambda)\, S_G(\lambda)\, d\lambda \tag{1b}$$

$$B_S = \int S(\lambda)\, S_B(\lambda)\, d\lambda \tag{1c}$$

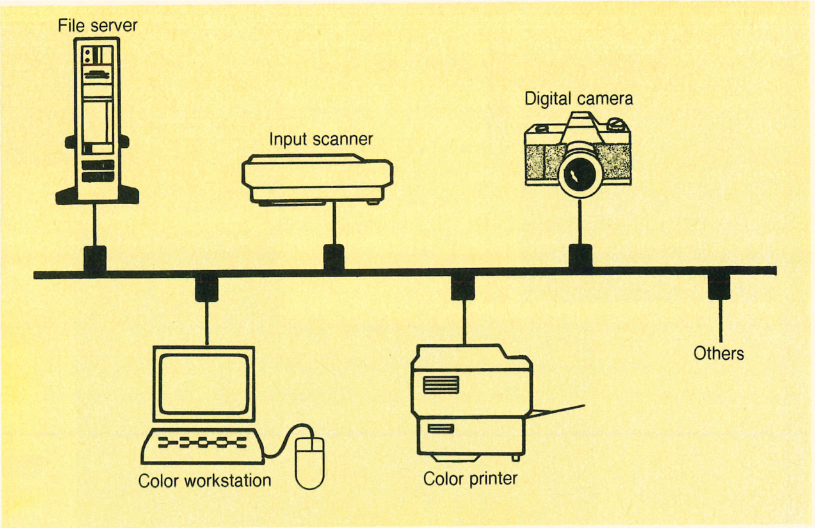
where $S_i(\lambda)$ denotes the spectral response function for each of the three sensors. Because different image capture devices generally have different spectral responses, a particular 3-tuple may represent different colors in different devices. With an appropriate set of spectral response functions, denoted by $V_R(\lambda)$, $V_G(\lambda)$ and $V_B(\lambda)$, equation 1 can also model the human visual system. The resulting 3-tuple is referred to as a tristimulus vector.

Mathematically, equation 1 defines a projection from the infinite-dimensional space of all possible spectral power distributions onto the three-dimensional subspace spanned by linear combinations of the three spectral response functions. Different power distributions might yield the same 3-tuple, so a device would interpret them as the same color; such power distributions are called metamers. Metamerism is a device-specific attribute: For devices with different spectral responses, different sets of colors are metameric. Of course, the human eye is the image capture device with which we ultimately are concerned. Ideally we would want the same sets of metameric colors for both the initial image capture device and the human viewer. This can be achieved only if the spectral response functions for the device are a linear combination of those for the human eye. Commercially available image capture devices generally do not meet this condition, limiting the accuracy of digital color reproduction.

The illuminant is another important factor one must take into account when considering metamerism. For a scanner the stimulus $S(\lambda)$ is the product of the spectral power distribution $I(\lambda)$ of the light that illuminates the copy material and the spectral transmittance or reflectance $C(\lambda)$ of the copy itself:

$$S(\lambda) = I(\lambda)\, C(\lambda) \tag{2}$$

Thus two different copy colors $C_1(\lambda)$ and $C_2(\lambda)$ may yield



Electronic imaging system with multiple image capture and display devices and a workstation for manipulating and processing images. **Figure 2**

the same tristimulus values under one illuminant but not under another. A similar situation arises in the case of a digital camera imaging a scene.

The set of colors described by all distinct 3-tuples of output values for a specific image capture device is referred to as the native color space for that device. Typically these values are integers from 0 to 255. All possible combinations of three of these values gives 256^3, or 2^{24} (approximately 16.7 million), distinct colors. To be able to process meaningfully the color imagery captured by different devices using a workstation or personal computer, it is necessary to be able to convert from the native space for each device to a common reference color space.

Representing color for reproduction

Display devices. On the output side of our electronic imaging system we are faced with a different interpretation for the 3-tuples used to describe color. Let us start by considering the display.[4] The display of a color image generally is made up of an additive mixture of three different color stimuli called primaries, each of which has a different spectral power distribution. Although the distributions may be broadband, each lies primarily in the long-, the medium- or the short-wavelength region of the visible spectrum. Like the photoreceptor response curves, these primaries are denoted as red, green and blue, respectively.

The primaries typically are arranged in a spatially nonoverlapping configuration, such as the red, green and blue phosphor-dot triad found in a cathode-ray-tube monitor. Thus the perceived additive mixture depends on a spatial averaging by the human visual system. A simple model for the resulting spectral power distribution $D(\lambda)$ is

$$D(\lambda) = R_{\mathrm{D}} D_{\mathrm{R}}(\lambda) + G_{\mathrm{D}} D_{\mathrm{G}}(\lambda) + B_{\mathrm{D}} D_{\mathrm{B}}(\lambda) \tag{3}$$

where $D_i(\lambda)$ denotes the spectral power distribution for each of the three primaries, and R_{D}, G_{D} and B_{D} are the tristimulus values, or intensities, of the primary stimuli associated with the color to be displayed. Here the 3-tuple represents the amount of the three primaries rather than the response of three sensors, as it does on the image capture side of the imaging system. The set of colors generated by all possible 3-tuple values for the primaries defines the color space of the display device. Because this space is generally different from the native color space of the image capture device, one must be able to convert between color spaces to reproduce a captured image faithfully at the display.

Hard-copy devices. In a typical image processing application, the user manipulates various sources of input imagery any number of ways, viewing the intermediate results on a display and basing the choice of the next processing step on the displayed results of the step just completed. When he or she has finished this sequence of tasks, the final step is often to generate a hard copy of the final result. Whether it is to be viewed in transmission or reflection mode, a hard copy achieves the impression of color through a subtractive rather than an additive process.[5] The resulting copy is made up of three or more colorants, each characterized by its spectral transmittance or reflectance function.

With transmission copy, such as a developed 35-mm slide, three colorants is the norm. In analogy to the additive color process, each colorant strongly absorbs in either the long-, medium- or short-wavelength region of the visible spectrum, resulting in the perceived color cyan, magenta or yellow. Roughly speaking, if one considers white to be the sum of red, green and blue in equal proportions, then one may regard cyan as being white minus red, or equivalently blue plus green; that is, $\mathrm{C} = \mathrm{W} - \mathrm{R} = \mathrm{B} + \mathrm{G}$. Similarly, $\mathrm{M} = \mathrm{W} - \mathrm{G} = \mathrm{R} + \mathrm{B}$ and $\mathrm{Y} = \mathrm{W} - \mathrm{B} = \mathrm{R} + \mathrm{G}$.

With reflection copy, such as the pages in this magazine, three colorants are also widely used; but often the output device uses a fourth colorant that is strongly absorbing at all wavelengths. This colorant is, of course, black (K). Using it produces a cleaner rendition of blacks than does a subtractive combination of C, M and Y and also may reduce the cost of the print, because black is often the least expensive colorant.

Modeling subtractive color is much more complex than modeling additive color. But for the idealized model of block colorants with unity transmittance (or reflectance) in all but nonoverlapping bands (that is, when only one colorant absorbs at any wavelength), it is possible to express the overall spectral absorptance $A(\lambda)$ for a copy containing the colorants cyan, magenta and yellow in the form

$$A(\lambda) = CA_{\mathrm{C}}(\lambda) + MA_{\mathrm{M}}(\lambda) + YA_{\mathrm{Y}}(\lambda) \tag{4}$$

where $A_i(\lambda)$ is the spectral absorptance function associated with each colorant, and C, M and Y are coefficients related to the amount of each colorant. When viewed under an illuminant with spectral power distribution $I(\lambda)$, the spectrum of the light transmitted through or reflected from the copy is $I(\lambda)[1-A(\lambda)]$. This model neglects the coloration contributed by the medium or substrate and any scattering that may take place as the light passes through the copy. Coloration and scattering play a particularly important role in the appearance of printed images.

Now we must get back to our user, who has been patiently waiting for the transparency or print while we discussed the basic elements of subtractive color. When that print appears, he or she is going to expect it to look like the image displayed on the workstation monitor. The user may even hold the print next to the workstation display to compare the images. Achieving identical appearance between soft and hard copy is a difficult and unsolved problem. Traditional colorimetry provides only a partial solution. Even assuring a colorimetric match is difficult, because the systems designer may not have control over the illumination under which the print is to be viewed. The problem is complicated by the effect of adaptation on color perception. (See the article by Robertson.)

Device-independent color

From the preceding discussion we see that each device, be it for image capture or display, has its own native color space, defined by the relationship between input or output colors and the corresponding 3-tuples used to represent them. To process color in an electronic imaging system, one must be able to convert readily between any native device space and a standard color space. Image capture devices now on the market provide 3-tuple outputs in their native spaces. Similarly, a display or printing device expects a 3-tuple input that is in its native space. Until recently, the user of an electronic imaging system had to perform his or her own characterization of the color spaces of the devices in the system and write his or her own software to perform the conversion between spaces. Because of the special instrumentation and tedious measurements required to do such characterization, it was common to forgo the entire process and simply accept unmatched colors throughout the system.

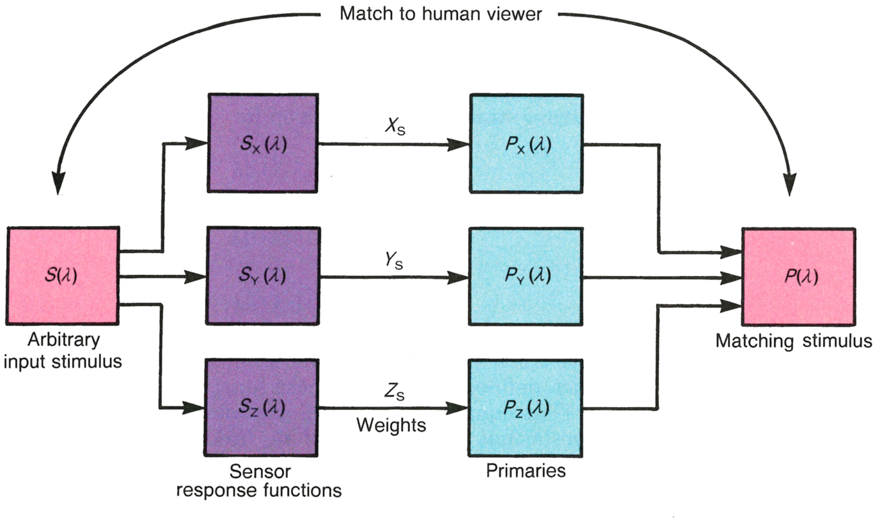
Color management systems and calibration tools that allow the user to work with color in a more device-independent manner are now becoming available. The color management system resides in the host computer and develops a characterization for each input or output device connected to the host by performing a series of automatic or semiautomatic measurements. Thus the user can work with color in a standard color space and have a fully color-calibrated system without having to understand the specifics of the native color space of each device in the system. A user who has images represented in a standard color space can freely transfer data between different systems and accurately capture, process, and display or print the images anywhere.

As significant an improvement as this is, it is only the first step toward device-independent color. At the next step, we expect each device to become its own color manager. It will continuously monitor its color characteristics and maintain its own equations and tables for conversion between its native space and a standard space. All communication among capture and reproduction devices and computers will use a standard color space, so that it will no longer be necessary for a computer or device to know the specifics of another device's native color space. At that stage, color will be truly device independent.

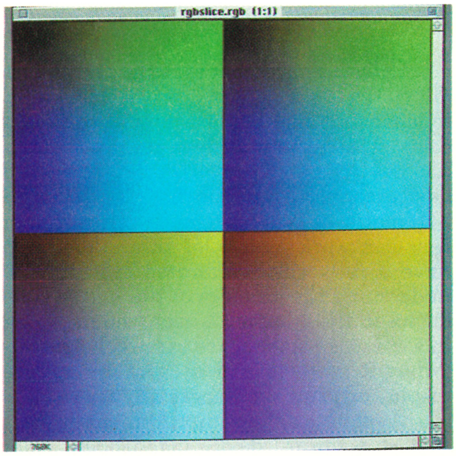
So far I have provided an overview of electronic imaging systems for processing color images, emphasizing the different forms in which color is represented throughout the system and the need for a way to tie the forms of color together. The other articles in this special issue describe in more detail each of the pieces of an imaging system, including the human viewer. In the rest of this article, I will focus on how the pieces go together, again emphasizing the digital representation of color throughout the system.

Linear transformations between color spaces

In seeking a common framework within which to represent color throughout an electronic imaging system, a logical place to start is with the human viewer, who, after all, will ultimately judge whether colors match at different points in the system. The trichromatic representation of color is based on the fact that one can match an arbitrary color corresponding to a spectral power distribution $S(\lambda)$



Relationship between sensors and primaries. One can recreate, or match, an arbitrary color stimulus $S(\lambda)$ "seen" by a sensor by multiplying three "primary" colors $P_i(\lambda)$ (i = X, Y or Z) by the appropriate weighting factors (X_S, Y_S or Z_S) generated by the sensor and additively combining the weighted primaries. To a human viewer the original stimulus $S(\lambda)$ and the synthesized stimulus $P(\lambda)$ are indistinguishable. **Figure 3**



a

b

Comparison of RGB and YC_rC_b spaces. **a:** Slices of red–green–blue color space at four different red values. Amount of red in slices increases from top left to top right to bottom left to bottom right. Within each slice, the origin of the green and blue axes is in the upper left corner, with green increasing to the right and blue increasing from top to bottom. **b:** Slices of YC_rC_b color space at four different Y values, with Y increasing the same way as red in **a**. Within each slice, the origin of the C_r and C_b axes is in the center, with C_r increasing to the right and C_b increasing from top to bottom. **Figure 4**

using an additive combination of three appropriately chosen primaries X, Y and Z with spectra $P_X(\lambda)$, $P_Y(\lambda)$ and $P_Z(\lambda)$ that yields an overall spectral power distribution

$$P(\lambda) = X_S P_X(\lambda) + Y_S P_Y(\lambda) + Z_S P_Z(\lambda) \tag{5}$$

Here X_S, Y_S and Z_S are the tristimulus values or weights required to match the original spectral distribution $S(\lambda)$.

If we take equation 1 as a model for the human visual system, substituting the spectral response functions $V_i(\lambda)$ for the sensor response functions $S_i(\lambda)$, then a match between stimuli $S(\lambda)$ and $P(\lambda)$ implies that the tristimulus values corresponding to these two colors must be identical; that is, $R_S = R_P$, $G_S = G_P$ and $B_S = B_P$, where

$$R_S = \int S(\lambda) V_R(\lambda)\,\mathrm{d}\lambda \tag{6}$$

and

$$R_P = \int P(\lambda) V_R(\lambda)\,\mathrm{d}\lambda \tag{7a}$$

$$= X_S \int P_X(\lambda) V_R(\lambda)\,\mathrm{d}\lambda$$

$$+ Y_S \int P_Y(\lambda) V_R(\lambda)\,\mathrm{d}\lambda + Z_S \int P_Z(\lambda) V_R(\lambda)\,\mathrm{d}\lambda \tag{7b}$$

G_S, G_P, B_S and B_P are defined by equations analogous to equations 6 and 7. These relationships may be compactly expressed with vector notation. Let the tristimulus vector corresponding to the response of the viewer to the original stimulus be $\mathbf{c}_S = (R_S, G_S, B_S)^T$ and the tristimulus vector corresponding to the matching stimulus (the weighting factors) be $\mathbf{w}_S = (X_S, Y_S, Z_S)^T$, and define a matrix $\boldsymbol{R}$ whose ijth element is the response of sensor i to primary j, so that

$$r_{ij} = \int P_j(\lambda) V_i(\lambda)\,\mathrm{d}\lambda \tag{8}$$

where j = X, Y or Z, and i = R, G or B. We may then write

$$\mathbf{c}_S = \mathbf{R}\,\mathbf{w}_S \tag{9}$$

This equation shows that the tristimulus vector $\mathbf{c}_S$ corresponding to the viewer's response to the stimulus $S(\lambda)$ is related by a simple linear transformation to the vector $\mathbf{w}_S$ of weights for the primaries $P_j(\lambda)$ required to additively match $S(\lambda)$.

The primaries $P_j(\lambda)$ are arbitrary as long as their projections onto the three-dimensional subspace of spectral power distributions defined by the human visual system span that subspace. They need not even be physically realizable, as is the case with the CIE Standard Observers, which define the standard XYZ color spaces. (See Robertson's article.)

Inverting equation 9 and using equation 6, we can find the tristimulus vector $\mathbf{w}$ in the XYZ color space corresponding to any stimulus $S(\lambda)$. To simplify this computation, we define color matching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$, which are simply the XYZ coordinates corresponding to a monochromatic stimulus with wavelength λ. Then, under the assumption of linearity, the XYZ coordinates for an arbitrary stimulus $S(\lambda)$ are given by

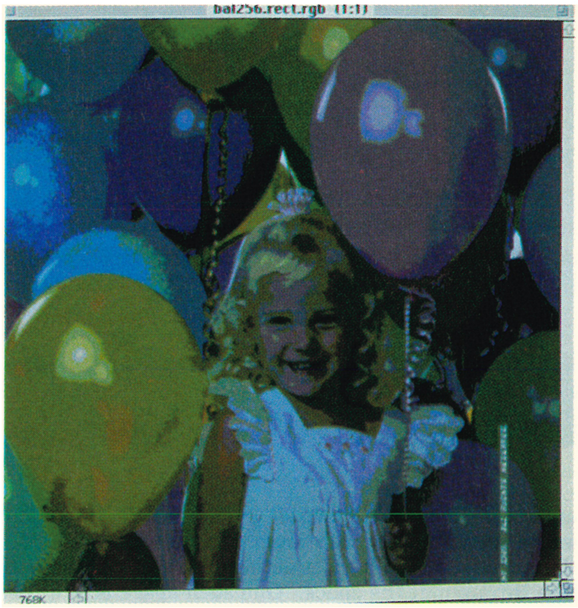
$$X_S = k \int S(\lambda)\,\bar{x}(\lambda)\,\mathrm{d}\lambda \tag{10a}$$

$$Y_S = k \int S(\lambda)\,\bar{y}(\lambda)\,\mathrm{d}\lambda \tag{10b}$$

$$Z_S = k \int S(\lambda)\,\bar{z}(\lambda)\,\mathrm{d}\lambda \tag{10c}$$

where k is a constant. The CIE standard XYZ spaces are, in fact, defined in terms of these functions.

That equations 1 and 10 are identical in form also suggests that the color matching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ may be interpreted as the spectral response functions for a new observer (referred to as the "Standard Observer"



a

b

Digital color image displayed with 256 colors that were selected from a set of 16 million colors to produce either an image-independent palette (**a**) or an image-dependent palette (**b**). The palette for **a** was obtained by uniformly sampling the RGB color space. For **b**, by contrast, the palette is made up of colors that are common in the image. (Original image courtesy of Eastman Kodak.) **Figure 5**

in the CIE standard XYZ spaces). Thus equation 9 provides the conversion between the color space corresponding to the human viewer and the space for a sensor whose spectral response functions are a linear transformation of those for the human viewer. Mathematically, one is merely transforming between different bases for the three-dimensional subspace of the infinite-dimensional space of spectral power distributions. (The three-dimensional subspace is defined by the human visual system.) As mentioned above, there is no such one-to-one transformation between the subspaces corresponding to two sensors whose spectral responses are not related by a linear transformation.

Finally, let me return to our original specification of the XYZ space in terms of a set of primaries with spectra $P_X(\lambda)$, $P_Y(\lambda)$ and $P_Z(\lambda)$ that one can additively combine (equation 5) to produce a matching stimulus $P(\lambda)$. We see that equation 5 has precisely the same form as equation 3. Thus one may also use a linear transformation such as that given by equation 9 to convert among the color spaces of additive output devices and the standard XYZ space.

To summarize, there is a unique relationship between sensors or color matching functions and primaries, as illustrated in figure 3. Given any three primaries defined by their spectral power distributions and subject to the above-mentioned conditions, there is a corresponding set of sensor response functions that yield tristimulus values. These values, when used to weight an additive combination of the primaries, match the input stimulus seen by the sensor. Conversely, given any three sensor response functions that are a linear transformation of those corresponding to the human viewer, there is a set of three primaries. These primaries, when weighted by the sensor outputs and combined additively, yield a stimulus that matches that seen by the sensor. The primaries are not unique, because each one may be replaced by another color that is its metamer. In the context of a specific image capture or reproduction device, either the set of sensor response functions or the set of primaries is real; the other is fictitious. In addition, conversion between any two color spaces corresponding to different bases for the three-dimensional subspace of spectral power distributions defined by the human viewer may be accomplished by a simple 3×3 matrix multiplication. Unfortunately, the validity of this statement rests on the validity of the simple models presented thus far. As I already mentioned, these models are not completely satisfactory; consequently, we need more complex transformations, such as nonlinear transformations, to convert between native device spaces and standard spaces.

Nonlinear transformations

To develop more accurate transformations between color spaces, one can take one of two approaches—or combinations of the two.[6] The first approach is model based and depends on successfully analyzing the sources of nonideality in device behavior. In some cases, this approach can be quite simple. For example one can accurately model the nonlinear relationship between electron-gun drive voltage and the resulting current density in CRT monitors using a simple power-law relationship. Because the analog drive voltage is proportional to the digital value input to the digital-to-analog converter for each gun, and the displayed intensity is proportional to the current density, the overall relationship between the digital input V and displayed intensity D for each gun is

$$D = k_i V^{\gamma_i} \tag{11}$$

where k_i and γ_i are constants and $i =$ R, G or B. This equation and a 3×3 matrix transformation form the basis for a color-space conversion often used for CRTs.[7]

We don't need to look far for a much more complex example. In the model for a color hard-copy device, I implicitly assumed that the colorant density, or concentration, could be varied continuously between zero and some maximum value. With many printing technologies, however, this is not possible. Instead, the printer lays a fixed-density array of colorant dots onto the medium. To vary

Digitally halftoned images displayed using eight different colors. **a:** Image halftoned with a dispersed-dot algorithm and displayed on a cathode-ray-tube monitor with 100-dot/inch resolution. **b:** Image halftoned with a clustered-dot algorithm and printed by a 300-dot/inch cyan-magenta-yellow-black thermal transfer device. (Image in **b** halftoned and printed by Ron Gentile, Adobe Systems Inc) **Figure 6**

a

b

the amount of a particular colorant, the printer varies the dot size. This method of achieving the rendition of tones from very light to very dark is called halftoning. Because the dot array is space-filling only for the darkest tones, transmission directly through the medium or reflection directly from the substrate also contributes to the overall color. Scattering can significantly affect overall printed color. For example, with reflection copy, light that enters the bare substrate may be scattered in such a manner that it exits under a colorant dot and is thus partially absorbed. This phenomenon, known as the Yule–Nielsen effect, results in a copy that is darker than would be predicted simply on the basis of the fraction of paper covered by colorant dots. In addition, if the printer lays down halftone patterns for the three colorants in three separate passes, the dot arrays may not be registered relative to one another, and dots from different colorants will overlap in a spatially varying manner that repeats with a period much larger than the period of the array itself. The resulting ring-like patterns, a type of moiré commonly referred to as rosettes, have long been familiar to workers in the graphic arts and printing industry. To minimize their visibility, the printer rotates the differently colored dot arrays by carefully selected angles. One can use the Neugebauer equations to predict the color that results from these overlapping halftone dot arrays.[5]

The difficulty of accurately characterizing these and similar effects motivates the second approach to performing nonlinear transformations between color spaces: One simply generates a table that provides the correct output 3-tuple for every possible input 3-tuple.

To specify the transformations between color spaces, be they model or table based, it is necessary to estimate the parameters of the model or the entries for the table. With input devices, this may be done by capturing an image of a test target containing patches with colors whose XYZ color-space coordinates have been measured. With output devices a test target is displayed or printed to produce a test copy, and the user measures the colors with a colorimeter or spectrophotometer. Once the measurement has been made, one can use a least-squares or similar technique to estimate the model parameters. The number of measurements required to characterize color for a particular device increases with the number of degrees of freedom in the model. A table-based transformation requires the largest number of measurements. Calibrating a printer may require hundreds of measurements.

One also uses nonlinear transformations to obtain color spaces for which the coordinates are correlated to the perceptual attributes of color—hue, saturation and lightness. (See Robertson's article.) The coordinates should also be such that the minimum displacement within the color space that produces a noticeable color change is the same throughout the space. One such space is YIQ, used in broadcast television. Here Y, called the luminance component, is correlated with lightness, and the remaining two coordinates, called the chrominance component, are correlated with hue and saturation. Because humans are less able to perceive fine spatial variations in chrominance then in luminance, the chrominance signal is allocated a smaller bandwidth in video broadcast transmission than is the luminance signal.

Another space of this type is YC_rC_b. Here Y is again the luminance component, and C_rC_b is the chrominance component. (See figure 4.) When color images are compressed in this form the chrominance information typically occupies only one fifth of the total space required to store the image. In addition, the CIE has defined two uniform color spaces, which can be used to help calibrate imaging systems and in perceptually based image processing.

Sampling and quantization effects

Although digital systems have moved color more fully into the information age, they have some limitations and disadvantages. These may be traced to the simple fact that a digital color image is represented by an array of 3-tuples, each of which is a finite-length binary number. Each 3-tuple represents the color within an a finite area of the original scene or copy centered at one point on a lattice of sampling points. Any variation or detail within this area is lost or distorted during sampling. The viewer sees this degradation as a jagged or blurred rendition of sharp edges or a loss of fine detail from the original.[8]

In addition, because each 3-tuple is represented by a finite number of binary digits or bits, only a finite number of distinct colors may be presented. The system must assign those colors from the input image that lie between representable colors (that is, those inside the gamut of the color space) to the closest representable color. If the number of distinct colors that may represented is too small, artifacts may appear. (See figure 5a.) For example, one may observe color contouring in areas that appear smooth in the original image. "Blocking" may occur where colors that should be distinct are rendered as being the same. In some cases, there may be noticeable color shifts: If an input color lies beyond the range of extreme colors that can be represented—that is, outside the gamut—it is replaced by a color that is within the color space. This may cause desaturation of highly saturated colors[9] in addition to the artifacts discussed above.

One may largely alleviate these problems by decreasing the spatial sampling interval or increasing the number of bits used to represent each sample. However, because this remedy requires input and output devices to have higher spatial resolution, greater dynamic range or both, it drives up the cost. In addition this approach increases the size of the image files, so the user must either accept slower system performance or pay for increased processor and channel bandwidth and larger storage capacity both in computer memory and in peripheral devices such as disk drives. To some extent these costs can be offset by storing and transmitting the images in compressed form and decompressing them as needed. However, compression adds to the computational burden of manipulating the images.

Until recently the typical size for digital color images (primarily dictated by the resolution of the displays and digital hard-copy devices commonly available during the past decade) was 512×512 pixels with 24 bits (3 bytes) per pixel, or 8 bits per color coordinate. With a total size of approximately 0.75 megabytes, these images taxed the ability of systems to quickly perform any but the simplest processing steps. However, with the advances in technology that have occurred in recent years, images of this complexity are now hardly a challenge.

Nonetheless, as we enter an era of full-color document processing, we find that system requirements have grown enormously. It requires over 25 Mbytes to store or transmit in uncompressed form a single 8.5×11-inch page scanned at 300 pixels per inch, 3 bytes per pixel! Compression algorithms now widely available can reduce the image size by a factor of 10 or so with little or no loss in image quality. Document processing systems represent pages consisting of a mixture of text, graphics and images with a page description language that provides a much more efficient characterization for text and graphics than the pixel-by-pixel representation or bit map. The page is converted to bit map form directly at the display or printer. If one inputs the document by scanning a hard-copy original, it can then be converted to a page description language through character recognition techniques.

Because of the finite word length used to represent color values, conversion between color spaces can cause additional degradation of image quality, especially if nonlinearities are involved. For a table-based transformation, a 24-bit output value must be provided for each of 2^{24}, or approximately 16 million, possible input colors. If the table is not stored at full resolution, generating it and implementing the transformation requires interpolation in a three-dimensional space.

At the system output, we are often faced with additional quantization. The display device may not be capable of representing all 256 levels (8 bits) of each stimulus, due to memory limitations or other constraints associated with the technology. If the only limitation is memory, it may be possible to provide a look-up table that allows selection of a small palette of colors from a much larger set.[10] In this case, the user may customize the palette to the particular image being displayed, yielding much better quality than would be possible with an image-independent uniform quantizer. (See figure 5b.) If the display colors are fixed and cannot be chosen from a larger set, then multilevel halftoning techniques are required for acceptable image quality.[11] In this case, the perception that a particular color is present at some point in the image depends on the human visual system's spatially averaging[11] over the colors actually displayed in the neighborhood of that point.[12] The proper choice of a dispersed-dot digital halftoning technique minimizes the loss of spatial detail and appearance of texture artifacts that result from halftoning. (See figure 6.) As discussed earlier, printing devices commonly employ halftoning. Here, however, the inability to reproduce fine texture patterns accurately may necessitate the use of a clustered-dot digital halftoning technique that is analogous to the traditional halftoning methods used in graphic arts and printing. Finally, because display devices quite often have a more limited gamut than is available in the rest of the system, it may be necessary to perform further gamut compensation prior to printing or display of the image.[9]

References

1. G. Wyszecki, W. S. Stiles, *Color Science: Concept and Methods, Quantitative Data and Formulae*, Wiley, New York (1982).
2. J. B. Cohen, W. E. Kappauf, Am. J. Psychol. **98**, 171 (1985).
3. H. J. Trussell, Color: Res. Appl., February 1991, p. 33; Summer 1985, p. 171.
4. R. W. G. Hunt, *The Reproduction of Color in Photography, Printing and Television*, Fountain, Tolworth, UK (1987).
5. J. A. C. Yule, *Principles of Color Reproduction*, Wiley, New York (1968).
6. J. Shiau, L. C. Williams, J. Imaging Sci. Technol. **36**, 211 (1992).
7. W. T. Hartmann, T. E. Madden, J. Imaging Technol. **13**, 103 (1987).
8. A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice–Hall, Englewood Cliffs, N. J. (1989).
9. R. S. Gentile, E. Walowit, J. P. Allebach, J. Imaging Technol. **16**, 176 (1990).
10. M. T. Orchard, C. A. Bouman, IEEE Trans. Signal Proc. **39**, 2677 (1991).
11. R. S. Gentile, E. Walowit, J.P. Allebach, J. Opt. Soc. Am. **7**, 1019 (1990).
12. R. Ulichney, *Digital Halftoning*, MIT P., Cambridge, Mass. (1987). ■