Putting Colour Displays to Work

David Travis
Human Factors Unit, BT Labs, United Kingdom

Abstract

The information superhighway is upon us. It is now commonplace to see business users manipulating pictures and video on their desktop machines, and there is increasing demand to network these machines to share information with others. Visual images will be the currency of the information age, and the spoils will go to companies that accurately reproduce these images.

A major barrier to image reproduction is accurate colour rendition. Current solutions require complex, expensive and lengthy calibration procedures. Yet the dream of the information superhighway is to provide ubiquitous access to multimedia information. To achieve this, systems must be easy to use: for colour displays, this means that device calibration must be virtually transparent. This paper describes a user-centred method to achieve this that is simple and quick, and eventually may be good enough for most mass-market multimedia applications.

Introduction

Colour displays have been described as the slip roads—the access and exit points—of the information superhighway. Yet at the same time, colour displays present a real barrier to mass-market multimedia. The issue is simple enough: colour reproduction is poor. Colour can be an important component in making decisions in the real world—choosing a tie, for example, or a new suit. But as well as having obvious implications for home shopping, accurate colour reproduction has implications for a range of services, including telemedicine. For example, in a project investigating teledermatology at BT, we have found that dermatologists use the variation in skin colour to make important diagnostic decisions. Slight errors in colour reproduction result in gross diagnostic errors.

Current display systems need to be calibrated so as to accurately reproduce colour. But this is a complex, expensive and lengthy process. We clearly need quick, simple and accurate methods of display calibration. This paper describes such a technique.

The technique is a novel method of display calibration that allows individual users to calibrate their displays in CIE space without using photometers or colorimeters. Luminance/voltage non-uniformities are corrected by asking subjects to match the brightness of an on-screen dithered patch with an adjacent undithered patch. With a small number of brightness matches for different levels of dither, the display can be calibrated within the error of a good photometer. Moreover, the method makes no assumption about the underlying function (*e.g.* gamma function) and can be applied to novel display technologies as well as CRTs. Colour calibration (within CIE space) is achieved by asking users to match Munsell patches by adjusting the three display primaries. This method works well in principle, but it depends on the ambient lighting and shows some error in colour matching; we are now developing a hand-held LED-based visual colorimeter that avoids both these problems.

Approach

In many areas of measurement we can calibrate a sample against a cheap and widely available physical standard. The ruler is a perfect example: it is cheap and simple to use. For measurements that require a high degree of accuracy, a more precise instrument might be needed (a vernier gauge for example) but for most purposes, a ruler is good enough. In contrast, there really is only one way to do colorimetric calibration, and this is expensive in both time and equipment, and it is not a simple procedure. Our objective was to devise a measurement procedure for colour that was simple, yet still retained sufficient accuracy to be good enough for most multimedia applications.

Display calibration is a two-step process. First, the luminance/voltage non-uniformities are modelled and corrected. The current industry standard is to make measurements with a photometer for a range of 'grey' levels. A graph is drawn relating luminance (grey levels) with the voltage applied to each of the three display primaries, and the system is modelled mathematically. The function relating luminance and voltage on display systems is non-linear. Since colour space manipulations work on the assumption that the underlying primaries obey laws of additivity, the non-linearities need to be measured and corrected. For some displays, the non-linearity can be calibrated by a gamma function, and hence this process is frequently called gamma correction. This is a straightforward, if laborious procedure, but needs equipment (a photometer) and a skilled user.

The second step is more complex. The colour coordinates of the three display primaries must be measured and the 'tristimulus matrix' derived. This is a mathematical description relating the device-dependent colour space formed by the three display primaries to an internationally standadised colour space. This is currently achieved by measuring the colour coordinates of the three display primaries with a colorimeter.

Luminance/Voltage Non-uniformities

Our psychophysical method for measuring the luminance/voltage non-uniformity uses dithering techniques. Two squares are presented to the subject. One square is dithered, so that a fixed percentage of pixels are on at their maximum value and the rest are off. For example, in theory a dithered square where 50% of the pixels are on at their maximum luminance and 50% are off should produce half the maximum obtainable luminance for that display. Other dithers were shown to our subjects, to produce luminances within the range 10% to 90%.

Figure 1 shows an example of the experiment. The square on the left is dithered and fixed in luminance. The square on the right is not dithered, and is adjusted in luminance by the subject to obtain a brightness match with the dithered square.

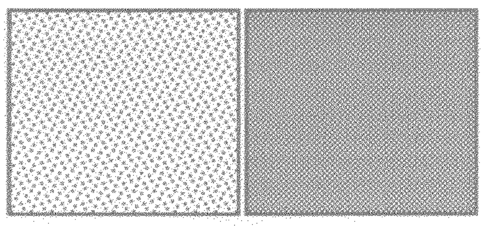

Figure 1. The psychophysical task used for measuring luminance/voltage non-uniformities

Figure 2 shows a set of data collected from seven subjects. The filled symbols show the average settings of the subject. The open symbols show the corresponding data measured using a photometer. Ideally, both sets of data would be concordant. For the photometric measurements, gamma is about 2.25; for the psychophysical measurements, gamma is about 2.10.

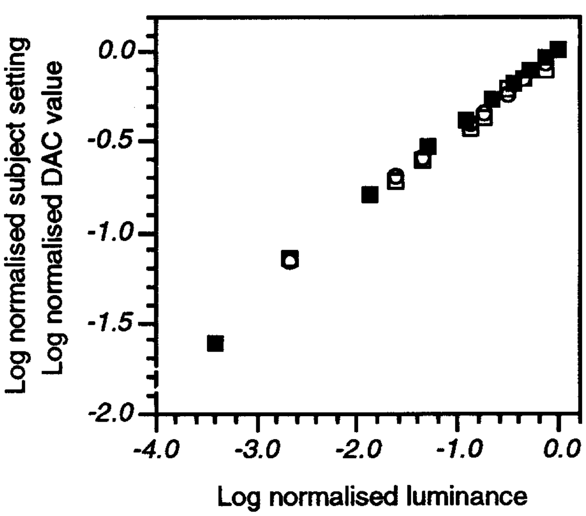

Figure 2. Results from the psychophysical task compared with physical measurements from the displays

If the function describing the monitor non-linearity is truly a gamma function, then it is necessary to measure only two points. This is because a gamma function plots as a straight line on a double logarithmic plot. Hence, it is feasible to psychophysically gamma correct a monitor by obtaining just *two* brightness matches. This makes the method very quick indeed.

On the other hand, if the function cannot be represented by a gamma function (as is the case with some liquid crystals) the curve can be derived by piecewise linear interpolation. Either way, the method accurately estimates the underlying function—yet without the need for a photometer.

Colour Calibration

The correction of luminance/voltage non-uniformities is barely half the solution to accurate colour reproduction on electronic displays. Having achieved linearity for the individual electron guns, there remains the substantial problem that the colour co-ordinates of the red, green and blue primaries vary significantly from one display device to another.

The product of a colour calibration is a tristimulus matrix; if we know the tristimulus matrix for a monitor, that monitor is calibrated. There is a physical approach to solving this problem which is in every way more difficult than the physical solution to the correction of luminance/voltage non-uniformities. The equipment necessary is more expensive, more complex and the method is more time consuming. Here we describe a psychophysical method which is simple, and once again places the emphasis on the user's judgements.

In our first instantiation, we provided subjects with three calibrated colour chips. We attached them to the display device and simply asked subjects to adjust the three display primaries to achieve a colour match.

First it is necessary to derive the tristimulus matrix for the display from three fixed physical samples. We know that:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_R & X_G & X_B \\ Y_R & Y_G & Y_B \\ Z_R & Z_G & Z_B \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (1)$$

where X_i, Y_i, Z_i (i=R,G,B) are the tristimulus values of the display primaries, and X, Y, Z are the tristimulus values of a physical sample matched by the three primaries R,G,B.

Now consider the equation relating *any* three physical samples (within the gamut of the monitor) reproduced on the screen and the RGB values that match:

$$\begin{bmatrix} X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \end{bmatrix} = \begin{bmatrix} X_R & X_G & X_B \\ Y_R & Y_G & Y_B \\ Z_R & Z_G & Z_B \end{bmatrix} \begin{bmatrix} R_1 & R_2 & R_3 \\ G_1 & G_2 & G_3 \\ B_1 & B_2 & B_3 \end{bmatrix} \qquad (2)$$

where X_i, Y_i, Z_i (i=1,2,3) are the tristimulus values of the physical samples and R_i, G_i, B_i (i=1,2,3) are the RGB values required to produce a match.

This can be expressed in simpler terms as:

$$\mathbf{C} = \mathbf{T}\,\mathbf{M}$$

where

C = 3 x 3 matrix of the tristimulus values of the physical samples
T = the tristimulus matrix of the display primaries
M= 3 x 3 matrix of the RGB values that match the physical samples

This can be manipulated to give:

$$\begin{aligned}\mathbf{CM}^{-1} &= \mathbf{TMM}^{-1}\\ &= \mathbf{TI}\\ &= \mathbf{T}\end{aligned}$$

Thus we have an expression for T in terms of the three physical samples and their colour matches. This T can be found and can calibrate the monitor for all colour reproduction purposes.

Method

We used three Munsell swatches (see Table 1) and attached these to an Apple 14" colour display. Room illumination was from standard fluorescent tubes. Subjects adjusted the colour of an on-screen colour patch using the standard colour Macintosh control panel. 5 matches were made to each of the three physical samples, and then these were averaged in CIE uniform chromaticity space.

Table 1: Munsell swatches used in the experiment. The x and y values were measured *in situ*.

Munsell Chip	CIE x	CIE y
10RP 5/4	0.493	0.3988
5G 6/6	0.42	0.4641
2.5 PB 6/8	0.3743	0.3716

In our experiment, we ran three subjects through the calibration procedure, and calculated a unique tristimulus matrix for each subject. It may seem more exact to call these matrices 'estimated tristimulus matrices', in other words assuming that the 'correct' tristimulus matrix is obtained by physical measurements of the display. In actual fact there are some errors in this assumption, since the display primaries interact with each other when these physical measurements are taken. The novel method described above partially accounts for these interactions (because it measures chromaticities produced by *mixing* the three primaries and not just the three primaries themselves), and lends itself to both physical and psychophysical calibration.

Results

The test for any tristimulus matrix is how well it can predict XYZ values from given RGB values according to eqn. (1), when the actual XYZ values are known. In our evaluation we first obtained the tristimulus matrix from psychophysical judgements, and then used this tristimulus matrix to predict XYZ values for 31 RGB patches. The actual XYZ values of these patches was measured in advance, and they were evenly distributed within the colour gamut of the monitor. Error was defined as the colour difference between the XYZ co-ordinate predicted and the XYZ co-ordinate measured. These values can be summed and a mean error can be calculated for each tristimulus matrix.

The results for the three subjects are shown in Table 2. This table shows the average error (in ΔE^*_{uv} units) for the psychophysical tristimulus matrix compared with that for the physically-measured tristimulus matrix.

Table 2. Average error (in ΔE^*_{uv} units) for the two tristimulus matrices

Subject	Physical	Psychophysical
HD	6.04	34.06
MS	6.04	20.97
SJE	5.30	27.31

Discussion

Colour reproduction in multimedia systems presents a different set of problems to the demands made in desktop publishing, printing and the textile industries. Our aim in this study was not to find an alternative to high-end colorimeters where this precision is truly needed. Instead, our aim was to develop a system that would result in acceptable errors in applications where colour reproduction is important, but some error can be tolerated. Earlier, we drew an analogy between colorimeters and vernier gauges, and argued that for some applications a ruler would be good enough. We introduce more error, but that is acceptable, because at the same time we produce a system suitable for quick, mass-market use.

What are the limitations of the procedure we use for correcting luminance/voltage non-uniformities? A key assumption is that lighting a fixed proportion of pixels results in the same fixed proportion of the maximum luminance. In pilot studies, we found no evidence to contradict this assumption. A second assumption is that subjects can do this quickly and easily. One problem we found with crude dither patterns was that the dither pattern itself became highly visible. This made it more difficult for the subject to set a luminance match because the texture of the two patches appeared different. In subsequent work, we have shown that placing a high-pass filter (for example, tissue paper) over the two patches reduces this effect and makes the matching procedure simpler still.

What are the limitations of the procedure we use for colour calibration? Understandably, the psychophysical errors are larger than those obtained from physical measurements of the display, and they are probably too large for most multimedia applications. However, the values in Table 2 are averaged; this means some matches had lower and some had larger errors. Indeed, in some regions of colour space, the psychophysical tristimulus matrix was in fact more accurate than the physically-measured one.

Some of the error is due to the fact that we did not factor room illumination in (or out of) our calculations. The design reported here requires three physical samples. It is unlikely that these could be fabricated to be metameric under various illumination conditions, and hence will not be practical for widespread use. However, we have now sketched a design for such a system based upon visual colorimeters that will avoid illumination problems. In this system, we use lights (LEDs) rather than paper. A three-way switch produces fixed and known mixtures (reddish, greenish and bluish lights) of the three primaries on the left half of the bipartite field. The source in the right half of the bipartite field comes from the display screen itself. The user adjusts the colour of the display screen source to match each of the three fixed light sources.

A second issue concerns the difficulty in colour matching experienced by some subjects. It proved hard to achieve a colour match across different media, perhaps due to textural differences. We expect our new design to provide more accurate colour matching (since users will be matching lights with lights and not lights with paper).

In conclusion, psychophysical colour calibration offers the promise of putting colour displays to work on the information superhighway. Luminance/voltage non-uniformities can already be corrected using psycho-physics. But work still needs to be done to improve the efficiency of psychophysical colour calibration.