Colorspace Interchange Using sRGB

By

Gary Starkweather

Color management technology for personal computers is now becoming available on a broad basis. Much has been written about how to get the best color, what transforms to use etc. Among the most recent additions to editorial content has been issues regarding the use of sRGB as a standard interchange colorspace. Several articles have been rather shrill in denouncing or detracting from sRGB's inherent value. This is unfortunate and this article is intended to explain both the benefits and concerns of using any colorspace for color data interchange in addition to using sRGB as that space. It is hoped that the material contained here will help the reader understand why sRGB has value as well as what its limitations are. No "standard," whatever it is comes without some limitations. The appropriate solution is that which offers the most benefit and the least trouble. It is felt that sRGB offers an acceptable solution at this time. The purpose of this brief paper is not a deep technical discussion which is available in detail along with other references at www.srgb.com but an explanation of the choices available and why the sRGB choice was made.

Color interchange spaces have the benefit that all color devices can conveniently exchange color data in a common format. This foregoes the need to constantly translate between color spaces due to difference data representations in the various devices. The CIE system (Commission Internationale de l'Eclairage or International Commission on Illumination) provides a solid scientific foundation for color data based on careful experimental measurements of human observers. However, there are a number of ways to utilize the CIE data. The CIELAB and CIELUV systems are but two systems intended to provide color data that at least approximates perceptual uniformity. However, two color patches having highly similar LAB values do not necessarily appear the same depending on visual conditions. The color of the illuminant (white point), the absolute level of the scene irradiance (visually the illuminance), the surrounding colors etc. all serve to defeat the "perfect match" unless the initial and test conditions are identical. Photographs are often quite acceptable to most users but rarely does one travel back out to the original scene of the photograph and check to see that the leaf e.g., and the photograph of the leaf are identical. The appearance of the leaf in the photo is what is critical in most cases. This will not necessarily be true for demanding commercial applications such as buying clothing, for example. One can usually benefit from the old axiom in graphic arts originally coined by Professor Miles Southworth of RIT. The axiom states "that "clean and bright is always right" and "dull and gray is not the way."

So why not just pick CIELAB, for example, to become the interchange space. First of all, this in and of itself would not suffice. One would need to specify a white point and illuminance level of the scene in order to have any reasonable chance of achieving a proper representation. The other problem with CIELAB is that getting data into and from it requires the use of cube roots or raising values to the third power. It would be nice to be able to store color data in a form that when one wanted to view it, it was indeed ready for viewing under most reasonable conditions. Storing data in CIELAB would require converting every pixel through a set of power function routines. This might not only be time consuming but could be fraught with possible errors depending on who was doing the conversion. Another approach that might satisfy those concerned with gamut issues (CIE carries the full gamut extent of the visual system), would be to let the data conform to a theoretical display such as a laser display. Laser's have virtually monochromatic output and therefore the primaries of a laser display would reside on the spectrum locus of the CIE diagram as shown in figure 1.

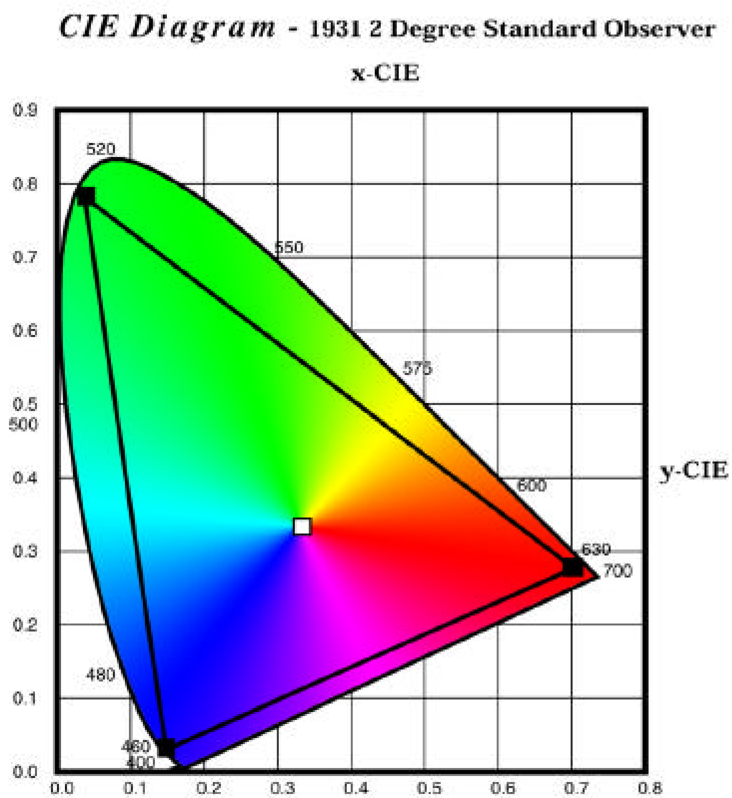

Figure 1 – Laser Display Color Gamut

Using the 633 nm. output of Helium-Neon, the 514 nm. output of Argon and the 442 nm. output of Helium-Cadmium would make for an impressive color gamut. Other laser lines such as those of Krypton etc. could also be considered but would not alter the conclusions of the argument.

First, in utilizing a laser display the data can be carried across devices without gamut loss. That is the data would stay around as positive numbers. No commercial displays, scanners or printers exceed or come close to exceeding the laser display gamut. The problem comes from the fact that no device normal users will get their hands on has this gamut. Thus data in the laser display colorspace will have to be converted for display and print. This requires additional calculation again, however, not as complex as with CIELAB. Gamut mapping would be required however for destination devices. Thus, we have the calculation and representation issues of CIELAB and the large gamut but non-device representative nature of the laser display. What other choices are available? Today, most of the devices on which images are viewed are displays. The predominant display technology today (~80%) are the cathode ray tube or CRT display. Color flat panels whether passive or active matrix type also try to come close to the color CRT model. High definition television or HDTV will use a set of phosphors known as ITU-R 709.BT. The color gamut of these phosphors is shown on a 1931 CIE diagram in figure 2.

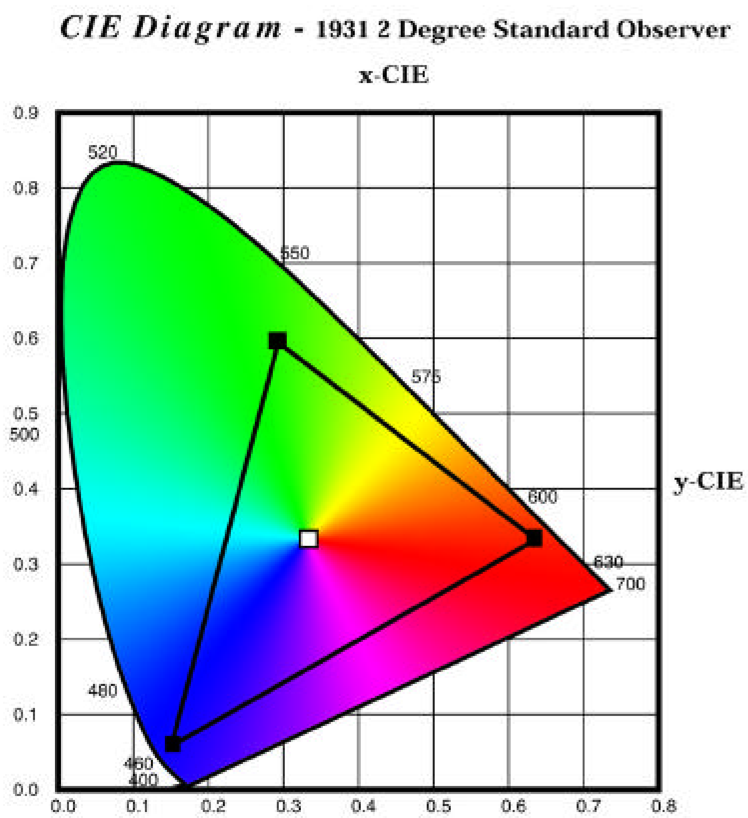

Figure 2 – sRGB Color Gamut

From the spectral and colorimetric characteristics of these phosphors, a standard known as sRGB was proposed and adopted by a number of companies such as Hewlett-Packard and Microsoft. The reader can visit the web site at www.srgb.com for additional information.

First, sRGB is not believed to be a perfect interchange space but it does possess a number of salient features that should be appreciated. First, as can be seen in figure 3, it does not differ much from the old Apple 13” RGB monitor gamut.

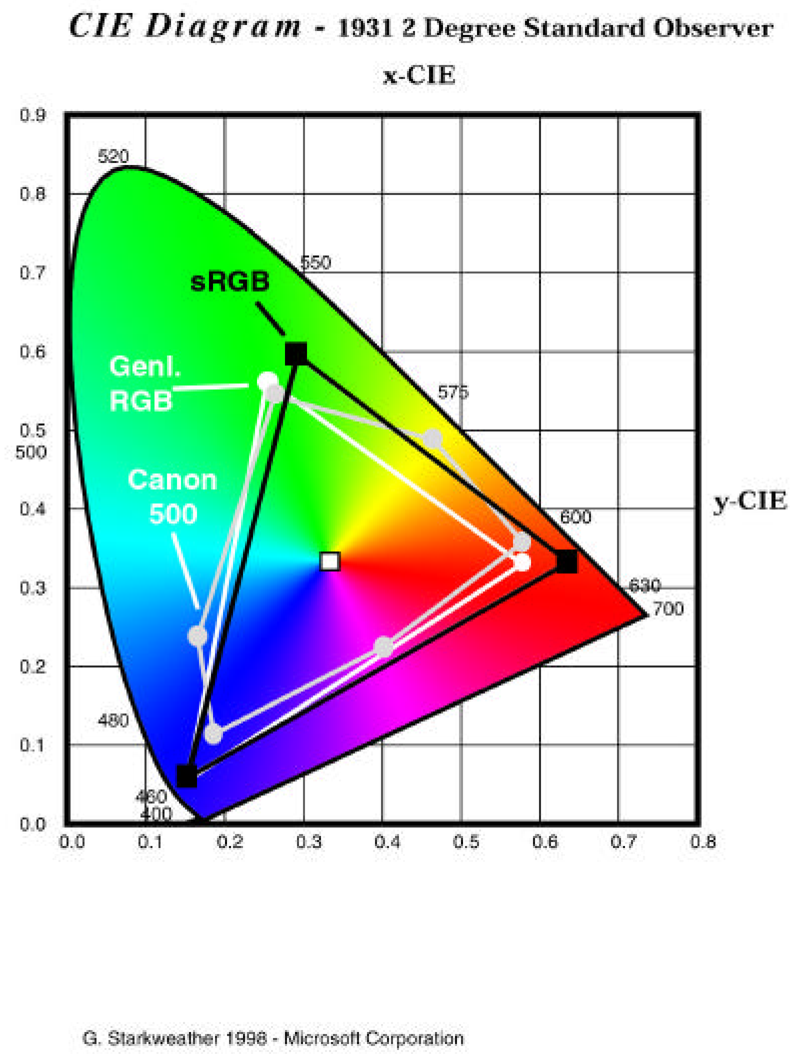

Figure 3 – Color Gamut Comparisons

Thus, those people imagining that this gamut is strikingly different from their CRT do not need to worry. Furthermore, the sRGB standard has built into it the illuminance levels and other visual environment features that are the most common among the consumer and general user community. Perhaps the most controversial parts of sRGB have to do with gamut limitations and the gamma value built into the standard. Let's take the gamma value first. The sRGB gamma value of 2.2 provides pictures that look darker on a monitor set to a lower gamma value, say 1.8 or even 1.4. This is an easy change of course but the gamma of 2.2 was chosen for good viewing conditions and an attempt at perceptual uniformity. Equal perceptual steps in a grayscale image would tend to follow a gamma of 2.2 although not perfectly. To utilize the 2.2 gamma of sRGB some users will have to modify how they have worked in the past and this always presents some difficulties. The gamut issues are also of concern to many. They claim that the sRGB gamut is too limited and hence "clips" their output device's ability to reach its full potential. This is only true if one permits this to be the case. For example, in figure 3, a gamut of the venerable Canon CLC500 color copier/printer is shown along with sRGB and the old RGB realized on the first PC color monitors. Note that while some of the cyan colors are limited by sRGB, the brightest greens and reds are output device limited not sRGB limited.

In reality, there are devices such as film scanners with their high quality images that exceed the sRGB gamut. This is not a reason to toss out the sRGB standard. When colors go out of gamut mathematically, they can either exceed an accepted maximum usable value like 255 or fall below an accepted minimum value such as zero. Going out of gamut does not cause the data to disappear from the universe unless the data is purposely "clipped" to an artificial lower and upper bound. If one keeps the data that goes out of gamut then when devices that exceed the sRGB must be dealt with, the data is available for use in the necessary calculations. Therefore, a major concern about any gamut that can be exceeded is avoided. One of the important benefits of sRGB is often overlooked or not mentioned by its detractors. That benefit is that most people using personal computers spend a great deal of time either viewing, editing or otherwise dealing with color data on their screens. Since the sRGB colorspace is or can be quite representative of a majority of displays, data in sRGB can be directly viewed on their monitors without modification. This is an important advantage in that purveyors of color data can place their content in sRGB and know that in most instances when it is viewed on computer monitors it is in a "ready-to-go" state. No cube roots, table look-ups or other processing is required. Until color data processing is so pervasive and fast as to be both spatially and temporally transparent to the user this is and will be a substantial benefit. If the phone system were like much of today's proposed color spaces, we would print and display phone numbers in a form that would require multiplication and division combined with addition and subtraction to get the number we needed to dial. Why not just give the user the 767-2345 number instead of a code they must be prepared to process? SRGB is aimed at providing an acceptable not all encompassing solution to users and information suppliers alike.

In summary, sRGB is not a perfect color space. However, it is representative of the majority of devices on which color is and will be viewed. Furthermore, sRGB is a profile that is compliant with ICC profiles and current color management systems. Carrying around out of gamut data would permit one to utilize the full color capability of their devices, albeit at some risk. The risk is that my red may be brighter than your red since we have different printers. Utilizing in-gamut sRGB data means that we can ensure a substantial degree of consistency as well as have our output expectations reasonably well set at the point of creation (the display). More sophisticated color management tools not usually in the interest domain of ordinary users would permit utilizing data outside the sRGB gamut for those brightest of colors for which the user may have paid a lot to realize. Thus, properly used, sRGB provides a consistency, speed and quality level that should satisfy the majority of consumers who want quality images with minimum hassle. For those for whom color is a highly polished skill from which they either earn a living or enjoy in its own right, sRGB does not foreclose their ability to add value to the imaging chain. Using sRGB is aimed at achieving the equivalent experience of dropping off photos for printing. You are not asked what temperature you desire for the first developer or how many stops you wish to "push" the development process. Aristotle did not likely comprehend the era of digital color but his statement that “simplicity is the truest elegance” could apply to color imaging as well. I think he would appreciate sRGB.