A new method for the optimum generation of real colours on CRT monitors

This content has been downloaded from IOPscience. Please scroll down to see the full text.
1997 J. Opt. 28 107
(http://iopscience.iop.org/0150-536X/28/3/003)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 147.156.83.172
This content was downloaded on 23/09/2014 at 17:06

Please note that terms and conditions apply.
A new method for the optimum generation of real colours on CRT monitors

A Palomares†∥, A Lorente†, V Arnau†, F Martínez Verdú§ and J M Artigas†

† Depto de Óptica, Universitat de València, C/Dr Moliner, 50, Burjassot, 46100 València, Spain
∥ Instituto de Robótica, Universitat de València, C/Dr Moliner, 50, Burjassot, 46100 València, Spain
§ Depto d’Optica i Optometria, Universitat Politècnica de Catalunya, C/Violinista Vellòs, 37, Terrasa, 08222 Barcelona, Spain

Received 26 August 1996, accepted 3 December 1996

Abstract. The use of computers for colour generation is widely extended in colour research. Sometimes the reproduction of colours whose chromaticity coordinates correspond to a specific real standard is fundamental. In this paper we present a review of the capabilities and limitations of CRT monitors to generate colours. Finally, we propose a method for the optimum generation of real colours based on visual perception.

Keywords: CRT colour control, colour calibration, colour gamut, visual perception.

1. Introduction

The use of CRT monitors for colour generation is at present widespread in colour research. Colour is also a basic element for computer engineers, although it is often used as a secondary factor. Thus, software designers find the option that is easier to use, but mostly on the basis of intuition or aesthetics. The use of colorimetric equations is commonly avoided, hence ‘...a visually uniform colour space is usually assumed in theory, although the RGB colour model is often used in practice...’

However, sometimes the exact reproduction of certain colours is fundamental as in the case of colours whose chromaticity coordinates correspond to specific real standards. Here it is important to know the chromatic capabilities of the monitor being used. It is well known that monitors have a set of limitations in the generation of colours: there are technical and physical limitations, and limitations due to the use of digital cards.

These limitations may vary with the capacity of new devices. However, colorimetry is a tool to provide standardization. If the devices change (a likely result of technical improvements) the colorimetry needs to be explicit. The ‘XYZ’ representation is known as the ‘device-independent’ representation [1]. We can reproduce
the original colours only with additional knowledge of the device’s colorimetric characterization.

In this paper, from a strictly colorimetric viewpoint, we propose a simple and practical general method which will allow the computer engineers and colour researchers to know the capabilities of their devices for the reproduction of real colours optimized on the basis of visual perception.

2. Technical limitations: working range

The limitations of CRT monitors in the generation of colours are well known. Post and Caihou [2] evaluated different methods for generating colours on the screen. Artigas et al [3] analysed the dependency of the chromaticity coordinates of the three primaries of the monitor on the corresponding digital values (R, G, B) and the influence of screen area. Schuchard [4] studied this problem by introducing visual threshold data in an attempt to unify the scales. In any case, it is always necessary to calibrate the monitor [5–8] before hand to establish a working range. Berns et al [9–11] showed in an extensive and detailed paper the requirements to perform the colorimetric characterization of CRT displays.

3. Physical limitations: colour gamut

It is clear that the generation of colours on a CRT monitor is mainly restricted by the following factors.

(1) It is only possible to generate colours with coordinates (x, y) within the triangle defined by the chromaticity coordinates of the three primaries of the monitor (xR, yR), (xG, yG), (xB, yB) (figure 1).

(2) The maximum digital value that each of the guns can achieve gives the maximum possible luminance and it limits the colours that can be generated.

These factors define the CRT’s colour gamut (the full range of colours an imaging device can generate).

4. Digital cards’ limitations: quantization errors

The digital cards which control the CRT monitors, only allow colours defined by integer values of (R, G, B). The colour space we can generate is therefore discrete. This introduces a quantization error that we minimize on the basis of visual perception.

5. Colorimetric characterization of the screen: matrix ‘A’

The colorimetric characterization of the monitor is achieved by calculating the matrix ‘A’ which allows us to change from normalized digital values with respect to the white point (RN, GN, BN) to normalized tristimulus values with respect to the white-point luminance (XN, YN, ZN) [12]:

\[
\begin{pmatrix}
X_N \\
Y_N \\
Z_N
\end{pmatrix} = A \ast \begin{pmatrix}
R_N \\
G_N \\
B_N
\end{pmatrix},
\] (1)

The calculation of ‘A’ involves:

(1) obtaining the chromaticity coordinates of the three display phosphor primaries of the monitor: (xR, yR, zR), (xG, yG, zG), (xB, yB, zB), and that of the white point (x0, y0, z0) under a standard illuminant, by measuring them with a colorimeter;

(2) determining the scaling factors (aR, aG, aB). These factors convert the (xR, yR, zR), (xG, yG, zG) and (xB, yB, zB) values into trichromatic units. It is usual to convert the white-point values into tristimulus values, and normalize them with respect to the white-point luminance:

\[
a_R x_R + a_G x_G + a_B x_B = x_0/y_0
\]

\[
a_R y_R + a_G y_G + a_B y_B = 1
\] (2)

or in matrix form:

\[
\begin{pmatrix}
x_R & x_G & x_B \\
y_R & y_G & y_B \\
z_R & z_G & z_B
\end{pmatrix} \ast \begin{pmatrix}
a_R \\
a_G \\
a_B
\end{pmatrix} = \begin{pmatrix}
x_0/y_0 \\
1 \\
z_0/y_0
\end{pmatrix}.
\] (2')

The inverse of the matrix in equation (2') gives the relationship between the scaling factors (aR, aG, aB) and the normalized tristimulus values of the white point (x0/y0, 1, z0/y0):

\[
\begin{pmatrix}
a_R \\
a_G \\
a_B
\end{pmatrix} = \begin{pmatrix}
x_R & x_G & x_B \\
y_R & y_G & y_B \\
z_R & z_G & z_B
\end{pmatrix}^{-1} \ast \begin{pmatrix}
x_0/y_0 \\
1 \\
z_0/y_0
\end{pmatrix}.
\] (3)

If, in general form, we add RN normalized units of red, GN normalized units of green and BN normalized units of blue, the resulting (XN, YN, ZN) tristimulus values will be expressed as in the following equation:

\[
\begin{pmatrix}
a_R x_R & a_G x_G & a_B x_B \\
a_R y_R & a_G y_G & a_B y_B \\
a_R z_R & a_G z_G & a_B z_B
\end{pmatrix} \ast \begin{pmatrix}
R_N \\
G_N \\
B_N
\end{pmatrix} = \begin{pmatrix}
X_N \\
Y_N \\
Z_N
\end{pmatrix}
\] (4)
The tristimulus values of their tristimulus values, as follows:

\[
\begin{bmatrix}
R_N \\
G_N \\
B_N
\end{bmatrix} = \mathbf{A}^{-1} \begin{bmatrix}
X_0 \\
Y_0 \\
Z_0
\end{bmatrix}
\]

and the maximum normalized luminances for each primary:

\[
\begin{align*}
Y_{\text{Max} - R} &= a_R y_R \\
Y_{\text{Max} - G} &= a_G y_G \\
Y_{\text{Max} - B} &= a_B y_B.
\end{align*}
\]

6. The ability to generate the colour of coordinates \((x, y)\)

The next step is to analyse the mixtures of the monitor’s primaries to determine the digital values \((R, G, B)\) that yield a colour of certain chromaticity coordinates \((x, y)\). The tristimulus values \((X, Y, Z)\) corresponding to the mixture of three primary colours are obtained by the sum of their tristimulus values, as follows:

\[
\begin{align*}
X &= X_R + X_G + X_B \\
Y &= Y_R + Y_G + Y_B \\
Z &= Z_R + Z_G + Z_B.
\end{align*}
\]

Therefore, bearing in mind that the equations which relate the tristimulus values \((X, Y, Z)\) of a colour to its chromaticity coordinates \((x, y)\) are the following:

\[
\begin{align*}
x &= \frac{x}{y} \\
y &= \frac{y}{y} \\
z &= \frac{z}{y}.
\end{align*}
\]

From the addition of equations (8) and the definition of chromaticity coordinates it follows:

\[
\frac{Y}{y} = \frac{Y_R}{y_R} + \frac{Y_G}{y_G} + \frac{Y_B}{y_B}
\]

substituting (9) into (8) and taking into account (10) we obtain a system of equations:

\[
\begin{align*}
\frac{Y_R}{y_R}(x - x_R) + \frac{Y_G}{y_G}(x - x_G) + \frac{Y_B}{y_B}(x - x_B) &= 0 \\
\frac{Y_R}{y_R}(y - y_R) + \frac{Y_G}{y_G}(y - y_G) + \frac{Y_B}{y_B}(y - y_B) &= 0 \\
\frac{Y_R}{y_R}(z - z_R) + \frac{Y_G}{y_G}(z - z_G) + \frac{Y_B}{y_B}(z - z_B) &= 0
\end{align*}
\]

which in matrix form is

\[
\begin{bmatrix}
\frac{x - x_R}{y_R} & \frac{x - x_G}{y_G} & \frac{x - x_B}{y_B} \\
\frac{y - y_R}{y_R} & \frac{y - y_G}{y_G} & \frac{y - y_B}{y_B} \\
\frac{z - z_R}{y_R} & \frac{z - z_G}{y_G} & \frac{z - z_B}{y_B}
\end{bmatrix}
\begin{bmatrix}
Y_R \\
Y_G \\
Y_B
\end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.
\]

The third row is a linear combination of the other two rows, since ‘\(x + y + z = 1\)’. Therefore this system has an indeterminate solution. The result is logical since whenever the proportions of the three primaries are maintained, a colour with the same chromaticity coordinates \((x, y)\), but with different luminances \(Y\), will be generated. The infinite solutions of this system may be expressed as follows:

\[
\begin{bmatrix}
Y_R \\
Y_G \\
Y_B
\end{bmatrix} = k
\begin{bmatrix}
\frac{x - x_R}{y_R} & \frac{x - x_G}{y_G} & \frac{x - x_B}{y_B} \\
\frac{y - y_R}{y_R} & \frac{y - y_G}{y_G} & \frac{y - y_B}{y_B} \\
\frac{z - z_R}{y_R} & \frac{z - z_G}{y_G} & \frac{z - z_B}{y_B}
\end{bmatrix}
\]

where \(k\) is an arbitrary constant.

If the three determinants of equation (13) are positive, then the colour of the chromaticity coordinates \((x, y)\) will fall within the chromatic triangle formed by the three primaries of the monitor and can therefore be reproduced. This does not imply that this will work with any luminance.

7. Limits to luminance: maximum luminance allowed for a colour

We know that, visually, a greater or lesser luminance does not simply mean that the colour is seen to be more or less bright, since the chromaticity perceived by a human observer varies with luminance (Bezold–Bruke effect) [13]. Therefore, in order to reproduce visually a colour as accurately as possible, its luminance must be specified.

Therefore, the following step is to study the range of luminances within which this colour may be generated. Once the required proportion of each primary is known, the brightness of the colour will be limited by the maximum luminance that the monitor is able to generate with each of the primaries (7). This means that for maximum brightness of the colour \(Y_{\text{max}}\), one of the primaries appears in the mixture with the maximum luminance allowed by the monitor, whereas the other two will have luminances below their maximum luminance value.

To calculate \(Y_{\text{max}}\), the constant \(K_{\text{Max}}\) which appears in equation (13) must be determined. The minimum of the three values which result from introducing (7) into the numerators of equation (13) is taken as the constant to solve (11). That is:

\[
K_{\text{Max}} = \text{Minimum}
\]

\[
\begin{bmatrix}
Y_{\text{Max} - R} \\
Y_{\text{Max} - G} \\
Y_{\text{Max} - B}
\end{bmatrix} =
\begin{bmatrix}
\frac{x - x_R}{y_R} & \frac{x - x_G}{y_G} & \frac{x - x_B}{y_B} \\
\frac{y - y_R}{y_R} & \frac{y - y_G}{y_G} & \frac{y - y_B}{y_B} \\
\frac{z - z_R}{y_R} & \frac{z - z_G}{y_G} & \frac{z - z_B}{y_B}
\end{bmatrix}
\]

This ensures that by using all the luminance permitted by the monitor for this primary, the luminance of the other two will be lower than the maximum. Once the constant
Figure 2. The digital cards which control the CRT monitors, only allow colours defined by integer values \((R, G, B)\). There are eight possible assignations for generating the colour. This introduces a ‘quantization error’.

\(K\) has been calculated (14), the values \((Y_R, Y_G, Y_B)\) can be obtained from equation (13), and then:

\[ Y_{\text{Max}} = Y_R + Y_G + Y_B. \] (15)

Therefore the colour \((x, y, Y)\) can only be generated if its luminance, \(Y\), is equal to or lower than \(Y_{\text{Max}}\).

8. Generation of the colour of coordinates \((x, y, Y)\)

After checking that the chromaticity \((x, y)\) and luminance \(Y\) of the colour can be generated, the values \((x, y, Y)\) are used to calculate its normalized tristimulus values \((X_N, Y_N, Z_N)\) with equations (9). Applying the standard transformation matrix \((A^{-1})\) for CRT monitors, the normalized values \((R_N, G_N, B_N)\) can be calculated by equation (6). The \((R, G, B)\) values can be calculated from the normalized values, as follows:

\[
\begin{pmatrix}
R \\
G \\
B
\end{pmatrix}
= (2^p - 1) * \begin{pmatrix}
R_N \\
G_N \\
B_N
\end{pmatrix}
\] (16)

where ‘\(p\)’ is the number of bits on the digital card.


Quantization error

The three values \((R, G, B)\) found in equation (16) are real whereas the digital cards, which control the CRT monitors, only allow integer values. Therefore we must decide which integer value is to be assigned to each of these three real values. There are eight possible assignations for generating the colour (figure 2). This introduces a new error called the ‘quantization error’. Here they are denoted by \((R, G, B)_i\) where \(i = 1, 2, \ldots, 8\) in order to formalize the operations below.

The problem now is to decide which of these eight values is colorimetrically the closest to the colour \((x, y, Y)\) to be generated. Given that the RGB space has no perceptual uniformity, the criterion of minimum distance is not appropriate. It is helpful to transform the space RGB into a uniform one, for example a CIELUV space, where the minimum distance criterion is colorimetrically more correct since the geometric distance corresponds, approximately, with the perceptual distance. We chose the uniform space CIELUV because it was originally defined for light colours (as is the case of the monitor). Had we been dealing with surface colours then we would have chosen CIELAB. Obviously to calculate the CIELUV coordinates, first its chromaticity coordinates have to be calculated in the XYZ space (normalized to 100 for the luminance of the white point). This is done for the eight points:

\[
\begin{pmatrix}
X \\
Y \\
Z
\end{pmatrix}
i = 100 
\frac{2^p - 1}{A} \begin{pmatrix}
R \\
G \\
B
\end{pmatrix}_i
\] (17)

The equations for transforming the tristimulus values \((X, Y, Z)\) to the CIELUV system \((L^*, u^*, v^*)\) proposed by CIE in 1976 [14] are the following:

\[
L^* = 116(Y/Y_0)^{1/3} - 16 \quad Y/Y_0 \geq 0.00885
\]

\[
u^* = 13L^*(u' - u'_0)
\]

\[
v^* = 13L^*(v' - v'_0)
\]
with

\[ u' = \frac{4X}{X + 15Y + 3Z} \]
\[ u'_0 = \frac{4X_0}{X_0 + 15Y_0 + 3Z_0} \]
\[ v' = \frac{9Y}{X + 15Y + 3Z} \]
\[ v'_0 = \frac{9Y_0}{X_0 + 15Y_0 + 3Z_0} \]

where \((X, Y, Z)\) and \((X_0, Y_0, Z_0)\) are the tristimulus values of the colour and the reference white [14]. The CIELUV difference, \(D_{uv}\), between two colours is calculated from:

\[ D_{uv} = \sqrt{(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2}. \]  

In the uniform colour spaces (CIELUV, CIELAB) a difference in the order of five units permits the visual differentiation of two colours (however, in certain experimental conditions it is possible to reach lower differential thresholds).

So, among the eight initial possibilities, we shall choose the digital values defining a colour whose CIELUV distance to the colour we want to reproduce is minimum. This distance indicates the colorimetric error which occurs when the given digital values \((R, G, B)\) are assigned to generate the colour \((x, y, Y)\).

10. Conclusion

We hope that this method will help software designers and colour researchers to assess and control the colorimetric capacity of their display system. This preliminary step will allow us to generate, where possible, any real colour defined by its chromatic coordinates. By using a uniform colour space it will also be possible to quantify the colorimetric error incurred. It will then be possible to impose tolerance margins depending on the colorimetric precision of the application.

References