Which color gamut metric best predicts human display preference?

James M Hillis, Jennifer Schumacher, James Thielen, John Van Derlofsky, John Wheatley, Gilles Benoit, Art Lathrop

3M, Saint Paul MN

Abstract

Advances in display technology depend on our ability to develop miniscule elements that emit a broad range of light intensities and colors. Photometry and colorimetry provide tools to help developers and manufacture chose these targets for future and existing technology. Uncertainty remains however as to which metrics provide the best guidelines. We examined the relationship between 8 color metrics and human preferences for displays that differed only in color gamut. We found that (1) volume metrics, computed from display luminance and color capacity, outperformed area metrics computed only from color and (2) of the color metrics we considered, CIECAM'02 saturation performed best.

Keywords

Display quality; colorimetry; CIECAM; quantum dots

1. Objective and Background

Advances in Quantum Dot technology significantly expand the range of colors available from LCDs. While consumers prefer larger color gamut displays, the relationship between display color specifications (e.g. gamut area, primary chromaticity coordinates) and preferences is poorly understood. This uncertainty comes in part from the absence of readily available summary metrics that accurately depict the relationship between color gamut (or, more specifically, the chromatic coordinates of the display primaries) and user experience. Improved metrics would allow developers and manufacturers to set more meaningful and progressive targets and provide better alignment between engineering specifications and visual experience. In this paper we review current practice and examine the ability of different metrics to predict consumer preferences.

Current practice in color display metrology and communication: To characterize the color properties of a display, a typical lab may measure chromaticity coordinates of the white point and color primaries. This is consistent with recommendations in the The Information Display Measurement Standard (IDMS) for full and boxed primary and secondary color measurements and the use of CIE u'v' area as a summary metric. They also recommend volume-color-reproduction-capability in the CIE L*a*b* coordinate system. This, however, is more difficult to compute and does not appear to be a regular part of communications within the display industry. While these measures are based on decades of improvements in colorimetry, the relationship between these values and visual experience remains uncertain. Further, it remains common practice to use outdated and biased colorimetric quantities. For example, CIE 1931 x,y chromaticity coordinates are often used for communication. While useful for this purpose, they are non-uniform with respect to perceived color differences; that is distances in x,y are not valid indicators of apparent color differences. A more general concern is that our systems of colorimetry are based on judgments of small uniform color patches presented on uniform backgrounds. How accurate and reliable this system is for predicting the apparent quality of natural images or text on current general purpose displays is not well understood.

Display metrology that does not validly or reliably reflect consumer experience can hamper developers and manufacturers from setting meaningful, progressive targets. Here we aim to illuminate more meaningful metrics. To do so, we examine metrics that emphasize the joint chromatic and luminance contrast capacity of a display. This emphasis is consistent with current scientific understanding of the fundamental importance of chromoluminance contrast for visual experience

Representing Display Capability as the Volume of Color Contrast: The importance of chromatic and luminance contrast is recognized by the color-volume-reproduction-capability recommendation in the IDMS. There are two features of this representation that are advantageous. The first is the volumetric representation: this captures the inherently 3 dimensional nature of human color vision and completely depicts the boundary of the perceptually relevant light producing capacity of displays. The second is the use of color contrast as the fundamental unit of measurement. The CIE L*a*b* representation improves on the earlier (CIE 1931) color representations by improving uniformity for the representation of apparent color differences (a unit difference corresponds more closely to perceptual unit). In this system, L* is proportional to the cubed-root of normalized luminance. This captures (1) the non-linear response of the Luminance channel in the human visual system and (2) adaptation of the visual system to average luminance. The chromatic a*,b* coordinates also incorporate adaptation and non-linear response so that jointly the coordinate system is more uniform with respect to perceived chromoluminance differences. To help demonstrate how this system better reflects visual experience, consider an example comparing the x,y,Y and L*a*b* systems.

We compared the color volumes of an LCD display and a 3M QDEF™ enhanced LCD display. The volume of each display in xyY coordinates was 16.34u3 & 15.75 u3, suggesting that image reproduction capacity of the LCD display is greater than the quantum dot enhanced display. This is not consistent with perceptual experience: People consistently report that large color gamut QDEF™ displays appear both brighter and more colorful. In Lab coordinates, the volume of each display in Lab coordinates was 86,063 u3 and 118,801 u3 indicating a significant perceptual advantage for the quantum dot enhanced display. This is entirely consistent with reports of typical viewers.

This result demonstrates that the volume of producible colors in L*a*b* is a better reflection than volume computed in x,y,Y coordinates. We now consider more recent developments in color metrics to address the question: What measure best predicts experience?

Since its inception, several shortcomings of the CIE L*a*b* color space have been identified. These include nonlinearities in hue and non-uniformities in color differences. The CIE Color Appearance Model adopted in 2002 (CIECAM 2002) corrects many of these shortcomings. It also incorporates predictions for different dimensions of color judgments (e.g. chroma, saturation, colorfulness) and contextual influences on color appearance. The calculations are, however, significantly more complex. For display manufacturers and developers, it is not clear that the additional complexity is worth the effort. Ultimately the judge of display quality is the consumer. If the additional complexity does not improve our ability to predict consumer preferences then it is difficult to justify. Further, if volume metrics do not improve prediction over area calculations, there may be no need to move to volume measures. Here we examine gamut area and volume metrics relationship to human display preferences.

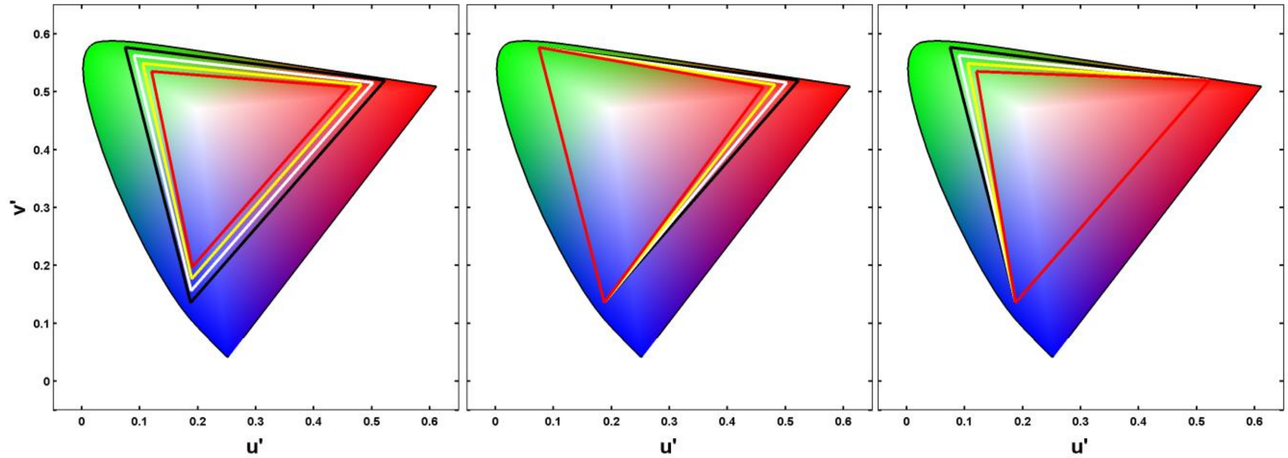

Figure 1: Left, middle and right panels show the RGB symmetric, red primary and green primary gamut conditions. The Small RGB condition was approximately 76% of the NTSC 1953 gamut standard in u'v' coordinates

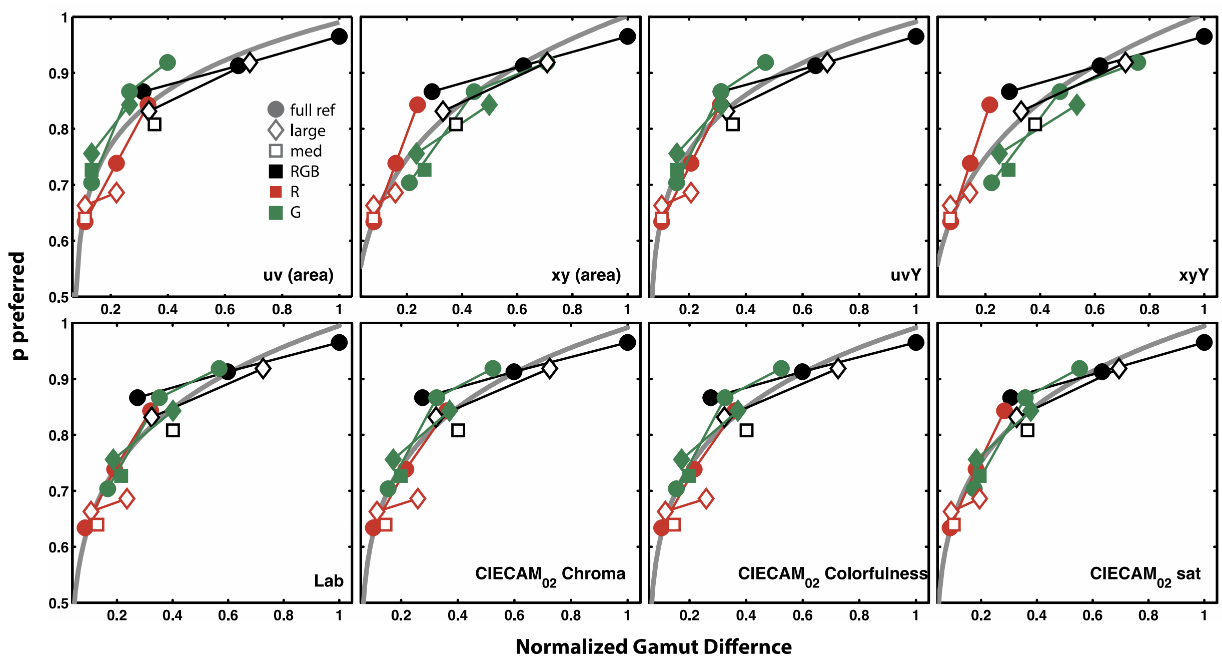

Figure 2: Proportion of large gamut images chosen over the small gamut differences as a function of the 8 different display metrics (labeled in each panel). Values on the x-axis were normalized after subtracting the minimum value.

2. Results

Preference data were collected using paired comparisons: observers were shown two images simultaneously presented on two adjacent calibrated high-gamut displays and asked to indicate which they preferred (see Schumacher et al. 2013 for details). Images were presented in the full gamut of the display and 9 smaller simulated gamuts. Three of these gamuts were achieved by moving the u'v' coordinate of the red, green and blue primaries uniformly toward the display white point. Three were achieved by moving only the red primary. And the final three were for similar shifts in the green primary. Figure 1 shows the gamut conditions in u'v' coordinates. We refer to the four gamut sizes as Full, Large, Medium and Small. Each observer made judgments on one of four representative highly colorful images and 2 judgments per pair of gamuts represented. So each participant completed 60 trials (10 paired comparisons, 3 gamut manipulations, 2 trials). Images were gamut mapped to preserve hue and luminance of each pixel.

Figure 2 shows observer preferences (proportion of times the larger gamut was selected over the smaller one) as a function of display metrics (normalized magnitude difference from the smallest gamut for each metric). Filled circles, open diamonds and open squares represent, respectively, comparison Full, Large and Medium against the Small gamut, Full and Large against Medium and Full against Large. Black, red and green symbols represent the RGB gamut manipulation, the red only and the green only manipulations. The gray curve in each panel is a logistic regression fit by maximum-likelihood criterion. The metrics shown are, from top left to bottom right: gamut area in u'v' and x,y, gamut volume in u',v',Y, x,y,Y, L*a*b*, CIECAM02, Chroma, Colorfulness and Saturation. Good prediction is demonstrated by close clustering of the points around the regression curve. We ranked the metrics using Akaike's Information Criterion (AIC). The top three metrics in order were: CIECAM02 saturation, u'v'Y and L*a*b* volumetric measures. ΔAIC (AIC relative to the top ranked model) between CIECAM02 saturation and u'v'Y was 2.77 suggesting that the CIECAM02 saturation volume performs reliably better. ΔAIC for L*a*b*, the third ranked metric and the one recommended in IDMS was 7.7 indicating it is notably less accurate than the CIECAM02 in predicting preferences. ΔAIC between CIECAM02 saturation and the remainder of the metrics was >8.25 indicating that this is a notably better choice for predicting preferences. Also of interest, the u'v'Y volume measure performed markedly better than the u'v' area metric indicating a relative deficiency in area metrics.

3. Impact

Our results reveal the value in applying more advanced, color appearance and color contrast spaces to characterize displays. In particular, the color volume capacity as measured in the CIECAM02 saturation index was the best predictor of preferences. This provides developers and manufactures with a target that is meaningful in terms of consumer preferences (i.e. develop displays to maximize volume as specified by CIECAM02 saturation and lightness). We are currently exploring other, more readily computable, summary metrics that perform as well. Two that have promise are distances from white point to primaries computed from the CIE 2000 distance formula (Luo, Cui &Rigg , 2000) and a non-linear product of peak luminance and chroma in L*a*b* coordinates proposed by Nakatsue (2013).

4. References

[1] E. A. Fedorovskaya, H. de Ridder, and F. J. J. Blommaert, "Chroma variations and perceived quality of color images of natural scenes," Color Research & Application, vol. 22, no. 2, pp. 96–110, Apr. 1997.

[2] M. R. Luo, G. Cui, and B. Rigg, "The development of the CIE 2000 colour-difference formula: CIEDE2000," Color Res. Appl., vol. 26, no. 5, pp. 340–350, Oct. 2001.

[3] T. Nakatsue, "Evaluation on the Colorfulness of Displays" The Society for Information Display Digest, p. 567, 2013

[4] M. Sakurai, R. L. Heckaman, S. E. Casella, M. D. Fairchild, T. Nakatsue, and Y. Shimpuku, "Effects of display properties on perceived color-gamut volume and preference," Journal of the Society for Information Display, vol. 16, no. 12, p. 1203, 2008.

[5] J. F. Schumacher, B. J. Stankiewicz, J. Van Derlofske, D. J. Lamb, Q. Sanford, K. C. Tyson, and S. Paul, "Toward a New Perceptual Quality Metric for Displays," 3M Techincal Document. .