Behavior Research Methods, Instruments, & Computers
1993, 25 (3), 371-383

COMPUTER TECHNOLOGY

Calibration of a color monitor for visual psychophysics

ANDREW B. METHA, ALGIS J. VINGRYS, and DAVID R. BADCOCK
University of Melbourne, Parkville, Victoria, Australia

It has become common for stimuli used in visual psychophysical experiments to be presented on high-resolution color cathode-ray tubes (CRTs) such as the Barco CDCT 6551. These enable a flexibility of color, spatial-frequency content, temporal-frequency content, duration, size, and position that is not provided by most other media. CRTs are, however, not perfect; they suffer from the effects of temporal instability, spatial variability, lack of phosphor constancy, gun interdependence, and gun nonlinearity. This paper describes methods of assessing these aspects of monitor performance with respect to how significant each may be in psychometric terms. Although every application of CRT use in visual psychophysics is different, some general rules can be formulated to help ensure that unwanted effects are kept to a minimum. For the CRT used in this study (Barco CDCT 6551), a warm-up time of 30-45 min is necessary before chromatic and luminous stability ensues. Restriction of individual gun outputs to within 10%-90% of the possible range ensures that the effects of gun interdependence and lack of phosphor constancy are negligible. Calibration methods dealing with the linearization of gun output are also discussed.

CRT (cathode-ray tube) technology offers a unique opportunity to display a variety of visual stimuli. The advantages offered by CRTs over other media include flexibility of color, size, position, and presentation duration as well as spatiotemporal frequency content of the stimuli. A major disadvantage of the CRT is that the maximum light output achievable is limited, compared with other display techniques such as Maxwellian view. This is still well into the photopic range, as is borne out by observing that images are in fact colored on CRT screens. Color CRTs have three channels, so it is possible and desirable to describe their output in terms of a mixture of three independent primaries, although, as will be shown, strict independence is not achieved. In order to determine the limitations that may exist in applying CRTs to the study of the human visual system, we have undertaken a thorough study of the colorimetric aspects of a Barco CDCT 6551 monitor that was used in our laboratory. We believe that this high-resolution color monitor is likely to be used in other laboratories and that the information provided in this paper will be useful to others. Moreover, the techniques described here can be applied to other brands and models of CRTs and should provide a useful reference for those unfamiliar with this technology or colorimetry.[1]

This work was supported by ARC Grant A79030414 and NH & MRC Grant 890568 to A.J.V. We thank George Smith for his consultation and helpful comments on the manuscript. We also thank two anonymous reviewers for their very useful comments. Address correspondence to A. Metha, Department of Optometry, University of Melbourne, Parkville VIC 3052, Australia.

In the present study, we consider several aspects of monitor performance: temporal stability, spatial variability, phosphor constancy, gun independence, and gun linearity. Some of these attributes are "hardwired" by the manufacturer and cannot be changed by the end user, but a knowledge of their limitations will allow adoption of appropriate strategies to prevent these limitations from corrupting results. Other attributes may be altered by the user, and methods are given to optimize output. In this paper, we first present a basic review of how colorimetric theory relates to CRTs, and then discuss each of the above considerations.

MATERIALS AND METHODS

The CRT used in our laboratory is a Barco CDCT 6551 RGB monitor controlled by a Cambridge Research Systems visual stimulus generator (VSG) card in a 386 PC clone. Radiometric measurements were taken with a Spectra-Pritchard 1980B telespectroradiometer under the control of a Hewlett-Packard 9826 computer. In order to replicate the practical environment, all measurements were made under the same conditions as those used for psychophysical studies. The telespectroradiometer was positioned coincident with the expected location of a human observer's eye. Note that the spectroradiometer's sampling rate can alias with the frame rate of the CRT (90 Hz) in its fast-scan mode; hence, measurements were obtained using the medium sampling speed.

COLORIMETRY AND COLOR CRTs

Colorimetry

The basic laws of colorimetry are essentially the rules of metamerism, and can be expressed in terms of the linearity and trichromacy of (foveal) color matches (Estévez & Spekreijse, 1982). The linearity of color matching is formulated as Grassmann's laws. These can be summarized as follows (the = sign indicates that a color match exists):

1. If $A = B$, then $\alpha A = \alpha B$. (Proportionality Rule)
2. If $A = B$ and $B = C$, then $A = C$. (Transitivity Rule)
3. If $A = B$, then $A + C = B + C$. (Additivity Rule)

The trichromacy of color vision ensures that colors can be matched by an additive mixture of at least three suitably chosen "primaries" over a large range of conditions (Wyszecki & Stiles, 1982). Any set of three different primaries can be used, but red, green, and blue are most often chosen. These primaries can be readily produced and used to create a large gamut of visible colors by their addition or superposition. Maximal color domains will only be achieved when highly saturated or spectral colors are used as primaries. Thus,

$$A = r\mathrm{R} + g\mathrm{G} + b\mathrm{B}, \tag{1}$$

where A is the desired color and r, g, and b are the scalar amounts of R, G, and B primaries, respectively.

If A falls outside the range constrained by R, G, or B, a match can still be achieved by judiciously adding one of the primaries to A and then matching this resultant with the remaining primaries. For example, if we need to add R to A, the equation above becomes

$$A + r\mathrm{R} = g\mathrm{G} + b\mathrm{B}, \tag{2}$$

using the components that have been defined in Equation 1.

In 1931, the Commission Internationale de L'éclairage (CIE) adopted a set of nonrealizable primaries, which they called X, Y, and Z. With positive amounts of these imaginary primaries, one can theoretically match any color perceived by a standard observer. The required amounts of each primary are called the *tristimulus values* of the match and are also denoted X, Y, and Z. These can be calculated for any sample color once the spectral power distribution, $P(\lambda)$, of the sample is determined from a spectroradiometric scan:

$$\mathrm{X} = \int P(\lambda)\, x'(\lambda)\, d\lambda \tag{3}$$

$$\mathrm{Y} = \int P(\lambda)\, y'(\lambda)\, d\lambda \tag{4}$$

$$\mathrm{Z} = \int P(\lambda)\, z'(\lambda)\, d\lambda \tag{5}$$

In the equations above, $x'(\lambda)$, $y'(\lambda)$, and $z'(\lambda)$ are called the *color-matching functions* or distribution coefficients, and they represent the amounts of each of the primaries X, Y, and Z that are required to match an equienergy white light, taken one small waveband, $d\lambda$, at a time. The color-matching functions are tabulated in Wyszecki and Stiles (1982, Table 1, pp. 725-735) and Hunt (1991, Appendix 3).

The tristimulus values can be normalized by their sum to give chromaticity coordinates x, y, and z:

$$x = \mathrm{X}/(\mathrm{X} + \mathrm{Y} + \mathrm{Z}) \tag{6}$$

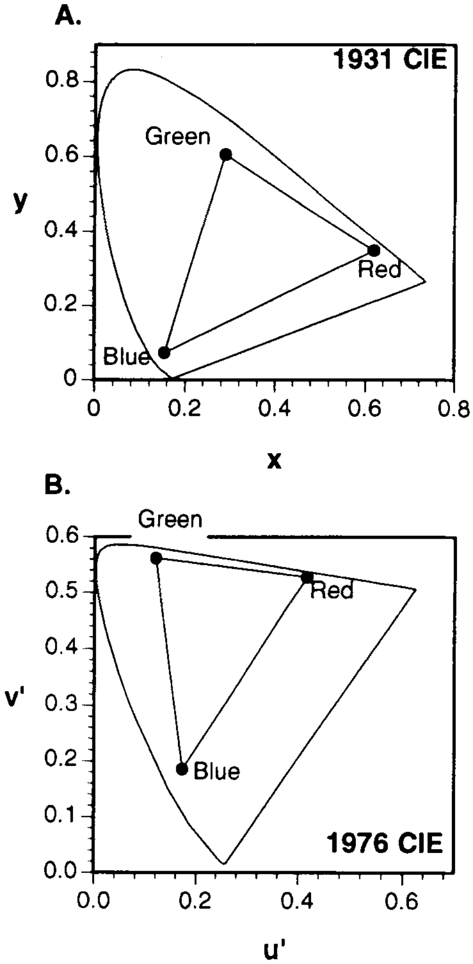

Figure 1. The gamut of colors available on a CRT screen in terms of 1931 CIE (x,y) space (A) and 1976 CIE (u',v') space (B).

$$y = \mathrm{Y}/(\mathrm{X} + \mathrm{Y} + \mathrm{Z}) \tag{7}$$

$$z = \mathrm{Z}/(\mathrm{X} + \mathrm{Y} + \mathrm{Z}) \tag{8}$$

Because $x + y + z = 1$, only two of these coordinates are required to generate a color space such as the 1931 CIE chromaticity diagram (see Figure 1A). A shortfall of this diagram is that perceptually equal color steps are not represented by equal distances. Attempts at a uniform color space resulted in the CIE's adopting the 1976 CIELUV system, which is a linear transformation of the 1931 XYZ system (see Figure 1B). In this system, the chromaticity diagram is based on the coordinates u' and v', defined as

$$u' = 4x/(-2x + 12y + 3z) \tag{9}$$

$$v' = 9y/(-2x + 12y + 3z), \tag{10}$$

where x, y, and z are as defined above.

From these chromaticity coordinates, color difference relationships in terms of just noticeable differences (JNDs) can be established. The reader interested in further detail is referred to Wyszecki and Stiles (1982), Billmeyer and Saltzman (1981), or Hunt (1991).

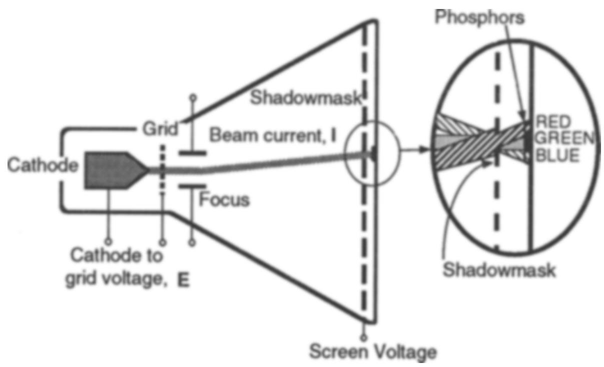

Figure 2. Schematic representation of a color CRT with one gun. The inset illustrates the shadowmask principle, whereby the spatial register of the guns and the shadowmask ensure that each gun excites only one phosphor type (modified after Cowan, 1987, and Rodieck, 1983).

Color CRTs

A color CRT consists of a very precise arrangement of three grids and cathodes (guns), a phosphor screen, and a shadowmask all contained within an evacuated glass tube. A schematic representation showing one channel of a CRT is displayed in Figure 2. The guns supply the electrons needed to activate the phosphors on the screen. These are focused and deflected in a precisely controlled manner by internal electromagnetic fields, so extraneous magnetic influences (e.g., power cables, other CRTs) should be kept to a minimum. The cathode supplies an electron beam current (I), which is regulated by the cathode-to-grid voltage (E), resulting in a power function relationship such that $I \propto E^{\gamma}$. This relationship is often referred to as the gamma function and will be discussed later. The beam is aimed by the focusing plates through the apertures of the shadowmask and onto the phosphor screen. The inset to Figure 2 schematically depicts how the three guns are arranged (usually triangular) so that each beam passes through a single aperture of the shadowmask to excite only one phosphor type.

Applying the rules of trichromacy to CRTs, it follows that we can treat the three gun/phosphor channels as independent primaries that are able to produce any color inside their operating domains. In the present paper, we will use the term *channel* to indicate the entire cathode-grid-phosphor complex needed to produce a color on the screen. Figure 1 shows the gamut of colors that can be displayed on the CRT in terms of the CIE color spaces (1931 CIE color space and 1976 UCS uniform color space). The filled circles represent the chromaticity coordinates of the red, green, and blue phosphors; any color within the triangle can be displayed. The CIE recommends that the 1976 (UCS) space be used for applications in which colored lights are mixed additively (CIE, 1978), such as for CRT applications. In this space, color is described by u' and v' coordinates in such a way that changes in the v' dimension roughly correspond to blue-yellow percepts, whereas u' changes approximate red-green percepts (Figure 1).

In concert with the ideas of Cowan (1983) and Cowan and Nelson (1986), an ideal CRT should provide primaries that:

1. are able to be controlled in a predictable and precise (preferably linear) manner, from zero to maximum intensity (gun linearity);
2. are independent, so that each channel is not influenced by the actions of any of the others (gun independence);
3. maintain a constant relative spectral distribution as intensity is varied (phosphor constancy); and
4. are widely separated in color space in order to provide a large gamut of visible colors.

Most of the factors mentioned above are preset by the maker and are beyond the control of the user. Precise gun linearity is difficult to achieve because the channels follow a gamma function; however, methods can be adopted to linearize this function. This will be discussed in detail later. In addition, any system used to investigate human vision needs to be able to do the following:

1. Spatially superimpose the primaries. Alternatively, a very fine matrix of phosphor dots could be used, providing that the matrix structure (triad) cannot be resolved by the observer. This aspect is limited by the spatial contrast sensitivity function (CSF) of the human eye, the size of the matrix element, and the viewing distance. For most modern high-resolution screens, a viewing distance of 50 cm should be adequate to prevent resolution of the matrix elements.
2. Refresh the screen image at a fast rate with suitably matched phosphor persistence so that the temporal resolution of the human eye can be studied. This aspect is limited by the temporal CSF of the eye, as well as phosphor rise and decay times. Although modern phosphors respond very quickly, Vingrys and King-Smith (1986) have shown that an asymmetry in the temporal characteristics of the red and green channels may produce an unwanted luminance transient detectable by the eye when changing from a white background to a maximal red presentation. For stimuli involving less than maximal changes, such transients should not be perceptible. In order to avoid such artifacts, it is advisable to drive the CRT at less than maximum levels (a point that will be elaborated upon later).
3. Provide an output that is constant over the spatial extent of the screen (spatial variation). This factor is unlikely to be achieved; hence, some level of spatial variation must be accepted. This spatial variation must not produce steps that are visible to the eye and should be similar in the three channels in order to prevent color or contrast artifacts. Misconvergence of the gun beams can also produce a spatial luminance artifact; Vingrys and King-Smith (1986) describe a simple photographic method to check for its presence. If misconvergence occurs, it can be minimized by the realignment of the magnetic coils (due to the high voltages inside a tube, this is best left to a trained technician).
4. Maintain a stable output during the duration of an experiment as well as from day to day (warm-up charac-

teristics). For the purpose of this discussion, we will assume that the screen is adequately protected from extraneous magnetic and electrical fields, although in our laboratory, apart from the obvious siting precautions, we do not go to any other lengths to shield the CRT. (The Barco has a degaussing cage to minimize the effects of unwanted magnetic fields.)

Some of these aspects are elaborated on in the following sections.

WARM-UP CHARACTERISTICS

If the output of a CRT does not quickly converge to asymptotic values for both luminance and chromaticity, its usefulness as a color stimulator for visual experiments is limited. Cowan (1987) has shown that the luminance of a CRT may vary by as much as ±1% over an 8-h period; this is an acceptable level of fluctuation in most cases, because the changes in output are not abrupt in time and therefore will not be perceived. However, it is good practice to allow a CRT to "warm up" for a period of time before it is used for experimentation. Visual inspection following startup reveals that the color of the screen changes with time. Such changes probably reflect differential characteristics in the thermal equilibration of the high-voltage cathodes and control grid accelerators of each channel. However, there is little guidance in the literature regarding the minimum warm-up period needed by CRTs. Cowan (1991) suggests that this period should be about 1 h, but provides no data to support his claim; it is difficult to determine the manner by which a warm-up may be quantified. In the following, we discuss some of the problems encountered in making such measures and give reasons for adopting our methods, which we believe are rational and consistent with human psychophysics. We do, however, acknowledge that it is not the only method for assessing such changes.

These methods have been adopted because, for psychophysical purposes, the primary question seems to be how long a CRT should warm up before it can be used confidently within some specified level of tolerance. Because of the application of the CRT, we felt that it was best to consider the consequence of any fluctuation in psychometric terms or the JND.

Data were collected on 2 separate days for 2- or 3-h periods following startup in order to gauge the day-to-day variance. If it is assumed that warm-up variation arises from thermal equilibration within the pregun circuitry or the guns themselves, this should be reflected by differences in the warm-up profiles of the CRT running from a "cold start" (off period > 14 h, overnight) to that found for a "warm start" (restart after 20 min off following a 3-h on period). These two conditions were adopted because they represent typical operating modes in our laboratory. We also felt that 3 h should enable thermal equilibration of the guns, whereas 20 min should substantially upset this condition. Two warm-start runs and two cold-start runs were made.

The voltage of the guns at which measurements are made influences the outcomes. Because it was impractical to investigate the warm-up phenomenon at many different relative voltages, it was decided that the most useful gun voltage (for psychophysics) should be studied exhaustively. For our purposes, this voltage corresponds to half the maximum output range, which is the adaptive background or "white point" often used during psychophysical experiments. It also usually allows the greatest dynamic range of modulation for each gun.

Measurements were taken every 3 min (the time taken to complete a radiometric scan from 370 to 730 nm in 10-nm steps; bandwidth 10 nm) during the first 30 min and thereafter at 15- or 30-min intervals. Cold-start readings were made for 3 h, and warm-start measurements were terminated after 2 h. Although the spectroradiometer that we used is a precision tool, we conducted a pilot trial to ensure that any variance found during a warm-up was due to the CRT and not to the spectroradiometer. This trial established that the spectroradiometer is stable to within ±0.01% as soon as it is able to start making measurements.

In Figure 3, the warm-up characteristics of the CRT are shown in terms of the luminance and CIE 1976 chromaticity coordinates (u', v') for both cold and warm starts. Asymptotic values were similar in all cases, but were achieved earlier for warm starts. Figure 3 shows that a small (6%) but not insignificant change was found during warm-up. A stable luminance was achieved only after some 60 and 150 min following warm and cold starts, respectively. This appears to be an extremely long period of time for stabilization. The greatest variability was found within the first 20 min of a cold start (see luminance and v' values in Figure 3); however, after this period the trends were similar. Figure 3 also confirms the subjective impression that the screen became more greenish-blue (decrease in u' and v') within the first 30 min after startup.

The practical consequence of the warm-up profile can be appreciated in one of two ways. The first would require CRT warm-up to proceed until some acceptable limit of variability in photometric or colorimetric measurements is achieved. Figure 3 indicates that there was as much as a 6% luminance variation over 3 h for a cold start and 4% for a warm start. The curves seem to asymptote at 2 h for the cold start and at 1 h for the warm start. Chromaticity coordinates appear to reach asymptotic values by about 20 min in all cases. Figure 3 shows that, for a cold start, there was an initial 10- to 20-min period of unpredictable output before stability ensued. This stable phase occurred consistently for both conditions, although for the warm starts, the data were displaced by about 40 min. This suggests that there are at least two stages of the warm-up process: the cold start exhibits both phases, whereas the warm start quickly passes through the initial stage and, by 5 min, displays characteristics of the slow phase of warm-up. Cowan reports that a CRT may fluctuate by ±1% over an 8-h period, even after warm-up (Cowan, 1987). Our data show that a 1% tolerance in luminance

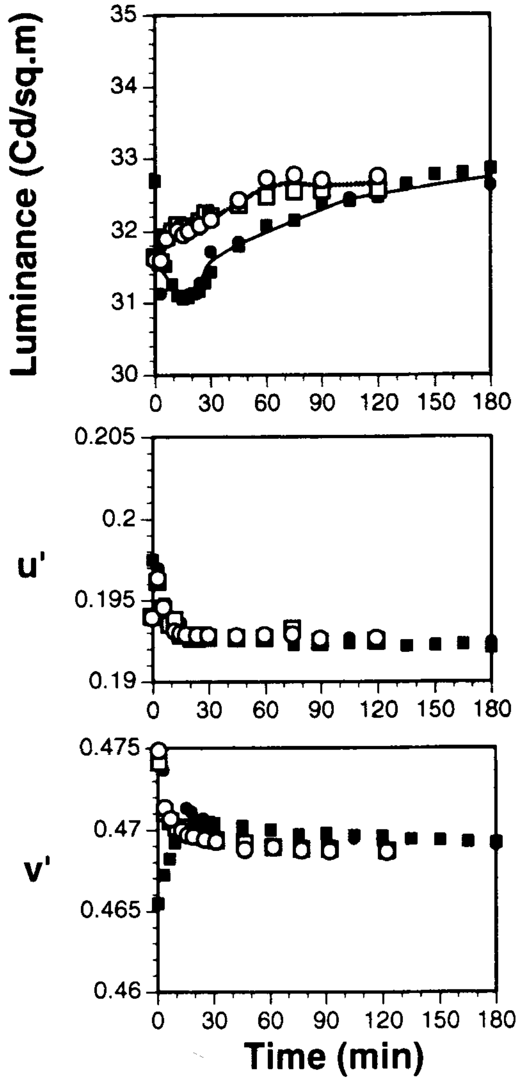

Figure 3. Colorimetric data of the warm-up process. The squares indicate results from the first data set and the circles represent data collected 2 days later. The filled symbols refer to "cold" starts, and the open symbols represent data for "warm" starts. Refer to the text for definitions.

of our final value was achieved 40 and 100 min after warm and cold starts, respectively, suggesting that this time interval is the minimum required to achieve stable long-term levels. However, although such a hardware-related logic seems reasonable, it disregards the visual capacity of the end user (human eye).

An alternative approach is to consider the warm-up in terms of psychophysical units. An appropriate index for this purpose is called the *total color difference* (ΔE^*_{uv}, Wyszecki & Stiles, 1982), which takes into consideration luminance and chrominance variations and is scaled in psychometric units or JNDs (Mukherjee & Venkatesh, 1986). The total color difference (ΔE^*_{uv}) between each measured color and the final asymptotic value (180 and 120 min for cold and warm starts, respectively) is plotted in Figure 4. Figure 4 shows that if we adopt the criterion that the CRT screen should not vary by more than 1 JND over 1 h of use (a typical time frame for a psychophysical experiment; see scale of Figure 4), then a warm-up of 30 min for a warm start is all that is required. On the other hand, the warm-up time needed for a cold start by this same criterion is about 45 min. Such warm-up periods should ensure that differences in output will remain imperceptible to the human eye over the period of use.

The preceding psychometrically determined warm-up time is conservative and should satisfy the most stringent criterion, especially in cases in which the CRT is used as an adaptive background upon which stimuli are presented. It is evident from Figure 3 that most of the change seen in the ΔE^*_{uv} index following 30 min of warm-up is a luminance change, because the u' and v' indices stabilize within 20-30 min. The eye is not a very good detector of absolute changes in luminance over an extended time scale (1 h) or spatial extent. Also, it has been shown that chromatic pathways adapt very quickly, so small changes in relative cone quantal catch rates may not affect the results of measured thresholds (Stromeyer, Kronauer, & Cole, 1985). However, significant changes may be observable at a "second site" and these may affect the results, regardless of the slow time course of the change. Applying the preceding criterion for warm-up should safeguard against such a possibility.

The variability of the CRT limits its use as a calibrated light source. For example, in absolute threshold experiments in which stimuli are presented on black backgrounds, or when the CRT is used as a reference source for material samples, a screen-generated color may become perceptibly different at a later time due to temporal changes in the CRT even after warm-up. For applications such as these, it is necessary to consider the time interval between successive comparisons. Therefore, the data pre-

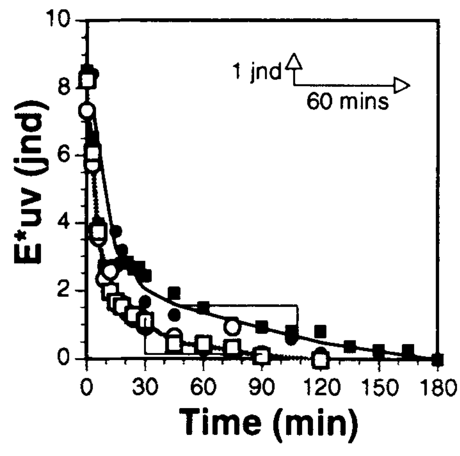

Figure 4. Change in total color difference (ΔE^*_{uv}) for both cold and warm starts with respect to the final color obtained at the end of each run. In order to achieve an imperceptible change ($\leq$ 1 JND) over the ensuing 60 min, a "cold" CRT must warm up for 45 min, whereas a "warm" CRT needs 30 min. The symbols are the same as those used in Figure 3. Refer to the text for details.

Table 1
Color Differences Between the Final Colors of Each Run

Condition	Lum	u'	v'	E^*_{uv}
Cold Day 1	32.85	0.1921	0.4692	0.43
Warm Day 1	32.57	0.1924	0.4688	0.41
Cold Day 3	32.63	0.1924	0.4691	0.18
Warm Day 3	32.75	0.1926	0.4687	0.59
M	32.68	0.1923	0.4690	0.00
SD	0.12	0.0002	0.0003	0.17

Note—Lum = luminance, *M* = mean, *SD* = standard deviation.

sented in Figure 4 become of limited use because they were generated with all guns operating at 50% level. Rather, our data serve as a warning that slow and predictable drifts of color output are likely over time and that even after a suitable warm-up period, stability may not achieve a desired level of tolerance. Performance characteristics should be established on an individual basis for such applications.

It is also useful to know how much CRT output varies from run to run over different days. In Figure 4, the total color difference (ΔE^*_{uv}) was calculated by using the final color as the reference. This is not to imply that the final resting color of the screen was identical in all cases, but only to show how the color changed with time for each run. In order to compare the final screen color and luminance obtained from our four different starts, we calculated the luminance and chromaticity coordinates for the final white point (see Table 1). This shows that the final color and luminance achieved after warm-up was very stable over time, deviating less than 1 JND from the average value and that luminance differences produced the greatest component of variability.

It is recommended that the Barco CDCT 6551 and CRTs in general be allowed to warm up for some period before the output is stable enough to collect experimental data. The duration of warm-up depends on the tolerance levels required, the mode of operation of the CRT during the experiment, and whether the CRT has recently been in use. For laboratories using the CRT to provide a steady background upon which carefully controlled stimuli are presented, we suggest a 45-min warm-up if the CRT is switched on after being off overnight, or 30 min if the CRT has been on as recently as 30 min, or operating at some level other than the intended white-point (background) level. Because luminance shows greater variability than chrominance during warm-up, we suggest that users ascertain their own warm-up characteristics by monitoring this parameter.

GUN INDEPENDENCE

For a CRT to be used as an additive color-mixing device, each gun must be able to function independently of the others. For instance, if we need to change output by increasing the red gun alone, then the output of the

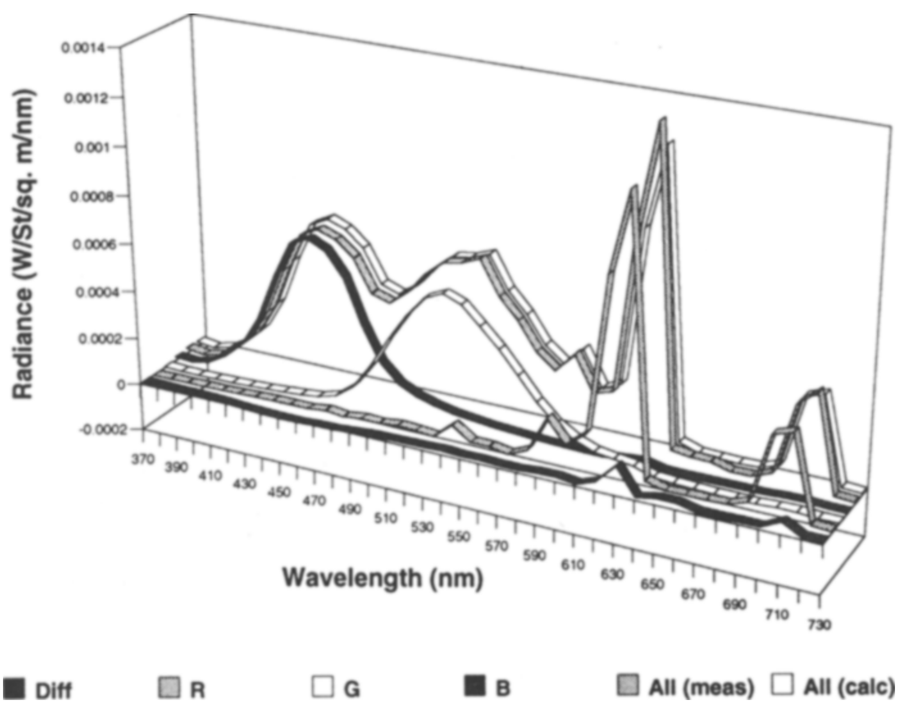

Figure 5. Gun independence for a single output level. Measured spectral output of individual and combined guns operating at the 50% level. If gun independence holds, then the calculated sum of the three individual gun outputs should be equivalent to the measured output for all three guns operating simultaneously. The legend lists the curves from front to back.

green and blue channels must remain constant. Gun independence must be shown to hold in the operating range of a CRT before it can be used in a predictable manner as a stimulus generator for psychophysical studies. Rodieck (1983) indicated that some departure from perfect independence should be expected for most CRTs. He argued that because output radiance is only linearly related to the beam current (I) for a given screen voltage (see Figure 1), changes in the beam currents may result in small screen-voltage fluctuations, causing a failure of current-radiance linearity, and hence independence between the guns. The extent of this nonlinearity depends on the sophistication of the internal circuitry of the device.

Cowan (1991) suggested that for gun independence to hold, the spectral power distribution of light emitted by the CRT when the three guns are operating together, $P_\lambda(E_r,E_g,E_b)$, should be equal to the sum of the contributions of each individual gun alone. That is,

$$P_\lambda(E_r,E_g,E_b) = P_\lambda(E_r,0,0) + P_\lambda(0,E_g,0) + P_\lambda(0,0,E_b) - 2P_\lambda(0,0,0), \quad (11)$$

where E_r, E_g, and E_b are the input voltages of the red, green, and blue guns, respectively, and the last term, $P_\lambda(0,0,0)$, is the "dark light," taking into account the internal noise and reflections from the face plate. Two dark-light units are subtracted from the equation because they appear once for each gun measurement (i.e., three times), even when the input is zero. One way of testing gun independence for a given CRT is to measure the output of each gun separately, and then see if the measured output of a combination of guns is equal to the sum of their individual contributions predicted from measurement. Although gun independence can be checked by using a photometer of limited spectral sensitivity as long as a broad range of the visible spectrum is covered, it should be noted that limited spectral filters may result in different indices of gun independence as emphasis is shifted to different spectral bands (see Figure 5). For this reason, we chose to analyze gun independence in terms of the total spectral radiance (Watts/Ster/m^2) across the visible spectrum (370–730 nm). Because this index spans the entire visible spectrum, it should be sensitive to interactions that may result in changes to psychometric units.

In Figure 5, our analysis is shown as a function of wavelength for the three guns operating at 50% of their maximal output. Individual gun measurements are added (with two dark-light values subtracted) and compared with the measured output of the three guns operating simultaneously. The difference plot in Figure 5 indicates that, even after allowing for the dark light, there is a lack of perfect agreement between actual and calculated outputs and that gun independence is not achieved.

However, a plot such as that shown in Figure 5 tells us little of the practical consequences of this finding. Figure 6A is a plot of the individual and combined gun radiances over a range of operating levels. The difference between calculated and actual radiance is shown in Figure 6A in absolute terms, and in relative terms [(measured − calculated)/measured] in Figure 6B. The relative plot shows that gun independence held to within ±3% for operating levels between 0% and 95%; positive values in Figure 6B indicate that measured values exceeded calculations. The slightly larger discrepancy found at the maximum operating level may have been due to blooming or spillover from each beam onto adjacent phosphor dots. Repeated measures made for the 50% operating level indicate that there is little variance in the data with the standard deviation indicated in Figure 6B. The data do not show any clear relationship between measured and calculated values. Cowan (1987) concluded that the measured output of a Tektronix 690SR CRT was less than the calculated output by approximately 1%. The magni-

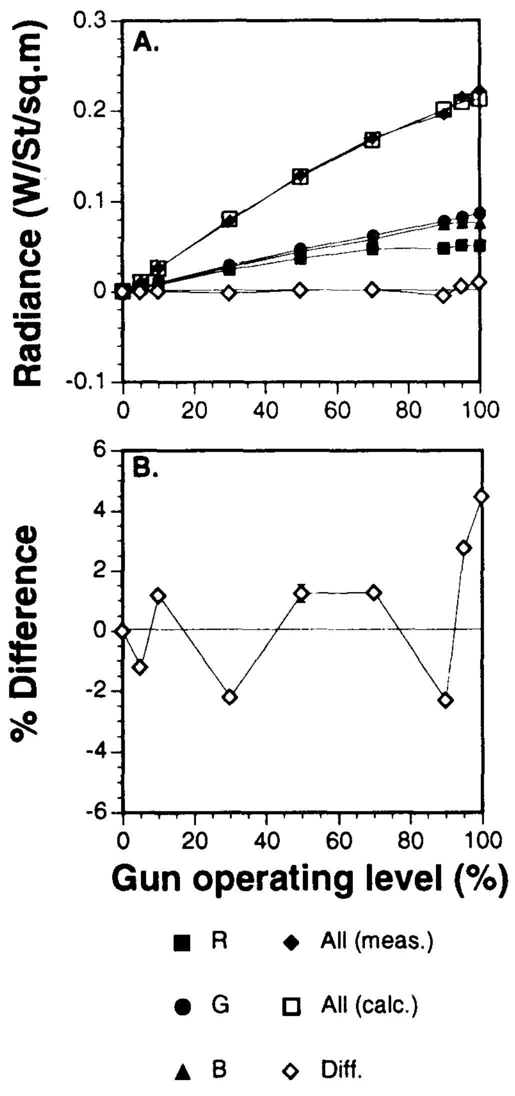

Figure 6. Gun independence as a function of operating level. (A) Radiance of individual and combined gun outputs at a range of operating levels. (B) The difference between calculated and actual radiances [(measured − calculated)/measured] expressed as a percentage of actual radiance values.

tude of the discrepancy in our findings is similar for the 5%-75% operating range and any lack of similarity may reflect design differences in the two CRTs. Violations of gun independence could be expected to be greatest with all three guns operating. However, interactions between pairs of guns may identify particularly bad combinations that could be avoided during operation. For example, if the addition of the blue gun has undesirable effects, perhaps the CRT could be used in a manner that would minimize blue gun usage. The data for such an analysis are not given, but they can be summarized by noting that there is no pair combination for which the interactions are any better or worse than for any other pair. The results were similar to those given in Figure 6. Gun independence is an aspect of CRT performance that is preset by design considerations and is not changeable by the user. The lack of independence reported for the Barco may serve as a guide for those who are interested in using this equipment or who require a set performance level for a specific use.

PHOSPHOR CONSTANCY

It is desirable for CRT primaries to have spectral power distributions that are independent of intensity level. When this condition holds, we can say that phosphor constancy has been achieved. One way of checking for phosphor constancy is to measure the spectral output of a channel at different operating levels. Figure 7A shows the spectral output of the red gun operating at 5% and 10%. For phosphor constancy to hold, we require the 10% curve to have the same shape but twice the area of the 5% curve. Replotting these data in logarithmic terms (Figure 7B) gives a better appreciation of the relationship; each curve should now be amenable to exact superposition by vertical displacement. Close inspection indicates that phosphor constancy does not strictly hold; the differences are small, but we need to determine how important they are.

The data of Figure 7 are difficult to analyze in terms of how well (or not) phosphor constancy is achieved and of the practical significance of any departures from this ideal state. The shape of the spectral curve of each phosphor is relatively constant as a function of channel output, so a simple check of phosphor constancy requires that the chromaticity coordinates remain invariant as output varies. Standard color difference formulae can then be used to infer the significance of any departure from phosphor constancy.

As for gun independence (Equation 11), the dark-light component needs to be allowed for before any meaningful analysis of phosphor constancy can be made. Figure 8 shows the chromaticity coordinates of each of the guns calculated in such a manner over a range of operating levels. In all cases, it is evident that red and green phosphors appear to be slightly desaturated when operated at low levels (< 10%), whereas the blue phosphor is most saturated at these low levels. This analysis indicates that the phosphors do not achieve constancy and change their spectral shape as a function of output.

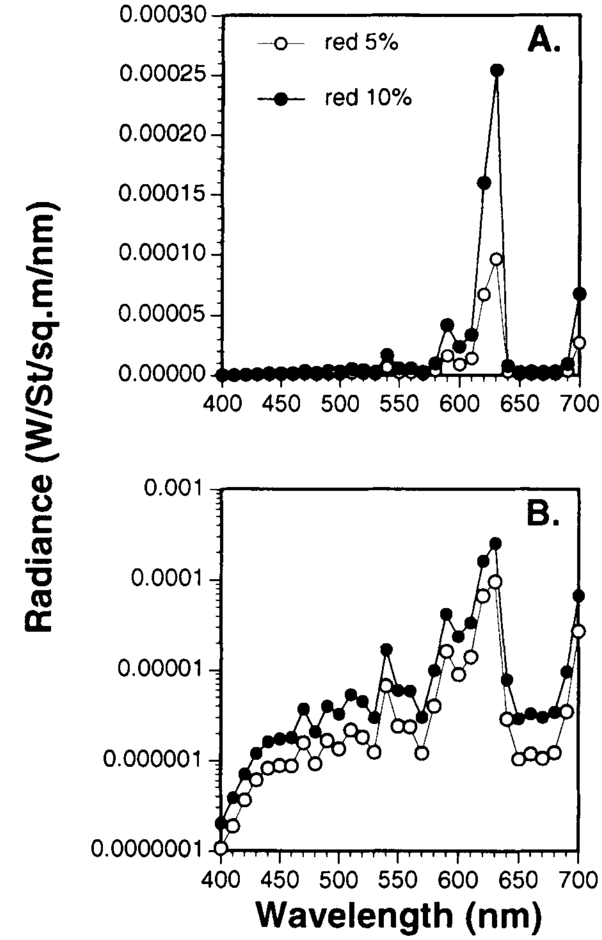

Figure 7. A check of phosphor constancy. (A) Spectral output of the red gun operating at 5% and 10%. (B) Same as A, but plotted on a logarithmic scale. If phosphor constancy holds, then the two curves in B should have the same shape and should be superimposed when slid vertically.

In order to determine the significance of the lack of phosphor constancy, Figure 9 shows the relative chromatic difference for each operating level from a nominal value (50% output) when only one gun is operating. Recall that in CIE 1976 (u',v'), space increasing u' relates to increasing redness, whereas increasing v' relates to increasing blueness. In Figure 9, the desaturating effect of low outputs on red and green is evident as the red phosphor becomes less red and yellow and the green phosphor becomes less green and yellow. On the other hand, the blue phosphor lost saturation as operating levels were increased. To analyze the significance of the chromaticity changes, we treated these separately, and assumed (for calculation purposes) that they were not accompanied by any corresponding changes in luminance. Once this was done, we analyzed chromaticity changes in isolation by considering C^*_{uv} (chromatic difference) against operating level (see Figure 10). This is scaled in terms of JNDs between the measured color and our nominal (50%) color.

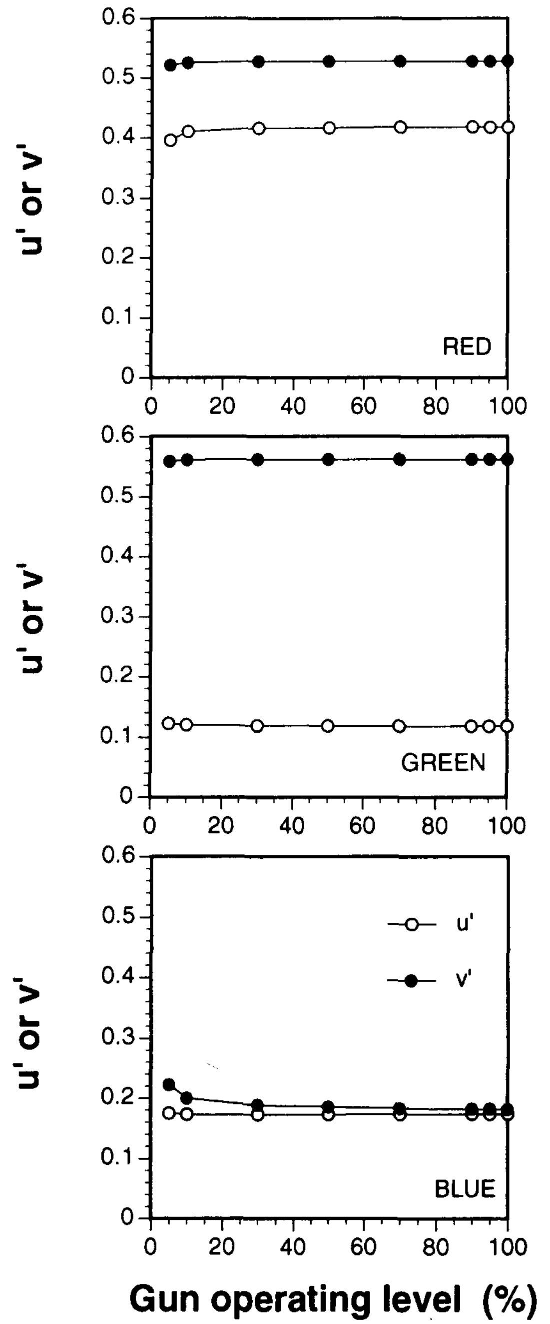

Figure 8. A check of phosphor constancy over a range of operating levels for each gun, using chromaticity coordinates (u' and v').

To get a feel for these data, it is worth noting that the color differences (C^*_{uv}) of adjacent caps in the commonly used Farnsworth D-15 panel test of color vision is about 12 JND (range 9–17) and that most color-normal observers find these discriminations easy to make. Although an analysis of this type disregards luminance changes, it alerts us to the possibility that deviations from phosphor constancy may be significant at operating levels of 10% or below, especially for the blue and red guns. This factor is beyond the control of the user; it is enough to be aware of the potential problem and to avoid using this low end of the CRT's dynamic range when color representation is important.

ACHIEVING GUN LINEARITY: GAMMA CORRECTION

Rodieck (1983) noted that a CRT does not behave as a linear device. This means that the voltage applied to the guns is not linearly related to their output in terms of

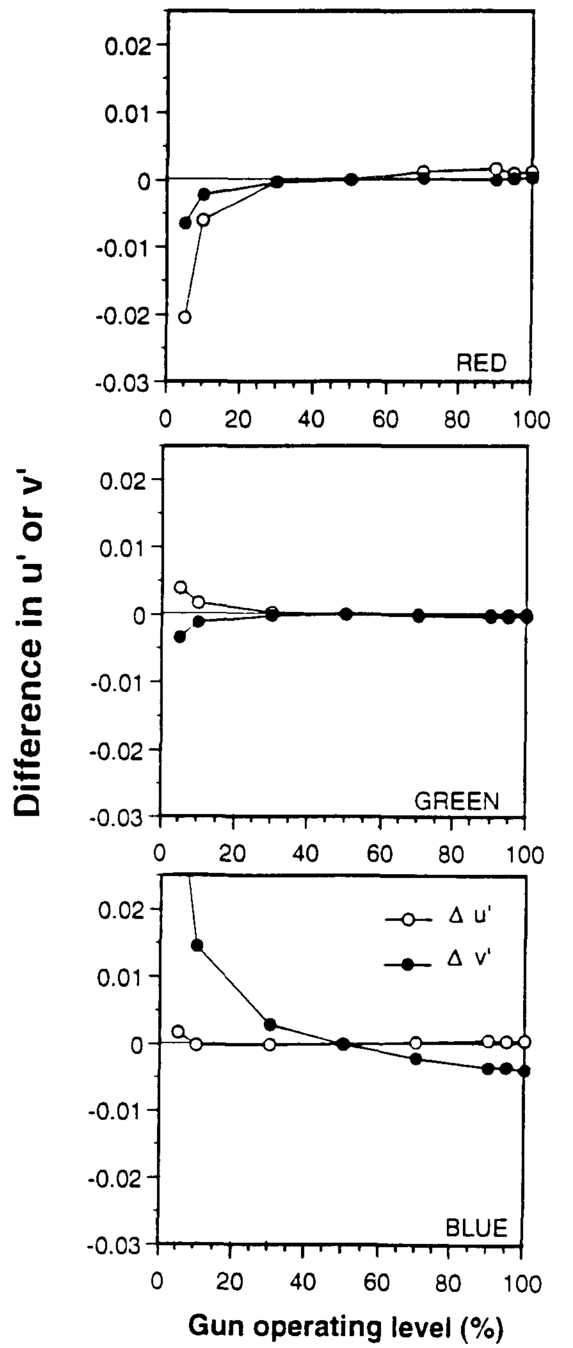

Figure 9. Differences in u' and v' from nominal values (50% levels) for each gun.

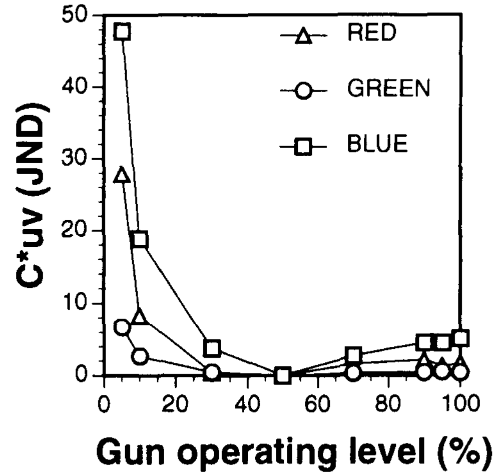

Figure 10. Calculated color difference (C^*_{uv}) plotted against operating level for each gun.

screen luminance or photons emitted. Considering any single gun in isolation (Figure 2), the beam current (I) is controlled by the cathode-to-grid voltage (E) by a power law relationship. This relationship is the basis for the nonlinearity and leads to the gamma relationship

$$\Phi(E) = \beta E^{\gamma}, \tag{12}$$

where $\Phi(E)$ is some physical measure of output (e.g., luminance), E is the applied voltage, β is a constant, and γ is the exponent of the power function. A different β and γ apply for each gun. The standard gamma relationship of Equation 12 fails to consider any dark-light effects—the output of the screen when zero voltage is applied to any gun (from ambient light reflected off the face plate, phosphorescence of the phosphors, etc.). To take this into account, we need to add another term to Equation 12 to yield

$$\Phi(E) = \alpha + \beta E^{\gamma}. \tag{13}$$

The term "α" is the dark light of the system. This form of gamma relationship (Equation 13) has not appeared in past discussions, probably because CRTs have traditionally been used at levels well above those that may be affected by dark light. Moreover, there is no simple transformation to linearize it and thus it is less amenable to curve fitting by using a simple linear least squares fit. However, we suggest that the dark-light term should be included in the gamma correction, especially if the ambient room illumination, which contributes to α, is not zero. This is often the case in visual psychophysical experiments in which an illuminated screen surrounds the CRT. We also propose that nonlinear methods should be used to fit this curve and give examples for our empirical data in Figure 12.

Figure 11 shows the relationship between CRT input (voltage) and output (measured luminance) for the red gun of our Barco system. The lower, hatched, dotted line joins actual measured data points. In the far left-hand corner, it can be seen that there was a measurable output from the screen, even when input voltage was zero (dark light). We suggest that the dark light should be assessed in the environment where the CRT is to be used in order to allow for local variations in ambient light. It also demonstrates how the power law applies up to maximum input voltage and that the output shows no evidence of saturation (similar curves were obtained for the other guns). The lower, thin solid line in this figure shows a curve fitted to the data using Equation 13 as a model. The parameters α, β, and γ have been chosen to minimize the chi-square statistic for the model. A computer program utilizing the iterative Marquardt-Levenberg procedure adapted from Press, Flannery, Teukolsky, and Vetterling (1988) performed this analysis. The estimated errors in these parameters were 0.02%, 0.3%, and 0.03%, respectively.

In most instances, the gun voltages are controlled by a computer graphics system that needs to be linearized with respect to output intensity. This gamma correction can be achieved in a variety of ways; it can be produced by hardware in the form of fixed electronic circuitry that assumes an a priori gamma relationship. However, hardware methods typically lack flexibility and precision. A more flexible approach involves the generation of compensating gamma "look-up tables." Look-up tables (LUTs) are data tables that are called upon at one extra stage of processing in order to achieve the gamma correction. Figure 11 demonstrates how this conversion is achieved so that the signals received by the digital-to-analog converters (DACs) produce linear outputs. The desired input ("a" in Figure 11) for each gun is calcu-

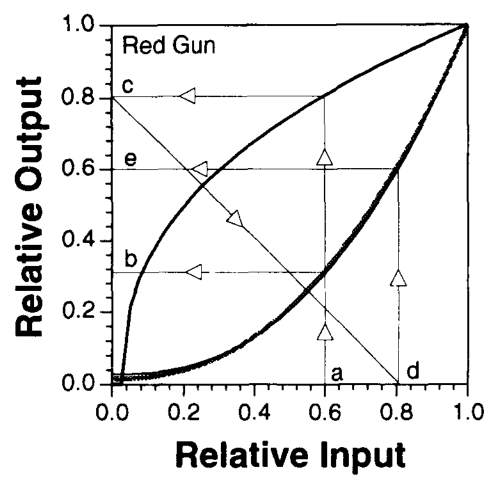

Figure 11. Input-output relationship (gamma function) and its inverse for the red gun. The lower curves are the actual and modeled gamma functions. The top curve is the inverse of the model. Without linearization, a call for 0.6 units would result in an output of 0.31 units (a→b). Using the correction look-up table, a call for 0.6 units (a) would be sent to the look-up table containing a value of 0.81 (c). Applying this number to the DACs (d) would result in an output of 0.6 units (e), as requested.

lated as if the guns were linear, and this set of numbers enters the LUTs, which generate new values ("d" in Figure 11) to be sent to the DACs. This procedure compensates for the nonlinearity of each of the three gun/phosphor combinations.

In Figure 11, we assume that the desired output of the screen is "e" or 0.6. If a linear system existed, the input would be 0.6 units, or "a." However, the nonlinearity indicates that an input of this magnitude to the DACs results in an output of about 0.31 units, or "b." In order to correct the nonlinearity, an inverse function (upper solid curve in Figure 11) needs to be determined to linearize the relationship; it is this inverse relationship that is entered into the LUT. When the desired value "a" enters the LUT, a value of 0.805, or "c," is returned. Indeed, we can now see that when the corresponding number "d" is sent to the DACs, the desired output of 0.6 is obtained at "e."

In our setup, the LUTs are composed of lists of 4,096 numbers (12 bit). There is one LUT for each gun, because each gun has its own characteristics. Because the calibration of a CRT has been shown to change over short periods of time (Cowan, 1987), it is necessary to have a procedure for maintaining accurate and up-to-date LUTs with minimal fuss. To this end, we suggest that an automated procedure be adopted. Our system has been automated so that each gun is operated for a short period of time at 409 levels, increasing from zero to maximum. A telephotometer takes 10 readings of central-screen luminance in rapid succession. These analog outputs are digitized by an analog-to-digital converter (AS-1, Cambridge Research Systems) and the mean and standard deviation of these data are used to determine the best-fitting parameters for the curve described in Equation 13. Once a best fit is obtained by the iterative procedure, the inverse function is obtained and a compensating LUT is generated. For the gamma function given by Equation 13, the appropriate inverse function needed to fill the LUT has the form

$$\mathrm{LUT}(x) = [(x - \alpha)/\beta]^{1/\gamma}, \qquad (14)$$

where x is the desired output, and α, β, and γ are as defined previously. During this procedure, no gamma compensation should be applied to the signals; this can be achieved by setting a linear LUT for that time when the data are recorded. Although the entire calibration procedure takes less than 15 min to complete, there is obviously a tradeoff between precision and the time needed for recalibration. Our experience indicates that weekly to monthly checks produce little change in the resultant gamma function.

The tolerances described earlier for the red gun are typical of those usually found and indicate that a good fit between the data and the model was obtained. Note that for the values where α is greater than x (when the desired output is less than the dark light of the system), the LUT function becomes undefined. This simply reflects the physical impossibility of the situation and values of zero should be entered into the LUT whenever this occurs. This will set a lower operating limit on the CRT, which may reflect the residual ambient light in a particular laboratory or the noise of a particular CRT.

Once gamma compensation is applied to the LUTs, the resulting input-to-output relationship is fairly linear, as shown in Figure 12. This relationship for gun input to output was determined 2 days after the LUTs were generated and demonstrates the stability of the calibration over this time period. The fact that a straight line fitted to the data does not intercept the output axis at zero for the upper curve (red gun) is evidence that there was some unaccounted ambient light present during this particular measurement. This example serves to illustrate the need to generate LUTs in conditions identical to those under

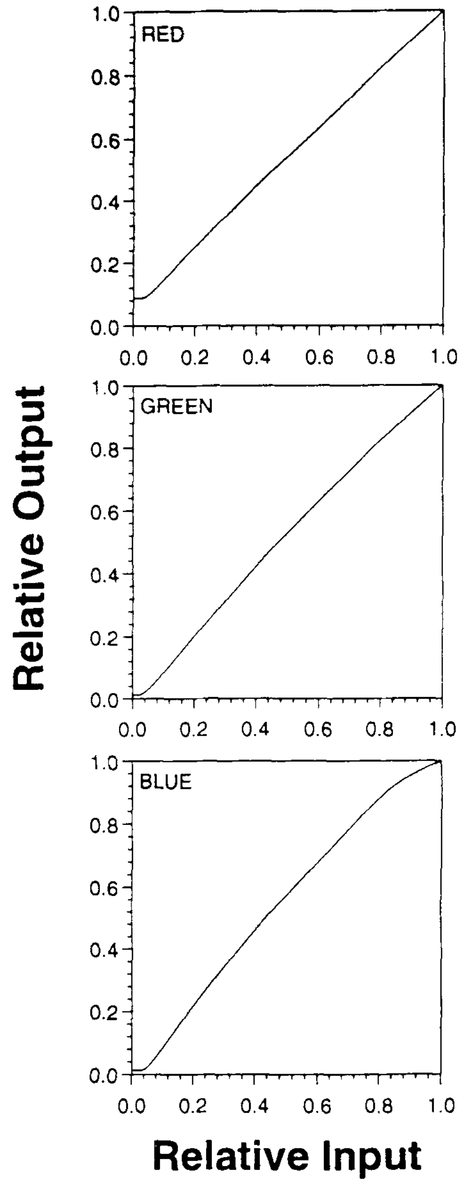

Figure 12. Input-output relationships for all guns 2 days after gamma correction had been applied. Note the "dark light" at the lower left of each curve.

which experimentation proceeds in order for the calibration to be valid. It also demonstrates that errors are likely to be insignificant if operations are kept within 10% and 90% of the operating range.

Note that Figure 12 shows a sharp departure from linearity at levels below 5% of the operating range due to the dark light of the CRT. This is present for all the guns and reminds us that we need to limit our operations to levels above 5% in order to avoid this type of error. The upper operating limit is set by some of the factors discussed previously, including gun interactions.

SPATIAL VARIATION

The light emitted from a CRT is most intense at the center of the screen and drops off significantly at more peripheral locations. Figure 13 illustrates this fact for our Barco CRT. In Figure 13, the white-point luminance and chromaticity coordinates (u' and v') were obtained at a different screen position for 50% gun output. Total color differences (ΔE^*_{uv}), which consider luminance changes and chrominance differences (C^*_{uv}), are also shown with respect to the central location. The reason underlying the luminance decrease can be explained in terms of the relative geometry of the electron beam and shadowmask. As the beam scans away from the center, the effective shadowmask hole becomes smaller due to the oblique angle of incidence (Cowan, 1987). In conjunction, the phosphors do not emit light in a Lambertian fashion. The emitted light has a marked polarity, with most of it being released in the same direction as the electron beam and away from an observer aligned with the center of the screen. As Figure 13 shows, the effect can be significant

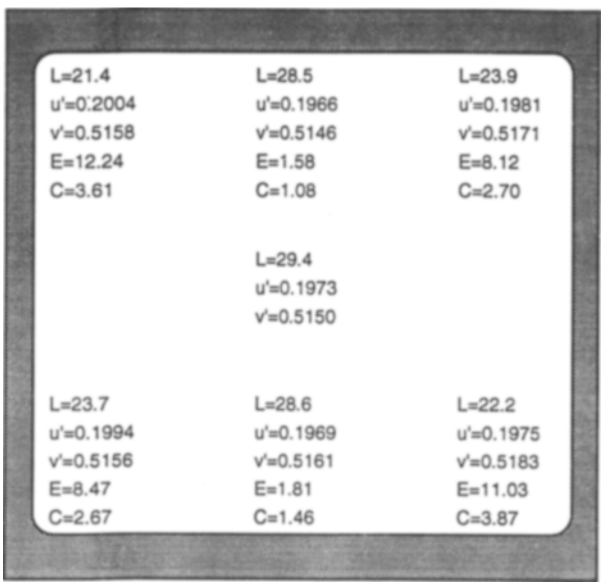

Figure 13. Luminance (cd/m^2) and colorimetric data for the white point at various screen locations. Total color difference (ΔE^*_{uv}) and chrominance difference (C^*_{uv}) are given with respect to the center of the screen.

if extended images are displayed on the screen. However, as most of the color difference is due to a luminance loss toward the edges rather than a chrominance change, this may not be so critical for investigations of color mechanisms because the total chrominance change is only just suprathreshold. Nevertheless, for best results, we recommend that images be restricted to small and symmetrical regions around the center of the screen. Spatial artifacts may also arise in the center of the screen, due to improper convergence of the guns. Vingrys and King-Smith (1986) describe a simple way to check for this artifact by utilizing black and white photography. Alternatively, the human eye could be used with a set of red and green filters to isolate the different phosphors.

RECOMMENDATIONS

A knowledge of the various aspects of CRT monitor performance allows us to make a series of recommendations regarding the general use of these displays. It needs to be stressed that the way in which a CRT is calibrated and actually used depends entirely on the application at hand. The following recommendations have been adopted by our laboratory for the purpose of presenting carefully controlled stimuli on the screen that are different from a white point operating at half the dynamic range of each gun. The CRT, in our case, is surrounded by a large screen that is illuminated to match the white point in luminance and chromaticity.

Regarding spatial variation across the screen, it is advisable to confine stimuli to small symmetric patches near the center of the screen. There is a payoff between spatial extent of the stimulus and viewing distance. Given that this distance needs to be longer than 50 cm to avoid resolution of the pixels, the maximum spatial extent of a circular patch confined to the central third of the screen is about 15°. The data of Figure 13 suggest that color constancy is rather good for the entire extent of the vertical meridian, and that only the corners need to be avoided. This problem of spatial variation may make it difficult to produce large stereograms, for instance, especially if they were meant to be isoluminous.

We recommend that the CRT be allowed to warm up for at least 45 min after being switched off overnight. This can be reduced to 30 min if the CRT has been prewarmed and switched off for under 30 min, or has been operating at some level other than the intended background.

In order to keep the problems arising from lack of gun independence and phosphor inconstancy at minimal levels, it is recommended that the operating range of the guns be maintained within 10% and 90% of maximal output. This would also tend to negate any problems arising from incomplete gamma correction, which is worse at the bottom end of the operating range, due to unstable dark-light contribution arising from changes in room illumination.

Finally, we suggest that the linearization or gamma correction of the gun outputs be done by fitting an equation of the form of Equation 13 to the input-voltage/output-intensity relationships of each gun, and by using Equation 14 to create compensation tables. This procedure needs to be performed at intervals individually established for each CRT by monitoring how these functions change. The data collection of the calibration procedure should be done in conditions identical to that used in the application of the CRT.

REFERENCES

BILLMEYER, F. W., & SALTZMAN, M. (1981). *Principles of color technology* (2nd ed.). New York: Wiley.

CIE. (1978). *Recommendations on uniform color spaces, color-difference equations and psychometric color terms.* (Supp. No. 2, CIE Publication No. 15 [E-1.3.1] 1971/[TC-1.3.]). Paris: Bureau Central de la CIE.

COWAN, W. B. (1983). An inexpensive scheme for calibration of a colour monitor in terms of CIE standard coordinates. *Computer Graphics*, **17**, 315-321.

COWAN, W. B. (1987, October). *Colorimetric properties of video monitors.* Notes for a short course presented at the meeting of the Optical Society of America, Rochester, NY.

COWAN, W. B. (1991). Colour television. In P. Gouras (Ed.), *Vision and visual dysfunction: Vol. 6. The perception of colour* (pp. 294-308). London: Macmillan.

COWAN, W. B., & NELSON, R. (1986). On the gun independence and phosphor constancy of color video monitors. *Color Research & Applications*, **11**(Suppl.), s34-s38.

ESTÉVEZ, O., & SPEKREIJSE, H. (1982). The "silent substitution" method in visual research. *Vision Research*, **22**, 681-691.

HUNT, R. W. G. (1991). *Measuring colour* (2nd ed.) West Sussex: Ellis-Horwood.

MUKHERJEE, A. & VENKATESH, Y. V. (1986). Digital color reproduction on color television monitors. *Computer Vision, Graphics, & Image Processing*, **36**, 114-132.

PRESS, W. H., FLANNERY, B.P., TEUKOLSKY, S.A., & VETTERLING, W. T. (1988). *Numerical recipes in C: The art of scientific programming.* Cambridge: Cambridge University Press.

RODIECK, R. W. (1983). Raster-based colour stimulators. In J. D. Mollon & L. T. Sharpe (Eds.) *Colour vision* (1st ed., pp. 133-144). London: Academic Press.

STROMEYER, C. F., III, KRONAUER, R.E., & COLE, G. R. (1985). Second-site adaptation in the red-green chromatic pathway. *Vision Research*, **25**, 219-237.

VINGRYS, A. J., & KING-SMITH, P.E. (1986). Factors in using color video monitors for the assessment of visual thresholds. *Color Research & Application*, **11**(Suppl.), s57-s62.

WYSZECKI, G., & STILES, W. S. (1982). *Color science* (2nd ed.) New York: Wiley.

NOTE

1. We must emphasize that the results presented in this paper refer specifically to the Barco CDCT 6551 monitor used in our laboratory. We feel that these observations should generalize to other makes of high-quality CRTs. This would not necessarily include all other color displays, such as VGA or SVGA, which are commonly interfaced with computer terminals. Our limited experience with these latter monitors suggests that their characteristics may be variable. Any potential user of a color monitor should apply the techniques outlined in this paper, especially at the extremes of the operating range, as part of a calibration procedure.

(Manuscript received August 18, 1992;
revision accepted for publication March 5, 1993.)