Issues in Color Matching

Joel Barsotti, Derek Smith and L. A. Heberlein

SMPTE Mot. Imag. J 2013, 122:37-40.
doi: 10.5594/j18278

Issues in Color Matching

By Joel Barsotti, Derek Smith, and L. A. Heberlein

To create a numerical description of color (e.g., X,Y,Z), one applies a color matching function (CMF) to spectral power distribution data acquired with an instrument such as a spectroradiometer. All the adjustments one makes to a video display or to video data depend on the accuracy of these numbers. The broadcast industry and others for whom color fidelity is crucial have long depended on the 1931 CIE CMF. Recent and continuing advances in display technology, however, have exposed serious deficiencies in this CMF. These deficiencies have long been known to academic researchers, who have in the intervening years proposed several alternative CMFs. This paper reviews the critical flaws that render the 1931 CMF no longer reliable and surveys the strengths and weaknesses of candidates for its replacement.

INTRODUCTION

Anyone who looks at an image on a video display needs to be able to know that the display is presenting the image the same way another display would show it. This need is particularly acute for professionals in the broadcast, production, and digital intermediate communities who make technical and creative decisions to alter the image, based on what they see. The professional video community has adopted standards (of which ITU-R BT.709 is an extremely prominent example) in an attempt to provide a reliable consistency in image presentation if all monitors are calibrated to the same standard. If the director of photography, editor, colorist, director, and recreational viewer have all set their monitors to a white point of D65, a pure white should appear the same on all the displays.

The technology for enforcing this consistency consists of reference pattern generators, which are tested and certified at the factory to produce a calibrated pattern of exactly the characteristics the standard specifies and color analyzers—tristimulus colorimeters and spectroradiometers—that measure what the display actually produces. Software such as the products produced by SpectraCal and others can direct the generator to display a pattern, direct the color analyzer to perform a reading, analyze the difference between the target specified by the standard and the actual result measured, present the difference graphically to a user, and explain what to adjust to resolve the disparity—or, in some cases, connect directly to a controller in the display to apply the correction automatically.

All of these measurements depend on a well-established body of color science that has been remarkably stable over many years.[1] Recent advances in display technology, however, have exposed problems with the most basic steps in color measurement. This paper reviews the problems that we have uncovered and surveys some possible resolutions for the problems.

Color Matching Functions

Specifying and adjusting color requires a metrical system. Sarkhar describes the goal this way: "At the heart of colorimetry is the concept of an ideal trichromatic observer, whose color-matching properties are expressed by three independent functions of wavelength."[2]

The first attempt to meet this goal was compiled by the Commission Internationale de L'Éclairage in 1931, based on color-matching experiments conducted in the 1920s.[3]

The experiments relied on split screens. A fixed color was projected on one side of the screen, and on the other side, the observer could attempt to match the fixed color by adjusting the luminance of three monochromatic primary lights. The field of view was two degrees of angular subtense, so the resulting color matching function (CMF) is usually referred to as the "CIE 1931 2° Standard Observer," represented in **Fig. 1**.

The results of the matching are reported as three values: X, Y, and Z. Once data are captured as XYZ tristimulus values, the data are usually converted arithmetically into another color description system, such as CIE x,y,Y or L*a*b* values for ease of understanding and presentation. Most graphical presentation of color information is in one of these derived formats.[4]

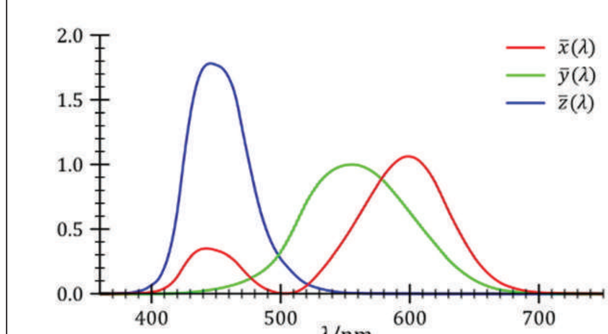

Figure 1. The CIE 1931 2° Standard Observer CMF.

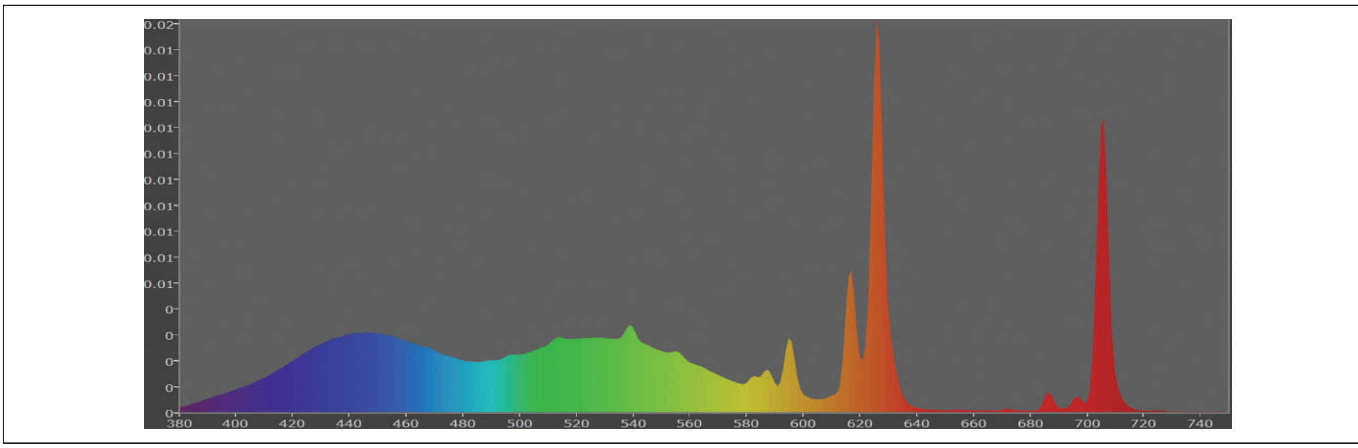

Figure 2. SPD of a typical CRT.

In association with the development of the 1964 CIE UV color definition, a 10° Standard Observer was added. Nearly all the colorimetric data in general use today derives from use of either the 2° or 10° Standard Observer. If you wished to measure color information about a video monitor, you might position yourself in front of the display and point a commercially available spectroradiometer (such as the instruments from Photo Research or Konica Minolta) at the display, and on the back of the instrument you would be able to read numerical values describing the color of the light emitted by the display. These values could be in a variety of formats, such as CIE x,y,Y, but in all cases, the displayed values are derived from X,Y,Z. The instrument may have selectable options that allow you to choose which CMF to use: 2° or 10°.

THEORETICAL DEFICIENCIES OF THE STANDARD OBSERVER

Judd reported in 1951 that the Standard Observer seriously underestimates human sensitivity to light with wavelengths below 460 nm, and he published improved CMFs that corrected for this error.[5] Stiles and Burch accumulated a dataset that documented the deficiencies of the Standard Observer and later proposed revisions to provide better results with their data.[6] Vos added further corrections to Judd in 1978,[7] incorporating the infrared color reversal described by Brindley.[8]

Yet, though the defects in the 1931 Standard Observer have been published in the scholarly literature for years, Standard Observers continue to be employed almost to the exclusion of any other CMFs. The reason is that the defects do not cause many practical difficulties in day-to-day use. If instruments that give repeatable, precise readings based on the Standard Observers are used, and the results are used to adjust the displays, the results perceived by human beings will, within the limits of the documented differences in perception from one human to another, generally agree with the results reported by the instruments. At least that has been the case up to the present day.

PRACTICAL DEFICIENCIES OF THE STANDARD OBSERVER

In 2011, the authors received calls from a knowledgable engineer at a major European broadcasting facility reporting anomalous results. If a display was adjusted using our software and a reference spectroradiometer, the result did not match other displays. Although the instrument reported that the displays were emitting light of the same color, human beings reported that the color was visibly different.

Support was given by first by double-checking the methodology of the observation, the correct use of the tools, then by supplying progressively more precise (and expensive) spectroradiometers to verify the results, and finally investigating further.

It was found was that the anomaly was caused not by any imprecision in the measurement methods, nor by the tools, but instead by the known flaws in the Standard Observer exposed by new display technology with very different spectral power distribution (SPD) compared to a tradtional CRT monitor.

Specifically, monitors are now asked to portray a much wider gamut of colors than before. The Digital Cinema P3 standard, for example, specifies primaries far outside the bounds of the Rec.709. It is more difficult for display manufacturers to achieve the wider gamut, and they are pushing their technologies beyond previous envelopes to hit the new standards. The result is SPDs that are different from SPDs of previous monitors. They are novel in several ways. The primaries may be narrower, sometimes even producing notable spikes in the SPD. There is also sometimes an unprecedented amount of energy at the very boundaries of human vision. **Fig. 2** shows the familiar SPD of a CRT, in contrast to which SPD of the novel display illustrated in **Fig. 3** shows a huge amount of energy at the high (red) end and a sharp narrow spike at the low (blue) end.

These novel SPDs cause the venerable CMFs to return anomalous results; that is, known flaws in the Standard Observer did not cause visible discrepancies when used to measure previous displays, but the new displays have sufficiently different SPDs that the flaws in the Standard Observer now create visible disparities.

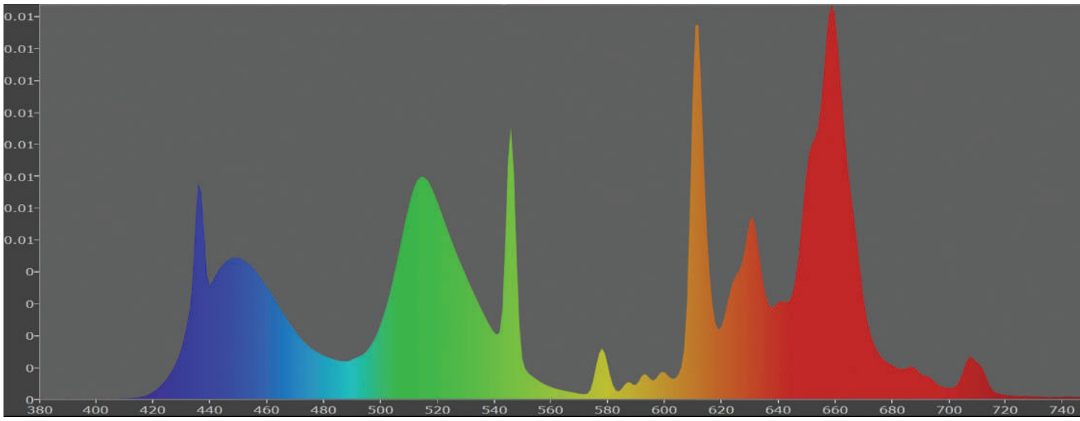

Figure 3. SPD of a novel wide gamut display.

If you calibrate the two displays with the SPDs illustrated in **Fig. 2** and **Fig. 3** using the most accurate spectroradiometers available, so that the spectroradiometers return identical XYZ values, a group of humans will agree that the two monitors are visibly different. Note that this is not (or at least not primarily) a problem of metamerism, the well-documented problem of differences in the color perception among a population. A metameric difference is one in which two individuals might report different color matches because of differences in their color-processing faculties. In a metameric problem, different humans report different results (in different directions as plotted on a graph), based on differences in their individual vision. In this problem, while, of course, each person will see things differently, they agree that the standard observer is wrong, and they agree on the direction of the error.

POTENTIAL RESOLUTIONS

It has been a long time since 1931. In the more than 80 years since the adoption of the 2° Standard Observer, investigators in universities, research institutes, and industry have proposed several alternative CMFs that ameliorate the known problems with the 2° Standard Observer. These include the work of Stiles and Birch, as well as the innovations of Judd and Vos reported above, in addition to proposals by Wyszecki[9] and Jagla Jägle.[10]

The CIE established a Technical Committee, TC 1-36, in 2000 with the goal of creating a "fundamental chromaticity diagram of which the coordinates correspond to physiologically significant axes," and in 2006 the committee produced the CIE 170-1 standard, with new CMFs that improved greatly on the 1931 Standard Observer.[11]

Peter Csuti and Janos Schanda of the University of Pannonia (in Hungary) have proposed new CMFs, informed by CIE 170-1, but designed specifically to provide correct results (that is, meter results that humans agree with) when measuring light sources such as LEDs on which new displays are based.[12]

A variety of displays have been measured with these alternative CMFs, and the alternate CMFs have been incorporated into software so that others can also accumulate a body of data.

That body needs to be larger before the color community feels warranted in abandoning a standard that has the advantages of being a long-established universal, despite its defects. Many displays have been measured many times over the past 80 years using the Standard Observers, and (because those displays have fairly similar SPDs), the measurements can with fair confidence be compared to each other, as they are based on the same fundamental formulas. If the community were to consider another standard, it would be important not to sacrifice this history.

The measurements that have been made since 1931 have not been merely passive. They have been used to adjust monitors that have been used to produce content. It would seem very dangerous to propose a new standard that says to a content producer, "That monitor on which you have mastered all your material for years is now wrong. Starting today, I am going to adjust that monitor to different settings, and all of your material will appear different on it."

Clearly, the most desired outcome would be to discover that one of the alternate CMFs produces the same results on all the old displays, but also produces correct results on the new displays. (*Correct* being defined as results that allow those displays to be calibrated to match the old displays.) If we calibrate an old reference CRT with the new CMF, it would be preferable not to change its settings or appearance. However, if a prototype display is received from a manufacturer, using previously unknown technology, producing an unusually shaped SPD, it would be desirable to be able to calibrate this novel display to the same standard as the old CRT. It would then be possible to call in a group of experienced professionals to observe the result and have them report that the two displays appear in the same color. There is currently vigorous debate as to whether this outcome is achievable.

CONCLUSION

The relentless innovation of display manufacturers, particularly the unprecedented SPDs created in an attempt to produce wider gamut displays in order to meet such new standards as P3, have exposed serious deficiencies in the CMFs that underlie all colorimetry as it is practiced today. If we use the old methods on the new displays, meters report results that humans dispute. The meters say the displays are producing the same color, but the humans agree that the colors are different. (That is, this is not just a problem of metamerism, but a more fundamental problem with the metrology.) We cannot continue using the old CMFs if we want to calibrate these new displays so that they are usable.

Yet the dangers of switching to new CMFs are daunting. No one wants to invalidate previous data or change our view of vaults of existing video material. We do not know enough today to confidently recommend an answer. However, the availability of alternate CMFs in contemporary color measurement software allows us to accumulate the data that will inform the decision that must be made.

REFERENCES

1. Charles A. Poynton, *Digital Video and HD*, 2nd ed.: *Algorithms and Interfaces*, Morgan Kaufmann: San Francisco, 2012.
2. Abhijit Sarkar, "Identification and Assignment of Colorimetric Observer Categories and Their Applications in Color and Vision Sciences," Ph.D. dissertation, Ecole polytechnique de l'Université de Nantes, 2011.
3. CIE No 15-2, "Colorimetery," 2nd ed., Commission Internationale de L'Éclairage, Vienna, Austria, 1986.
4. Deane B. Judd and Günter Wyszecki, *Color in Business, Science and Industry*, 3rd ed., John Wiley: New York, 1975.
5. D. B. Judd, "Report of U.S. Secretariat Committee on Colorimetry and Artificial Daylight," *Proc. Twelfth Session of the CIE*, p. 11, 1951.
6. W. S. Stiles and J. M. Burch, "Interim Report to the Commission Internationale de l'Eclairage Zurich 1955, on the National Physical Laboratory's Investigation of Colour-Matching," *Optica Acta*, 4:168-181, 1955.
7. J. J. Vos, "Colorimetric and Photometric Properties of a 2° Fundamental Observer," *Color Res. Appl.*, 3:125-128, 1978.
8. G. S. Brindley, "The Colour of Light of Very Long Wavelength," *J. Physiol.*, 130:35-44, 1955.
9. Günther Wyszecki, "Proposal for a New Color-Difference Formula," *J. Opt. Soc. Am.*, 11:1318-139, Nov. 1963.
10. L. T. Sharpe, A. Stockman, W. Jagla, and H. Jägle, "A Luminous Efficiency Function, V*(λ), for Daylight Adaptation," *J. Vision*, 11:948-968, 2005.
11. CIE No 170-1, Fundamental Chromaticity Diagram with Physiological Axes, Commission Internationale de L'Éclairage, Vienna, Austria, 2006.
12. Peter Csuti, János Schanda, Raghuram Petluri, Kelly McGroddy, Gerard Harbers, "Improved Color Matching Functions for Better Visual Matching of Led Sources," Proceedings of 27th Session of the CIE, 10-15. July 2011, Sun City, South Africa, ISBN 978 3 901906 99 2.

Presented at the SMPTE 2012 Annual Technical Conference and Exhibition, Hollywood, CA, 23-25 October 2012. Copyright © 2013 by SMPTE.

THE AUTHORS

Joel Barsotti is director of software development at SpectraCal, Inc. As a veteran of the software industry, Barsotti spent the first several years working at a graphic design studio managing color critical work stations. He then began writing his own software to calibrate PCs and home theater computers, and was then hired away to SpectraCal. In his current position, he has presided over the development of one of the most sophisticated color management packages available. The research he has done developing this engine makes him uniquely suited to discussing color matching functions and their role in video calibration.

Derek Smith wrote CalMAN video calibration software when he found there was no available software that would properly calibrate his own home theater, and then went on to found SpectraCal when he realized others might be interested. In the previous quarter century he developed an extremely wide variety of software for nearly every kind of device you can connect to a computer, from the most rudimentary UART to peak performing race cars. An extreme skier and Tae Kwon Do instructor, he brings an intense passion for excellence to everything he undertakes, and for several years, his primary focus has been the practical application of color science to ensuring image fidelity in a broad range of industries, from medical imaging to post-production.

L.A. Heberlein is the president of SpectraCal, Inc., leading provider of video calibration software to home theater, commercial A/V, medical imaging, broadcast, video production, and post-production. Presented at the SMPTE 2012 Annual Technical Conference and Exhibition, Hollywood, CA, 23-25 October 2012. Copyright © 2013 by SMPTE. has led teams of software developers for more than 30 years, repeatedly undertaking first forays into new areas of expansion for computer science. These have included implementation of the first commercial operating system on RISC and the first multi-user services for Microsoft operating systems.