Reproducing and synthesizing colour in computer graphics

GARY W MEYER

The use of colour display devices in computer graphics is explored. First, some principles of colour science are reviewed. Next, *RGB* and NTSC monitors, the two most common colour reproduction media used in computer graphics, are discussed. Finally, some of the colour synthesis problems encountered in reproducing colour collections, doing proper anti-aliasing, and making synthetic images are addressed.

Keywords: colour display devices, computer graphics, colour science, colour monitors, anti-aliasing, synthetic images

Computer graphics provides an excellent mechanism for exploring the colour reproduction characteristics of typical display devices and for investigating the properties of colour spaces that are used to solve colour reproduction problems on these display units. The equipment employed in computer graphics allows precise spatial and intensity control of output devices such as CRTs and film recorders. The trichromatic nature of human colour vision leads to three-dimensional colour spaces for describing colour sensations, and computer graphic algorithms for rendering three-dimensional objects can be used to depict these colour spaces. This paper explores the role that computer graphics can play in both synthesizing and reproducing colour on display devices.

A typical (but by no means completely general) computer graphics colour reproduction environment is shown in Figure 1. The CIE *XYZ* colour notation system is the switch point through which colour input and output mechanisms communicate with one another. The diagram shows four separate video frame stores (or frame buffers). These may in fact be the same physical device, but for performing colour space transformations it is conceptually a different device depending on what video equipment is attached to it.

The next section in this paper reviews some colour science concepts that are important for performing colour reproduction. The two sections that follow after that describe several of the colour transformation paths in Figure 1 in greater detail. On the output side of the CIE *XYZ* switch point, *RGB* monitors and NTSC encoders are considered. Synthetic imaging and standard colour collections will be discussed as producers of CIE *XYZ* tristimulus values. Where appropriate, efficient algorithms will be given for bypassing the CIE *XYZ* transformation and performing the colour synthesis calculations directly in the primary system of the target colour reproduction medium.

Department of Computer and Information Science, University of Oregon, Eugene, Oregon 97403, USA

PRINCIPLES OF COLOUR REPRODUCTION

The observer's frustrum of vision defines the solid angle through which the electromagnetic energy from the environment is received. A picture plane oriented perpendicular to the observer's line of sight is pierced by the rays of light which compose the frustrum. Colour reproduction techniques seek to create a colour sensation at each point of the reproduction which is identical to the colour sensation produced by the electromagnetic energy distribution $E(\lambda)$ of the ray which intersects the picture plane at the same relative location. Although this technique reproduces the colour at each point independently, the colour of each point must also look correct when viewed in the context of all the points which surround it.

What is known as 'exact colour reproduction'[1] can be

0141—9382/89/030161–10 $03.00

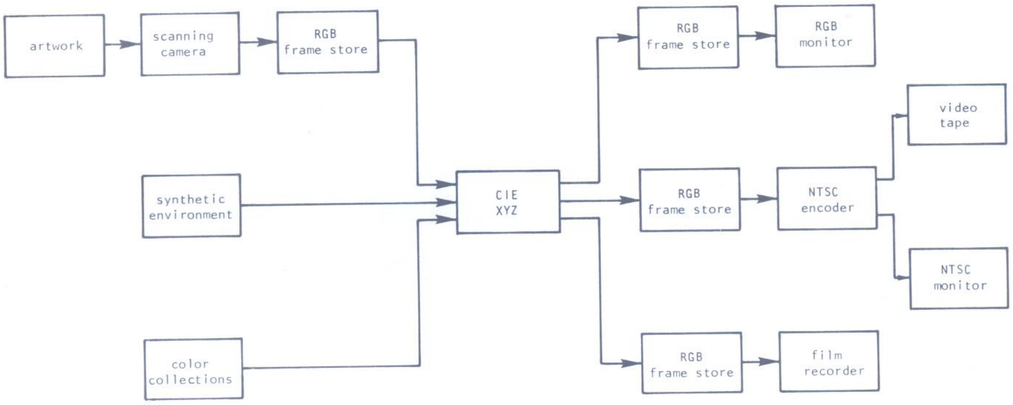

Figure 1. Typical computer graphics colour reproduction environment

obtained by using the principles of colour science. The absolute spectral energy distribution $E(\lambda)$ (W cm^{-2} sr^{-1} nm^{-1}) at each point of the picture plane is resolved to its CIE XYZ tristimulus values by use of the relations:

$$\begin{aligned} X &= C \int E(\lambda)\bar{x}(\lambda)\mathrm{d}\lambda \\ Y &= C \int E(\lambda)\bar{y}(\lambda)\mathrm{d}\lambda \qquad (1) \\ Z &= C \int E(\lambda)\bar{z}(\lambda)\mathrm{d}\lambda \end{aligned}$$

where C is equal to 683 lm W^{-1} cm^2 sr. These tristimulus values can be expressed in terms of RGB reproduction primaries by means of a linear transformation:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} \quad M \quad \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad (2)$$

where the elements of the matrix M can be determined given the colorimetric properties of the RGB primaries expressed in terms of the CIE XYZ system.[2] A detailed derivation of this transformation is given in the Appendix. Each point of the reproduction is then created by mixing together the proper amount of light from each primary. The composite spectral energy distribution created by additively mixing RGB primaries can also be produced by using three filters to selectively subtract energy from an equal energy distribution. No matter which approach is taken, the match obtained by considering each point isolated from the rest of the image continues to hold when the point is replaced into the context of the picture.

To scale the results of 'exact colour reproduction' so that they fit within the dynamic range of existing colour reproduction devices, 'colourimetric colour reproduction' is necessary. Assume that the primary source of illumination in the scene is provided by a light source with spectral energy distribution $H(\lambda)$. This light reflects off surfaces having spectral reflectance $\rho(\lambda)$ (ignoring transparency) to yield spectral energy distributions $E(\lambda) = \rho(\lambda)H(\lambda)$ at the picture plane. The tristimulus values are found by using Equation (1) with:

$$C = \frac{100}{\int H(\lambda)\bar{y}(\lambda)\mathrm{d}\lambda} \qquad (3)$$

This ranks the luminance of all objects in the scene relative to the brightness object (a perfectly reflecting surface) which has luminance of 100. Because the chromaticity remains constant no matter how C is selected, it is often useful to express this tristimulus colour specification as chromaticity (x,y) and relative luminance Y. The amount of each RGB reproduction primary necessary to produce this colour is found through the use of Equation (2). To take full advantage of the colour reproduction device's gamut the units of the RGB primaries are arranged so that $R = G = B$ = maximum corresponds to the XYZ tristimulus values produced by the perfectly reflecting surface.

Of the above two types of colour reproduction, 'colourimetric colour reproduction' is the most relevant for computer graphics work. The basic notation for expressing colour is then the CIE XYZ tristimulus values, with Y representing the luminance relative to the luminance of a perfect reflecting or transmitting surface in the scene. Spectral energy distributions are expressed in the CIE XYZ notation system using Equations (1) and (3), and the primaries of various colour reproduction devices are reached via Equation (2).

COLOUR REPRODUCTION IN COMPUTER GRAPHICS

Figure 1 shows three of the major computer graphics colour reproduction media. The simplest is an RGB monitor, the colorimetric properties of which are determined by its phosphors and the relative gains of its three guns. An NTSC encoder, which creates a signal for a tape recorder or an NTSC monitor, has a fixed set of primaries in which its inputs are to be expressed. Film is the third media shown in Figure 1 and is by far the most complex. The RGB values used to expose the film are produced by a nonlinear transformation from CIE XYZ space. The silm recorder may be driven by a video signal as shown in the diagram or it may be fed discrete values for each position on the film plane.

This section discusses the colorimetry of *RGB* monitors and the NTSC encoding scheme.

RGB monitor

A colour television monitor is an additive colour reproduction system which employs three independent primaries to generate colour. Unlike the colour matching experiments where the light beams from the three primaries were all coincident, in a colour television tube there are discreet phosphor dots for each primary. The phosphor dots are deposited as triads spaced 0.6 mm apart for coarse resolution and 0.3 mm apart for high resolution. From the normal viewing distance, the observer's visual system is unable to resolve the dots and the light from each triad is merged by the time is reaches the surface of the retina.

Given the chromaticity coordinates of a monitor's phosphors and the chromaticity and luminance of the monitors white point, the transformation between CIE *XYZ* space and the *RGB* space of the monitor can be determined.[2] The rays in CIE *XYZ* space which correspond to the *RGB* primaries of the monitor are determined by the chromaticity coordinates of the monitor phosphors. The three sides of the monitor gamut closest to the origin are determined by the planes which are parallel to the three different pairs of these three rays. The absolute position of the monitor gamut sides closest to the origin is determined by the point on each primary ray which corresponds to the tristimulus value produced when the primary is at its minimum setting. The three sides of the gamut furthest from the origin must each be parallel to one of the planes containing the primaries. The absolute position of the planes furthest from the origin is determined by the point on each primary ray which corresponds to the tristimulus value produced when the primary is at its maximum setting. The position and shape of a typical RGB monitor gamut in CIE *XYZ* space is shown in Figure 2.

'Balancing' a monitor involves adjusting the minimum and maximum settings of the primaries so that a consistent chromaticity is produced whenever equal amounts of the primaries are requested. Because the chromaticity coordinates of each phosphor are known, the chromaticity and luminance (x,y,Y) and minimal points on each ray can be determiued by using a photometer to measure the luminance produced by the display. However, low-intensity luminance measurements may be difficult to make, and some error is introduced by measuring each primary while the other two are turned off. Alternatively, one can use a colour comparator to adjust the maximum and minimum setting of each primary until a 'white' colour of known chromaticity and luminance is matched. A colorimeter which measures chromaticity and luminance can also be used for this purpose.

The above procedure correctly defines the boundary of the monitor gamut but does not account for the nonlinear relationship between the signal sent to the monitor and the luminance produced by the display. Asking for half the maximum signal level does not produce the tristimulus value which lies half the distance between zero and the maximum tristimulus value for that primary. Since the chromaticity coordinates for the primary are known, one can determine the tristimulus value which *is* produced by measuring the luminance with a photometer. Experiments show that a power-law relationship exists for CRTs between the signal voltages R_v, G_v, B_v and the R, G, B position that they generate along the appropriate ray in CIE *XYZ* space. This can be expressed as:

$$\begin{aligned} R_v &= R^{1/\gamma_r} \\ G_v &= R^{1/\gamma_g} \\ B_v &= R^{1/\gamma_b} \end{aligned} \tag{4}$$

where γ_r, γ_g, and γ_b are usually between 2.0 and 3.0. γ_r, γ_g, and γ_b often have close to identical values and are therefore referred to collectively as the 'gamma' of the monitor. Some issues to consider when determining the 'gamma' of a monitor are discussed by Cowan.[3] The nonlinear distribution of CIE *XYZ* tristimulus values produced by linearly varying R_v, G_v, B_v values is shown in Figure 3.

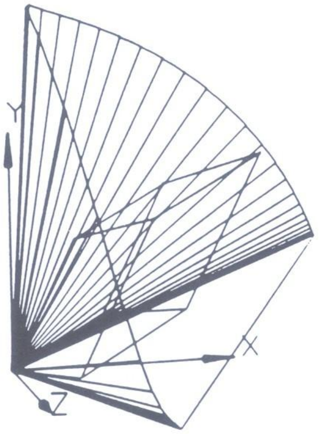

Figure 2. Monitor gamut and cone of realizable colour in CIE XYZ *space*

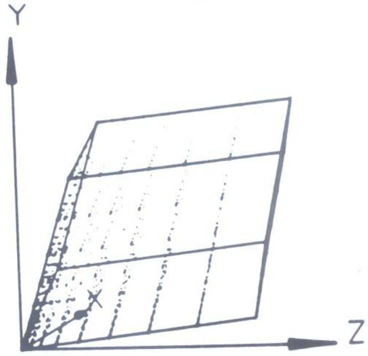

Figure 3. Nonlinear distribution of tristimulus values caused by linear variation in signal voltage

Figure 4 shows a typical computer graphics equipment configuration and how the above colour space transformations might be accomplished in this environment. The CIE *XYZ* tristimulus values are linearly transformed into linear luminance *RGB* values based on the phosphor chromaticities and the white point setting of the monitor connected to the frame store. These values must be adjusted to compensate for the nonlinear characteristics of the CRT. This can be done before the values are loaded into the frame store, or the linear luminance *RGB* values can be loaded directly into the frame store with the correction accomplished by a look-up table which lies between the main memory of the frame store and the digital-to-analogue converters which produce the monitor signal voltage. The bit depth of the look-up tables must be greater than the bit depth of the main memory if the frame store's effective bit depth is to match its actual bit depth. For example, a 256 position 8 bit look-up table wil have approximately 173 unique values stored in it when loaded to compensate for a monitor with gamma of 2.5. A frame store memory with 8 bit per storage location indexing into this 256 position table can come away with only 173 unique values. This means that the effective bit depth of the frame store is less than the actual bit depth.

NTSC encoding

The National Televison Systems Committee (NTSC) encoding scheme was designed to produce a colour television signal suitable for broadcast. The major objectives were to minimize bandwidth and to achieve compatibility with black-and-white television receivers. This section will discuss how the transmission primaries Y_t, *I*, and *Q* of the system were selected, the relative position of the transmission primaries and *RGB* receiver primaries in CIE *XYZ* space, and the colour space transformations necessary to convert between these primary systems.

To handle monochrome broadcasts where only the luminance at the picture plane is known, the NTSC decided that only the colour transmission primaries, Y_t, should have a numerical value proportional to luminance. The transmission primaries *I* and *Q* are therefore zero for a monochrome broadcast with all of the information carried by the Y_t primary. A black-and-white receiver uses the Y_t portion of the signal to modulate the intensity of its single fixed chromaticity phosphor. The Y_t was also assigned chromaticity coordinates of $x = 0.310$, $y = 0.316$ by the NTSC. The position of this primary in CIE *XYZ* space is clearly distinct from the *Y* axis of the space and explains the use of the symbol Y_t (for $Y_{\text{transmission}}$) in this article instead of the symbol *y* proposed by the NTSC (see Figure 5). A colour television receiving a monochrome signal modulates the intensities of its three primaries so as to produce CIE *XYZ* tristimulus values which lie along the Y_t locus in CIE *XYZ* space.

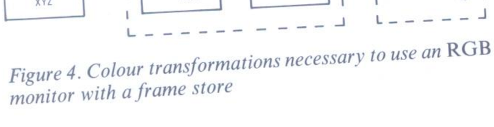

Figure 4. Colour transformations necessary to use an RGB *monitor with a frame store*

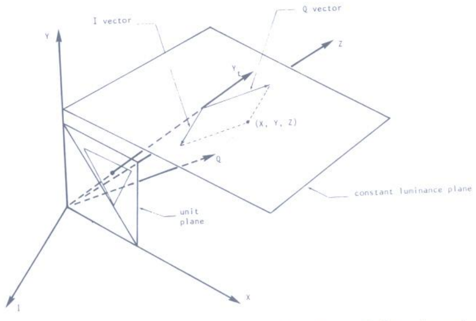

Figure 5. Vector addition of the Y_t, I, *and* Q *primaries in CIE* XYZ *space*

To transmit colour broadcasts where all three CIE *XYZ* tristimulus values at the picture plane are known, the NTSC introduced the *I* and *Q* primaries. Because the luminance is completely determined by the Y_t primary, the *I* and *Q* primaries must lie in the *XZ* plane of CIE *XYZ* space with directions determined by factors to be discussed below. To reach a CIE *XYZ* tristimulus value by vector addition of the Y_t, *I*, and *Q* primaries, one first reaches the proper constant *Y* plane by using the Y_t primary and then moves to the proper location on that plane using the *I* and *Q* primaries (see Figure 5). A colour television receiver uses its three primaries to produce the CIE *XYZ* tristimulus value indicated by the Y_t, *I* and *Q* signals. A black-and-white television receiving a colour signal ignores the *I* and *Q* portions and matches the brightness of the colour by modulating the luminance of its single phosphor according to Y_t.

The direction of the *I* and *Q* axes on the *XZ* plane were selected, in part, to minimize the bandwidth necessary to transmit these two primaries. Roughly speaking, the bandwidth of the signal determines the number of pictures elements that can be placed in a unit horizontal or vertical distance. The objective is to select a bandwidth which produces no more picture elements than the eye can resolve. Colour Plate 1 demonstrates that the bandwidth requirements for the *I* and *Q* axes finally selected are less than those for the Y_t axis. The final bandwidth requirements were determined not by the method in Colour Plate 1 but by viewing complex scenes[4].

In formally stating the CIE *XYZ* to NTSC Y_t *IQ* transformation, a colour television receiver with specific colorimetric properties is assumed. Given the chromaticity coordinates of that receiver's phosphors and the chromaticity and luminance of that receiver's white point, the transformation between CIE *XYZ* space and $Y_t IQ$ space can be determined.[5] The position of the RGB_{NTSC} monitor in Y_t *IQ* space is shown in Figure 6.

Figure 7 shows how the above linear colour space transformations can be accomplished using a video frame store, an NTSC encoder, and a hypothetical NTSC monitor. The CIE *XYZ* tristimulus values are matrixed

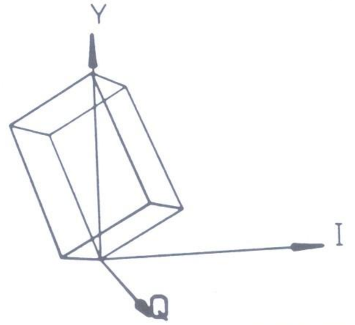

Figure 6. Typical monitor gamut in Y_t*TQ space*

into the RGB_{NTSC} primaries and loaded directly into the frame store with no provision made to accommodate the gamma of the receiver. The output of the frame store is passed to an NTSC encoder which performs both the matrixing to $Y_t\,IQ$ primaries and the bandwidth limitation. The ideal NTSC monitor matrixes the $Y_t\,IQ$ primaries back to RGB_{NTSC} primaries, corrects for the gamma of the monitor, and applies the proper voltage to a CRT which has NTSC phosphors and is adjusted to the proper white point.

Unfortunately, this colorimetrically perfect situation is impossible to obtain. The NTSC decided to centralize the gamma correction circuitry in the broadcast stations rather than incorporate it in every home television receiver. This means that the RGB_{NTSC} primaries must be adjusted as follows

$$\begin{aligned} R_{volt} &= R^{1/\gamma} \\ G_{volt} &= G^{1/\gamma} \\ B_{volt} &= B^{1/\gamma} \end{aligned} \qquad (5)$$

before being matrixed to the Y_tIQ primaries. A γ of 2.2 is typically used to compensate for both the nonlinearities of the CRT and the effect of viewing conditions.[1] If it were not for the difference in bandwidth between the Y_t, I, and Q channels, the sequence of gamma correcting from RGB_{NTSC} to $RGB_{NTSC_{volt}}$, matrixing from $RGB_{NTSC_{volt}}$ to Y_tIQ, and then matrixing from Y_tIQ to $RGB_{NTSC_{volt}}$ would cause no errors in the colorimetry of the reproduced image. However, the nonlinear gamma correction means that some luminance information is present in the I and Q signals and is therefore subjected to the stronger bandwidth limiting of these channels. This causes a distortion of the colours for fine area detail. In addition to the problems caused by gamma correcting before transmitting, modern television receivers use phosphors different from the NTSC phosphors, which further complicates the colorimetry of the system.

Figure 8 shows how practical NTSC encoding is accomplished. The major difference between this and Figure 7 is that the gamma correction is done in the frame store before $Y_t\,IQ$ encoding takes place. It should be emphasized that many NTSC monitors matrix from the $Y_t\,IQ$ primaries to the $RGB_{NTSC_{volt}}$ primaries instead of the primaries for the CRT phosphors actually used.[6] Also note that the RGB values in the frame store often need to be adjusted in order to avoid NTSC signals with video amplitudes that exceed the permitted voltage range.[7]

Colour Plates 2, 3 and 4 demonstrate some of the concepts from the preceding paragraphs.

In Colour Plate 2, the Y_t, Y_tI, and $Y_t\,Q$ portions of a $Y_t\,IQ$ signal are shown next to the full colour version. A demonstration of the relative resolution in the Y_t, I, and Q channels is shown in Colour Plate 3. This image was created by fast Fourier transforming an original monochrome image, filtering the transformed image with filters that approximate the effect of the Y_t, I, and Q band limitation, and then inverse fast Fourier transforming the image.[8] Colour Plate 4 shows the successive degradation of a full colour image until each of the Y_t, I, and Q channels has reached its full NTSC band limitation.

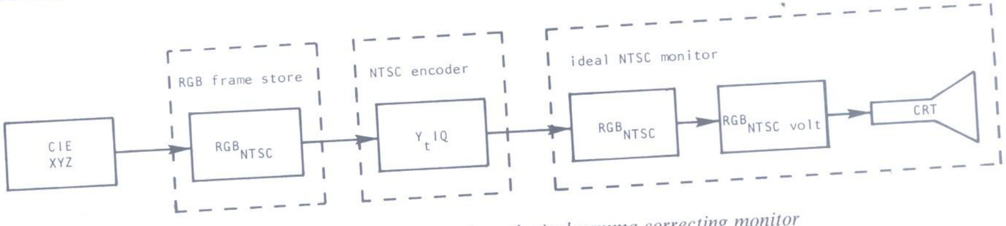

Figure 7. NTSC encoding from a frame store using a hypothetical gamma correcting monitor

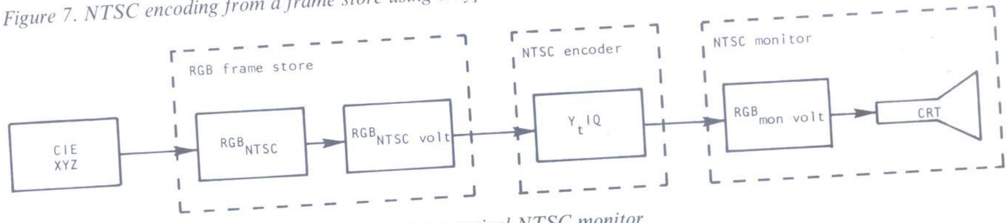

Figure 8. NTSC encoding from a frame store using a typical NTSC monitor

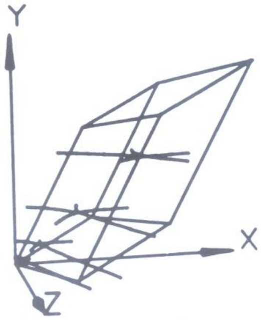

Figure 9. 'Spiders' of constant Munsell value and their position relative to the monitor gamut

COLOUR SYNTHESIS IN COMPUTER GRAPHICS

Figure 1 shows the three principle means by which colour specifications are generated in computer graphics. A colour may be selected from a colour collection (like the *Munsell Book of Colour*) for which the CIE XYZ tristimulus values are already known. A synthetic imaging scheme may yield a spectral energy distribution which can be converted to CIE XYZ tristimulus values using Equations (1) and (3). Finally, the spectral energy distribution leaving the surface of a piece of artwork can be passed through filters with transmittance based on the standard observer matching functions and the resultant spectral energy distribution integrated by a television camera to yield (after matrixing) the CIE XYZ tristimulus values. In all three cases, simplifications are possible so that the colour specification can be immediately expressed in terms of the RGB primaries of a specific colour reproduction device. This section discusses the colour synthesis problems involved in reproducing colour collections, anti-aliasing, and synthetic imaging.

Reproducing of colour collections

Colour collections, of which hardware-store paint swatches are a familiar example, have been developed for a variety of reasons. Many simply serve as palettes from which colour selections are made. Some exhibit 'perceptual uniformity' which makes them useful in evaluating colour differences. Other contain colour comparison standards for sorting and grading materials like fabric and soil samples.

If a colour collection is expressed in terms of CIE XYZ tristimulus values, then it can be reproduced on a digitally controlled colour television monitor. An example of such a collection is the *Munsell Book of Colour*. The Munsell system is a perceptually uniform colour organization scheme expressed in cylindrical coordinates with dimension hue (angle), chroma (radius), and value (height). In the early 1940s, the CIE tristimulus values for the *Munsell Book of Colour* were determined in terms of CIE illuminant C.[9] Figure 9 shows 'spiders' of constant Munsell value and their position within a typical colour television monitor gamut. Each spider arm has constant hue and the dots on the arm represent equal chroma spacings. In Colour Plate 5 the CIE XYZ definition of the Munsell colour solid has been used to produce a computer-generated picture of it. Additional details concerning the reproduction of this and other perceptually uniform colour collections can be found in Reference 10.

A number of hue, saturation, and brightness colour specification systems have been proposed which are approximations of the Munsell system (saturation and brightness being used as synonyms for chroma and value, respectively) and which transform efficiently into the RGB primaries of an additive colour reproduction system. The position of the constant value Munsell 'spiders' in the RGB gamut of a typical additive colour reproduction device is shown in Figure 10. If the cube is placed on end so that the neutral axis through the centre of the spiders is vertical, then the points in the cube can be referenced using the cylindrical coordinates of the Munsell system. Brightness (or value) corresponds to height, saturation (or chroma) is the radial distance from the neutral axis, and hue is the angular position around the neutral axis. Each of the proposed hue, saturation, and brightness transformations of the RGB gamut incorporates portions of this geometric relationship. Colour Plate 6 depicts the hexcone system described

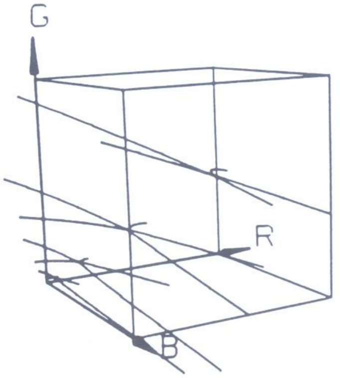

Figure 10. Constant value Munsell 'spiders' in the RGB gamut of an additive colour reproduction device

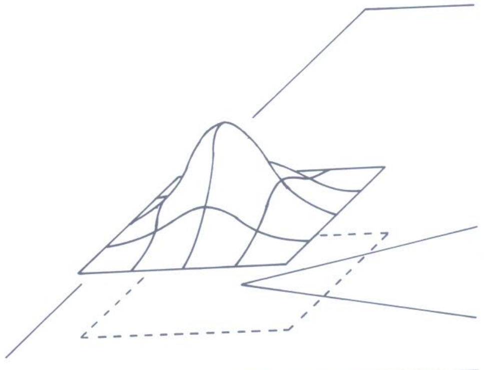

Figure 11. Gaussian shaped convolution mask being used to sample polygonal environment

by Smith.[11] Other examples of such transformations can be found in the literature.[12–15]

Anti-aliasing

The continuous version of a computer graphics image must be sampled so that it can be loaded into a digital frame store. Each sample should contain some average representation of the area surrounding it so that when the image is reconstructed from the discrete samples the best approximation to the original continuous version is obtained. This section addresses the colour mixing problem inherent in this operation.

Because perfectly square corners are possible in a mathematically defined computer graphics environment, the Fourier transform of an image to be sampled is usually rich in high-frequency components. This means that even if the reproduction device is capable of perfect reconstruction, it is practically impossible to sample an image with a grid fine enough to satisfy the sampling theorem and avoid aliasing. Hence, one is forced to low-pass filter the image in such a way as to reduce the sample spacing required and to produce the best approximation possible to the original image when reconstructed by the frame store and display device. This low-pass filter can be inverse transformed from the frequency domain to the spatial domain and convolved with the image to obtain the same result directly in the spatial domain. Because the subsequent sampling will only occur at discrete locations in the image, the convolution need only be performed at these points. Thus, the convolution and sampling can be performed together in one step.

The optimum shape and size for the convolution mask to be used at each sample point has been the topic of much research in computer graphics.[16–19] Suggested cross sectional shapes include rectangular, gaussian, and trapezoidal. More complicated filter shapes have also been proposed based on optimization techniques which take the shape of the CRT spot into account.[20] To avoid aliasing, adjacent convolution masks must touch one another and in most cases must be wide enough to overlap to their neighbour's centreline. A gaussian-shaped convolution mask is shown in Figure 11.

To perform the convolution, a weighted average is taken of the tristimulus values for the image plane polygons which lie under the convolution mask. The resultant tristimulus coordinates constitute the sample value for that location. This process can be visualized as vector addition in an additive colour space. The colour vectors determined by the tristimulus values for each polygon are first scaled according to how much of the convolution mask's volume lies above each polygon. (The total volume of the convolution mask is assumed to be unity.) Then vector addition for all the component polygons yields the resultant tristimulus values. Colour Plate 7 shows a pixel that contains pieces of several polygons and Colour Plate 8 shows the colour that results after the above colour mixing calculation has been performed.

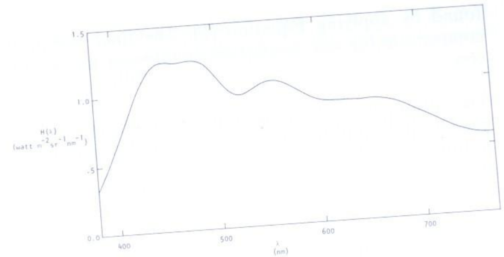

Figure 12. Absolute spectral energy distribution of light source

The coordinate system in which the vector addition is performed must exhibit a linear variation of luminance throughout. In the case of colour television, the most general approach would be to perform the vector addition in CIE *XYZ* space, matrix to the *RGB* primaries of the monitor, and then compensate for the nonlinear relation between the *RGB* primaries and the signal voltages necessary to reproduce the primaries at the surface of the display. Because the transformation from CIE *XYZ* space to the *RGB* primaries is linear, the vector addition could be done directly in the linear luminance *RGB* primary system of the monitor. No matter what colour reproduction medium is used, the final step of compensating for any nonlinearities in the process is crucial if satisfactory results are to be achieved.[21]

There are two assumptions in this approach:

- This integration process yields the same sensation as was produced during the colour matching experiments while viewing overlapping coloured lights on a white screen.
- Since the discrete phosphor dots in colour television behave in this manner, these assumptions seem to be well justified.

Application to synthetic imaging

The principles of tristimulus colorimetry have direct application to synthetic image-generation problems in computer graphics. After optical phenomena have been modelled, colorimetry can be used to generate a display for an appropriate colour reproduction technology. In the most general case, one computes the spectral energy distribution and then converts it to the signals necessary to drive the display device. Sometimes, this process can be simplified and the calculations can be performed directly in the primary system of the reproduction medium. This section covers the general case of working with spectral energy distributions and the simple image synthesis problems of diffuse reflection.

Spectral energy distributions

In the most general approach to image synthesis, spectral energy distributions are detemined by modelling optical phenomena and the laws of tristimulus colorimetry are used to convert the spectral energy distributions to the signals necessary to drive a display device. The first step in converting the spectral energy distribution is to determine the CIE *XYZ* tristimulus values by using Equations (1) and (3). If colour television is the final medium, the amount of each reproduction primary necessary is found by applying Equation (2). The final step is to compensate for any nonlinearities inherent in the process.

The colour of a perfect radiation emitter, known as a blackbody[22], will be determined as an example of this process. Using quantum arguments, Max Planck has shown that the spectral energy distribution for any blackbody radiator is given by:

$$E(\lambda) = \frac{2\pi C_1}{\lambda^5(e^{C_2/\lambda T} - 1)} \tag{6}$$

where λ is wavelength and T is absolute temperature. The constants C_1 and C_2 have the values 0.59544×10^8 Wμm^4/m^2 and 14 388 μmK, respectively. Note that the emitted radiation only becomes noticeable to the human eye when the temperature exceeds 977°F and the emitted radiation enters the visible range.

The colour scale which results from blackbody radiation at temperatures between 2 200 and 6 000 °K is shown in Colour Plate 9. The spectral energy distribution for a blackbody radiator at these temperatures was found from Equation (6), the signals necessary to drive a colour television monitor were determined using the process described above, and a photograph was taken of the resulting display. The colour scale starts at 2 200 °K because this is the lowest temperature which yields chromaticities that lie within the gamut of the monitor.

Besides being limited by the gamut of chromaticities which the monitor can produce, the reproduction of blackbody radiation colour is also limited by the maximum amount of light energy which the monitor can generate. All of the colours in Colour Plate 9 were scaled so that the colour for the highest temperature was produced at the maximum monitor intensity. The result is that the colours for low temperatures become very dark, making the perception of their hue quite difficult. Colour Plate 10, which is the colour for the blackbody at 2 200 °K scaled to the maximum intensity at which the monitor can produce the colour, demonstrates that if the monitor had more dynamic range the colour scale in Colour Plate 9 would look quite different.

Diffuse reflection

Consider a planar polygonal object illuminated by a light source with absolute spectral energy distribution $H(\lambda)$ shown in Figure 12. If the surface scatters an incident ray of light equally in all directions, then it is said to reflect diffusely and to satisfy Lambert's (cosine) law. The absolute spectral energy distribution $E(\lambda)$ leaving a facet is then determined from the expression $E(\lambda) = \cos(\theta)H(\lambda)\rho(\lambda)$ where $\rho(\lambda)$ is the spectral reflectance of the surface (see Figure 13) and θ is the angle between the surface normal and a ray from the light source to the surface. The absolute spectral energy distribution for facets with several different orientations is shown in Figure 14.

Equations (1) and (3) are used to find the CIE *XYZ* tristimulus values for each facet. It can be seen that once the tristimulus values X_0 Y_0 and Z_0 for the surface where $\theta = 0$ have been calculated, the tristimulus values for a facet with arbitrary orientation θ can be found from the expressions:

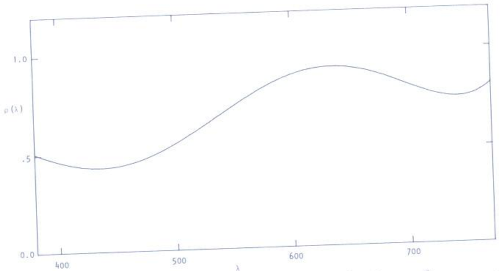

Figure 13. Spectral reflectance distribution of opaque material

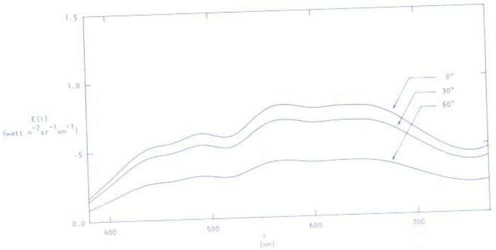

Figure 14. Absolute spectral energy distributions for facets with different orientations

$$\begin{aligned} X &= \cos(\theta)X_0 \\ Y &= \cos(\theta)Y_0 \\ Z &= \cos(\theta)Z_0 \end{aligned} \tag{7}$$

The locus of these points is a ray from the origin to the point (X_0, Y_0, Z_0). Thus, the chromaticity coordinates for all facets of a diffusely reflecting object will be the same.

If the reproduction medium is colour television, the *RGB* tristimulus values can be found by using Equation (2). Since this is a linear transformation, once R_0, G_0 and B_0 are found from X_0, Y_0 and Z_0, the *RGB* tristimulus values for the facets can be found from:

$$\begin{aligned} R &= \cos(\theta)R_0 \\ G &= \cos(\theta)G_0 \\ B &= \cos(\theta)B_0 \end{aligned} \tag{8}$$

In fact, in many synthetic imaging applications one simply specifies the colour of an object by selecting an R_0, G_0, and B_0 (or X_0, Y_0, and Z_0 as was the case in Colour Plate 5) without considering spectral energy distributions. For all cases, the *RGB* tristimulus values from Equation (2) or Equation (8) must be adjusted to compensate for any nonlinearities in the display device.

Complex illumination models

This section has taken a brief look at the problem of modelling optical phenomena on a wavelength-by-wavelength basis. A blackbody radiator was offered as an example of first using the laws of physics to synthesize a spectral energy distribution and then determining the tristimulus values necessary to create the equivalent colour sensation on a colour television monitor. This result could be useful in portraying an object at an elevated temperature. Simple diffuse reflection was given as an example of a reflection model that is based on real physical principles yet can be expressed directly in terms of the primaries of a specific colour reproduction device.

Local-reflection models of far greater complexity than the above two examples have been developed and applied to synthetic image generation. Examples of these models and the results that can be obtained using them can be found in Cook[23], Hall[24]. As the complexity of a model increases, so does the computational expense of using it. One way of decreasing the amount of calculation that is necessary is to apply the reflection model at as few wavelengths as possible across the visible spectrum. Meyer[25] provides one approach to solving this problem.

CONCLUSIONS

Computer graphics can be a significant tool for those concerned with the design and use of colour display devices. A digital frame store connected to a CRT (or other output hardware) provides a precise way of controlling both the spatial and intensity characteristics of the device. In this paper, the use of both *RGB* and NTSC monitors as computer graphic output devices was investigated. Computer graphic algorithms for creating pictures of three-dimensional objects can be used to both explore colour space and synthesize colour in realistic images. In this paper, a representative colour space (the Munsell colour solid) was reproduced, spectral energy distributions were calculated, and a simple illumination model was described. Colour display devices are an essential piece of computer graphic hardware, and computer graphics can play an important role in the design and use of these devices.

APPENDIX

In this appendix, the terms of the matrix M in Equation (2) are derived. This transformation of primaries is a standard part of many colorimetric calculations.[2]

The transformation between the *RGB* colour space of one set of colour reproduction primaries and 1931 CIE *XYZ* space can be expressed as:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} C_r x_r & C_g x_g & C_b x_b \\ C_r y_r & C_g y_g & C_b y_b \\ C_r z_r & C_g z_g & C_b z_b \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (9)$$

where R, G and B are each on a range from 0.0 to 1.0 and represent the coordinates of the colour in the monitor's *RGB* colour space; X, Y, and Z are the coordinates of the colour in 1931 CIE *XYZ* space; and x_r, x_g, x_b, y_r, y_g, y_b, z_r, z_g and z_b are the chromaticity coordinates of the *RGB* primaries in CIE *XYZ* space. Given the luminance Y_w and chromaticity coordinates x_w and y_w of the colour produced when $R = G = B = 1$, the constants C_r, C_g, and C_b in the above expression can be found from the relations:

$$C_r = \frac{Y_w}{y_w} \frac{x_w(y_g - y_b) - y_w(x_g - x_b) - x_g y_b - x_b y_g}{D}$$

$$C_g = \frac{Y_w}{y_w} \frac{x_w(y_b - y_r) - y_w(x_b - x_r) - x_r y_b - x_b y_r}{D}$$

$$C_b = \frac{Y_w}{y_w} \frac{x_w(y_r - y_g) - y_w(x_r - x_g) - x_r y_g - x_g y_r}{D}$$

$$D = x_r(y_g - y_b) + x_g(y_b - y_r) + x_b(y_r - y_g)$$

The matrix M in Equation (2) can be found by inverting the matrix in Equation (9).

ACKNOWLEDGEMENTS

Some of the work presented here was performed at the Cornell University Program of Computer Graphics.

REFERENCES

1 **Hunt, R W G** *The reproduction of colour* John Wiley and Sons, New York, USA (1975)

2 **Judd, D B and Wyszecki, G** *Colour in business, science and industry* John Wley and Sons, New York, USA (1975)

3 **Cowan, W B** 'An inexpensive scheme for calibration of a colour monitor in terms of CIE standard coordinates' *Comput. Graph.* Vol 17 (1983) pp 315–321

4 **Fink, D G (Ed.)** *Colour television standards* McGraw-Hill, New York, USA (1955)

5 **Pearson, D E** *Transmission and display of pictoral information* John Wiley and Sons, New York, USA (1975)

6 **DeMarsh, L E** 'Colorimetric standards in U.S. color television' *J. SMPTE* Vol 83 (1974) p 1

7 **Odgers, C R** 'An introduction to video signal terminology, concepts and standards, and their influence on frame buffer design' *Siggraph 81 tutorial: Two-dimensional computer animation* (1981)

8 **Gonzalez, R C and Wintz, P** *Digital image processing* Addison-Wesley Publishing Co., Reading, UK (1977)

9 **Newhall, S M, Nickerson, D and Judd, D B** 'Final report of the OSA subcommittee on the spacing of the munsell colours' *J. Opt. Soc. Amer.* Vol 33 (1943) p 385

10 **Meyer, G W and Greenberg, D P** 'Perceptual colour spaces for computer graphics' in **Durett, H J (Ed.)** *Colour and the computer* Academic Press, Boston, USA (1987) pp 83–100

11 **Smith, A R** 'Colour gamut transform pairs' *Comput. Graph.* (1978) pp 12–19

12 **Davis, S and Metrick, L** 'An easy language for talking with colour machines' *Tekscope* Vol 10 (1978)

13 **Anon** 'Status report of the graphics standards committe' *Comput. Graphs.* Vol 13 (1979)

14 **Joblove, G H and Greenberg, D P** 'Colour spaces for computer graphcs' *Comput. Graph.* (1978) pp 20–25