The XYZ Colour Space

The XYZ colour space has the unique property of being able to express every colour that the human eye can see which in turn means that it can express every colour that can be captured by a camera and hence every colour that anyone might ever want to reproduce in video.

This property led to the adoption of the XYZ colour space by the Digital Cinema Initiative (DCI) as its standard colour space. Expressing colours using in the XYZ colour space supports at least as much depth of colour as the film against which digital cinema is pitched.

However, while it is possible to build a camera that comes close to recording colours in the XYZ colour space – a so-called perfect camera – it is impossible to build a monitor that works in the XYZ colour space because X, Y and Z are not real colours. It is also impossible to select a set of colour primaries for a monitor that covers all the different colours that can be expressed in XYZ. So compromises have to be made.

This white paper gives the background to the introduction of the XYZ colour space and explains the compromises introduced in converting digital cinema material for transmission and display. It also introduces the facilities offered by the OmniTek OTM and OTR waveform analysis systems that aid the post-production task of achieving the least possible degradation in any transmission from the ideal as expressed by the XYZ source.

Why RGB isn't good enough

We are taught at school that any colour can be produced by mixing light from red, green and blue display primaries and indeed CRTs, LCD displays and plasma screens all generate coloured images by applying different intensities to arrays of red, green and blue sub-pixels. However a significant range of colours can't be produced in this way. In particular, it isn't possible to produce a wide range of blue-green colours – colours that are perfectly possible to capture in a camera and reproduce on film.

Lying behind the loss of these colours is the response of the 'red', 'green' and 'blue' cones in the human eye to light of different wavelengths. This response is illustrated in *Figure 1*. The $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ curves outline the responses of the red, green and blue cones respectively. These curves show how each type of cone responds to the wavelengths of the visible spectrum. (The bar at the top shows the colours associated with these wavelengths.)

Notice how these ranges overlap. Also notice that, while the responses for blue and green cones show a single peak over this range of wavelengths, the red cones have a second smaller peak in the blue part of the spectrum.

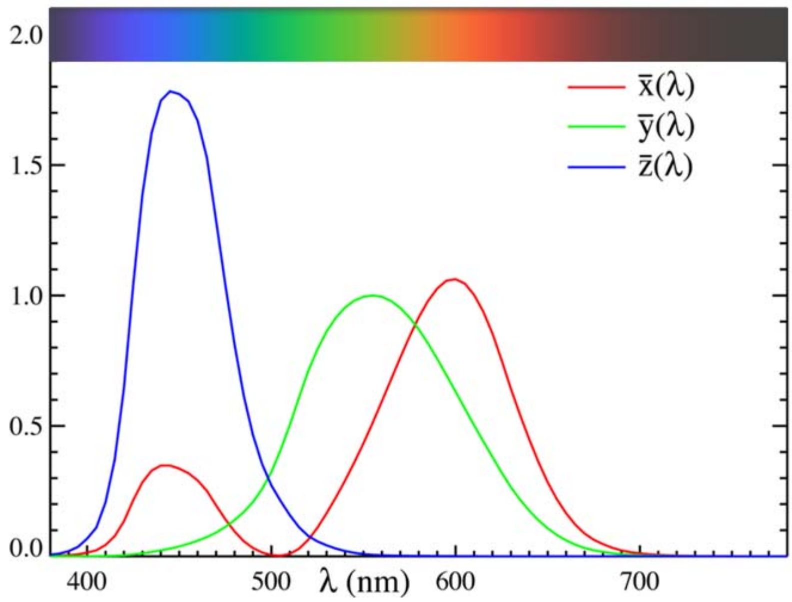

Figure 1: Normalised responses of the red, green and blue cones of the eye to different wavelengths.

(The details of these responses were deduced from results by John Guild of the National Physical Laboratory and W. David Wright of Imperial College, who in separate experiments in the 1920s asked subjects to adjust the intensities of the supplied red, green and blue light sources to match light of chosen wavelength. Different people have different red, green and blue cone responses so the curves shown above are for a nominal standard observer.)

The colour an object appears to the observer is the eye's overall response to the particular mixture of wavelengths reflected by the object. The component at any particular wavelength causes a signal in each of the red, green and blue cones that is equal to the product of the intensity of the light at that wavelength and the cone's response to light of that wavelength. The response to colour in humans is (approximately) linear and so this can be expressed mathematically as:

$$X = \int I(\lambda).\overline{x}(\lambda)d\lambda = \text{Signal from Red cones}$$
$$Y = \int I(\lambda).\overline{y}(\lambda)d\lambda = \text{Signal from Green cones}$$
$$Z = \int I(\lambda).\overline{z}(\lambda)d\lambda = \text{Signal from Blue cones}$$

The colour that the viewer perceives a pixel on a monitor to be is similarly a suitably-weighted sum of the eye's response to light of different intensities from the monitor's display primaries.

For many of the colours you want to reproduce, it is possible to find a combination of intensities for the display primaries that will trigger the desired response from the viewer's red, blue and green cones. However this is not true for all colours. A particular difficulty is found with blue-green colours. A true blue-green colour such as that corresponding to a wavelength of 500nm just triggers the blue and green cones. However, as the following diagram shows, any output from either the blue or the green primary (represented by the vertical blue and green lines) will cause a response from the red cones as well as from the blue/green cones. This makes it impossible to produce a combination of outputs from the primaries that just triggers a response from the blue and green cones.

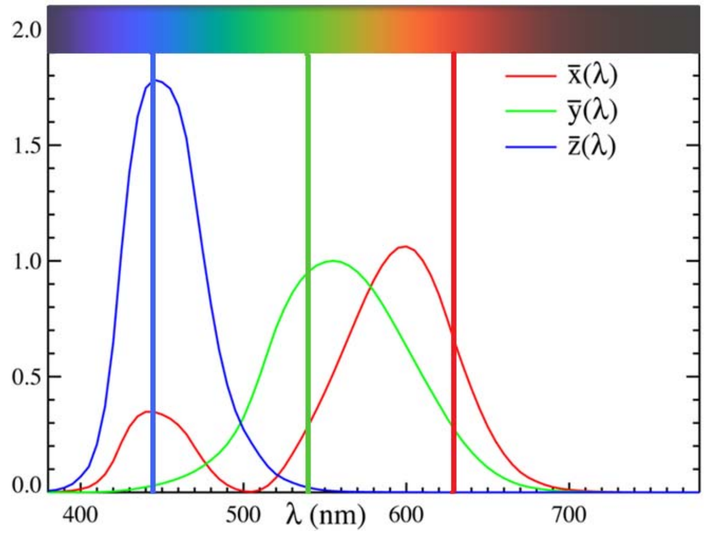

Figure 2: Eye response diagram highlighting the responses to a representative selection of red, green and blue primaries.

The Birth of the XYZ Colour Space and its adoption by the DCI

In a bid to express any colour unambiguously, researchers set out to define a colour space that could express the full range of colours visible to the human eye. This colour space had to meet the following criteria:

- That three independent and positive variables are necessary and sufficient to specify any colour
- That only the tri-stimulus values of the sources were relevant in expressing any colour, not their spectral composition
- That if one or more of the sources are changed gradually, the resulting tri-stimulus values also change gradually.

The obvious choice to meet these criteria was a colour space based on the individual responses of the red, green and blue cones and so was born the XYZ colour space in which:

$$X = \int I(\lambda).\overline{x}(\lambda)d\lambda = \text{Signal from Red cones}$$

$$Y = \int I(\lambda).\overline{y}(\lambda)d\lambda = \text{Signal from Green cones}$$

$$Z = \int I(\lambda).\overline{z}(\lambda)d\lambda = \text{Signal from Blue cones}$$

The XYZ representation of a colour is completely unambiguous, and any colour in the visible spectrum is represented by a combination of positive X, Y and Z values. However, while it is possible to build a camera with the response characteristics needed to record colour in XYZ, X, Y and Z are not real colours and moreover have some negative RGB components. So it isn't possible to build a display that works in the XYZ colour space. Instead, monitors continue to work with RGB primaries and moreover the results produced by any monitor depend on exactly what red, green and blue primaries it uses.

In the world of film, the cameras and the chemicals used in capturing and developing images shot on film allow the capture of a high proportion of the visible spectrum. The objective for digital cinema was therefore to offer the widest possible range and depth of colour.

Rather than link the standard to any particular technology, the Digital Cinema (DCI) specification did two things. Firstly, it promoted the use of the XYZ colour space in video for transmission. Secondly, it specified as its reference system for display a hypothetical projector which uses the set of RGB primaries that offer the largest colour range that could possibly be achieved. Taken together, these give digital cinema the twin advantages of the best possible position in relation to film while at the same time future-proofing the DCI standard against future improvements in technology.

Converting between XYZ and RGB

The XYZ representation of a colour is completely unambiguous, and any colour the eye can see is represented by a unique combination of X, Y and Z values. However, monitors and TV displays operate in a world of RGB primaries. So while the XYZ colour space provides an excellent way of communicating the desired colour, we also need a method of converting from XYZ to the appropriate RGB.

Fortunately, observation shows that the human response to colour is (approximately) linear. This empirical rule is known as Grassman's Law and it means that any tri-stimulus value in one colour space can be converted to the corresponding tri-stimulus value in another colour space by applying a linear transform (i.e. one involving just multiplication and addition). Thus any XYZ value can be converted into an RGB value by applying a linear transform. This is represented mathematically as:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} & & \\ & \text{3x3} & \\ & & \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

(Similar matrices are used in converting from RGB to the XYZ colour space.)

The mathematics involved in determining the 3 x 3 matrix to apply in any particular case is fairly complex but it is given, along with a worked example, in SMPTE RP 177 (and reproduced as an appendix to this white paper).

There are however two important issues with conversions from XYZ to RGB.

Firstly, a significant proportion of the colours that can be seen by the human eye (and therefore able to be expressed in XYZ colour space) transform to negative R, G or B values, which in turn means that they cannot be reproduced using the RGB colour primaries. The only thing to do with any negative value is to replace it by the smallest permitted value (which is typically 4). Similarly, the transformation can also produce values that exceed the upper limit of the accepted range which therefore have to be replaced by the largest permitted value (which is typically 1019). The effect of these adjustments is to de-saturate the colours in the image. This de-saturation is simply an inevitable consequence of colour space conversion, though it is one that exercises the brains of post-production colorists!

The other notable aspect of these conversions is the different ranges of colours that the different RGB systems work with. Not only do different monitors use different devices as the sources of their colour primaries, but even the PAL, NTSC and HD transmission standards define sets of RGB primaries that are subtly different from each other (see *Table 1*).

Video Standard	Red		Green		Blue	
	x	*y*	*x*	*y*	*x*	*y*
PAL	0.64	0.33	0.29	0.60	0.15	0.06
NTSC	0.63	0.34	0.31	0.595	0.155	0.07
HD	0.640	0.330	0.300	0.600	0.150	0.060
DCI	0.680	0.320	0.265	0.690	0.150	0.060

Table 1: Display primaries defined by the different video standards.

These differences mean that colours that are perfectly valid in one RGB system can fall outside the valid colour gamut in another RGB system.

Another thing to be aware of is that video that uses the XYZ colour space is commonly delivered as gamma-corrected X'Y'Z'. This means that, before it can be converted to another colour space, the gamma correction that has been applied prior to transmission has to be reversed.

XYZ and the OmniTek OTM/OTR Systems

In the light of Digital Cinema's adoption of XYZ as its standard colour space, support for the XYZ colour space is offered on both OmniTek's OTM system and on its OTR systems.

This support is offered through the VIEW_XR_DCI software option which is aimed at Digital Cinema applications. As well as offering support for the XYZ colour space, the VIEW_XR_DCI option also offers support for 12-bit data, Gamut Histogram displays and a unique 'CIE Chart display' (described in a separate white paper).

On a superficial level, the addition of XYZ support is seen in the addition of XYZ to the choices offered at particular points of the OTM/OTR application. For example, the Waveform View becomes able to display XYZ waveforms alongside the RGB and YCbCr waveforms that are provided as standard. XYZ also becomes offered as an option for the main Gamut display.

The various conversions that are needed in switching between RGB/YCbCr representations of pixel colour and the equivalent XYZ values are all carried out automatically for you, along with the reversal of the gamma correction applied to X'Y'Z' source data. All you need to supply are details of the colour primaries to use in interpreting XYZ signals (from a choice of SMPTE C Gamut; EBU Gamut, Rec 709 Gamut or DCI Gamut), and the standard to follow in converting XYZ to YCbCr (from a choice of SMPTE C, EBU or Rec 709).

Other displays enable the user to see the effect of the conversions that are applied. In particular, the CIE Colour Chart provides a very precise way of identifying the colours that fall outside the available RGB range.

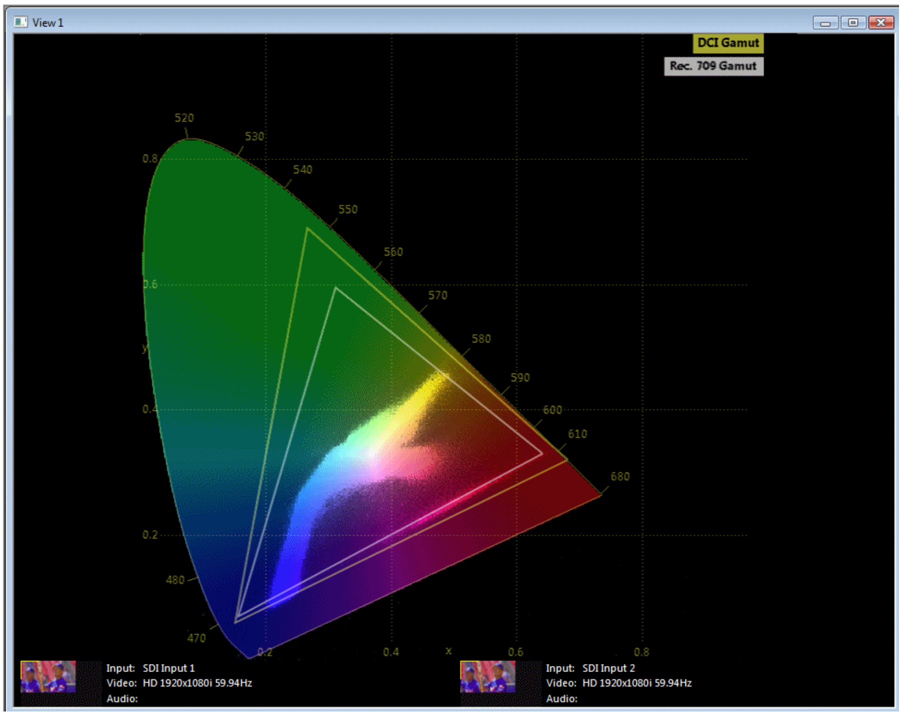

Figure 3: Example OmniTek CIE Colour Chart Display

The triangles that are overlaid on this display are especially important. Each triangle is formed from the points representing the three colour primaries associated with a particular video standard and encompasses all the colours that that video standard supports.

Two triangles may be displayed – a yellow triangle and a white one. The yellow triangle marks out the range of colours supported in the source video format. The white triangle marks out the range of colours supported by a destination video format selected by the user.

The bright points on the display each represent the colour of a pixel in the frame that is currently being analysed.

The chart shows directly how the colour will be changed on conversion to the selected destination video format simply by looking at those bright spots that fall between the two triangles because any pixel where the colour falls inside the yellow triangle but outside the white triangle on the CIE Chart display will have to be modified in order to conform with the new standard.

Inevitably these modifications result in some de-saturation of the image as the move from the original colour to the final colour is always a move towards the centre of the chart and hence a move towards white.

Appendix: Deriving the Matrix for Converting XYZ to RGB

SMPTE RP (Recommended Practice) 177 includes the following description of how to derive the matrix used to convert between the XYZ representation of a colour and its representation using any particular set of RGB colour primaries.

3.3 General procedure

The general procedure for deriving the matrix relating normalized linear RGB signals to CIE XYZ tri-stimulus values is described in this clause and an example derivation is given in annex B. The RGB signals are normalized such that reference white has the values R=G=B=1.0. The step-by-step process is as follows:

3.3.1 Obtain the CIE *x,y* chromaticity coordinates of the reference white (D_{65} for television) and of the RGB primaries.

3.3.2 Compute the *z* coordinate for the reference white and each of the RGB primaries:

$$z = 1 - (x + y)$$

3.3.3 Form the following matrix and column vector from the x y z numerical values of the reference primaries and white:

$$\mathrm{P} = \begin{bmatrix} x_r & x_g & x_b \\ y_r & y_g & y_b \\ z_r & z_g & z_b \end{bmatrix} \qquad \mathrm{W} = \begin{bmatrix} x_w / y_x \\ 1 \\ z_w / y_w \end{bmatrix}$$

Note that the W vector, representing the reference white, has been normalized so that white has a luminance factor of 1.0; i.e., Y = 1.0. This is necessary so as to cause the video reference white signal (R=G=B=1) to produce the reference white with a unity luminance factor.

3.3.4 Compute the coefficients C_i on the left side of the equation below by multiplying the W vector by the inverse of the P matrix. Note the notation P^{-1} indicates the matrix inversion operation. These coefficients are normalization factors which normalize the units of the RGB primaries such that a unit amount of each combine to produce the white point chromaticities with a luminance factor of 1:

$$\begin{bmatrix} C_R \\ C_G \\ C_B \end{bmatrix} = \mathrm{P}^{-1} * \mathrm{W}$$

3.3.5 Form the diagonal matrix from the coefficients C_i computed in 3.3.4:

$$\mathrm{C} = \begin{bmatrix} C_R & 0 & 0 \\ 0 & C_G & 0 \\ 0 & 0 & C_B \end{bmatrix}$$

3.3.6 Compute the final normalized primary matrix NPM as the product of the P and C matrices:

$$\text{NPM} = \begin{bmatrix} X_R & X_G & X_B \\ Y_R & Y_G & Y_B \\ Z_R & Z_G & Z_B \end{bmatrix}$$

3.3.7 This matrix, NPM, is the final result and relates television linear RGB signals to CIE XYZ tri-stimulus values as follows:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_R & X_G & X_B \\ Y_R & Y_G & Y_B \\ Z_R & Z_G & Z_B \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

3.3.8 The luminance equation for this set of primaries is the second row of the NPM matrix:

$$Y = Y_R(R) + Y_G(G) + Y_B(B)$$

In some cases, the NPM matrix values rounded to four digits may result in a luminance equation whose terms do not sum to 1.0. In that situation, the NPM matrix should be column normalized to force the second row to sum to 1.0.

3.3.9 Computations of color-difference signal coefficients should use all 10 digits of the luminance equation as determined above. These data should be multiplied by applicable scaling factors before rounding. Round to four decimal places and/or four digits, whichever extends the number further.

In some cases, the coefficients of the color-difference equations may not sum to zero after rounding. In that situation, the coefficients should be renormalized to force the coefficients of each equation to sum to zero.

Annex B (informative)

Example derivation of normalized primary matrix

B.1 Given the reference white chromaticities:

$$x = 0.3127$$
$$y = 0.3290$$

and a set of reference primaries:

$$x_R = 0.640 \qquad y_R = 0.330$$
$$x_G = 0.300 \qquad y_G = 0.600$$
$$x_B = 0.150 \qquad y_B = 0.060$$

B.2 The following values of C_i are derived:

$$C_R = 0.6443606239$$
$$C_G = 1.1919477979$$
$$C_B = 1.2032052560$$

B.3 The values for the NPM matrix before rounding to four digits are:

$$\text{NPM} = \begin{bmatrix} 0.4123907993 & 0.3575843394 & 0.1804807884 \\ 0.2126390059 & 0.7151686788 & 0.0721923154 \\ 0.0193308187 & 0.1191947798 & 0.9505321522 \end{bmatrix}$$

B.4 The luminance equation is:

$$Y = 0.2126390059(R) + 0.7151686788(G) + 0.0721923154(B)$$

and rounded to four digits:

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

in which the coefficients sum to 1.0.

OmniTek
Intec 2 Unit 3, Wade Road, Basingstoke, Hants RG24 8NE, UK
Tel: +44 (0)1256 345900 Fax: +44 (0)1256 345901
Email: support@omnitek.tv Web: www.omnitek.tv