Image Quality Issues for High Resolution TFTLCDs

Steven L. Wright, Robert W. Nywening and Steven E. Millman
IBM Watson Research Center, Yorktown Heights, New York

James Larimer, Jennifer Gille* and Jeff Luszcz*
NASA Ames Research Center, Moffett Field, California
****Raytheon Systems, Ames Research Center, Moffett Field, California***

Abstract

Image quality issues for high resolution TFTLCDs are discussed. For text, very high resolution is needed. For color images, the intrinsic color quality of TFTLCDs has the potential to exceed CRTs. As the spatial resolution of TFTLCDs increases, the number of gray levels can be reduced without degrading image quality. This will reduce the bandwidth required to drive these displays.

Introduction

Thin Film Transistor Liquid Crystal Displays (TFTLCDs), used in notebook computers, are appearing in the desktop monitor marketplace at prices currently a factor of two greater than similar CRT displays. Cost reduction for desktop TFTLCDs is linked to manufacturing volume and to the number of display panels which can be made per sheet of mother glass. Compared to CRTs, TFTLCDs are thinner, lighter and require less power. They do not flicker, produce less glare, and have low susceptibility to electromagnetic interference. TFTLCDs are now beginning to provide image quality which exceeds CRTs.

The largest performance drawback of TFTLCDs has been the narrow acceptable viewing angle for TN-mode liquid crystal (LC) cells. This is important for large-area desktop displays, where the corner-to-corner subtended angle for a single user can be large. However, dramatic improvements in viewing angle have been achieved with in-plane-switching cells[1] and vertically-aligned LC cell modes,[2,3] and there is continued development of compensation films for all modes.[4,5,6] The viewing angle of desktop TFTLCDs will continue to improve, not only for gross effects such as image reversal, but also for subtle aspects of luminance gamma and color fidelity.

Another problem area for TFTLCDs is video quality, which is generally recognized to be poorer than CRTs. There is little agreement on the severity of the problem. Different scan conversions required for each type of display provide different quality video signals to the display, resulting in noticeably different image quality. DVD video on TN-mode notebook computer displays looks good.

The reconstruction of continuous motion and flicker are mutually exclusive characteristics. The fast response and decay of CRT phosphors causes flicker, which requires a high frame rate to hide. The slow response speed of LC cells and pixel latching leads to low flicker. Frame rates as low as 20 Hz have been applied to notebook TFTLCDs in our laboratory to reduce power consumption, without objectionable flicker. The slow response of LC materials, however, does reduce the contrast of moving portions of the image sequence, and needs improvement. This is particularly true for wide-viewing-angle technologies such as in-plane switching. Improvements in LC materials and data protocols will lead to improved video on LC displays.

Recent comparisons of TFTLCDs and CRTs are given in reference 7, which can be updated by the developments mentioned here. TFTLCDs can economically provide very high pixel densities within a digital drive environment. This is a strong advantage over CRT technology where similar features are prohibitively expensive or impossible. CRTs have high quality color,[8] but there are aspects of TFTLCD technology which suggest that even this quality can be exceeded. For example, it has been proposed to utilize high pixel density to accurately render colors with a limited palette,[9] potentially important for web-based commerce.

Pixel Density and Pixel Count

For many computer applications, displaying more information on a single screen[10] improves user efficiency and utility. Information content can be in any form, such as text, numbers, graphics, or images. The maximum displayable information content depends upon screen size, pixel density, pixel independence, color gamut, pixel architecture, number and spacing of grayscale levels, peak brightness, and contrast. Some of these factors affect the amount of data on the screen, others affect the visibility of this information. Applications include text and document processing, engineering design, command and control, medical imaging, insurance processing, pre-publication review, and web-based commerce. The highest pixel density and pixel count in a display to date has been achieved in a prototype TFTLCD from IBM[11,12] This display has a pixel pitch of 0.126 mm, corresponding to 201 dots per inch (dpi), and a 16.3" diagonal. The format is QSXGA (2560x2048), corresponding to 5.24 million color pixels, or 15.7 million subpixels, with 18-bit color. For comparison, the standard image format on a Kodak PhotoCD is 3072x2048, nearly adequate to faithfully render 35mm film images.[13] The Sony DDM-2802 28" CRT has a pixel format of 2048x2048, dot pitch of 0.31 mm (82 dpi), luminance of about 80 cd/m^2, and weight of about 230 pounds. Hitachi 21" monitors have a 1600x1200 format with a smaller dot pitch of 0.22 mm (115 dpi). CRT dot pitch is not a particularly meaningful parameter, since the beam spot size can be a factor of two larger than this.

Displays are needed that can render two full pages of text (and images) with paper-like quality. With increased information content, productivity improvements in the range of 20 to 30% have been obtained in CAD drawing tasks by simply reducing the amount of scrolling and zooming necessary to perform a specific task. The image quality of characters rendered with CRTs is poorer than paper,[14] and this causes reading speed to be less with CRTs than paper for small fonts. As the CRT pixel density increases, reading speed increases.[14] Letter-counting and reading tasks have also been used to quantify differences in subject performance and preference using TFTLCDs with different pixel densities as compared with CRTs.[15] The effective acuity of the display is determined largely by pixel density and sharpness, and this also correlates with subject preference for high pixel densities for displayed text. Small letters can be rendered much more sharply on a high resolution TFTLCD than on a CRT. However, the discrete nature of TFTLCD pixels creates "jaggies" in letters when the individual pixels can be perceived by the viewer. Also, some font designs may have a thin appearance when rendered with discrete and high density pixels, since the line widths can only be integer multiples of the pixel widths. For large fonts, the fuzziness of CRT pixels actually ameliorates these kinds of problems. However, for smaller fonts, discrete TFTLCD pixels are at a distinct advantage, provided the pixel density is high enough, because the eye also blurs the pixels.

The preferred viewing distance depends upon a number of factors, of which image quality is only one. The application and other ergonomic factors play a large role in determining the viewing distance. Newspaper print sizes have evolved so that there is about a factor of two safety margin, that is, typical newspaper fonts are just legible when the paper is held at a distance about a factor of two larger than typical reading distances.[16] For color prints viewed at 25 cm, it has been estimated that 250 pixels per inch are sufficient for high-quality continuous tone prints, or 250 lines per inch with 10x10 microdots for halftone prints.[17] Surprisingly, there is no simple mathematical framework by which to compare what can be observed in high quality prints and high-definition RGB stripe pixels in a display. Given the demand for higher resolution in print technology, it is dangerous to underestimate what can be distinguished in an image, particularly by the younger population, having an average visual acuity of about 20/16. A person with 20/20 vision can resolve features which subtend about 1 minute of arc, but vernier hyperacuity is about a factor of ten sharper than this, as demonstrated by the vertical lines in Figure 1. The vertical line to the left has a displacement or "kink" which can be perceived more easily than a dot of the same size, to the right.

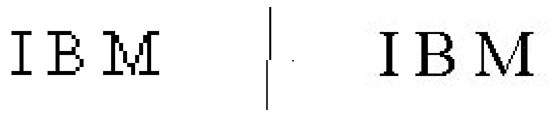

Figure 1. Jaggies and vernier hyperacutiy. At a distance of 100 cm, these letters appear as 12-point letters would at 50 cm on 100 and 200 dpi displays. At 100 cm, the line displacement and dot subtend about 40 arcsec.

A number of display sizes, formats, and pixel densities have been simulated using a screen capture and print approach. Figure 1 illustrates the jaggies which appear for two pixel densities. Although small letters can be rendered much more sharply on a high resolution TFTLCD than on a CRT, pixel densities exceeding 200 dpi may be required to eliminate visible jaggies at preferred viewing distances. Another problem is accurate kerning for WYSIWYG, which depends on the number of horizontal bins available for letter placement. An SXGA display, which contains 1280 columns, is insufficient to properly kern a page of text.[18] If 8.5" wide paper were divided into 1280 bins, this would correspond to 150 dpi. Horizontal pixel counts of at least a factor of two larger than this are needed to render a single page of text without noticeable kerning artifacts.

Very high resolution displays will not provide greater utility to the user, however, unless other system aspects are improved as well. The display drive architecture must be suitable and both operating system and software application improvements are needed for proper icon and font scaling. High resolution displays are best characterized by screen size and pixel density just as printers are today. Display formats, such as XGA, SXGA, etc. lose utility when dot pitch exceeds 130 dpi.

Luminance and Gamma

Most of the information content in an image is contained in the spatial luminance signal. This characteristic of human vision forms the basis of data compression schemes, such as JPEG, and television broadcast bandwidth allotment. The display transfer function, the relationship between screen luminance and applied digital level, can dramatically alter the appearance of images. If a plot of log luminance versus log level follows a straight line, then the slope is defined as gamma. In most applications, however, the transfer function deviates from this power law, usually at low intensity levels. Before color quality can be seriously considered, the luminance characteristics of the display must be thoroughly understood and optimized. The luminance-voltage characteristics of liquid crystal displays are different than CRTs, but the luminance to digital level relationship may or may not be similar.

For CRTs, the phosphor luminance results from the current of the electron beam impinging on the phosphor. The beam current-voltage relationship generally follows a triode 5/2 power law, where the voltage on the electron gun is determined by the combination of the electronics in the display adapter card contained within the computer and high voltage amplifiers contained within the display. All of the signals in the path from the display adapter card to the gun are analog. The digital levels from the computer are converted to analog voltages, usually in a linear fashion, with an offset, i.e. a digital level of zero results in a non-zero analog output.[19,20] The amplifiers in the CRT need not be linear, either. Although much of the CRT gamma relationship is determined by the electron gun, the overall display gamma is a function of all components, including the display adapter card.

For TFTLCDs, the pixel transmittance depends on the polarization state of the liquid crystal layer, which depends on the RMS ac voltage across the cell. For normally white TN-mode cells, the transmittance follows an inverted “S-shaped” curve. Typical structures are fully transmitting (bright) for a pixel voltage of 1 V or less, and fully opaque (dark) for a voltage of about 5 V. The relationship between digital level and data voltage applied to the columns in the array is determined by the set of reference voltages supplied to the driver chip. These are set by a resistor chain with voltage taps. There is a nonlinear relationship between digital level and data voltage. For 6-bit column drivers, the first 3 bits are determined by 9 reference voltage values; the remaining 3 bits are done by linear interpolation. The reference voltages control the luminance-level relationship, which can be chosen over a wide range. For notebook displays, these values are chosen largely to optimize operating characteristics rather than the quality of still images. For example, it is desirable to achieve maximum brightness, and minimal mouse cursor disappearance with movement. Typical notebook TFTLCDs have gamma values between 3 and 4, much larger than CRTs. Different computer manufacturers have adopted different CRT gammas; for PC, Macintosh, and SGI, the standard gammas are 2.2, 1.8, and 1.0 respectively. Considering all factors of human visual characteristics, computer and display system environment and practice, a gamma value for TFTLCDs of 2.2 is a reasonable choice.[17,20,21] If needed, the gamma could be changed in the display “on-the-fly” for image data files which have been created for a particular computing environment but are being displayed in another environment. Given the current variability and evolution of TFTLCD characteristics, some form of ICC monitor profile and operating system specification is needed.

Another important area of concern for TFTLCDs is luminance contrast. Under bright ambient conditions, the contrast ratio of TFTLCDs is much higher than CRTs, due to differences in screen reflectivity and peak brightness. However, under dark viewing conditions, CRTs have dimmer dark states, and the contrast ratio of CRTs can be very large. This contrast can be utilized to better render images which contain detail in dark areas of the image. Optimal rendering of images on TFTLCDs viewed in dark ambients will require improved dark states. The white state luminance can be increased by increasing the backlight intensity, commonly done for avionic and military applications. This will not improve the contrast ratio in dark ambients, but it does improve apparent color saturation, peak brightness and contrast in bright ambients. Peak brightness is important for rendering image highlights. To avoid image artifacts due to additional TFT photoleakage,[22] improved black matrix structures are needed in the array. The black state can be improved by cell structures that don’t contain spacer balls, with better cell gap uniformity, and improved compensation films. For TN-mode cells, the spacer balls leak light in the dark state. On average, there is about one spacer ball per pixel. A variation in cell gap de-tunes the optimal extinction ratio in the black state. Crossed linear polarizers typically have an extinction ratio of about 600, which falls off for viewing angles far from normal incidence.[23]

Color Gamut, Stability, and Fidelity

The color gamut of TFTLCDs is determined largely by the backlight spectral power distribution, which is matched by the color filter absorption characteristics. For many notebook displays, the color gamut has been intentionally reduced by broadening the spectral transmission through the color filter. This makes the colors less saturated, but increases battery life. Any light absorbed in the color filters represents a loss in efficiency that raises the backlight flux required to reach a desired brightness level. For desktop TFTLCDs, the color gamut can be increased, since power is not a major design consideration. One way to improve the color gamut of TFTLCDs is to replace present cold-cathode fluorescent lamps (CCFLs) in the backlight with light-emitting diodes (LEDs). LEDs emit a narrower band of wavelengths compared to phosphors. Recent materials advances have led to the development of blue and green LEDs with chromaticities close to the corners of the spectral locus, as shown in Figure 2. The color gamut of LED primaries is considerably larger than CRT P22 phosphors or the phosphors used in CCFLs.[24] LEDs can also be rapidly switched, enabling field sequential color operation without color filters.[24]

In spite of the good color performance of CRTs, and advances made in calibration and color management, surprisingly little use is made of CRTs in situations where accurate color is required. Some of the problem has to do with viewing ambient, the color temperature of the monitor, and other fundamental problems with comparing emissive display images with printed paper. Due to the high reflectivity of phosphor materials, changes in the viewing environment can have a large effect on CRT dark colors and image quality. CRT colors drift over time, due to changes in the analog and high voltage electronics, and to phosphors aging at different rates.[25] The color stability of TFTLCDs is potentially better than CRTs since the spectral characteristics of color filters and LC are stable. The phosphors used in backlights are similar to CRT phosphors and exhibit the same aging properties. Most importantly, the drive environment for TFTLCDs can be made largely digital, which is inherently more stable than analog. A major factor limiting most TFTLCD monitors is the use of an analog signal path to preserve CRT compatibility. The digital signals for TFTLCDs need only be converted to analog at the driver chip outputs on the glass, located close to the edge of the array.

Another important monitor characteristic is the chrominance shift of the primaries and gray point with drive level. The chromaticities of a typical notebook computer TFTLCD are shown in Fig. 2, as a function of drive level. In tristimulus XYZ color space, if all gray colors (R=G=B) from black to white lie on a line through the origin, then these grays project to the same point in the xy chromaticity plane. As with printing, this characteristic is desired to avoid hue shifts in light and dark regions of the image. For the primary colors, chromaticity values shift from the saturated values at level 256 down to the black point of the display, at level (0,0,0). If the luminance of black is negligible, then the chromaticity of a primary will slowly converge to the chromaticity of the black state. Brighter black states result in faster convergence. Crosstalk between neighboring sub-pixels, correlated with drive level, will result in deviations from a linear locus in the xy chromaticity plane.

The rate at which the chromaticity values converge toward the black point can make accurate rendering of dark colors difficult, but differences in dim colors are difficult to perceive. For properly adjusted CRTs, the chromaticity of primaries shift very little until the level falls below about 20 out of 256, with little chrominance shift of the gray state.[26] For TN-mode TFTLCDs, there is a wide range of behavior, primariliy due to designs optimized for notebook applications. Generally, there is a chrominance shift for the gray state with level, due to different transmission characteristics of the primaries. The hue shift for the primaries is consistent with capacitive crosstalk in the array. TN-mode pixel chromaticities fall much more rapidly toward the black point than CRTs due to brighter black states. This rapid shift can begin to occur for levels as high as 120. This characteristic is a veiling gray effect: the primary colors loose saturation as the level is reduced. These chrominance shifts degrade the color performance for dark colors, however, with suitable corrections the perceptual color error can be reduced to acceptable values.[26] As the dark state luminance is decreased, this problem will diminish.

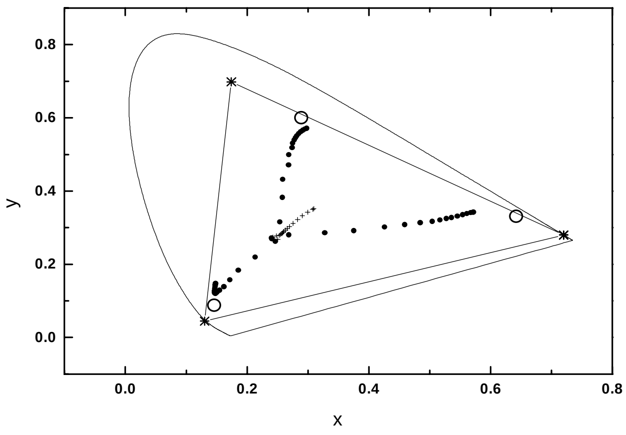

Figure 2. Chromaticities of primary and gray colors of a typical notebook TFTLCD, for all drive levels. Shown are CRT primary chromaticities (O) *and the color gamut for LED primaries* (*).

One characteristic shown in Fig. 2 which can occur is a peak in the chromaticity locus for the blue primary, corresponding to a peak in z or Z occuring below level 256. This is a consequence of a cell gap thickness which is optimized for green pixel operation and minimal power. The Gooch-Tarry criterion is met only for a narrow band of wavelengths, and this guarantees that the transmission-voltage characteristics of all three primaries will not be exactly the same.[27]

Utilizing High Resolution to Improve Color

High resolution TFTLCDs can accurately render colors in natural images by dithering a limited color palette. This reduces the size of the data file required as spatial resolution increases and reduces system costs. Dithering to obtain intermediate colors works because as dot density increases the eye's ability to resolve individual pixels diminishes. The limiting resolution of the eye acts as an averaging filter. It becomes possible to generate colors missing from the pixel palette by dither with the limited palette and subsequent blur filtering by the observer's eye.

The major web browsers have adopted a minimal color palette of 216 RGB triplets, but questions remain regarding what spatial resolution will be required with this palette and dithering to produce all colors without introducing noticeable artifacts in the rendered image. It has been proposed[9] that spatial dithering with the browser-safe colors is a good strategy for images at spatial resolution of 150 dpi or greater. Experimental validation of this approach was done using a prototype 157 dpi SXGA TFTLCD from IBM,[28] TN-mode, with 6 bits/color. Three different grayscale spacings were tried: geometric, linear, and power law with an exponent of 2.2.

Calibration

The luminance of each primary was measured at all 64 levels, and the chromaticities for each of the primaries and their combinations at maximum and midpoint drive levels. To remove the veiling gray effect from the experiment, for each primary, the minimum drive level was chosen such that the luminance was 4% of the maximum. Although not strictly true, it was assumed for image processing that the chromaticity did not change with drive level. This resulted in 51,50, and 62 useable levels for the red, green, and blue primaries. The 7 digital images used in the experiment were translated to luminance images by assigning the minimum luminance for each primary to digital count 0, the maximum luminance to 255, and the remaining values as if they had been generated for a Macintosh computer, with a system gamma of 1.8.

Stimuli

Four natural images and three computer-generated images were used. Three of the four natural images were digitizations of photographs, chosen to have qualities such as light and dark regions, smoothly-varying, and highly textured regions. Two of the computer-generated images were built from digital-count ramps of color mixtures, one with smooth corner-to-corner shading over a square, and the other with four horizontal strips with ramps of black-to-white, blue-to-yellow, green-to-magenta, and red-to-cyan. The third computer-generated image was a passage of ramped text, where both the text and background were ramped black-to-white. The resulting text had high contrast in the upper right and lower left corners, but of opposite polarity, and no contrast on the upper left to lower right diagonal. All seven images were 500x500 pixels, originally expressed in 8 bits/color (256 levels).

For the standard, high grayscale resolution images, Floyd and Steinberg error diffusion was used to dither to the actual, calibrated luminances. For the test images, with lower grayscale resolution 2-16 levels per primary, the images were similarly dithered in luminance, but the attainable luminance values were chosen in the following ways: 1) Linear luminance values, by dividing the luminance range into equal steps. 2) Power-law luminance values, in which equal steps in fractional drive level were raised to the power 2.2, multiplied by (L_{max}-L_{min}), then added to L_{min}. 3) Geometric luminance values, in which the steps depend logarithmically on luminance. In all cases, the closest calibration values to the ideal values were used. The linear steps are large in the dark portion of the range and small in the light region. The geometric and power-law steps are small in the dark portion and large in the light region

Procedure

A Method of Adjustment procedure was used to determine threshold grayscale resolution for each of the seven images rendered with each of three spacings, viewed in a dark ambient, at distances of 50, 62.5, and 75 cm. There were 3 observers, with normal (corrected) vision. Images were 8.1 cm square, centered on a gray background, and further masked to produced a diffuse surround. The presentation of stimuli was controlled using the computer keyboard.

For each image at a given distance, the observer was told to search for the image in a sequence that looked "just good enough", by increasing or decreasing the number of graylevels until the threshold number of levels was determined. Each adjustment began at a random position in the sequence. The image was presented for 3 seconds, after which the observer chose whether to increase or decrease or leave unchanged the number of gray levels until the image threshold was determined. Three replications in a row for each image were performed, and the entire data set replicated over days.

Results

The shallow ramps were markedly different from the other images, requiring more grayscale levels at all viewing distances. The averaged results of the two ramp images are shown in Fig. 3. At the shortest distance for these stimuli, only the full resolution images using over 50 levels per primary were "good enough". At all other levels for these stimuli the observers could see objectionable artifacts. These were typically banded regions, possibly Mach Bands, that appeared. The results for other images are shown in Fig. 4. Geometric spacing yields lower or equivalent grayscale levels relative to all other spacing for all stimuli used in this experiment. This suggests that geometric spacing of luminance (logarithmic spacing of levels) is best as expected from Weber's Law.[29] At the near distance the minimal number of levels that on average looked "good enough" was 20, or 4.3 bits for geometric spacing. Linear spacing is clearly inferior to the other spacing rules in that it requires more levels to look "good enough" for all stimuli and all viewing distances.

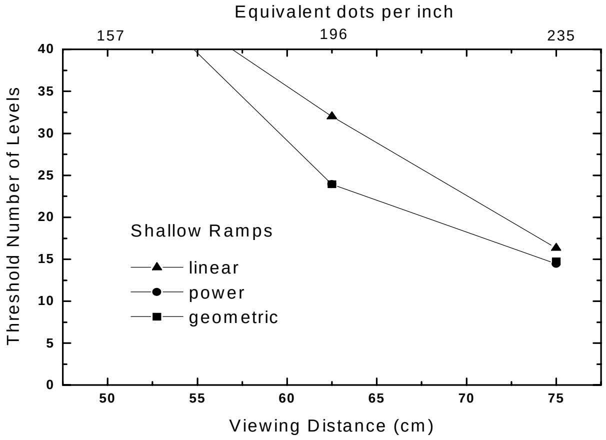

Figure 3. Minimum number of grayscale levels for shallow ramp images, averaged over three observers.

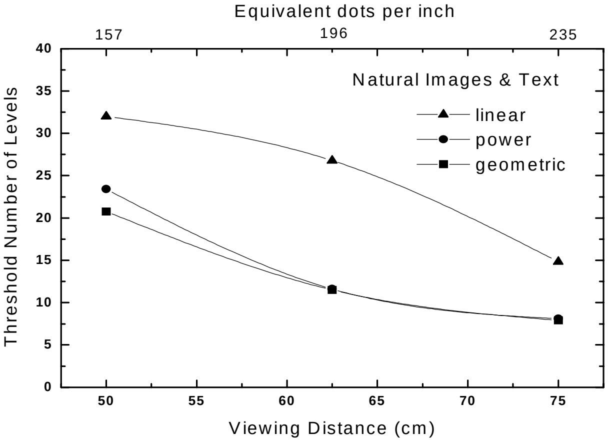

Figure 4. Minimum number of grayscale levels for natural and text images, averaged over three observers.

Conclusions

The potential for TFTLCDs for high image quality applications has been largely ignored due to the known viewing angle and video problems inherent in this technology and the high cost of flat panel displays. Both of these problems are rapidly diminishing, and the inherent advantages of TFT-LCDs for high resolution and pixel count, pixel sharpness, brightness, and improved color are now emerging for desktop monitors. Given the current rate of cost reduction and performance improvement of TFTLCDs, this technology may soon dominate the high-end monitor market. As higher resolution devices begin to appear, the bandwidth required to address these displays will also increase. The results of our experiment show that a savings is possible by reducing grayscale resolution and dithering as spatial resolution increases. It is once again confirmed that geometric spacing is best. However, the traditional power law spacing is only slightly inferior to logarithmic spacing, so no change in the conventional power law spacing seems warranted.

References

1. M. Oh-e, M. Ohta, S. Aratani, K. Kondo, Int. Display Res. Conf. , p.577, (1995).
2. A. Takeda et al., SID '98 Int. Symp. Digest, p.1077.
3. A. Lien et al., SID '98 Int. Symp. Digest, p.1123.
4. T.A. Sergan, J.R. Kelly, SID '98 Int. Symp. Digest, p.694.
5. Y. Saitoh et al., SID '98 Int. Symp. Digest, p.706.
6. H. Yuan et al., SID '98 Int. Symp. Digest, p.714.
7. L.D. Silverstein, pp. 119-132, in Display Systems, pub. Wiley & Sons, 1997; and L.D. Silverstein, SID '98 Seminar Lecture Notes, F-5 (1998).
8. W. Cowan, pp. 294-307, Ch.15, in The Perception of Colour, by P. Gouras, pub. CRC Press, 1991.
9. J. Gille, J. Luszcz, J. Larimer, Proc. SPIE **3299,** 368 (1998).
10. P. M. Alt, K. Noda, IBM J. Res.Dev. **42**, 315 (1998).
11. IBM Technology Fair, Sept. 14-17, 1998, Toyko, Japan
12. K. Schleupen et al., Int. Display Res. Conf., (1998).
13. R.W.G. Hunt, Proc. 5th Color Imaging Conf., p.1 (1998).
14. J.E. Sheedy, J. Electronic Imaging **1**, 405 (1992).
15. I.L. Bailey et al., joint U.C. Berkeley/IBM study, *unpublished.*
16. I.L. Bailey, U.C. Berkeley, School of Optometry, *personal communication.*
17. The Reproduction of Colour, by R.W.G. Hunt, pp. 689-690, 5th edition., pub. Fountain Press, 1995.
18. S. Millman, IBM Research Report RC-20524, Aug. 1996.
19. R.S. Berns, M.E. Gorzynski, R.J. Motta, Color Res. Application, **18**, 315 (1993).
20. A Technical Introduction to Digital Video, by C.A. Poynton, pub. Wiley, 1996.
21. Color Science: Concepts and Methods, Quantitative Data and Formulae, by B. Wyszecki and W.S. Stiles, pub. Wiley, 1982.
22. M. Kodate et al., Proc. Int. Display Res., L-62, (1997).
23. J. Chen et al., SID '98 Int. Symp. Digest, p.315.
24. F. Yamada, Y. Sakaguichi SID '98 Int. Symp. Digest, p.149.
25. R. Cappels, SID '96 Int. Symp. Digest, p. 875.
26. A. Cazes, IBM Watson Res. Center, *personal communication*
27. Liquid-Crystal Displays Addressed by Thin-Film Transistors, by T. Tsukada, pub. Gordon and Breach, 1996.
28. E. Colgan et al., IBM J. Res. Dev. **42**, 339 (1998).
29. J. Larimer et al., Int. Displ. Res. Conf., p.262, (1997).