Ecole Doctorale STIM
Sciences et Technologies de
l'Information et de Mathématiques

Spécialité: Automatique
Nom du Laboratoire : IRCCyN
Equipe: Images et Vidéo Communications

MODERN DISPLAYS: A CHALLENGE TO TRADITIONAL COLORIMETRY, AN OPPORTUNITY TO MOVE BEYOND THE STANDARD

Sarkar, Abhijit
Email: abhijit.sarkar@technicolor.com

Abstract : A fundamental basis of applied colorimetry is that all color normal observers can be represented, with reasonable accuracy, by a single "standard" colorimetric observer model. However, variations in observers' color vision characteristics can lead to significant differences in their color perception in modern display applications. These differences result from the peaky color primaries (commonly red, green and blue) that are often used to achieve more vivid and saturated colors. Such differences pose a technical challenge in color-critical professional applications. This paper presents an overview of this problem, and summarizes the experimental results from the current study that highlight the significance of the effect of observer variability. Also outlined is a novel experimental method using modern displays to classify various color-normal observers into a small number of categories, which, as per initial results, achieved perceptibly better color matches for several observers as compared to those obtained by using the average colorimetric observer.

Keywords : *Color Vision, Cone Fundamentals, Color Matching, Observer Variability, Displays.*
Collaborations : This doctoral research is conducted at Technicolor Research & Innovation, Rennes, under the supervision of IRCCyN-IVC, Polytech' Nantes, and is approved by the ANRT/CIFRE contract #0272/2008

1. Introduction

Color is defined as the perception that depends on the response of the human visual system to light and the interaction of light with objects. Thus, the perception of color requires three components, a light source, an object and an observer. The most commonly used mathematical way of describing color is through the CIE tristimulus values, X, Y and Z. To compute these values, contributions of relative spectral power of a CIE standard light source (S_λ), the reflectance factor of the object (R_λ), and the spectral sensitivities of color sensors (called color matching functions) of an average, standard observer ($\bar{x}_\lambda$, $\bar{y}_\lambda$ and $\bar{z}_\lambda$, read as x-lambda-bar, y-lambda-bar etc) are integrated over a range of wavelengths (λ). Mathematically, the integration is done by multiplication followed by summation, as shown in Eq. 1. The result is then normalized by a factor k, so as to assign the luminance (given by Y) of white an arbitrary value of 100. More details on the fundamentals of colorimetry are available in [1].

$$X = k\sum_\lambda S_\lambda R_\lambda \bar{x}_\lambda \Delta\lambda; \quad Y = k\sum_\lambda S_\lambda R_\lambda \bar{y}_\lambda \Delta\lambda; \quad Z = k\sum_\lambda S_\lambda R_\lambda \bar{z}_\lambda \Delta\lambda$$
$$k = \frac{100}{\sum_\lambda S_\lambda \bar{y}_\lambda \Delta\lambda} \tag{1}$$

The above equation is the basis of all colorimetric computations in basic and applied colorimetry. Practically, all mathematical representations of color in applied colorimetry originate from this equation. The standard observer color matching functions (CMFs) represent average color matching functions of a given group of observers, measured through visual experiments. Two sets of standard observer CMFs are commonly used in conventional colorimetry, CIE 1931 standard observer functions for 2° visual field, and CIE 1964 standard observer functions for 10° visual field. It is assumed that the standard observer represents the average of the whole population of color normal observers. However, for various physiological reasons [2], color vision characteristics of any real observer can be significantly different from the standard observer functions, which can result in misrepresentation of the color perception of a real observer. The difference between the CMFs of a standard observer and those of a real observer gives rise to another important and fundamental issue, as described next.

2. Observer variability and observer metamerism

When two color stimuli produce the same visual response, a visual match is obtained, where XYZ values corresponding to two stimuli are identical. Since XYZ values are computed by wavelength-wise integration, two stimuli with very different spectral power distribution can give rise to identical cone response for a given observer, leading to a "metameric" color match that is obtained by using a combination of a specific illuminant and a specific standard observer. Such a match established by one observer can, and quite often does, lead to a mismatch for a different observer, as the second observer has a different set of CMFs than the former. This phenomenon is commonly termed as observer metamerism.

The relevance of this issue is obviously quite dependent on the application context. When a display color is compared with its printed version on paper, the significance of observer variability is questionable [3]. However, the topic of observer metamerism has sparked renewed interest in the recent years with the proliferation of wide color-gamut displays. Whether based on LED (Light Emitting Diode) or employing laser primaries, all these displays compete with each other in achieving more vivid, more saturated and brighter colors. On the flipside, these displays are particularly susceptible to observer variability [4][5], since their narrow-band (i.e. peaky spectral characteristics) primaries cause noticeable shift in chromaticities of perceived colors with relatively minor change in the visual characteristics of the observer.

In a theoretical analysis by the present authors [6], it was found that the CIE 10° standard observer functions, as well as the average observer functions from a recent physiological model (proposed by CIE in 2006 [2]), do not accurately predict real observers' CMFs averaged within various age-groups. Because of this prediction error, colors on a narrow-band, wide-gamut display and a broad-band (smooth spectral characteristics) CRT display (Cathode Ray Tube), which are supposed to be matches based on real observer data and actual display spectral characteristics, were predicted to have significant colorimetric differences. The experimental data used in the study came from the most comprehensive color matching experiments till date, performed by Stiles and Burch in 1959 [7], on which the CIE 10° standard observer is based. This analysis showed that when it comes to modern displays, the issue of observer variability can cause conventional colorimetry to fail. The extent of this failure will depend on the spectral characteristics of the display, the specific colors that are being reproduced on the display, as well as on the CMFs of the observer viewing the display. Similar failure of CIE colorimetry has been observed when narrow band RGB-LEDs were matched with broadband lights [8]. Now, the question is: when is this failure inconsequential, and when it is serious enough to necessitate a solution?

3. Entertainment industry applications where observer variability can be a serious problem

Observer variability and metamerism can be a nontrivial issue in critical color matching tasks, for example in post-production applications. Suppose, for example, the color adjustment process (called color grading) of the raw movie content at the post-shooting stage. The Colorist has to work with the Director of Photography (DP) to adjust the colors in the original content so as to achieve color coherence and homogeneity throughout various scenes, while maintaining the artistic expressions originally envisioned by the Film Director and the DP. Further, the film may have to be converted to a version suitable for television or DVD (a process known as digital mastering). Processes like color grading and digital mastering are color critical, requiring high-fidelity color reproduction, often involving displays. Even though film studios have principally relied upon reference CRTs with broadband primaries, a rapid market adoption of wide-gamut, high-definition displays and projectors and gradual discontinuation of manufacturer support on CRTs may soon require the studios to employ these modern displays for post-production operations.

However, if the Colorist and the DP have different color vision characteristics, they will perceive colors differently, and the colors that look similar to one will look perceptibly different to the other. Conventional colorimetry will fail to account for this difference.

Very recently, studios have started offering remote color grading services, which means multiple devices being used by various professionals at multiple locations for color grading, a trend that is sure to make the issue of observer variability even more pertinent in the media and entertainment industry.

Thus, it is of interest to study the effect of observer variability in color matching across conventional and modern displays, and to acquire experimental data in such a context, which can be subsequently used to better model the observer variability, and to find solutions to the associated problems.

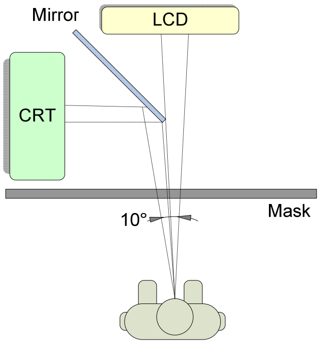

Figure 1. Experimental setup

4. A color matching experiment

A preliminary set of color matching experiments [9] were performed on a modern, wide-gamut LCD (Liquid Crystal Display) with narrow-band primaries, and a conventional studio CRT with broadband primaries. Because of the significant difference in the spectral characteristics of the CRT and the LCD, color matches were expected to vary significantly from one observer to the other.

The displays were placed perpendicular to each other, as shown in fig 1. A front-surface reflection mirror was placed in front of the CRT, and a mask was placed between the observer and the displays such that the observer's visual field consisted simply of a 10° circular area divided vertically in two equal parts (called bipartite field). The right half of the bipartite field was the LCD screen, and the left half was the CRT screen, seen through the mirror. The observers were asked to adjust the color on the left half of the bipartite field (CRT) to match the color on the right half (LCD). After each match, the spectra of the colors on the two displays were measured. Each observer matched nine colors.

Fig 2 plots the perceptual color differences between the two displays after each of the seven observers obtained a color match. There are nine colors shown along the x-axis. The color differences along the y-axis are represented in the form of bars, grouped by the seven observers. The color differences are calculated using 10° standard observer and an advanced color difference formula denoted by ΔE^*_{00} (read as delta E 2000). In other words, if colorimetry worked perfectly in spite of observer variability, all these color differences would have been zero. As a thumb rule, more than 1 unit of ΔE^*_{00} color difference can be visually perceived by most observers under carefully controlled viewing condition as in these experiments. The mean, maximum and the 90th percentile values of the standard observer-predicted color difference of individual observer color matches were close to 1.4, 3.3 and 2.6 ΔE^*_{00} respectively. An average color match prediction error of 1.4 ΔE^*_{00} over all colors and all observers is still acceptable, confirming that the 10° standard observer is a reasonably good representation of an average observer. However, the maximum and the 90th percentile ΔE^*_{00} values between individual observer matches predicted by the 10° standard observer are rather high. This indicates that for some colors, color match prediction by an average observer results in significant color match errors for many individual observers. In the same way, the results showed that colors on the two displays that were predicted by colorimetry as a match were in fact significantly different from real observer matches. In color critical applications involving modern displays, expert observers will likely find such differences unacceptable. The degree of the prediction error is dependent on the spectral characteristics of the display, and also on the observer-specific color matching functions. Based on our preliminary results, the problem seems to be nontrivial.

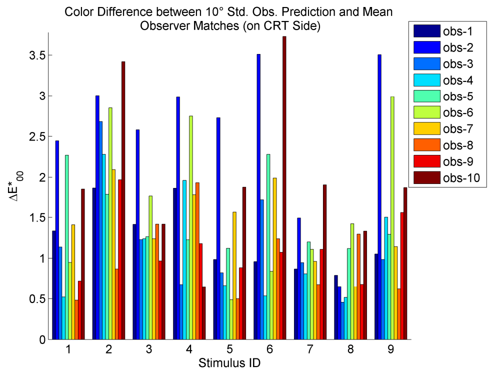

Figure 2. ΔE^*_{00} Color difference between CRT and LCD observer matches as predicted by 10° Standard Observer

Individual variability is ignored when a single standard observer is used to represent a whole population of real, color-normal observers. In color-critical applications involving displays with narrow-band primaries, this can result in significant error for many expert observers.

5. An experimental method for classifying color-normal observers

A multi-step theoretical analysis [10] was performed in order to derive several representative observer categories, which can possibly cover most of the color-normal observer population. First, a cluster analysis was performed on the most extensive set of experimental and computational data on color matching available to date. Next, an iterative algorithm was implemented, which involved several statistical measures of ΔE^*_{00} color differences between real observer data and the data predicted by the observer categories. Seven distinct observer categories were identified, which could potentially be used to classify real observers with normal color vision.

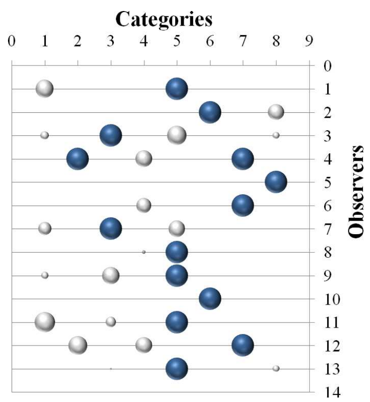

Figure 3. *Observer categories as determined through the observer classification experiment*

Next, an observer classification method was implemented using similar experimental setup as shown in fig 1. In paired presentations on the two displays, eight color-matches corresponding to the CIE 10° standard observer and the seven observer categories were shown in random sequences. Thirteen observers evaluated all eight versions of color-matches for fifteen test colors, pre-selected for high variance of tristimulus values corresponding to various observer categories. For each color, the observers were asked to assign the eight categories into one of three groups, namely, *unacceptable (U)*, *acceptable (A)* and *satisfactory (S)*. Based on the overall results for a given observer's choices for all fifteen test colors, a score S was assigned to each category: S = 0.8S + 0.2A – 1.0U. For majority of the observers, only one or two categories consistently produced either acceptable or satisfactory matches for all colors. The results are graphically represented in fig 4, where the observers are shown along the vertical axis, and eight versions are shown along the horizontal axis. The area of a bubble indicates the score of the corresponding category. The shaded bubbles are the assigned categories for various observers. It was possible to identify for each observer a specific category as the most appropriate. In other words, each observer was classified as belonging to one of these categories.

The results revealed that for eleven observers, CIE 10° standard observer (category 1), the current industry standard, was not among the two most preferred categories, and for three observers, it was rejected as an unacceptable match for all fifteen test colors. For all three observers, the preference was between categories 6 and 8. On the other hand, the observers who were closer to the standard observer always rejected category 6 and 8. From the results, it is clear that observers assigned to categories 6 or 8 are distinctly different from those assigned to categories 1 or 5, and also those assigned to 2 or 7. These groups of observers are mutually distinct. Note however that this discussion is not about an average match for all observers, for which the 10° standard observer will probably still be reasonably good [10].

Possible ways of further validating the observer classification results are currently being explored. The initial results are promising, considering the consistent and stable classification for most observers across a wide range of test colors.

6. Conclusion

When colors are compared on two displays with very different spectral power distributions, relying on the standard observer based colorimetry can possibly lead to unacceptable color matches for many color-normal expert observers. This is particularly an issue in highly color-critical applications. The results from the first phase of observer classification experiment are definite confirmation of this issue, but more importantly, they also show that such displays can be exploited to better predict the variability in individual observers. With that, a key hypothesis of this work is confirmed, that real, color-normal observers can be classified into a small number of categories by means of a practical experimental setup suitable for industrial applications. The new method for observer classification developed as part of this ongoing work is the first step in achieving our final goal of developing an *observer-dependent color imaging* method, where color workflow in a display device can potentially be tuned to one of several observer classes. Accomplishing that goal will open the doors to new industrial applications.

7. References

[1] R.S. BERNS, “Billmeyer and Saltzman’s principles of color technology”, Willy, New York, USA, 2000

[2] CIE, “Fundamental Chromaticity Diagram with Physiological Axes – Part I”, *CIE Technical Report*, 170-1:2006, 2006.

[3] B.OICHERMAN, M.R. LUO, B. RIGG, AND A.R. ROBERTSON, “Effect of observer metamerism on colour matching of display and surface colours”, Color Research and Application, vol. 33 (5), pp. 346-359, 2008.

[4] M.D. FAIRCHILD AND D.R. WYBLE, “Mean Observer Metamerism and the Selection of Display Primaries”, Final Program and Proceedings-IS&T/SID Color Imaging Conference, pp. 151-156, Albuquerque, NM, USA, 2007.

[5] R. RAMANATH , “Minimizing Observer Metamerism in Display System”, Color Res. & Appl., vol. 34(5): pp. 391-398, 2009.

[6] A. SARKAR, L. BLONDÉ, P. LE CALLET, F. AUTRUSSEAU, J. STAUDER, P. MORVAN, “Study of Observer Variability on Modern Display Colorimetry: Comparison of CIE 2006 Model and 10° Standard Observer”, Final Program and Proceedings, the 11th Congress of the International Colour Association (AIC) 2009, Sydney, Australia, 2009.

[7] W.S. STILES AND J.M. BURCH, “N.P.L. colour-matching investigation: final report”, *Optica Acta*, Vol. 6, pp. 1-26, 1959.

[8] P. CSUTI AND J. SCHANDA, “Colour matching experiments with RGB-LEDs”, Color Res. & Appl., vol. 33(2): 108-112, 2008.

[9] A. SARKAR, L. BLONDÉ, P. LE CALLET, F. AUTRUSSEAU, P. MORVAN, J. STAUDER, “A color matching experiment using two displays: design considerations and pilot test results”, Final Program and Proceedings, CGIV 2010 conference, Joensuu, Finland, 2010.

[10] A. SARKAR, L. BLONDÉ, P. LE CALLET, F. AUTRUSSEAU, P. MORVAN, J. STAUDER, “Toward reducing observer metamerism issue in industrial applications: colorimetric observer categories and observer classification”, submitted to the 18th Color Imaging Conference, 2010.