Desirable Improvements in Displays

Robert W.G.Hunt; University of Leeds; Leeds, LS2 9JT, United Kingdom

Abstract

Properties of displays in which improvements are needed include: greater consistency between pictures displayed from the same signals on different devices; greater colour gamut; less impairment by ambient illumination; and better resolution. Some of the ways in which these needs may be met are reviewed.

Current displays

It is convenient to divide current displays into medium-size (from about 20 cm up to about 125 cm diagonal (such as are commonly used in homes and offices), small (up to about 15 cm diagonal, such as are commonly used for individual seat displays in aircraft, and for displays on camcorders, digital cameras, and video telephones), and large (such as are used in projection devices and in out-door presentations). The shadow-mask cathode-ray tube (CRT) has dominated the medium-size market until recently, but liquid crystal displays (LCD) are now dominant for computer monitors, and, together with plasma panel displays (PDP), are beginning to dominate the domestic television market. Small displays are nearly all LCDs, but organic light-emitting diodes (OLED) are making an impact also. Large displays include banks of CRTs, projection devices based on CRTs, LCDs, or digital mirror devices (DMD), and light-emitting diodes (LED); some of the projection devices use the field-sequential principle to reduce costs.

Desirable improvements in displays

Areas where improvements in displays would be welcome include the following: consistency (of picture quality from one type of display to another), colour gamut, freedom from the effects of ambient light, and resolution.

Consistency

It is a common experience that the picture quality produced from the same signals on different devices is significantly different. A lecture prepared on a CRT monitor, looks different when viewed on a laptop LCD, and different again when displayed by an LCD or DMD projector. The factors in the devices that contribute to these differences include: the luminance dynamic response function, the colour of the white, the colours of the primaries, the angle of viewing in the case of older LCDs, and the luminance of the surround. (Although consistency can be improved by using colour-management profiling tools, this places an additional burden on the user and can be only partially successful; it is therefore desirable to remove the causes of the inconsistencies.)

A proposed standard default colour space, known as sRGB [1] includes a luminance dynamic response function based on broadcast television standards and offers a target for displays generally. The luminance dynamic response function of CRTs is affected mainly by the offset and the gamma, and these need to be set to correspond to the standard; television receivers using CRTs were, in the past, quite variable in picture quality from one set to another, but the consistency is now much better as a result of careful setting of offset and gamma during manufacture. Other display devices need to have factory settings that produce results that are close to the standard.

The colour chosen for the white in displays is not consistent, some using the standard D65, and others using the bluer 9300 K or whites of even higher correlated colour temperatures. In some displays it is possible to achieve greater luminance by using bluer whites, and this has been a reason for their adoption; however, the D65 white is included in sRGB, and its universal use in displays would remove a serious source of inconsistency.

The chromaticities of the red, green, and blue primaries used in displays affect the colours in the images. There are many factors involved in the choice of the primaries in different devices, such as cost, permanence, efficacy in producing light, ease of manufacture, and availability. Generally speaking, the greater the colour saturation of the primaries the lower the picture luminance and vice versa. In 1990 the CCIR, in its Recommendation 709 [2] adopted a set of chromaticities for high-definition television, and these are included in sRGB. The universal use of these primaries would promote consistency of colour rendering between different types of device.

In the case of older LCDs the appearance of the picture may change with the angle of viewing, so that, for consistency, the viewing should then always be as nearly as possible at right angles to the centre of the display

The lower the luminance of the surround to a picture, the lower is its apparent contrast. Compared to the average luminance of the picture, surround luminances may be similar (as in viewing typical monitors or reflection images), appreciably lower (as in viewing typical domestic television), or extremely low (as in typical cinemas); consistency of appearance in these different viewing situations requires pictures of different gammas to be displayed; these can now be determined by using the CIECAM02 Colour Appearance Model [3].

Colour gamut

The gamut of chromaticities that a display can produce is limited to the area within the triangle formed by the three points representing the red, green, and blue primaries in a chromaticity diagram, and this is usually smaller than the gamut of colours typically occurring in the real world [4]. It is very important to plot these gamuts on an approximately uniform chromaticity diagram, such as the CIE u',v' diagram; the use of the older CIE x,y diagram can give very misleading results. All practicable red primaries lie near the spectral locus, and green primaries are usually chosen that lie near it in the yellowish-green region; this ensures that the red, orange, yellow, and yellow-green colours that are prevalent in nature can be reproduced without appreciable error. But it means that the gamut is restricted in the cyan direction, and, although saturated cyan colours are not very common in nature, it does mean that when they do occur, as for instance in the colour of some lakes, they are outside the reproducible gamut of colours and are reproduced too pale. If a fourth, cyan, primary were added to the display, the gamut could be enlarged to contain these cyan colours. A way of increasing the colour gamut that avoids the undesirable signal-processing complication of using more than three primaries is to use more saturated red, green, and blue primaries; these can be provided by lasers, by LEDs, or by increasing the colour saturation of phosphors or filters, or by using filters with phosphors. The use of more saturated primaries usually results in a loss of light, and an additional white 'primary' is sometimes used to offset this loss, but this can result in distortions of colour rendering.

Effect of ambient illumination

High ambient illumination can have two detrimental effects on displays: first, because of simultaneous contrast, it can make the display appear too dark; and, second, it can result in the addition of white light more or less uniformly over the whole picture area (veiling glare or viewing flare) making the display appear lacking in colour saturation and of low contrast and low luminance dynamic-range (because of the addition of white light to the image light). The extent to which the display is darkened by simultaneous contrast depends on the ratio of the luminances of the light from the display and from the ambient light surrounding it. The extent to which the image on the display is reduced in colour saturation, contrast, and luminance dynamic-range, depends not only on this ratio but also on the inherent dynamic range, and on the surface reflection, of the display device itself. These factors result in the environment for CRTs being limited to medium levels of indoor daylight, and for LCDs and PDPs to higher levels of indoor daylight; only LEDs can be used in outdoor daylight.

The best solution to the problem of ambient light is to use reflecting displays; high ambient light then becomes an advantage, instead of being a handicap (unless the ambient light level is very low, and cannot be supplemented, when self-luminous displays are still necessary). The photographic and photomechanical printing industries have, of course, used reflection prints as their main display option for very many years.

The split-pixel transflective technique used in small displays to improve their visibility in outdoor daylight, provides a reflective display in addition to a self-luminous display. However, it is difficult to achieve high reflectance for several reasons. First, the pixels produce images over only part of the display area; this 'aperture ratio' may be only about 80%. Second, only some of this part, perhaps about 70% is usually devoted to the reflective image. Third, because of the presence of red, green, and blue filters, and other light-absorbing components, the reflective area may have a maximum reflectance of only about 15%. These factors, 0.8 x 0.7 x 0.15 = 0.084, result in a maximum system reflectance of only about 8% in diffuse illumination; in specular illumination, as a result of the incorporation of metallic reflecting elements, the maximum reflectance of the reflective area can be about 60%, resulting in a maximum system reflectance of about 30%. These split-pixel transflective displays provide useful improvements in the visibility of small displays in high ambient illumination, such as sunlight; but their low basic reflectance (8% in diffuse, and 30% in specular, illumination) is too low to achieve the higher level of picture quality over a wide range of viewing angles which is expected with medium-size and large displays.

To achieve reflective coloured electronic images, either the additive or the subtractive principles can be used [5,6]. In the additive form, a single layer of grey colorant, the density of which can be varied (for instance by electro-phoresis), can be used behind a red, green, and blue, filter array. This suffers from the disadvantages of reduction of resolution unless smaller pixels are used, and absorption of light in the filter array. By careful choice of the red, green, and blue filters used in the array, their total absorption can be limited to about 70%, so that a maximum reflectance of about 30% becomes possible. For a short while in the 1930s, Ilford Limited marketed a print system for the amateur photographic market which used an additive (Dufay type) reflective colour photographic paper; the maximum reflectance was about 30%, which, although significantly higher than the 8% (in diffuse illumination) achieved in the transflective systems, was still not high enough, and the product had only a short life. The absorption in the filter array therefore remains a major problem in additive reflective systems of this type. To try to overcome this problem, multi-layer devices are being developed, in which each layer contains a material, such as a polymer-dispersed liquid crystal (PDLC), that can be switched either to absorb as little light as possible or to use Bragg (interference) reflection to provide various amounts of red, green, or blue light [7]; achieving high maximum reflectance in this technology is not easy but, compared to additive systems using side by side coloured sub-pixels, it offers greater resolution for a given pixel size; in one application it is used to provide reflective billboard displays. In the subtractive form of coloured reflective electronic images, three superimposed layers of colorant have to be provided, one layer containing a cyan colorant, another a magenta colorant, and the third a yellow colorant. This subtractive form of display awaits a viable technology, but, again, provides greater resolution for a given pixel size; it also offers the possibility, at least in theory, of maximum reflectance similar to that of ordinary paper; the reduction to successful practice of a subtractive reflective display is a major challenge to the display industry.

An interesting application of reflection prints to displays in high ambient illumination was made in the 1964-1965 New York World's Fair [8]. Very large reflection prints were mounted on a tower and illuminated with high-powered xenon lamps; in sunlight the lamps provided an extra 50 000 lux, and after sunset an extra 10 000 lux. The effect was to give the impression of illuminated transparencies, this appearance being enhanced by the prints being surrounded by very black borders. In the case of reflective video displays, in suitable situations, supplementary illumination can be used to offset any rather low reflectivity of the materials, such as would occur particularly in the additive form. In this case, the maximum reflectance of 8 to 30% achieved in transflective displays, could result in effective presentations.

Resolution

The apparent resolution of a picture is scene dependent; any loss of resolution tends to be very noticeable when the picture contains a lot of important detail, as, for instance, with a group of people; but, in the absence of important detail, such as with a close up of one person's face, losses of resolution may be unnoticeable. A good example of this is when a drama is broadcast on television; scenes of close-ups of the actors can appear to be very sharp; but, when the whole cast is portrayed as a group, the lack of resolution can be very obvious.

The basic resolution of a system is determined by its number of lines, and the number of available pixels along each line. Display devices cannot improve this basic resolution, but they may sometimes reduce it, for instance by limitation of the addressability of the pixels, or, in projection devices, by lens aberrations.

In LCD, Plasma, OLED, and LED, displays, arrays of red, green, and blue sub-pixels are normally used side-by-side. If, in such arrays, a triad of primary colour sub-pixels is present for each pixel provided by the signals, then the resolution need not be reduced by the array; this can be the case for these displays, but it requires that the area of each colour sub-pixel is only one third of the area of a signal pixel. If it were possible to produce the red, green, and blue light in areas that, instead of being side by side, were on top of one another, then the colour pixels could be the same size as the signal pixel; the interference material, and the subtractive electronic paper, displays, described above, and a three-layer OLED display, would be examples of such devices.

In projection CRT, LCD, and DMD, displays, if full signal resolution is applied to the light-modulating components, no resolution need be lost; and, as there are no colour arrays, the cell size can, again, be the same as that of the signal pixel.

In shadow-mask tubes, if the number of colour sub-pixel triads is not as high as the number of signal pixels, some resolution is lost; furthermore, to avoid instability of the display in small areas, the electron spot size may be somewhat larger than the triad pitch (the distance between adjacent holes or slots in the mask), and this can reduce the resolution further; in some shadow-mask monitors, to improve the display of alpha-numeric characters, the electron spot size may be equivalent to up to about twice the triad pitch, so that further reduction in resolution may occur. The resolution of some shadow-mask tubes could be improved by increasing the number of colour sub-pixel triads and decreasing the electron spot size.

Because chrominance needs less resolution than luminance, one way of achieving full resolution with fewer pixels is to provide separate luminance and chrominance displays [9,10]. In broadcast television, because of gamma correction, the luminance signal is not a true luminance signal and some of the luminance information is carried by the chrominance signals; this means that the chrominance resolution should be restricted to only a quarter (a half both horizontally and vertically) of the luminance resolution as is the case in high definition television (HDTV). But when a true luminance signal is used, chrominance needs only one sixteenth (a quarter both horizontally and vertically) of the resolution required by luminance [11], so that a chrominance display produced by a red, green, and blue side-by-side array with a number of triads of colour sub-pixels equal to one-sixteenth of the signal resolution would be adequate; a luminance-attenuating array could then provide full signal resolution. If the chrominance display were produced by a superimposed, instead of by a side-by-side, device, then the number of (signal-size) pixels in each layer need be only one-sixteenth of the number in the luminance attenuator. Hence, a chrominance-luminance display could realise a substantial reduction in the total number of pixels for a given resolution, and thus hopefully reduce the complexity and cost of the device. (Greater reduction in the resolution of the chrominance signal carrying predominantly blue-yellow information, as is the case in the N.T.S.C. system, is not usually adopted.)

The resolution of filter-mosaic displays can be improved by using only two differently coloured filters on the array, the third colour being produced sequentially [12]. One example of these mosaiquential or STColor™ displays uses a cyan and magenta filter mosaic with yellow light to provide the red and green pixels for two-thirds of the colour-field time, and blue light for one third to provide the blue pixels. Advantages of this arrangement include: better resolution and less fixed-pattern noise compared to RGB mosaics; and lower colour-field frequencies and less colour break-up compared to RGB sequential displays. Configurations using alternate magenta and green, or alternate cyan and red, or alternate cyan and yellow, colour-fields, are also possible.

Conclusions

Areas where improvements to current display devices are desirable include: consistency of picture quality between different devices, and between different types of device; enlargement of the colour gamut; reduction in sensitivity to the effects of ambient light; and higher resolution. Consistency could be improved by adhering to the conditions contained in the sRGB Colour Space. The colour gamut could be enlarged by using an additional cyan primary or preferably by increasing the colour saturation of the primaries such as can be achieved by using LED or laser light sources. Reduction of the effects of ambient light could best be achieved by using reflective displays. Higher resolution for a given colour pixel size could be achieved by using multi-layer devices (such as might be provided by subtractive reflective displays, by interference materials, or by OLEDs); a given resolution could be achieved with a smaller total number of pixels by displaying chrominance and luminance separately, and the resolution of filter-mosaic displays can be improved by using the mosaiquential principle. A recent proposal is to use a real time two-dimensional reflective hologram on an LCOS microdisplay to provide a novel projection device [13].

Future possible displays of particular interest include those using single-array field-sequential, mosaiquential, reflective, or luminance-chrominance technologies. The use of compact optics for projection displays is also of considerable interest because it can make projection displays much more convenient in shape [14].

Acknowledgements

For helpful comments received during the preparation of this paper, I am grateful to Bill Crossland, Michael Pointer, Richard Salmon, Lou Silverstein, Dan Sun, Adrian Travis, and Ian White.

References

[1] Anderson, M., Motta, R., Chandraseka, S., and Stokes, M. Proposal for a Standard Default Color Space for the Internet – sRGB. *Fourth Color Imaging Conference: Color Science, Systems, and Applications,* 238-246. IS&T and SID, Springfield, Va, U.S.A. (1996).

[2] CCIR. Recommendation ITU-R BT.709 (formerly CCIR Recommendation 709) adopted 1990 . Basic parameter values for the HDTV standard for the studio and for international programme exchange. ITU, Geneva ,(1990).

[3] CIE Publication 159:2004. *A colour appearance model for colour management systems*. CIE (2004).

[4] Pointer, M.R. The gamut of real surface colours. *Color Res. Appl.*, **5**, 145-155 (1980).

[5] Crawford, G.P., Silverstein, L.D., and Fiske, T.G. Reflective color LCDs based on H-PDLC and PSCT technologies. *Journal of SID*, **5**, **1**, 45-48 (1997).

[6] Sunohara, K., et al. A reflective color LCD using three-layer GH-mode. *SI Digest of Technical Papers*, 103-106 (1996).

[7] Silverstein, L.D. Reflective color displays for imaging applications. *Third Color Imaging Conference: Color Science, Systems, and Applications,* 52-58. IS&T and SID, Springfield, Va, U.S.A. (1995).

[8] Bartleson, C.J., Reese, W.B., MacBeth, N., and James, J.E. Lighting Details of the Kodak Photo-tower… New York World's Fair. *Illuminating Engineering*, **59**, 375-381 (1964).

[9] Lee, M., Ko, H., and Song, D. US Patent 5 629 743. Video signal processor for two-panel liquid crystal projector (1997).

[10] Hubel, P., and Hunt, R.W.G. Configurations for color displays by the use of lenticular optics. US Patent Application, (2001).

[11] Hunt, R.W.G. The strange journey from retina to brain. *J. Roy. Television Soc.*, **11**, 220-229 (1967).

[12] Silverstein, L.D. Hybrid spatial-temporal method of color synthesis for enhanced display image quality. *Twelfth Color Imaging Conference: Color Science, Systems, and Applications,* Late breaking news paper abstract. IS&T and SID, Springfield, Va, U.S.A. (2004).

[13] Lawrence, N. Light blue optics. *Image Technology*, **87**, 5 (2005).

[14] Travis, A., Large, T., and Buckingham, M. Image quality in flat projection wedges. *Society for Information Display, 2004 International Symposium, Digest of Technical Papers,Volume XXXIV ,Washington State Convention and Trade Center, Seattle, Washington, May 26-27 2004*, paper 20.4, pages 850-85 (2004).

Author Biography

Robert Hunt received his B.Sc. in physics at Imperial College, London University, in 1943, and his Ph.D. in 1953 and D.Sc. in 1968, both from London University. He worked in the Ministry of Supply from 1943 to 1946, and at the Kodak Limited Research Laboratories from 1946 to 1982, finally as Assistant Director of Research. He was a visiting professor at the City University from 1967 to 1998, at the University of Derby from 1994 to 2004, and is currently a visiting professor at the University of Leeds. He has published two books, The Reproduction of Colour*, now in its 6th edition, and* Measuring Colour*, now in its 3rd edition.*