Wide Gamut Device-Independent Colour Image Interchange

C. A. Poynton
Sun Microsystems Computer Corporation, U.S.A.

In this article I will discuss some aspects of colour reproduction in CRT cameras and displays, photography and offset printing. I will explain how colour data can be coded in calibrated, nonlinear *RGB* values in order to take advantage of accurate colour interchange, and how colours outside the gamut of *RGB* primaries can be conveyed using an extension of *RGB* coding.

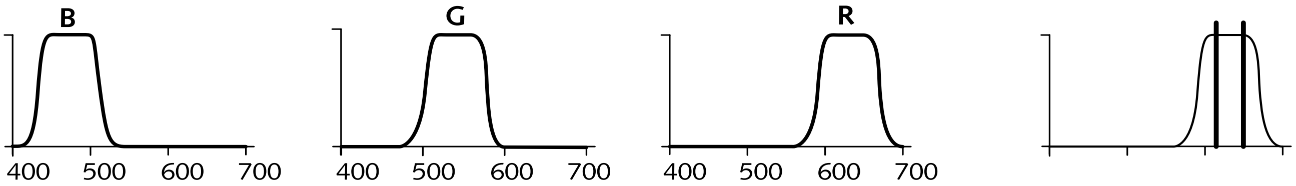
Video Cameras

The colour interpretation of red, green and blue data from a camera depends on the spectral characteristics of the optical filters that are used to separate the wavelengths of visible light into red, green and blue components, and on the spectral characteristics of the photosensitive elements in the camera. If the concatenated spectral characteristics do not closely approximate the spectral response functions of the CIE Standard Observer, then the camera will fail to "see" colours the same way that we do. This is illustrated in the sketch at the bottom of the page.

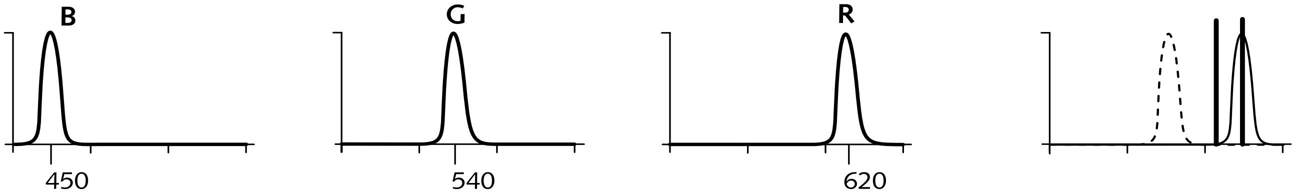
In practice, the choice of camera sensitivities is a compromise among color accuracy, sensitivity and cost. In broadcast-quality cameras, the linear-light signals from the photosensors are usually processed through a linear matrix, prior to gamma correction of the red, green and blue signals.

When spectra are sensed according to the standardized CIE weighting functions, colors can be plotted in the CIE *x, y* coordinates of the *chromaticity diagram* on the next page.

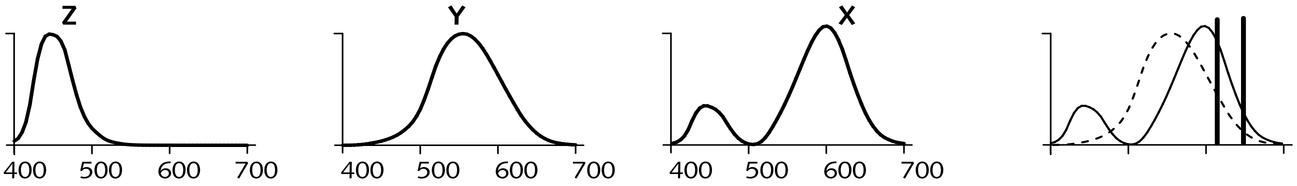
To achieve adequate performance with eight bits per colour component requires nonlinear coding. Refer to my article [1] for details.



Wideband filter set

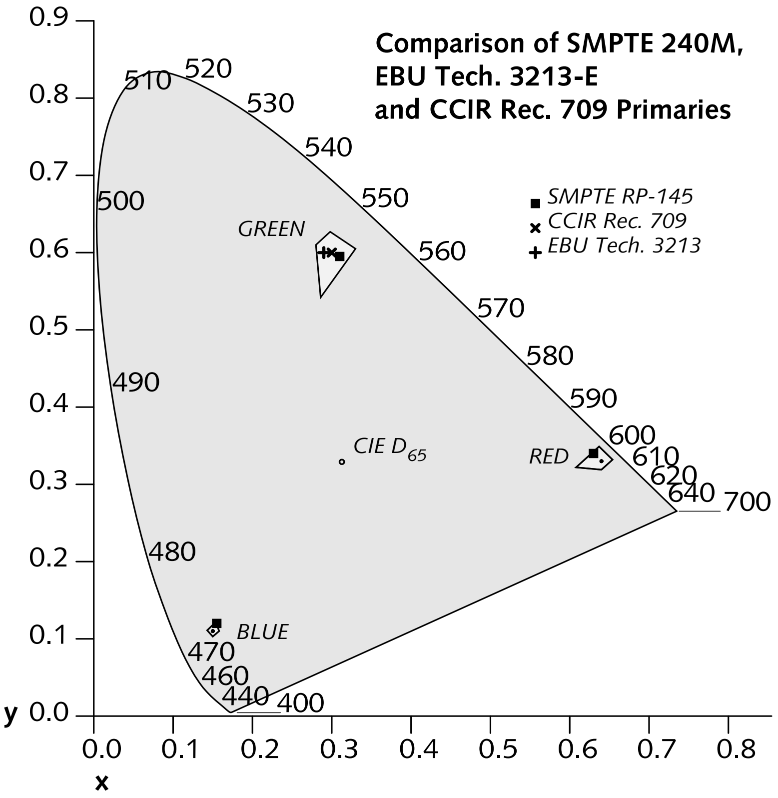


Narrowband filter set



CIE-based filter set

Reprinted from *Proceedings of the International Broadcasting Convention*, 16–20 September 1994, Conf. Publ. No. 397, © IEE, 1994. The author is no longer with Sun, and can be reached at <www.inforamp.net/~Poynton>.



CRT Displays

Cathode ray tube (CRT) monitors have red, green and blue phosphors. Consequently, video signals are universally based on RGB colour components. However, there is no universal, objective definition of what colours constitute "red", "green", "blue", or even "white." For RGB data that indicates saturated red – 100% red, 0% green, 0% blue – is the red intended to be scarlet, reddish-purple, or reddish-orange? In practice, the colour reproduced depends upon the interpretations given to each of the primaries by a particular device. In a CRT monitor, the colours produced by the phosphors used in the manufacture of the CRT – the *phosphor chromaticities* – determine the colours of red, green and blue. The balance of power among the three electron beams – the *white point* – determines the colour assigned to white. These parameters are different for different monitors, and colour reproduction is not predictable without control of these parameters.

In contemporary television systems, a reference white chromaticity of CIE D_{65} is universally used. However, two primary chromaticity standards are in use: SMPTE RP-145 in North America and Japan, and EBU Tech. 3213 in Europe. A third standard CCIR Rec. 709 has been agreed internationally for HDTV, but is not yet deployed. The CCIR Rec. 709 chromaticities are as follows:

	CIE x	*CIE y*
Red	0.640	0.330
Green	0.300	0.600
Blue	0.150	0.060
White	0.3127	0.3290

All three standards are very close in their technical parameters, and for practical purposes the CCIR set is an excellent compromise between SMPTE and EBU. The graph at the top of this page shows the red, green and blue primaries of these three standards, plotted in CIE *x, y* coordinates. The tolerances for the EBU set are defined in *u′v′* coordinates related to the sensitivity of human vision to colour differences.

Contemporary computer displays have primaries that approximate Rec. 709, although various white references are in use. Emerging colour management systems will soon make it relatively easy for computer systems to compensate for display devices having different primaries and white points.

Printing and Photography

When ink or other transparent coloured material is deposited in successive layers, the colour mixing is described as *subtractive*. The widest range of colours is produced when the colours of the pigments are yellow, cyan and magenta. Photographic film employs dye layers of these colours.

Process colour involves overprinting carefully-chosen amounts of four inks coloured cyan, magenta, yellow and black (CMYK). The black ink is not necessary in theory, but is used in practice because black ink is cheaper than coloured ink. When CMY information is to be printed, the process of *grey component replacement* finds the minimum amount of ink common among all three CMY components, and replaces the common component with black. The resulting mixture contains less total ink than if just CMY were used, so not only is the CMYK ink mixture cheaper than the equivalent CMY, but the printed paper dries faster and the press can be run at a greater speed.

The exact colour reproduced by a subtractive system depends on the spectral characteristics of its pigments (inks or dyes) and its media, and upon the illumination with which the reproduction is viewed. Characterizing a subtractive reproduction system is much more complicated than an additive system, because the overlap among the spectra of the pigments causes the colour combinations to be nonlinear. For example, yellow ink should ideally absorb only blue light, but a practical yellow ink absorbs a significant amount of green as well. Achieving correct reproduction of mixtures that involve green requires knowledge of this interference. There is one highly advantageous aspect of the so-called “unwanted absorption”, though: the nonlinearity of mixing produces a gamut for subtractive reproduction that is wider than would be the case for ideal, non-overlapping pigments. The gamut of everyday printing inks, and everyday photographic dyes, extends beyond the CRT gamut in certain regions of color space.

Once an image data is in CMY or CMYK form it is in some sense married to a particular set of pigments. Once an image is in CMYK form, to move the printing job from one plant to another is difficult or impossible. Also, reliance on the CMY colour model makes it difficult to use the RGB devices, software and images that are ubiquitous in the computer domain.

Device-dependent Colour

From the descriptions above, you can see that the reproduction of colour in these various systems – CRT displays, cameras, offset printing and photography – is highly device-dependent: if RGB values from one device are sent to another without consideration of different device characteristics, then poor colour reproduction will result.

The age of highly-specialized, single-function systems is over. Images are originated on film and transmitted on video. Images from computer desktop scanners or from photographs make their way to video. Video images find their way to print. People expect to move images among diverse equipment, and expect to retain the colour integrity of the images.

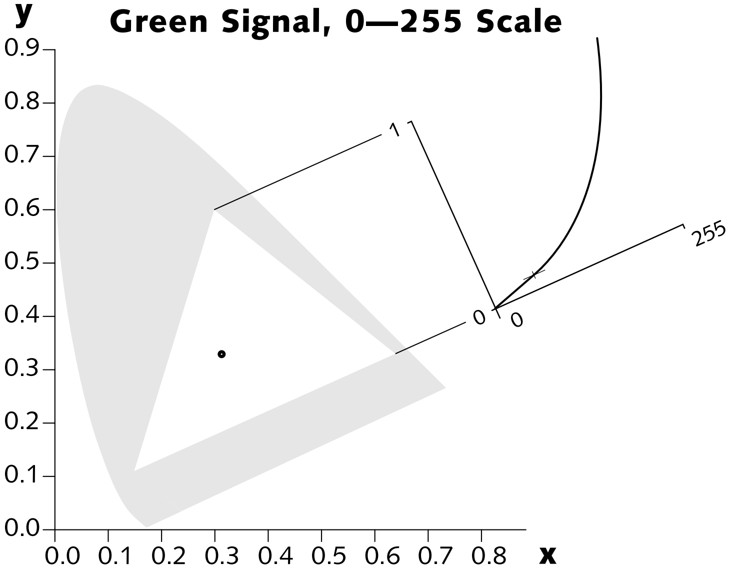
Although it is possible in theory to compensate for different colour systems by performing local correction to an image when it is imported or exported, the task of exchanging images will be simplified if we can identify a preferred interchange space.

Colour Image Interchange

I propose that a calibrated nonlinear RGB colour space can be exploited for digital image interchange to gain the advantages of device independence. In my opinion, very broad range of users could be completely satisfied by a single well-chosen RGB space.

For a single space to achieve acceptance it must be *objective*, that is, it must have a tightly-defined relationship to the CIE standards for colour.

To be practical, the chosen space must achieve very good image quality with three components of eight bits each. This means that the space must be reasonably perceptually uniform, in order to maximize the visual utility of the available codes. CIE *XYZ* space is not perceptually uniform, and requires about 16 bits in each



component to achieve excellent image quality. Consequently, CIE *XYZ* is not a good choice for an image interchange space. The CIELAB and CIELUV spaces have been successfully used to convey image data in a device-independent manner using three eight-bit components. However, these spaces have the disadvantage of computational complexity.

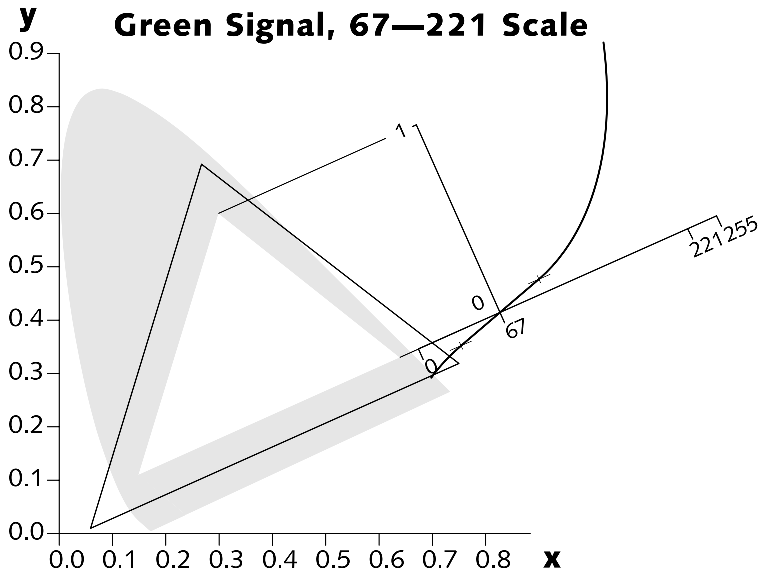
The chosen space should be capable of good interactive performance. Users will demand reasonably fast display of newly acquired image data, without a computationally prohibitive colour transform calculation. The computational complexity of spaces such as CIELUV and CIELAB prevents interactive performance.

CCIR Rec. 709 Calibrated, Nonlinear RGB

These constraints suggest a calibrated, nonlinear RGB space. We are fortunate to have obtained in April 1990 unanimous worldwide agreement on a calibrated nonlinear RGB space for HDTV production and program exchange: CCIR Recommendation 709. I suggest that the parameters of CCIR 709 form an excellent basis for a preferred colour interchange space for digital pictures.

Note that I have been careful to avoid using the words *standard colour space*: nothing in this approach precludes use of a different colour coding if it is more appropriate to a particular application. However, in my opinion, the identification of a well-chosen *default* or *recommended* colour coding has the possibility of sidestepping the creation of multiple, functionally-equivalent systems, with the attendant long-term conversion headaches.

Although they were developed for the purpose of measuring and specifying colour differences, the CIELAB and CIELUV spaces have been successfully used to convey image data. Roughly the same amount of computation is required to transform any of CIELAB, CIELUV or CCIR 709 RGB to print. However, CCIR 709 was designed to be extremely efficient for driving today's CRT displays: the transform from the CCIR 709 interchange space to a CRT's *RGB* primaries is extremely simple, because it must be performed at real-time pixel rates in low-cost consumer equipment! CCIR 709's suitability for CRT displays does not come at the expense of applicability to alternate display technologies, however: the acceptance of CCIR 709 as the worldwide HDTV colour standard ensures that it will thrive as an interchange space for new display technologies – in particular, flat panels – now under development. In addition, CCIR 709 is suitable for prepress: the ANSI IT8 committee is exploring use of the CCIR 709 primaries for exchange of device-independent *RGB* image data, although IT8 expects to use the CIE D_{50} white point that is ubiquitous in print work rather than the D_{65} standard that CCIR 709 inherited from television.



Wide Gamut Reproduction

Most physical devices – scanners and CRT displays, for example – employ *RGB* values that are physically constrained to be nonnegative. The gamut of colours that can be specified by all-positive *RGB* values is restricted.

In computing it is standard to code digital *RGB* data with zero for black and the all-ones code for white (for example, 255 in an eight-bit system). The graph at the top of the opposite page shows the excursion in the CIE *x, y* diagram when the green primary varies between 0 and 255, while the other two components vary from 255 to zero. The gamut is contained in a triangle whose vertices are the primary chromaticities.

In component digital video signals in the broadcast studio, the excursion from black to white is standardized as 16 through 235. The headroom and footroom were originally provided to accommodate filter undershoot and overshoot.

If a wide colour range is to be accommodated, extensions to the traditional *RGB* coding range are necessary. One way to extend the gamut of an *RGB* system is to allow *RGB* values that are negative or greater than one. For example, Kodak's PhotoCD™ system uses a coding where black is placed at code 67 and white is placed at 221. The CIE luminance of any code combination in the PhotoCD system is constrained to the range zero to unity, but providing that constraint is met, colors can be expressed that have one or two components below zero or above unity. The graph at the top of this page shows this coding range plotted on the CIE *x, y* diagram. The gamut is considerably larger than if the system is limited to excursions between zero and unity. PhotoCD is designed to reproduce all the colours that film can reproduce.

As an example of an extended colour, the saturated cyan of a Salem cigarette carton cannot be expressed as an all-positive combination of Rec. 709 primaries: that colour is outside the *gamut* of Rec. 709. However, that cyan can be represented as a set of *RGB* values where red is somewhat negative and blue and green are both slightly larger than unity. Upon conversion for a CRT, the out-of-range signals will be clipped, and the display will produce the most saturated cyan available. But the wide-gamut coding has the great advantage that the true colour is maintained for subsequent devices, for example if the colour is later to be reproduced on photographic film, or on a dye-diffusion printer.

References

[1] Poynton, Charles A., "Gamma and Its Disguises", in *SMPTE Journal*, Vol. 102, No. 12 (December 1993), 1099–1108.