The CIE XYZ and xyY Color Spaces

Douglas A. Kerr

Issue 1
March 21, 2010

ABSTRACT AND INTRODUCTION

A Color Space is a completely-specified scheme for describing the color of light, ordinarily using three numerical values (called *coordinates*). An important color space, defined by the International Commission on Illumination (CIE, the initials of its French name) is the CIE XYZ color space. It is widely used in scientific work, and color descriptions in other color spaces are often related to their representation in this space. A derivative of this color space, the CIE xyY color space, is often used as a way to graphically present the chromaticity of colors.

The XYZ color space itself has a fascinating genesis. Its nature, history, and role in both theoretical and practical color science are described in this article, along with that of its cousin, the CIE xyY color space.

The article begins with a review of several important technical concepts that are involved in the story.

BACKGROUND

Our work on this topic requires an accurate grasp of a number of fundamental matters in the area of colorimetry. Here I will review some of those topics and other critical matters.

Color

Color is a property that distinguishes among different kinds of light. It is defined wholly in terms of human perception. If two instances of light appear to a viewer to be the same color[1], they **are** the same color.

Color, as we use the term in the technical sense, is usually recognized by the viewer as having two aspects:

[1] I have to add, for rigor, "if observed under the same conditions". The subtlety that this honors, the matter of *chromatic adaptation*, will not play a role in this article.

- L*uminance*, which we can think of for our purposes as an indication of the "brightness" of the light.[2]

- C*hromaticity*, the property that distinguishes red from blue, and red from pink. (This is the property that lay people typically think is meant by "color", but the technical concept of color embraces luminance as well.)

What determines the color of light?

Color is not a primary physical property like the temperature or pressure of a gas. We can, however, ascertain the color of an "instance" of light by physical measurements which will predict for us the eye's response to it.

The physical property of the light that gives it its color is its "spectrum",[3] the "plot" of distribution of the power in the light over the range of wavelengths that can affect the eye (the "visible wavelengths").

The "shape" of the plot determines the *chromaticity* of the light; its overall "vertical scale" determines its *luminance*. That is, if we have two different instances of light, whose spectrums have the same shape, but for one instance is proportionately "stretched" vertically, the two instances have the same chromaticity, but the second one has a greater luminance.

In the other direction, things are not nearly so tidy. We can have several instances of light with different spectrums which nevertheless have the same color. In fact there are an infinity of spectrums that will have any given color.

This situation is called *metamerism*, and different spectrums having the same color are called *metamers* of the color.

Color spaces

It has been long recognized that, as perceived by most human observers, any color of light can be specified by stating three numerical values. That is, color is *three-dimensional* in the mathematical (not geometric) sense.

---

[2] There is a subtle but important formal distinction between *luminance* and *brightness*, but for our purpose here we can ignore it.

[3] The formal name of this is the *power spectral distribution* (PSD) of the light.

There are, however, many different schemes of coordinates under which these three numerical values can be defined. These schemes, when fully specified as to their details, are called *color spaces*.

There are several basic schemes, and for each, there are usually several different standard, fully-specified forms.

One important scheme is the *additive* scheme. Additive color spaces work by stating the amounts of three kinds of light (called the *primaries* of the color space), having different, and precisely specified, chromaticities, which if actually added together would create light of the color being described. This can be thought of both as a recipe for actually creating light of that color (as in a display screen) and as an abstract way to describe that color.

The three amount values are said to be the three *coordinates* (under the specific color space) of the color of interest.

Tristimulus color spaces

Additive-scheme color spaces can be called *tristimulus* color spaces, since they work by defining the amounts of three stimuli provided to the eye (the three primaries). However, it is a widespread custom to reserve the name “tristimulus color space” for a particular color space of this class, the CIE XYZ color space, the overall subject of this article.

How the eye determines color

It has been determined that (for fairly substantial luminance) the eye observes each tiny element of the image on the retina with three kinds of “cones”, which are “photodetectors” [4]. Each kind has a different *spectral response*, by which we mean a curve that tells how much “output” the cone delivers from light of a fixed “potency” at each wavelength over the visible range.

When an area on the retina is bathed in light with a certain spectrum, in effect, for each of the three kinds of cones:

- The spectrum of the light is multiplied by the spectral response of the cone, meaning that, for each wavelength, the “potency” of the light at that wavelength is multiplied by the value of the spectral response at that wavelength.

[4] There are a few humans (“tetrachromats”), all women, who have four kinds of cones. Accordingly, their perception of color requires four values to describe.

- All these products are added together[5], giving the output of the cone.

The three types of cone are called “L”, “M”, and “S”, referring to the fact that the peaks of their spectral responses are at different wavelengths, which we arbitrarily consider to be “long”, “medium” and short. The spectral response curves of the three types of cone are called $\underline{l}$, $\underline{m}$, and $\underline{s}$ [6] (the usual typography is an overbar, but that is a pain to produce in this word processor, so I will underline them instead).

We cannot actually determine these curves (formally described as response *functions*, since they are functions of wavelength). We can derive related curves by tedious visual experimentation (a process that will be discussed at length shortly), and from these indirectly come to conclusions as to the curves themselves. But different researchers have come to different conclusions.

Figure 1 shows a popular conclusion as to these three response curves (scaled so that their peaks are all at 100 units). They are labeled here with the cone type names, L, M, and S, not the function names.

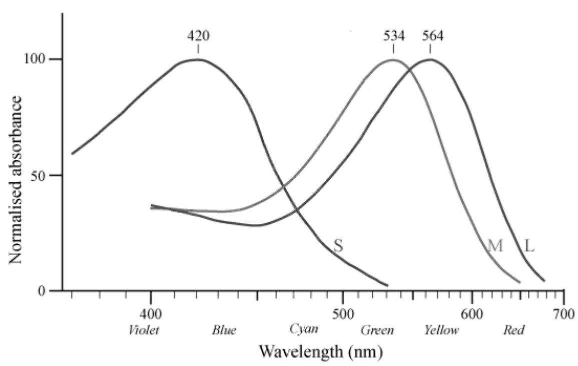

Figure 1. Eye cone response curves

THE COLOR MATCHING EXPERIMENTS

We cannot directly determine the three response functions, $\underline{l}$, $\underline{m}$, and $\underline{s}$, of the human eye, mainly because we can’t put a “meter” on the outputs of the cones.

[5] Since both the spectrum of the light and the spectral response function of the cone are continuous, the process is actually *integration*, not *summation*, but the concept is identical.

[6] The mathematician would write then as $\underline{l}(\lambda)$, $\underline{m}(\lambda)$, and $\underline{s}(\lambda)$, reminding us that all of them are functions of wavelength, λ.

But color scientists were able to precisely characterize the response of the human eye in a way that is almost as useful to us as knowledge of these curves. This was done by a tedious series of experiments with actual human observers, done in around 1931. The basic tests used a “tristimulus” concept with three actual primary light sources. A small screen was arranged so that one half was illuminated by light of a “test color”, while the other half was illuminated by light composed of adjustable amounts of the three primaries. The three primaries used had spectrums comprising only a single wavelength.[7] For the basic tests, the “test color” was also light whose spectrum comprised only a single wavelength, which was changed from one observation to another.

With a particular test light in place, the user was asked to adjust the amounts of the three primaries until there was an exact visual match at the boundary between the two halves of the screen. The amounts of the three primaries for which that occurred were recorded, and then another test light was put into place.

In some case, it turned out that a negative amount of a certain primary was needed for a match., How could that be done?

In such a case, that primary was added to the “test light” on the reference half of the screen, which was equivalent to a negative contribution of that primary to the “matching light”.

Multiple runs were made with multiple observers, and a vast mass of data accumulated.

Analysis of this data produced a consolidated model of the matching of light of known color by amounts of the three primary lights.

This was presented in the form of three curves, one for each primary, which showed the amount of that primary needed in the “mix” to make light whose color was the same as that of “monochromatic” test light of a certain wavelength, as a function of that wavelength. These were called the “color matching functions”, designated $\underline{r}$, $\underline{g}$, and $\underline{b}$. They are seen in figure 2.

---

[7] These are described as *spectra*l chromaticities, since when we decompose light by wavelength, as with a prism or diffraction grating, the entire display is spoken of as “the spectrum” (one of several uses of that term), and the light at any point along it only comprises one wavelength. Such chromaticities are also often called *monochromatic*, meaning “one color”. That term is really a misnomer; any instance of light has only “one color” (as we today understand the term). Still, I will use that term, since it is so commonly understood in technical writing.

They are not the eye cone response functions, $\underline{l}$, $\underline{m}$, and $\underline{s}$. But there is a mathematical relationship between any such set of matching functions and the eye cone response curves, which we can exploit in many beneficial ways.

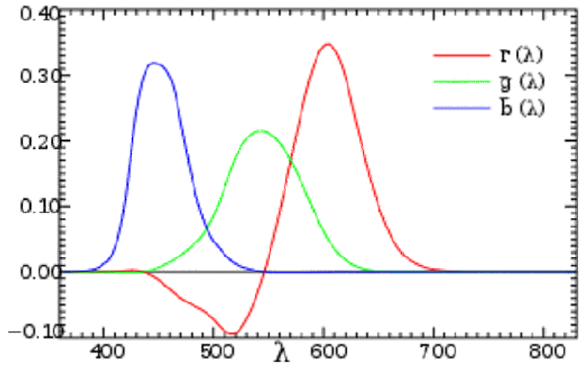

Figure 2. CIE RGB color matching functions

The three primaries chosen for test program had spectrums which (as we had previously noted) comprised only a single wavelength. Those wavelengths were 700 nm, 546.1 nm, and 435.8 nm. The latter two were chosen since they could be generated accurately by emissions from a mercury-arc lamp. The first one could not be (at the time) generated directly in any such precise way, but was chosen because the response of the eye in that spectral region made it not so critical that this primary be generated precisely.

The hues of these three primaries could be described as “red”, “green”, and “blue”, respectively. Of course those terms do not correspond to any specific hue—they are just quantitative generalizations.

Because this test protocol worked on the basis of the addition of light from three primaries, it actually defined a *bona fide* additive color space. That is, we can describe any color (not just the “spectral” colors used in the basic test runs) in terms of the amounts of the three primaries needed to produce light of the same color. Because of the “popular” names of its three primaries, this was said to be the CIE RGB color space.

MOVING ON

A disadvantage of this characterization of human color response is that the “red” matching function has negative values in part of the wavelength range. In fact, it can be demonstrated that, for a “color space” with “real” primaries (that is, ones we can actually generate, and see), at least one of the curves **must** have negative portions.

Analytical work done with the curves involves multiplications of light spectrum values by the curve values at different wavelengths, something that is today done easily by a computer, but which, done with desktop calculators and paper, was very tedious. There was concern that the need to keep track of both positive and negative values could increase the risk of a mis-step.

Thus the workers decided there was need to develop another color space whose underlying matching functions were "nowhere negative".

This did not call for another round of zillions of human observations. We can take a description of a color in terms of values of the three coordinates of one color space and convert it into values of the three coordinates of another color space—coordinates that revolve around a different set of primaries (assuming that both color spaces are of the *additive* genre and thus both **have** primaries).

And then we can, by another mathematical manipulation, determine what the corresponding set of matching functions would be for the "new" color space.

Now, there are an infinity of color spaces all of whose matching curves would be nowhere negative. So the wonks decided, while they were at it, to pick one of those that had some other useful property.

The process of transforming coordinate values in one color space (one set of primaries) to corresponding values in another color space (another set of primaries) turns out to involve three linear equations. We will show an example here. In it, the primaries of the "first" color space are called R, G, and B (fits what we are doing here), and the corresponding coordinate values are called *R*, *G*, and *B*[8][9]. I will call the primaries of the "second" color space X, Y, and Z (as in a typical algebra lecture), with the corresponding coordinate values called *X*, *Y*, and *Z*.

Here are the transforming equations:

$$X = \mathrm{m}_{11}R + \mathrm{m}_{12}G + \mathrm{m}_{13}B \qquad (1)$$

$$Y = \mathrm{m}_{21}R + \mathrm{m}_{22}G + \mathrm{m}_{23}B \qquad (2)$$

---

[8] This is consistent with the mathematical typographic practice of using italic letters to represent variables, which these are.

[9] Note that these are "linear" indications of the amount of the primary, not the nonlinear forms, also called R, G, and B (but no italics), in the actual "output" of such color spaces as sRGB.

$$Z = \mathrm{m}_{31}R + \mathrm{m}_{32}G + \mathrm{m}_{33}B \tag{3}$$

where the nine coefficients (constants), m, define the transformation. Their values depend on the chromaticities of the primaries of the "RGB" and "XYZ" color spaces.

This can be written in matrix form, thus:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \mathrm{m}_{11} & \mathrm{m}_{12} & \mathrm{m}_{13} \\ \mathrm{m}_{21} & \mathrm{m}_{22} & \mathrm{m}_{23} \\ \mathrm{m}_{31} & \mathrm{m}_{32} & \mathrm{m}_{33} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{4}$$

or, more symbolically, as:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [\mathrm{M}] \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{5}$$

where [M] represents the entire transform matrix.

Now before I proceed, let me reveal that X, Y, and Z (and *X*, *Y*, and *Z*) are not just illustrative names I used to illustrate the math as we do in algebra class. They will be the actual names used for the primaries (and coordinates) of this "new" color space, which will thus be called the "CIE XYZ" color space. (Thus life imitates lecture.)

Back to the story, realizing that we are free to choose from among many new color spaces in our quest for one with "everywhere non-negative" matching functions, we realize that for any one we choose, there will be a certain set of the nine values, m_{11}-m_{13}, that make up the transformation matrix to this space from the CIE RGB space.

And we have a lot of flexibility in doing that.

Now, it turns out that the luminance (*L*) of a color we have represented by its coordinates in an additive color space (whose coordinates happen to be called *R*, *G*, and *B*, just to help the story move along) is given by:

$$L = \mathrm{k}_R R + \mathrm{k}_G G + \mathrm{k}_G B \tag{6}$$

If we choose a "new" color space such that:

$$\mathrm{m}_{21} = \mathrm{k}_R\text{, } \mathrm{m}_{22} = \mathrm{k}_G\text{, and } \mathrm{m}_{23} = \mathrm{k}_\mathrm{B} \tag{7}$$

then *L* will always be equal to *Y*.

That is, the value of the coordinate of the Y primary in our new color space will always be the luminance of the color!

And in fact, that is why, in the field of colorimetry, the symbol for luminance is not the normal scientific symbol, L, but rather is Y.[10]

It may seem strange, then, that the primaries Y and Z do not contribute to the luminance of the color. How can that be? Can they be luminously-impotent? Yes. What kind of light could that be? Well, an imaginary kind, for one thing.

Remember that this entire XYZ story is a mathematical fiction, and these primaries, like superheroes in comic books, can have any properties the wonks wanted to give them, so long as the math works out properly in connecting the story to the real world.

In any case, by a process that involves this clever strategy regarding Y, and various other considerations, the wonks agreed on a set of primaries for this new color space (yes, its is our hero, the "CIE XYZ" color space). That directly (albeit by way of tedious mathematical manipulation) implied a corresponding set of matching functions.

Figure 3 shows these matching functions (labeled just X, Y, and Z):

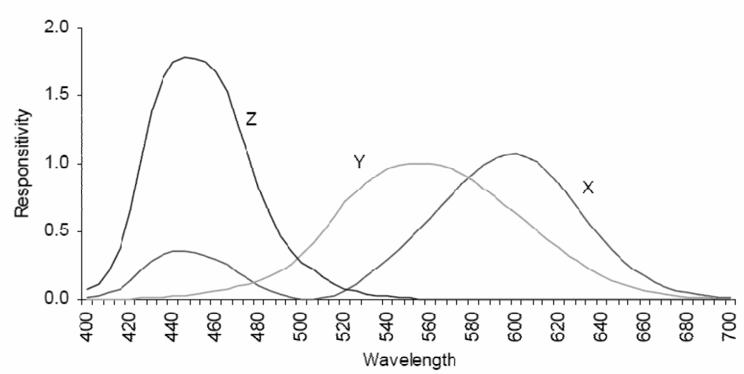

Figure 3. CIE XYZ color matching functions

I can't "show" you right now what the primaries are, because:

- I would show it on the CIE x-y chromaticity diagram, which won't be invented for another page or so.
- They are "imaginary"; they can't actually be generated, and if they could be, we couldn't see them. (More about this later.) This is an

[10] Students of this often say, "How sad that the wonks chose the same symbol, *Y*, for two different things: luminance, and the coordinate *Y*. This is very confusing." But, by definition, they are the same thing.

unavoidable corollary of the fact that the three response functions are nowhere negative (which is how we got here anyway).

We'll "see them" in a little while.

A, or maybe, "the", tristimulus color space

As noted earlier, any color space that revolves around the premise of representing a color in terms of amounts of three primaries—three *stimuli*—fed to the eye as one way to create the sensation of a certain color can be properly called a *tristimulus color space*, and this is true of the CIE XYZ color space.

But, by historical custom (traceable back to the time of the work leading to this color space), the term "tristimulus color space", without further elaboration, implies the CIE XYZ color space.

This is ironic, because in this color space, the three "stimuli" (its primaries) cannot really be emitted, and so in fact cannot be used as actual stimuli to the eye!

A COUSIN COLOR SPACE

The CIE XYZ color space is an *additive* color space: as we have seen, it revolves around the concept of describing a color by stating the amounts of three primaries that would be combined to make light of that color. (Can we make visible light, with a real color, by mixing together "imaginary" primaries? In the laboratory, no. On paper, mathematically, yes. We'll see how a little later.)

Another important genre of color space is the luminance-chrominance color space genre. These relate more directly to the human outlook on color than the additive spaces. In a luminance-chromaticity color space, one coordinate tells us the luminance of the color, while the other two, together, tell us the chromaticity (and there can be several ways this can be organized).

To allow the benefits of this in our scientific work, the wonks devised a luminance-chromaticity color space based on the CIE XYZ color space. Its chromaticity coordinates are called *x* and *y*. They are defined this way:

$$x = \frac{X}{X+Y+Z} \tag{8}$$

$$y = \frac{Y}{X+Y+Z} \tag{9}$$

Recall that Y, the coordinate of the Y primary, is in fact the luminance of the color, so we use Y (as is) as the luminance coordinate of this new color space.

It is called the **CIE xyY** color space.

The CIE x-y chromaticity diagram

Since the coordinates *x* and *y* tell us the chromaticity of a color, we can use a chart with *x* and *y* as its axes to plot points that indicate chromaticity, usually called the CIE x-y chromaticity diagram. We see it in figure 4.

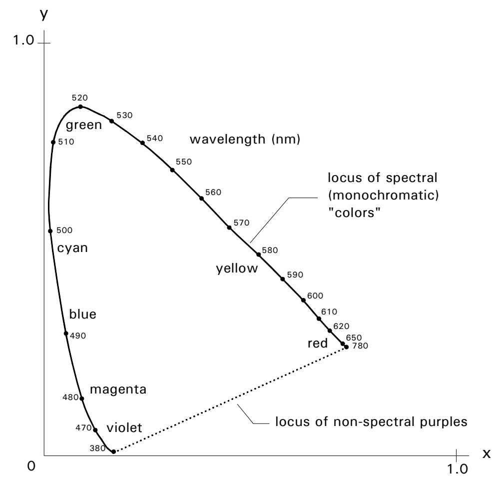

Figure 4. The CIE x-y chromaticity diagram

Any chromaticity corresponds to a point on this diagram. The “horseshoe” curve consists of the chromaticity points of every color of light whose spectrum consists of only a single wavelength. These are the “spectral” chromaticities of which we spoke earlier.

The dotted line at the bottom completes the region “enclosed” by the horseshoe. The chromaticities along it are not “spectral” (there is no light with only a single wavelength component that exhibits a color with such a chromaticity). They are called the *nonspectral purples*.

All colors of visible light have chromaticities represented by points inside the region bounded by the horseshoe (and the locus of nonspectral purples).

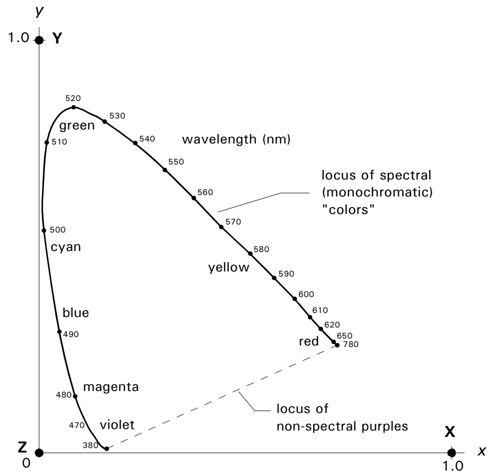

Figure 5. The CIE primaries X, Y, and Z

The imaginary primaries, X, Y, and Z

Figure 5 shows, on the x-y chromaticity diagram, the chromaticity of the three imaginary primaries of the CIE XYZ color space, X, Y, and Z.

Let me mention that when we speak of the “chromaticity” of these imaginary primaries, we are taking “poetic license”. Chromaticity is a property of color, and color only exists for light that can be seen by the eye. So these primaries don’t really have a chromaticity. But the poetic license here is fully excusable. Avoiding it would require some very cumbersome alternative terminology.

Even though the X, Y, and Z primaries are imaginary, they can participate in the mathematical model by which two or three primaries, added together, can define a color. With regard to the chromaticity of that color, we can see how this works by geometric construction. We often do this for color spaces with real primaries, or even for real colors that are not primaries, but it works just fine with these imaginary ones.

We’ll begin a demonstration on figure 6.

Before we get going, I’ll use this figure to illustrate an important point. The *chromaticity gamut* of the xyY color space (and thus of its parent, the XYZ color space)— the range of chromaticities it can represent—is contained within the triangle X-Y-Z. You will note that this embraces the entire chromaticity gamut of visible light (but just barely—the wonks made it come out that way).

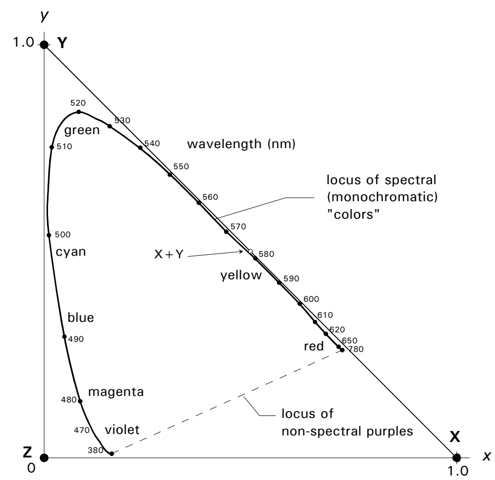

Figure 6. Addition of X and Y primaries

We start with a geometric construction to find the chromaticity of the color that **would be produced** by adding equal amounts of primary X and primary Y **if we could actually do that.** That chromaticity will lie on the line joining the points representing those primaries, and (since we postulated equal quantities of both) at a point equidistant from both—at the center of the line. The point representing the chromaticity of the compound light is labeled "X + Y" on the figure. Note that it is outside of the visual gamut (although just a bit). So we have made another imaginary "color". Oh, great!

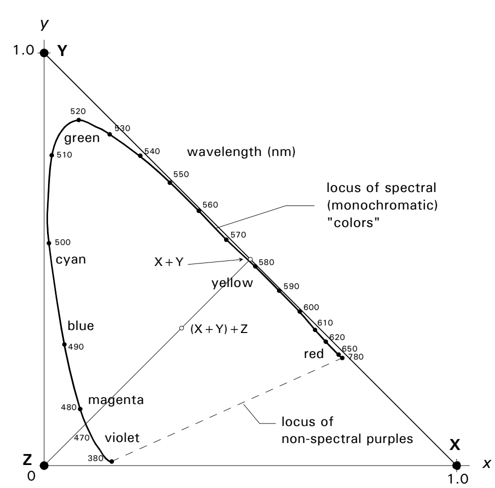

Figure 7. Adding in primary Z.

In figure 7, we take the next step, adding in that same amount of primary Z to our intermediate product, color "Z + Y". The final result will lie on the line joining **those** two chromaticities (as always when we add light of two different colors).

Since there is twice as much "stuff" in color "X + Y" (one ration each of primaries X and Y) as in one ration of Z, the result will be only half the distance from "X+Y as it is from Z (2/3 of the distance along the line from Z to "X+Y".

We see that chromaticity point labeled "(X+Y)+Z" (to recognize the procedure we used to find it), but of course it is also "X+Y+Z". And it is not imaginary—it lies well inside the visible light region. Reality at last!

Thus, even though primaries X, Y, and Z have no physical reality, they participate just fine in the mathematics by which a set of *X*, *Y*, and *Z* values define a color. They have performed this useful service for us for (as of this writing) almost 80 years. The more spiritual among us may be gratified by this technical example of "faith in things unseen" (and unseeable).

The white point

The resulting point in our exercise, labeled X+Y+Z, is in fact the *white point* of this color space. The white point is defined as the chromaticity of the color indicated when the three coordinates of an additive color space (*X*, *Y*, and *Z* in this case) have equal values. On the x-y chromaticity diagram, the chromaticity coordinates of the XYZ white point are $x=1/3$, $y=1/3$.

Is that actually "white"? Well, there is no chromaticity that uniquely qualifies to be called "white". There are a number of *standard illuminants*, with specified chromaticities (in fact, specified spectrums) that can be thought of as "specific kinds of white". The white point of the XYZ color space is exactly the chromaticity of one of those, *illuminant E*. Its specified spectrum is entirely flat—equal power in any wavelength increment of a given size. The "E" is evocative of "equal energy", a traditional way to describe that spectrum (although of course it is power, not energy, that is at issue).

Its chromaticity is very nearly that of another standard illuminant, D55, which is representative (in actual spectrum as well as chromaticity) of "mid-morning/mid-afternoon daylight".

USAGE OF THE CIE XYZ COLOR SPACE

General scientific work

The CIE XYZ color space (or its cousin, the CIE xyY color space) is the most common way in technical colorimetric work to describe the color of light. For example, in the specifications for other additive color spaces (such as sRGB), the chromaticities of the standard primaries are defined in terms of their x and y coordinates.

When equations are given to allow the description of a color to be transformed from expression under one color space into expression under a different color space, most commonly the transform is given in two steps, first a transform from the "source" color space to the CIE XYZ color space, and then a transform from the CIE XYZ color space to the "destination" color space.

ICC Color profiles

In the careful management of color representations in photographic imagining, *color profiles* are used to characterize the actual (often locally- and currently-measured) responses of such things as image scanner chains, computer display chains, or printing chains[11]. In fact, these profiles define a "parochial" color space for the device, since what interests us is (in the case of a printer chain) what numerical values to present to the printer to make it print a certain color—the ultimate significance of a color space.

In software that deals with this matter (such as image editing software, image "viewing" software, and certain browser software), the image is essentially transformed from representation under the color space associated with the arriving image data (perhaps sRGB, as is normally used in a JPEG image file) into the "parochial" actual color space of, for example, the display chain or printer chain being used at the moment.

These profiles are stored in files in a format defined by the International Color Consortium, and are called "ICC color profiles". These (usually) define the color space of interest (standard or parochial) in terms of transforms to and from the CIE XYZ color space.

At least conceptually, the process works this way. Suppose we will be starting with the image represented in the sRGB color space (which has its own profile; remember, a profile is really a very detailed description of a color space) and we must transform it to representation in the actual color space of the printer chain leading to a certain photo printer.

For each pixel of the image, the software, guided by the ICC profile for the sRGB color space, transforms its sRGB coordinates into XYZ coordinates under the CIE XYZ color space. Then, immediately, guided by the ICC profile for the printer chain, these coordinates are

[11] The term "printer chain" (for example) is used to make clear that the properties of interest are not just those of the printer itself but also come from the color-handling behavior of the printer driver software.

transformed into coordinates of the printer chain color space (which will probably be in the format of coordinates of the sRGB color space).

In actual practice, the software, which will need to do this for millions of pixels per image, probably for several images in this session, may combine the “sRGB to XYZ” transform (defined by the ICC sRGB profile) with the “XYZ to printer” transform (defined by the ICC profile for the printer in use) into a single transform, which it will then use to transform each pixel of each image in one operation.

The “intermediate” color space (CIE XYZ in most, but not all, situations) is functionally called the “profile connection space” (PCS).

FOR MORE INFORMATION

More detailed information on this topic, on some of the collateral matters discussed above, and on several related issues will be found in the article, “Digital Camera Sensor Colorimetry”, by this same author, probably available where you got this article.

#