Mean Observer Metamerism and the Selection of Display Primaries

Mark D. Fairchild and David R. Wyble, Rochester Institute of Technology, Munsell Color Science Laboratory, Rochester, NY/USA

Abstract

Observer metamerism is an unavoidable fundamental limitation of colorimetry. While standard observers might provide accurate estimates of population-mean color matching functions, they can never predict accurate matches for all individual observers. This limitation is minimized in applications where the spectral difference (or potential for metamerism) is small. This is typically when original and reproduction both have slowly varying spectral power distributions. In the world of displays, that would be represented by broad-band primaries. As display manufacturers attempt to design displays with wider color gamuts and greater luminance contrast, narrow-band primaries (in some cases monochromatic laser primaries) are becoming more common. Such displays enhance the potential for significant differences in color perception and matching across individual observers. The CIE has recently published a method for computing color matching functions for mean observers of various ages and for various field sizes. This paper examines the interactions between these various color matching functions and display primaries. It is shown that the magnitude of mean-level observer metamerism can be significant (on the order of 4 CIELAB units) for broad-band primaries and a factor of two or more larger (on the order of 10 CIELAB units) for narrow-band primaries. The potential disagreement between individuals is even larger and the discrepancies are particularly noticeable for large fields and near-white colors (as in trying to match the white point of two metameric projection displays or a display and metameric surround).

Introduction

Observer metamerism refers to differences in metameric matches when made by different observers. Identical spectral matches will match in color for all observers. However, when the spectral power distributions of the two stimuli differ, and only metameric matching is possible, a match made by one observer will typically not match for other observers. This is caused by differences in the color matching functions of the various observers. Among color-normal observers, these differences are caused by variation in macular pigment density, pre-retinal filtering in the optical media (cornea, lens, and humors), and differences in cone spectral responsivity and photopigment density. Higher-level mechanisms of chromatic adaptation and color appearance do not impact metameric matching. Color differences due to observer metamerism have been shown to be quite large (up to 20 CIELAB units) for cross-media image-reproduction applications.[1] This paper examines a new CIE procedure for computing color matching functions as a function of field size and mean-observer age[2] in conjunction with observer metamerism for displays with broad-band and narrow-band primaries.

Observer metamerism has been quantified in *ad hoc* procedures by comparing match predictions made with the CIE 1931(2°) and CIE 1964 (10°) standard colorimetric observers. For example, when a match is made for the 1931 observer, the color difference between the two stimuli using the 1964 observer can be computed to get a sense of the potential differences for other observers. In 1989, the CIE published a formal technique for evaluation of observer metamerism and a set of color matching functions for a so-called standard deviate observer.[3] This procedure suffered from a rather significant underprediction of observer metamerism potential due to the nature of its formulation and derivation from normalized color matching functions.[1]

A solution to these problems was proposed much earlier by Nimeroff et al.[4,5] They suggested creating a full system of color matching functions that included not only the mean functions, but spectral covariance functions as well. Such a system could be used to predict and formulate metameric matches along with confidence ellipsoids representing the distribution of mismatch for the population as a whole. Unfortunately such a system has not been completely formulated and implemented. However, ongoing research on the modeling of observer variability, combined with recent physiological, psychophysical, and genetics research shows promise for creating such a system in the future. A first step is represented by the recent publication of a system for computing cone fundamentals (color matching functions) for mean observers with a specified field size from 1° to 10° and age from 20-80 years.[2,6] This procedure captures the main physiological sources of variability and is based on the best available experimental data. It is analogous to the well-known CIE method for computing the D-series of daylight illuminants. At this point the *CIE2006* method (as it will be referred to in this paper) only predicts mean color matching functions for a given age and field size. It does not address variability at the individual level, but by examining the functions for various ages and field sizes it can provide very useful, conservative, measures of observer metamerism potential.

The potential for observer metamerism is a function of the physical stimuli as well as the observers' color matching functions. Non-metameric (spectral) matches will match for all observers. As the spectral differences between color-matched stimuli increase, the potential for observer metamerism also increases. This is because large spectral differences that cancel out one another for one observer might be critical to another. This concept was well illustrated by Alfvin et al.[1,7] who measured very large inter-observer differences in simple color matches when comparing CRT displays with photographic-dye prints illuminated by fluorescent daylight simulators. While this is not new, issues of metamerism are becoming more important in various display applications due to the variety of technologies being utilized. The historical importance of observer metamerism was highlighted by Wintringham[8] in his classic work on colorimetry in color television who stated the following. “There is another reason why it may be desirable to use desaturated primaries in a television receiver. It has been found in direct colorimetry that observer differences can be minimized by making the color triangle of the primaries no larger than is necessary to include the variation of chromaticities to be measured.[9,10]” The first reason he discussed was increased luminance due to more energy in the desaturated primaries. Wintringham’s direct colorimetry is what would today be called visual colorimetry (with indirect colorimetry being computation of matches using color matching functions and spectral power distributions).

Wintringham’s observations have come to light in recent applications such as digital cinema where projectors using xenon lamps with film primaries are being replaced by DLP projectors with different filters and some are looking toward application of laser primaries in projectors. All the while, filmakers are comparing the projected images with “proofs” on CRTs, LCDs, and plasma displays. The potential for metamerism is huge and it has already been noted that CIE 1931 colorimetric matches don’t hold for individual observers, particularly when setting white points and comparing large fields (e.g. half of the screen one technology with the other half another technology).[11] Other applications such as the color measurement of LED lighting [12] and the comparison of LED surround illumination with LCD display colors [13] have illustrated significant potential for observer metamerism and the potential improvements in colorimetry available through the use of variable sets of color matching functions.

CIE2006 Procedure

CIE TC1-36, was created in 1991 with the terms of reference to “establish a fundamental chromaticity diagram of which the co-ordinates correspond to physiologically significant axes.” F. Vienot of France is the committee chair and they have recently published the first part of their work defining the best set of color matching functions and a procedure to compute cone fundamentals (also color matching functions) for field sizes from 1° to 10° and a range of ages.[2] The CIE technique was used in this research to compute a wide range of mean color matching functions as cone fundamentals. Lacking an official designation, these color matching functions are referred to as CIE2006(field size,age). For example CIE2006(2,32) would refer to cone fundamentals computing using the CIE procedure for a 2° field size and mean observer age of 32 years.

The CIE2006 procedure begins by computation of the maximum density of the macular pigment as a function of field size (along with a stated assumption that it does not vary significantly with age). This is used to scale a defined relative spectral density function for the macula as illustrated in Eq. 1. Next the spectral optical density of the lens and other optical media is computed as a two-part function of age (it is not a function of field size since all light passes through the same ocular media). This density function is also inserted in Eq. 1 in the appropriate place. The low-optical-density absorbance spectra of the visual photopigments are derived and defined in tabular form.[2] These are then scaled by the peak visual pigment densities that are functions of field size (since cone shape changes across the retina) to obtain the cone absorptance spectra, α, used in Eq. 1. There is insufficient data to define changes in visual pigments with age. Finally, the cone absorptance spectra are multiplied by the transmittance of the macula, and ocular media to obtain cone fundamentals at the corneal level as illustrated in Eq. 1. Also, it should be noted that conversion from quantal units to energy units is required as a last step for most traditional colorimetric computations.

$$\begin{aligned}
\bar{l}(\lambda) &= \alpha_{i,l}(\lambda)\cdot 10^{\left[-D_{\tau,\max,macula}\cdot D_{macula,relative}(\lambda)-D_{\tau,ocul}(\lambda)\right]} \\
\bar{m}(\lambda) &= \alpha_{i,m}(\lambda)\cdot 10^{\left[-D_{\tau,\max,macula}\cdot D_{macula,relative}(\lambda)-D_{\tau,ocul}(\lambda)\right]} \\
\bar{s}(\lambda) &= \alpha_{i,s}(\lambda)\cdot 10^{\left[-D_{\tau,\max,macula}\cdot D_{macula,relative}(\lambda)-D_{\tau,ocul}(\lambda)\right]}
\end{aligned} \tag{1}$$

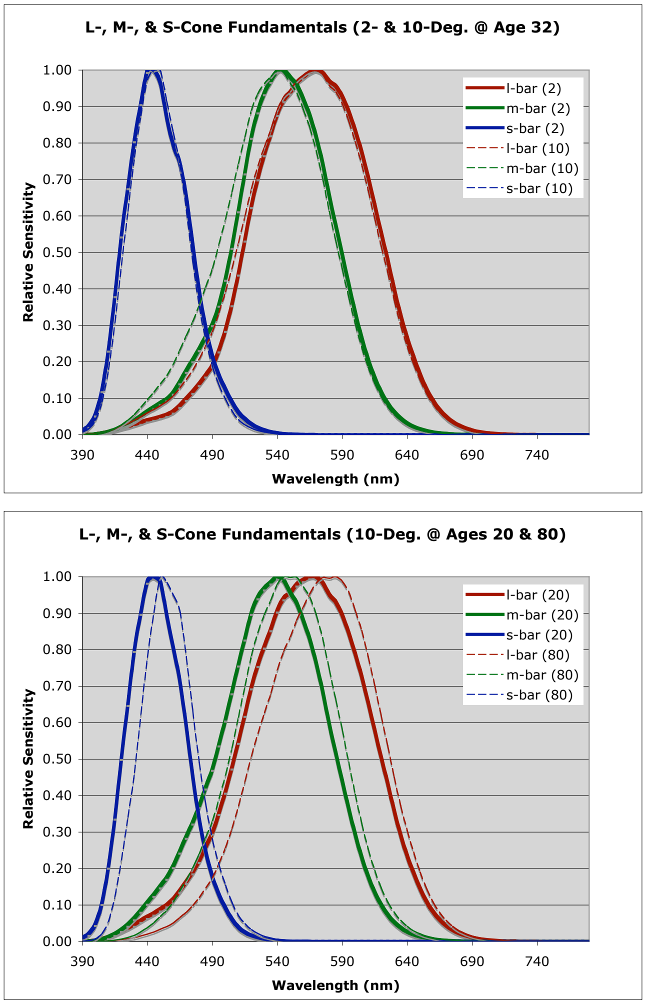

Figure 1. CIE 2006 cone fundamentals (color matching functions) for 2° and 10° field sizes at age 32 (upper panel) and ages 20 and 80 for a 10° field size (lower panel).

Microsoft Excel spreadsheet and MATLAB implementations of the CIE2006 procedure that allow computation of cone fundamentals for any desired combination of age and field size have been made available on the RIT-MCSL website at <mcsl.rit.edu/online/cie.php>.

Figure 1 includes four examples to illustrate the CIE2006 procedure. In the upper panel, the differences between 2° and 10° field sizes for age 32 are shown. The CIE2006(2,32) functions are essentially the standard color matching functions of the CIE2006 procedure.[2] The lower panel illustrates the effect of age by comparing the CIE2006(10,20) and CIE2006(10,80) functions. It is clear from Fig. 1 that the effects of age and field size on color matching functions are significant enough to impact practical applications such as display design and color management.

Experimental Procedure

The general concept of the experiment was to use the variable CIE2006 color matching functions to examine the magnitude of differences in mean-observer responses for hypothetical additive displays with broad-band and narrow-band primaries. The computational procedure is illustrated in Fig. 2.

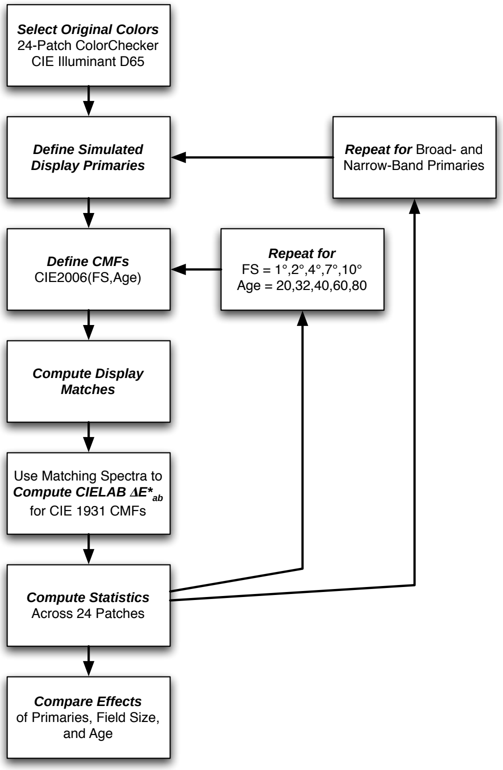

Figure 2. Flow chart of the computational procedure to evaluate the magnitude of mean-level observer metamerism as a function of display primaries and CIE 2006 field size and observer age.

The CIE2006 technique was used to derive 25 sets of cone fundamentals representing 25 different "standard" observers. These were for field sizes of 1°, 2°, 4°, 7°, and 10° with ages of 20, 32, 40, 60, and 80 years. The CIE 1931 and CIE 1964 observers were also considered, but since the CIE 1931 observer was used to evaluate and quantify the match variability, it always had the same performance (zero ΔE^*_{ab}). A set of test colors was established as typical measurements of the 24 patches of the standard GretagMacbeth ColorChecker Chart under CIE Illuminant D65.

Two hypothetical additive display systems were defined. The first was a broad-band display with primaries defined as approximate Gaussian spectral power distributions with peak energy at 450, 540, and 610nm. The widths of the Gaussian-like functions were such that the primaries sum to approximately an equal-energy distribution across the middle part of the spectrum. In other words, the 450 and 610nm primaries reach zero near the peak of the 540nm primary. The narrow-band primaries were selected to be close to the so-called prime wavelengths (which serves to minimize the effects of observer metamerism). They peaked at 450, 540, and 610nm and were effectively 5nm-wide rectangle functions in the tristimulus integration since the CMFs are sampled at 5nm. Figure 3. illustrates the spectral power distributions of the match to the ColorChecker white patch for both of the display systems.

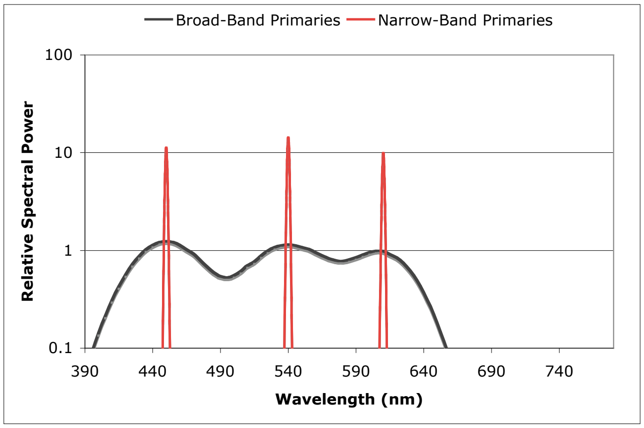

Figure 3. Spectral power distributions (log axis) for matches to the ColorChecker white patch under D65 by simulated broad-band-primary and narrow-band-primary displays.

While comparing broad-band and narrow-band primary systems will produce the greatest potential for observer metamerism, it is worth reiterating that the selection of the narrow-band primaries as prime wavelengths represents a realistic display design (from the perspective of gamut volume, efficiency, and brightness) and is the best-case for minimizing observer metamerism with a narrow-band display. Thus, again, the results presented in this paper are, if anything, conservative.

The computational procedure followed the flow chart in Fig. 2. Using a given set of cone fundamentals, the amount of RGB required to match each ColorChecker color using both sets of primaries was computed. This was repeated for all 25 sets of CIE2006 functions and the two CIE standard observers. For a given set of color matching functions three matching stimuli were obtained, the ColorChecker patches under D65 and the required RGB quantities for broad-band and narrow-band displays. These color matches were then evaluated using the CIE 1931 standard observer to compute CIELAB color differences between the ColorChecker and display colors. This was repeated for each of the 24 patches. These values were averaged across the 24 colors provide an indication of how different the given CIE2006 color matching functions are from the CIE 1931 observer for each simulated display.

Results and Discussion

The obtained color differences are summarized statistically (mean and standard deviation for each simulated display) in table 1. Even a cursory examination of these data shows that the potential for observer metamerism is significant, clearly varies with age and field size, and can be up to 5-6 times greater for the narrow-band display. A few examples are examined in more detail.

Table 1. Mean and standard deviation (SD) CIELAB color differences (2°)for the 24 colors of a GretagMacbeth ColorChecker Chart under D65 matched using trichromatic displays with broad-band (BB) and narrow-band (NB) primaries for CIE2006 standard observers of various ages and field sizes.

Age	***Field Size (°)***	***Mean (BB)***	***SD (BB)***	***Mean (NB)***	***SD (NB)***
20	1	1.21	0.72	1.91	0.79
20	2	1.02	0.75	1.57	0.79
20	4	2.18	1.25	4.00	2.60
20	7	3.24	1.92	7.61	4.57
20	10	3.89	2.37	10.06	5.70
32	1	1.38	0.70	1.63	0.73
32	2	0.74	0.59	1.33	0.63
32	4	1.79	1.00	3.92	2.36
32	7	2.84	1.65	7.47	4.22
32	10	3.48	2.08	9.88	5.41
40	1	1.56	0.79	1.53	0.79
40	2	0.64	0.51	1.24	0.56
40	4	1.54	0.84	3.88	2.17
40	7	2.56	1.47	7.36	3.98
40	10	3.20	1.89	9.72	5.19
60	1	2.10	1.18	1.52	1.18
60	2	0.87	0.50	1.24	0.53
60	4	0.90	0.47	3.73	1.73
60	7	1.87	1.01	6.98	3.39
60	10	2.48	1.40	9.19	4.51
80	1	4.44	2.92	3.26	2.67
80	2	3.12	2.16	2.29	1.76
80	4	1.83	1.36	2.84	1.15
80	7	0.97	0.82	5.04	1.96
80	10	0.72	0.60	6.78	2.65
CIE-31	2	0	0	0	0
CIE-64	10	2.53	1.68	4.30	2.39

Figure 4 shows a slice through the data for CIE2006 functions at age 32 with various field sizes. Plotted are the mean color differences across the 24 patches as a function of field size for each display type. The potential for observer metamerism generally grows with field size (indicating the difficulty in matching uniform areas on large displays, such as in digital cinema applications) and is always larger for the narrow-band system. Figure 5 is a similar illustration for a 10° field size and various ages. Again the potential for observer metamerism is always greater for the narrow-band display, but tends to decrease slightly with age. The average color differences always tend to be about 6 CIELAB units larger for the narrow-band display. For comparison purposes, the bottom two rows in table 1 provide similar data for the CIE 1931 and 1964 standard observers. The color differences are always zero for the CIE 1931 observer since the matches are computed and analyzed with the same color matching functions and there is no observer metamerism. For the CIE 1964 observer, there are significant differences, which are nearly twice as large for the narrow-band display.

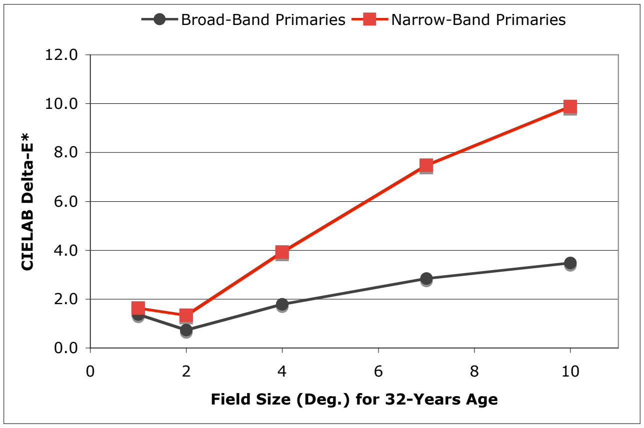

Figure 4. Mean CIELAB ΔE (CIE 2°) between 24 ColorChecker patches under D65 and the simulated displays for CIE 2006 observers at age 32 with various field sizes.*

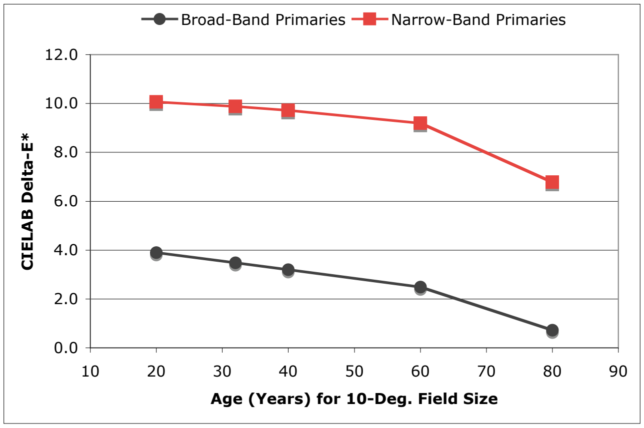

Figure 5. Mean CIELAB ΔE (CIE 2°) between 24 ColorChecker patches under D65 and the simulated displays for CIE 2006 observers at various ages with 10° field size.*

A closer look for a single color is provided in Fig. 6 where the CIELAB a*b* coordinates are plotted for each match to the ColorChecker white patch. This is indicative of the potential observer metamerism that might be encountered when attempting to white balance various displays with different types of primaries. It can be seen that the differences are substantially larger for the narrow-band display (lower panel), sometimes in excess of 10 CIELAB units (which would be a clearly perceptible difference in white balance, especially in side-by-side display comparisons). The results for the CIE 1964 observer are also plotted and can be seen to be most similar to the CIE2006(7,20) observer for this particular color. On average, as seen in table 1, the CIE 1964 observer performs most closely with the CIE2006(4,20) functions, but this would be a function of the particular metameric matches being examined and should not be considered of great importance.

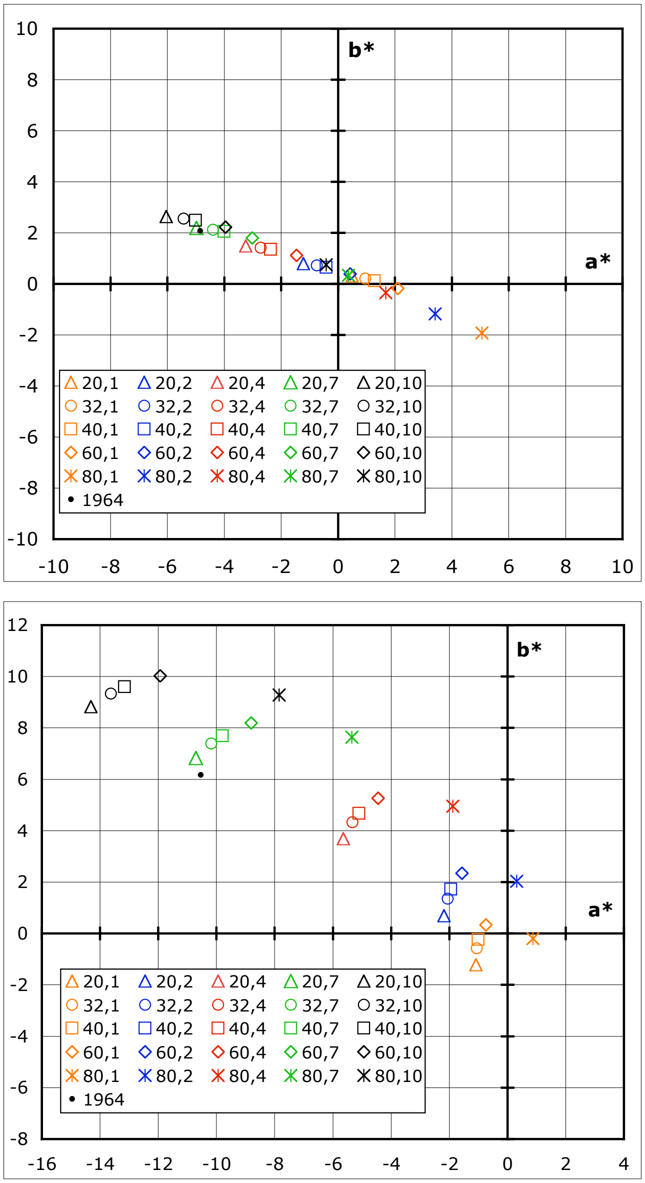

Figure 6. CIELAB a-b* values (CIE 2°) for the ColorChecker white patch when matched by various CIE 2006 color matching functions. Upper panel is for a broad-band-primary display while the lower panel is for a narrow-band-primary display. Note: The CIE 1931 2° observer would place both matches at a*=b*=0.0.*

Visualization of Impact

Figures 7-10 provide a form of visual simulation of the potential differences due to observer metamerism for a selection of CIE2006 color matching functions. They were created by computing the RGB values needed to drive each simulated display in order to match the given colors or images to one another and then using those RGB values on a single sRGB display. In other words, they are an illustration of how much the device-dependent color might need to be changed in order to compensate for observer metamerism. They are provided as an aid in understanding the visual importance of the data summarized in the previous section.

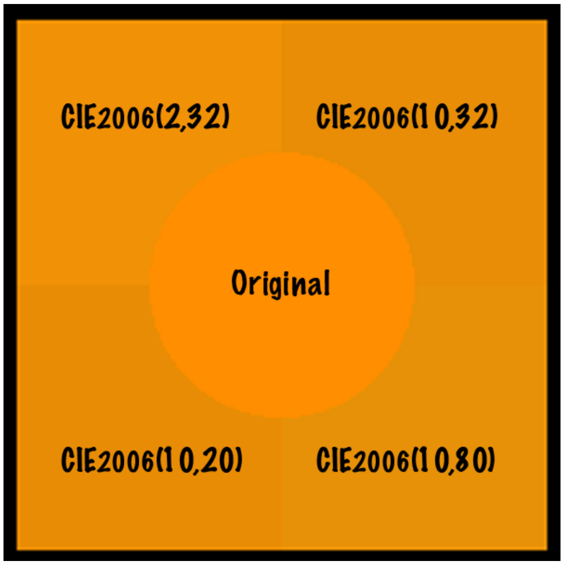

Figure 7. Key to the arrangement of figures 8-10. The central patch is an original color (Orange Yellow ColorChecker Patch) on the simulated broad-band-primary display. Surrounding patches are matches on the narrow-band-primary display for four different CIE2006 observers.

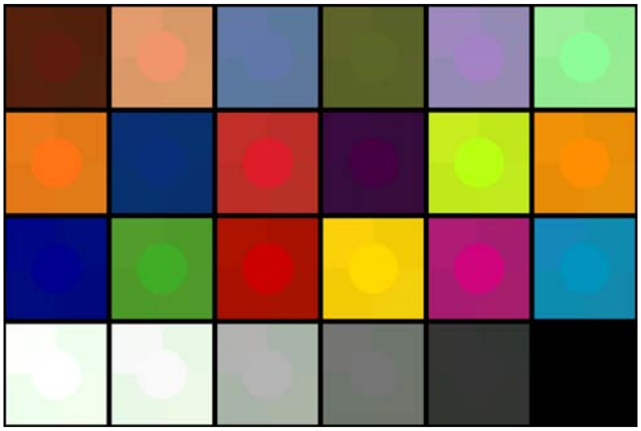

Figure 8. Visualization of differences in ColorChecker patches for a selection of color matching functions. CIE2006 with (fs,age) = (2,32), (10,32), (10, 20) and (10,80). Central circles are for the broad-band-primary display and the four quadrants are for the narrow-band-primary display with the selected color matching functions.

Figure 7 provides a legend of the arrangement of the color patches in Fig. 8 and pictorial images in Figs. 9 and 10. In each figure, the central image (or circular patch for the Fig. 8) represents original RGB values on the broadband display. The four surrounding images, or patches, are matches computed on the narrow-band display for four representative sets of color matching functions. From upper-left to lower-right, the CIE2006(fs,age) functions (2,32), (10,32), (10,20), and (10,80) were used. These images are designed to provide some sense of how different images that match to one observer might appear to another. Of particular note are the relatively small differences between the matching colors (images) for the (10,20) and (10,80) color matching functions. One might reasonably expect larger changes with age. However, this result is due to a fortunate choice of narrow-band primaries. It happens that the shifts in the color matching functions, though large, tend to result in the approximately equal values at the wavelengths of the primaries. This result helps illustrate the important interactions between primary selection and observer metamerism.

Conclusions

A computational analysis of potential for observer metamerism in broad-band and narrow-band display technologies was completed using the recently-published CIE procedure for computing standard color matching functions for various field sizes and observer age. The results clearly showed that display systems with narrow-band primaries have significantly greater potential for observer metamerism. Thus, despite their potential for other benefits, display manufacturers might want to consider forgoing the use of narrow-band primaries for displays in color-critical applications. Instead, it might be better to improve the perceived gamut of future displays by using multiple broad-band primaries with enhanced dynamic range. It should also be noted, that the potential for observer metamerism illustrated in this paper is for average observers at each age and field size combinations. Individual observers are likely to (and indeed have been shown to [1]) have signficantly larger degrees of observer metamerism and observe even larger differences between highly-metameric color matches such as those found in displays of significantly different technology.

Figure 9. Example differences in image rendering on a narrow-band-primary display for CIE2006 color matching functions with (fs,age) = (2,32), (10,32), (10, 20) and (10,80). Central image represents broad-band display original.

Figure 10. Example differences in image rendering on a narrow-band-primary display for CIE2006 color matching functions with (fs,age) = (2,32), (10,32), (10, 20) and (10,80). Central image represents broad-band display original.

References

1. R.L. Alfvin and M.D. Fairchild, Observer variability in metameric color matches using color reproduction media, *Color Res. Appl.* **22**, 174-188 (1997).
2. CIE, *Fundamental Chromaticity Diagram with Physiological Axes - Part 1*, *CIE Pub. 170-1:2006,* (2006).
3. CIE, Special Metamerism Index: Change in Observer, *CIE Publ. No. 80*, Vienna (1989).
4. I. Nimeroff, Propagation of errors in tristimulus colorimetry, *J.Opt.Soc.Am.* **47**, 697 (1957).
5. I. Nimeroff, J.R. Rosenblatt, and M.C. Dannemiller, Variability of spectral tristimulus values, *J. Res. NBS* **65**, 475-483 (1961).
6. A. Stockman and L. Sharpe, Physiologically-based colour matching functions, *Proc. ISCC/CIE Expert Symp. '06, CIE Pub. x030:2006,* 13-20 (2006).
7. M.D. Fairchild, M.R. Rosen and G.M. Johnson, Spectral and metameric color imaging, *RIT-MCSL Technical Report*, (2001).
8. W.T. Wintringham, Color television and colorimetry, *Proc. of the IRE* **39**, 1135-1172 (1951).
9. H.E. Ives, A color-match photometer for illuminants, *J. Opt. Soc. Am.* **7**, 243-261 (1923).
10. D. Nickerson, Color measurement and its application to the grading of agricultural products, *U.S. Dept. of Agriculture Misc. Pub. 580* (1946).
11. G. Demos, Personal communications from the work of the ASC Advanced Technology Committee, (2006).
12. C. Liu and M.D. Fairchild, The surround color and color matching functions, *IS&T/SID 14th Color Imaging Conference*, Scottsdale, 203-208 (2006).
13. P. Csulti and J. Schanda, Colour matching functions based on fundamental spectral sensitivity functions, *Proc. ISCC/CIE Expert Symp. '06,* CIE Pub. x030:2006, 21-24 (2006).

Author Biographies

Mark D.Fairchild is Professor of Color Science and Director of the Munsell Color Science Laboratory (MCSL) in the Chester F. Carlson Center for Imaging Science at the Rochester Institute of Technology. David R. Wyble is a Color Scientist at MCSL and recently received his Ph.D. in color science from Chiba University.