Journal of Imaging Science and Technology® 58(3): 030402-1–030402-20, 2014.

Modeling Observer Variability and Metamerism Failure in Electronic Color Displays

David L. Long and Mark D. Fairchild▲
Rochester Institute of Technology, Program of Color Science, 70 Lomb Memorial Drive Rochester, New York 14623
E-mail: david.long@rit.edu

Abstract. *The electronic display industry has begun a migration towards higher color gamut devices driven by LED, OLED, quantum dot and laser technologies capable of generating near monochromatic color stimuli in the traditional red, green, blue three-channel paradigm. The use of highly selective spectral stimuli, however, poses a risk to the consistency of visual experience amongst a group of disparate, but otherwise normal, color observers. Several models of spectral color vision have surfaced in recent research and are helping investigators to better understand the implications for color experience variability. The present research serves to summarize various color difference indices that may be useful in predicting the magnitude of observer response inconsistencies and applies them to simulations of current electronic displays as examples of potential concerns these new high-gamut technologies might raise. In particular, various laser-based displays are shown to perform with significantly increased observer variability versus traditional ITU-R Rec. 709 and SMPTE 431 RGB-primary displays utilized in the cinema industry. Further, observer metamerism can be reduced significantly with proper optimization of a multichannel projection system comprising seven explicitly designed primary spectra.* © *2014 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.ImagingSci.Technol.2014.58.3.030402]

INTRODUCTION

In designing digital color management strategies for still photography, computer graphics or motion picture imaging systems, the principal model employed for color vision comes from the International Commission on Illumination (CIE) 1931 2° standard observer.[1] This single trichromatic model summarizes a mathematical representation of the spectral sensitivity of the three integrated channels of human color vision isolated to the 2° field-of-view of the fovea. These color response curves were derived from bipartite field color matching experiments executed by Guild and Wright in the 1920s, involving 17 observers and validated by the CIE as representative of the worldwide population of normal color observers.[2] The published standard observer spectral responses represent an intentional transformation of the actual average data collected from these experiments to a form based on non-realizable primaries yielding $\bar{x}_\lambda$, $\bar{y}_\lambda$ and $\bar{z}_\lambda$ color matching functions (CMFs), summarized in Figure 1. The transformation is designed such that each spectral curve contains all positive values (a necessity of colorimeter hardware developed concurrent to the standard) and such that the 1924 photometric response curve, V_λ, can be matched by the $\bar{y}_\lambda$ function.

In 1964, the CIE sanctioned the addition of a wider field standard observer to be used in colorimetry of larger field-of-view stimuli.[2] The data were collected in 1959 in separate experiments at high illumination levels with 49 observers by Stiles and Burch[3] and at low illumination levels with 27 observers by Speranskaya,[4] with each experiment subtending a 10° visual field. Designated as $\bar{x}_{10\lambda}$, $\bar{y}_{10\lambda}$ and $\bar{z}_{10\lambda}$ and shown also in Fig. 1, these response curves have a firmer statistical grounding than the 1931 set. However, the 10° observer has no mathematical connection to modern photometry or the universally-used V_λ response, and most imaging industries have continued to employ system design based on the older narrower field observer.

Concerns for both the 1931 and 1964 CIE standard observers surround their derivation from limited demographic populations and their declaration of average behavior for all color-normal observers. In the 1980s, the CIE attempted to address inadequacies in models of observer variability and observer metamerism by introducing the Standard Deviate Observer.[5] These color matching functions were computed from differences amongst the original 1959 Stiles and Burch data and permitted confidence limits to be calculated for any colorimetric calculation. Unfortunately, subsequent research with this observer set has found it to grossly underpredict real observer variability.[2]

More recent research has generated greatly improved understanding of the anatomical and optical disparities amongst color-normal human observers. The CIE 2006 model (from the work of CIE TC1-36) summarizes a prediction of fundamental cone sensitivities and corresponding CMFs as dependent on observer age and field-of-view.[6] The general form of predicted $\bar{l}_\lambda$, $\bar{m}_\lambda$ and $\bar{s}_\lambda$ cone fundamentals is summarized in Eq. (1). Specifically, cone absorptivities, α_λ, and maximum macular density, $D_{\tau,\mathrm{max,macula}}$, are treated as field-size dependent, based on anatomical studies associated with each. Ocular media densities, $D_{\tau,\mathrm{ocul}}$, do not vary with field-of-view but are known to vary with observer age. The cone fundamentals can be further transformed to CMFs via matrices recommended by CIE TC1-36 and used in calculating colorimetry and color difference values for compared stimuli. Specifically, CIE TC1-36 defines an LMS-to-XYZ matrix for converting the 32-year-old observer

▲ IS&T Member.

Received May 1, 2014; accepted for publication Aug. 20, 2014; published online Oct. 9, 2014. Associate Editor: Marius Pedersen.
1062-3701/2014/58(3)/030402/20/$25.00

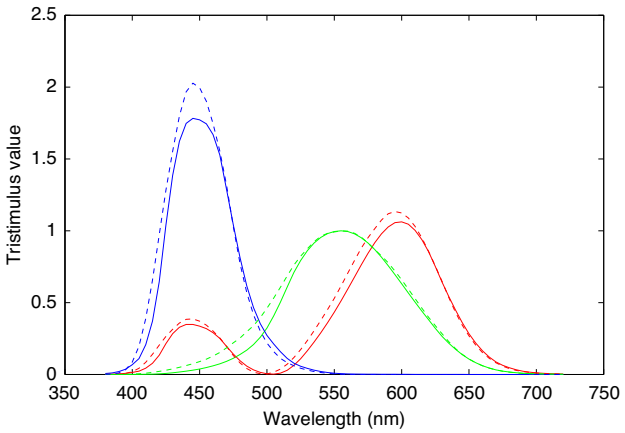

Figure 1. The CIE 1931 2° (solid) and 1964 10° (dashed) standard observer color matching functions.

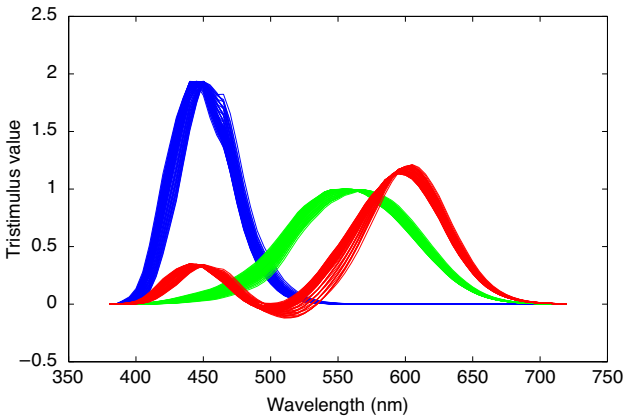

Figure 2. The CIE 2006 color matching functions for observers ranging from 20 to 80 years of age and across 1° to 10° field-of-view.

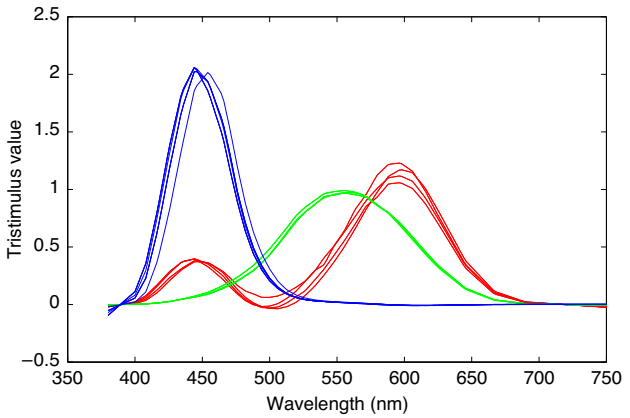

Figure 3. The Sarkar/Fedutina color matching functions.

in the 2° cone fundamental model to best match the 1931 CIE standard observer. A second matrix is used to transform the 32-year-old/10° cone fundamentals to the 1964 standard observer. CIE TC1-82 is currently refining the methodology to convert cone fundamentals from any age and field-of-view definition to an appropriate CMF. In the present work, however, the absolute variability of observer response is a key attribute analyzed. In an attempt to not diminish or exaggerate this variability from established cone fundamental models for which there are no corresponding CMF data, only the 2° LMS-to-XYZ matrix is considered for all transformations. Figure 2 summarizes a sampled collection of modeled observer CMFs spanning ages 20 to 80 and fields-of-view from 1° to 10°. Several researchers have pointed out that the CIE's model is imperfect in predicting the spectral behaviors of any single real observer but that the models generally encompass the ranges expected in a normal population.

$$\begin{aligned}\bar{l}_\lambda &= \alpha_{i,l,\lambda}\cdot 10^{-D_{\tau,\max,\text{macula}}\cdot D_{\text{macula relative},\lambda}-D_{\tau,\text{ocul},\lambda}},\\ \bar{m}_\lambda &= \alpha_{i,m,\lambda}\cdot 10^{-D_{\tau,\max,\text{macula}}\cdot D_{\text{macula relative},\lambda}-D_{\tau,\text{ocul},\lambda}},\\ \bar{s}_\lambda &= \alpha_{i,s,\lambda}\cdot 10^{-D_{\tau,\max,\text{macula}}\cdot D_{\text{macula relative},\lambda}-D_{\tau,\text{ocul},\lambda}}.\end{aligned} \quad (1)$$

In computational models, Sarkar et al.[7,8] have statistically grouped 47 of the Stiles and Burch observers into seven general base CMF sets by minimizing colorimetric prediction errors. The full candidate CMF sets were originated from 125 permutations derived from five distinct $\bar{l}_\lambda$, $\bar{m}_\lambda$ and $\bar{s}_\lambda$ cone fundamentals each. The five discrete fundamentals per cone type originated from cluster analysis on the Stiles and Burch data set together with 61 variations calculated from the CIE 2006 models for observer ages between 20 and 80. Sarkar used the categorization approach to successfully identify the primary color matching function descriptor of 30 real observers in a highly metameric matching experiment. Fedutina et al.[9] further confirmed the viability of the generalized Sarkar observers but refined the fundamental set to eight candidates using more metameric classification stimuli. Figure 3 summarizes the final CMFs, which were again each produced via transformation of cone functions using a single optimized LMS-to-XYZ matrix for all candidates. An observer calibrator apparatus was also constructed with narrow-band LED test primaries to classify any real observer into one of the fundamental CMF categories.[10]

Alfvin and Fairchild[11] as well as Fairchild and Heckaman[12] have utilized Monte Carlo models to generate color matching functions for likely observers based on real quantified anatomical variability in spectral lens transmission, macula density and $\bar{l}_\lambda$, $\bar{m}_\lambda$ and $\bar{s}_\lambda$ cone sensitivities. In the Heckaman examples, age-dependent transmission characteristics of the crystalline lens as described by Pokorny et al.[13,14] and Xu et al.[15] are taken and used to randomly generate transmission characteristics against US census data. Next, the macula density function described by Bone et al.[16] is similarly normally varied in peak density to conform to standard deviation values suggested by Berendschott and van Norren.[17] Finally, the cone fundamentals of Stockman et al.[18,19] are varied according to genetic models suggested by Neitz and Neitz,[20] and selections of cone response with distributions in L- and M-type peak absorptions are made comprising the final modeled physiology. A heuristic Monte Carlo collection of 1000 fictitious observers is generated and made available

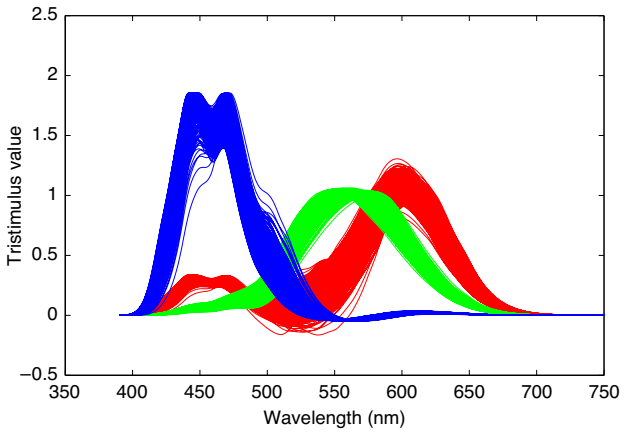

Figure 4. The Fairchild and Heckaman 2° color matching functions.

to compute probable distributions of observer variability and metamerism for real colorimetric match scenarios. Heckaman has generated CMFs via this method using a single LMS-to-XYZ matrix to center responses against the 2° 1931 observer and using a second matrix to center all CMFs about the 10° 1964 observer. The 2° set of 1000 CMFs considered in the present work is shown in Figure 4.

While none of these techniques are able to characterize precise color matching functions of any single actual observer, they each present an extent of response potentials useful in analyzing metameric failures in reproduced imagery on displays. Or, in the case of Sarkar and Fedutina, they offer potential for compartmentalizing real observers into broader populations of reasonably similar color sensitivity, permitting discrete display customization in color-critical applications. An example would be observer-dependent calibration of electronic displays for mixed-media color comparators used in print publishing.

Another example pertinent to motion picture workflows would be display calibration for mastering color-corrected content across multiple distribution platforms (film, ITU-R Rec. 709 television displays, SMPTE 431 digital cinema projectors, ITU-R Rec. 2020 monochromatic primary displays, etc.). What is most important in considering future digital color management paradigms is that advanced display technologies will necessarily challenge the utility of a single standard observer model to represent best practice color mastering. Creative professionals with one particular color response function may be generating aesthetic choices interpreted in very different ways by a full population of observers viewing content on narrow-spectra wide-gamut color displays.

OBSERVER METAMERISM INDICES

The quantification of observer metamerism for critical analysis demands attention to two different attributes of disparate CMF populations, color mismatch magnitude and observer variability. The former addresses traditional issues of color calibration where a device is tuned to deliver a color response against aim as defined by standard colorimetry employing intentionally chosen CMFs. The CIE has published three color difference formulas used widely in contemporary color industries, ΔE_{ab}, ΔE_{94} and ΔE_{00}, which are each derived from the 1976 CIELAB color space. The 1994 and 2000 permutations address failures of perceptual uniformity in CIELAB and the Euclidean ΔE_{ab} vector length calculation. Still, the premise of the CIELAB space and its validity as base index for metamerism quantification remains sound. The CIELAB coordinate system acts as an elementary color appearance space, defined in orthogonal axes of lightness perception, approximate red–green hue/chroma perception and approximate blue–yellow hue/chroma. The a^*b^* plane can be further considered as a circular coordinate system with appearance attribute hue represented as angle and chroma as distance from origin. Accepted appearance phenomena represented in the CIELAB encoding include a CMF-relevant chromatic adaptation, a reference white lightness adaptation and exponential radiometric scaling associated with visual perception uniformity. CIELAB itself is derived via input of XYZ tristimulus coordinates. By varying the CMF chosen to compute XYZ, CIELAB can serve as a reasonable appearance model for a specific theoretical observer, and thus color difference indices calculated can be presumed to be appearance relevant for that same observer. This practice is common, for example, in interchanging the 1931 and 1964 standard observers into CIELAB calculations as warranted by different applications. Ohsawa et al.[21] have inferred that such interchange is useful for interrogating observer statistics in cases where the field size is not even a practical factor. The models of CIE 2006, Sarkar/Fedutina and Heckaman all support general demographic analyses with their observer CMFs. In evaluating distributions amongst observer CMFs within a population, this tactic becomes critical for providing a uniform translation of color error when the spectral responsivity is intentionally varied.

Turning to observer variability, gross observer response inconsistencies are less an issue of absolute magnitude of color difference perception and more an issue of the variance of color differences experienced by a group of defined observers. The two are decoupled in the example where the overall color difference from the reference for each of a set of disparate observers is large but the shared experience amongst the observers relative to one another is similar. The opposite scenario is also possible, though to a lesser significance, where each observer may experience a small perception of color difference from the reference but the population of observers perceive significantly different experiences in hue, chroma or lightness error from one another. Several indices of observer response variability can be described by treating color difference not as a directionless quantity in CIELAB space but instead by breaking error vectors into their constituent axial components in the three-dimensional space. Use of ΔL^*, Δa^* and Δb^* designations (where the origin of the color space represents a perfect colorimetric match) permits the creation of an error ellipsoid in CIELAB whose volume is proportional to the magnitude of observer variability in assessing test and

reference stimuli. Again, each observer contributes a unique CMF in computing the full set of $\Delta L^*a^*b^*$ vectors, but the magnitude and direction of error from the reference are deemed relatable by treatment of CIELAB as a uniform color appearance space for small magnitude differences.

In the present research, the following indices are used to quantify observer metamerism magnitude and variability. Stimuli pairs may derive from any established reference spectrum and a corresponding reproduction spectrum.

$$\mathrm{OM}_x = \max(\overline{\Delta E_{y,P,i}}), \tag{2}$$

$$\mathrm{OM}_{x,\max} = \max(\Delta E_{y,P,i}), \tag{3}$$

where OM_x refers to observer metamerism magnitude based on CMF sets from $x =$ Sarkar/Fedutina (s), CIE 2006/TC1-36 (c) or Heckaman (h). Color difference values between a reference stimulus and the test sample are computed for $y = \Delta E_{ab}$ (ab), ΔE_{94} (94) or ΔE_{00} (00) for each patch in a patchset P for each observer i in the CMF set. The observer metamerism magnitude is the maximum individual observer average patchset color difference across all the patches in P. In this manner, the observer metamerism represents the on-average poorest color matching observer from the population of CMFs for the patchset. A slight variation of this index, $\mathrm{OM}_{x,\max}$, is based on measurement of the worst color difference patch across all observers in the given CMF set. This is thus the worst color match achieved across a full set of stimuli in the patchset considering all candidate observers. To minimize either of these indices suggests a move towards improving the color match between two stimuli for all observers in a population and thus a minimization of observer metamerism magnitude.

Observer variability indices are summarized by Eqs. (4) and (5):

$$\mathrm{OM}_{x,\mathrm{var}} = \overline{\mathrm{Vol}(\Delta\,(L^*a^*b^*)_P)}, \tag{4}$$

$$\mathrm{OM}_{x,\mathrm{varmax}} = \max\left(\mathrm{Vol}\left(\Delta\left(L^*a^*b^*\right)_P\right)\right), \tag{5}$$

where $\mathrm{OM}_{x,\mathrm{var}}$ refers to observer metamerism variability, the mean CIELAB ellipsoid volume constructed from CMF-based error vectors in L^*, a^* and b^* from each patch in a patchset P. The index is again dependent on the CMF set chosen as above. For the present work, covariance analysis is used to construct the ellipsoid volumes from individual observer CIELAB error vectors with a 90% statistical significance. $\mathrm{OM}_{x,\mathrm{varmax}}$ is the maximum ellipsoid volume from all patches in the patchset and is thus the particular stimuli pair with the broadest observer variability.

Fairchild et al. have documented a methodology used to evaluate observer metamerism in additive electronic displays employing the CIE 2006 color matching function models for observers of varying ages and subtending various angular fields of view.[22] Primary drive amounts needed to enforce a metameric match between aim spectra and the generated reproduction are calculated using a chosen CIE 2006 color matching function. Once matched for that particular observer, the resultant modeled spectra of each system are assessed for subsequent colorimetric match assuming the 1931 2° standard observer and resulting color difference values are tallied. This methodology maintains the benefits of using a single CMF color space for all determined color difference indices and also allows RGB color rendering of differences for visualization. The method, however, does not permit a summary of the color difference experienced by any particular disparate observer within the context of their own CMF, and so the previous indices summarized are preferred in these analyses.

Other traditional indices of color difference for a pair of stimuli invoke assessment of the spectral power distribution of the samples. As the spectral signatures for the compared colors become more similar, all attributes of perceived color difference, regardless of observer CMF, will shrink to zero. Two spectra may be compared by assessing the root mean square of spectral differences (RMSE) across a defined range of wavelengths or by assessing the maximum spectral error at any wavelength between the two samples. Many researchers prefer the latter because it is plausible for the RMSE to be small while a single wavelength may experience a large and consequential error, but the opposite is seldom true. In the present research, all errors are scaled as a fraction of the reference stimuli maximum radiometric power prior to the RMSE or maximum error computation. This permits analysis in relative spectral power output for comparing significance amongst stimuli of variable absolute spectral power. It also permits comparison of spectra in a more perceptually uniform context.

Finally, any stimuli pair may also be compared by accepted color difference formulas for a standard observer. The present research utilizes the 1931 standard observer, common to imaging system color evaluations. As appropriate, ΔE_{ab}, ΔE_{94} or ΔE_{00} is considered.

The various indices previously defined offer candidate response treatments for quantifying color error and color response variability amongst a group of observers interacting with colors reproduced on different additive electronic displays. However, such an analysis requires a sensible color reproduction objective for each evaluated display to be defined. In the present research, cross-media metamerism is evaluated by forcing a best match of spectral or colorimetric display output to a series of conventionally illuminated reflective test patch aims. The patchsets considered include

(1) MacBeth Color Checker (24 samples),
(2) MacBeth Color Checker DC (240 samples),
(3) US Patent No. 5,582,961 "Kodak/AMPAS" test spectra (190 samples),
(4) Munsell sample spectra (1269 samples),
(5) select high metamerism color set (65 samples).

Luminous spectral stimuli are produced via a model of these patchsets under CIE D65, CIE Illuminant A, a measured hydrargyrum medium-arc iodide (HMI) motion picture studio lamp and CIE fluorescent illuminant F2. Although comparison of different displays in metameric match to one another is common practice in motion picture workflows, an analysis encompassing metameric match

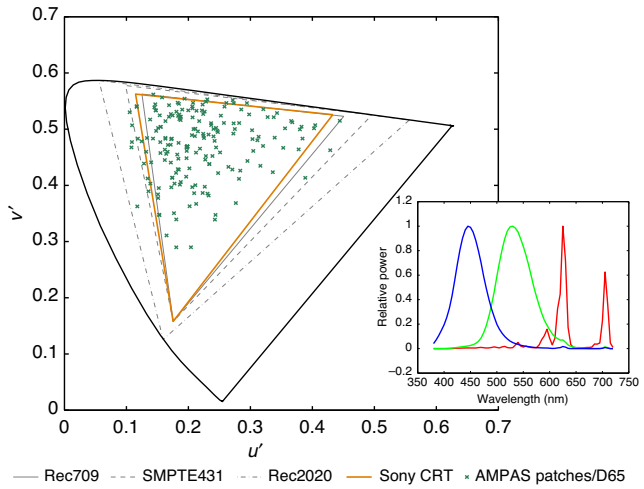

Figure 5. Sony PVM 14L2 CRT chromaticity gamut and peak-normalized primary spectra; color points representing Kodak/AMPAS color patches illuminated by CIE D65 are also included.

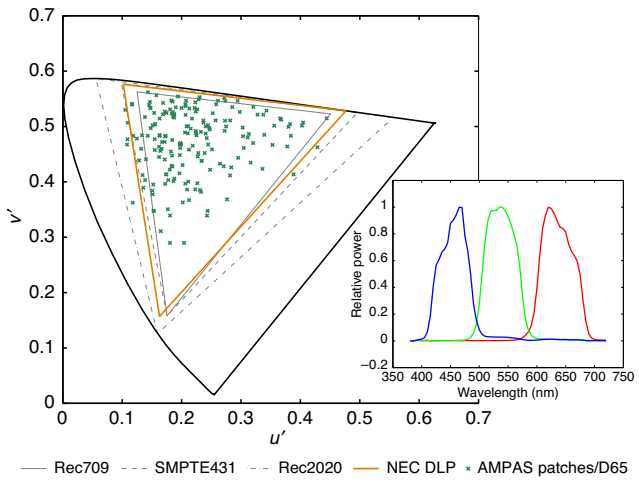

Figure 6. NEC 3000 digital cinema projector chromaticity gamut and peak-normalized primary spectra; color points representing Kodak/AMPAS color patches illuminated by CIE D65 are also included.

to real surface colors offers a broader interpretation of experiment results. Specifically, color and spectral matching of real scene stimuli on the display screen bridges the workflow between image acquisition and reproduction, setting an expectation for exhibition color reproduction control that exceeds current trichromatic convention and permits evolution to future spectral color correction models.

OBSERVER METAMERISM SIMULATIONS

To simulate observer metamerism in additive displays, six different systems were chosen and their primary spectra collected:

(1) Sony 14L2 PVM-class professional CRT;
(2) NEC3000 3-DLP SMPTE 431 professional digital cinema projector;

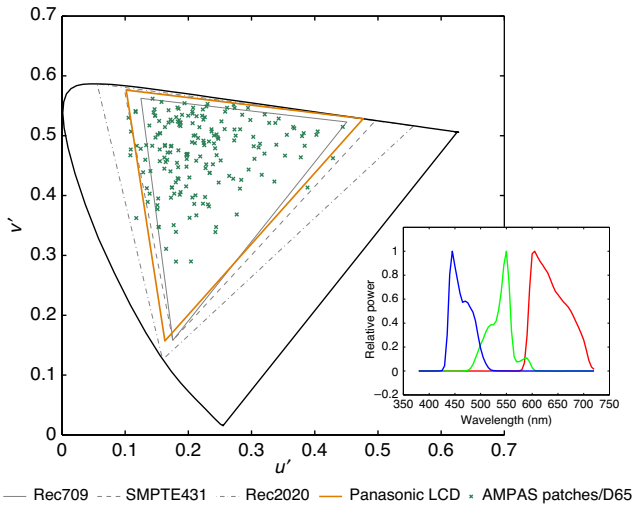

Figure 7. Panasonic PTAX200U LCD cinema projector chromaticity gamut and peak-normalized primary spectra; color points representing Kodak/AMPAS color patches illuminated by CIE D65 are also included.

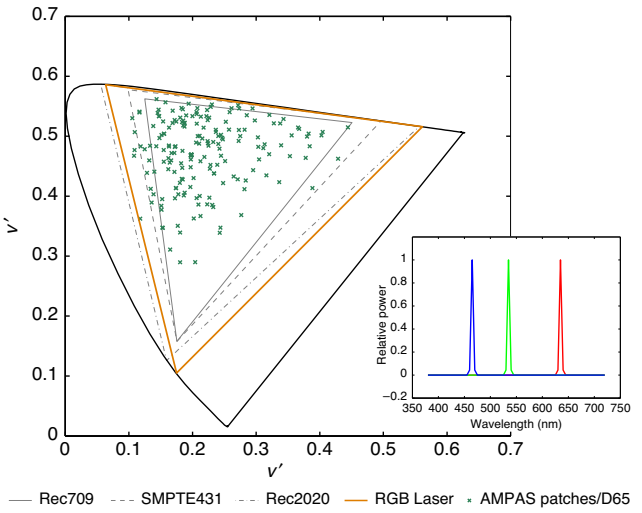

Figure 8. Example ITU-R Rec. 2020 RGB laser projector chromaticity gamut and peak-normalized primary spectra; color points representing Kodak/AMPAS color patches illuminated by CIE D65 are also included.

(3) Panasonic PTAX200U 3-LCD SMPTE 431 HDTV consumer projector;
(4) Prototype ITU-R Rec. 2020-compatible laser cinema projector;
(5) chromaticity-gamut-area-optimized eight-primary laser projector;
(6) metamerism-optimized seven-channel projector,

The $u'v'$ chromaticity-space gamuts of each of the displays are shown in Figures 5–10 along with normalized plots of measured spectra for each of the system color channels. Also included for gamut perspective are the chromaticity coordinates of the Kodak/AMPAS color patchset illuminated by CIE D65 and the boundaries of standard display gamuts defined by ITU-R Rec. 709 and Rec. 2020 and SMPTE 431's Digital Cinema Reference Projector, 'P3'. Systems (1)–(3) were chosen as representative of current motion picture

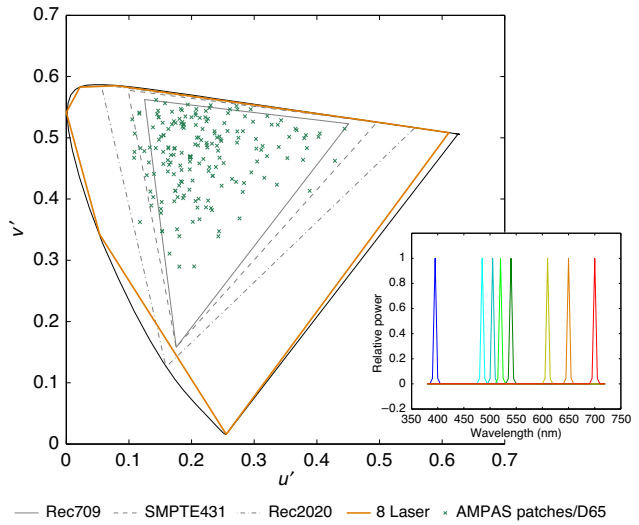

Figure 9. Maximized chromaticity-area eight-primary laser projector chromaticity gamut and peak-normalized primary spectra; color points representing Kodak/AMPAS color patches illuminated by CIE D65 are also included.

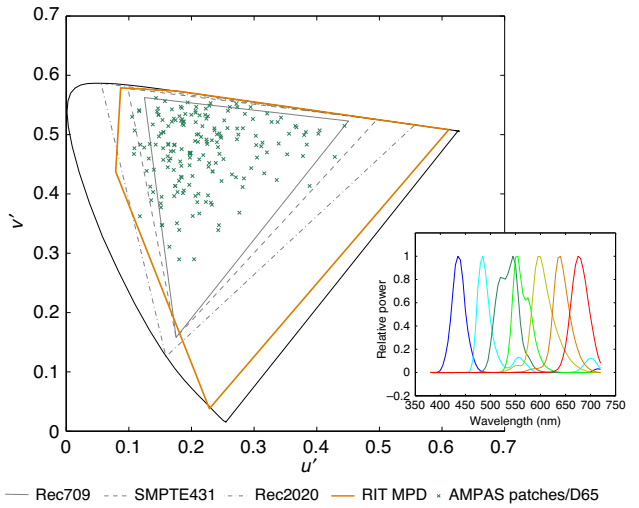

Figure 10. RIT seven-channel projector chromaticity gamut and peak-normalized primary spectra; color points representing Kodak/AMPAS color patches illuminated by CIE D65 are also included.

industry three-channel primary standards, including current HDTV video and current digital cinema exhibition. Rec. 2020 represents a next-generation laser display standard with wavelengths of 467, 532 and 630 nm. The gamut-optimized laser projector was modeled based on maximizing the polygon area of the display's *xy* chromaticity gamut versus the CIE spectral locus, using eight channels. The wavelengths thus determined were 395, 485, 505, 520, 540, 610, 650 and 700 nm. Chromaticity-plane color gamut is often touted in professional electronic display marketing materials and so this hypothetical multiprimary system with absolute maximum performance was conceived for comparison with the actual display systems. The metamerism-optimized display represents the color characteristics of a prototype multiprimary display built at Rochester Institute of Technology (RIT) to evaluate models of observer variability. This display was designed explicitly to generate a reduced observer metamerism according to Sarkar/Fedutina CMF models and to further the prior work of Koenig et al.[23]

For initial assessment, the chosen displays were color managed to match the various reference stimuli under the various illuminants according to 1931 standard observer color difference indices. Because systems (5) and (6) are over-specified in this objective (due to eight and seven adjustable primaries, respectively), these displays were co-optimized to constrain an exact metameric match to the stimuli as determined by the 1931 standard CMF set while subsequently minimizing OM_x. This optimization was not run for the 1269 Munsell color patches due to extreme calculation times in the simulations. For some color patches on these two displays, the color stimuli were outside the reproducible gamut of the device and so observer metamerism minimization alone was employed. For similar out-of-gamut failures on the three-channel displays, a minimization of the 1931 standard observer color difference was used rather than an observer metamerism optimization so as to faithfully maintain the original color management intent for an RGB system. A summary of observer metamerism indices for each display modeled to reproduce the color of each candidate patchset under each illuminant is presented in Tables I–IV. In each of these assessments, the Sarkar/Fedutina CMF set is used to generate the computed metamerism index values. A maximum 1931 2° color difference (CIE2000) of 0.0 for a given patchset in these tables is evidence that all patches were within the given display's gamut and rendered colorimetrically perfect to the standard observer according to the simulation intent employed. Again, where these color difference maxima are greater than 0.0, not all patches within the set were in gamut and an alternate optimization was executed for those patches.

An investigation of results for the D65-illuminated stimuli reveals very consistent performance across the five patchsets for the six modeled displays. In each case, the metamerism magnitude, OM_s (based here on simple ΔE_{ab}), is best for the RIT multiprimary display and worst for the eight-laser system by a ratio of at least 10:1. The Rec. 709 CRT and SMPTE DLP and LCD projectors represent the current display technologies used for cinema applications and so set the baseline for comparison with the other devices. In general, the professional-grade digital cinema projector from NEC is better than the consumer-grade Panasonic device (Figures 6 and 7 reveal how each delivers near-exact P3 chromaticity gamut with notably different primary spectra) and the CRT performs reasonably close to both. Each of these legacy systems though is deficient versus the RIT prototype by a factor of 2× to 3×. Models of the Rec. 2020 laser gamut projector yield a significant drop in color match versus the legacy equipment, although the performance is still not as poor as the eight-laser system. Delving deeper into the maximum color error amongst the eight Sarkar/Fedutina observers and amongst all the patches in each set, $OM_{s,max}$, very similar trends in both rank order and magnitude of

Table I. Sarkar/Fedutina observer metamerism indices for various displays relative to test patchsets illuminated by CIE D65 (1931 2° colorimetry match).

CIE D65	OM_S	$OM_{S,max}$	$OM_{S,var}$	$OM_{S,varmax}$	Mean RMSE	Mean peak err	Max DE00(31)
AMPAS190							
Sony CRT	2.77	17.13	4.6E–03	1.8E–01	0.44	1.92	6.43
NEC DLP	2.48	14.40	1.8E–03	1.3E–01	0.25	0.55	4.83
Panasonic LCD	2.71	10.20	2.4E–03	5.3E–02	0.27	0.77	3.35
Rec2020 Laser	5.50	11.47	3.8E–01	4.7E+00	2.07	9.41	0.00
8-laser	10.78	26.83	2.5E+02	1.9E+03	1.95	10.22	0.00
RIT MPD	0.79	6.35	1.0E–05	3.6E–04	0.28	0.63	0.00
MacBeth24							
Sony CRT	2.15	8.77	2.6E–03	4.7E–02	0.44	1.95	0.44
NEC DLP	1.83	8.52	2.8E–04	2.7E–03	0.25	0.52	0.00
Panasonic LCD	2.49	5.20	1.0E–03	5.5E–03	0.27	0.76	0.00
Rec2020 Laser	5.50	10.44	2.6E–01	1.3E+00	2.18	9.66	0.00
8-laser	11.61	27.31	3.1E+02	2.0E+03	2.08	11.01	0.00
RIT MPD	0.78	2.43	6.2E–06	7.5E–05	0.31	0.66	0.00
MacBeth DC							
Sony CRT	2.55	32.39	2.6E–02	2.4E+00	0.49	2.15	14.64
NEC DLP	2.28	25.36	8.2E–03	6.4E–01	0.30	0.60	11.21
Panasonic LCD	2.60	25.00	1.6E–03	1.3E–01	0.31	0.88	11.32
Rec2020 Laser	5.57	14.38	4.0E–01	2.7E+00	2.41	10.21	1.66
8-laser	11.53	27.89	2.8E+02	1.2E+03	2.35	12.34	0.00
RIT MPD	0.81	9.77	3.5E–04	8.1E–02	0.38	0.77	7.39
Big Metamers							
Sony CRT	5.57	24.47	5.1E–02	1.1E+00	0.40	1.65	8.60
NEC DLP	4.69	21.71	1.8E–02	2.2E–01	0.23	0.53	7.18
Panasonic LCD	4.26	16.83	8.3E–03	2.6E–01	0.25	0.71	5.90
Rec2020 Laser	5.38	16.02	3.3E–01	2.8E+00	1.57	7.40	2.22
8-laser	8.21	26.83	1.2E+02	1.9E+03	1.46	7.57	0.00
RIT MPD	0.71	2.84	1.7E–05	3.7E–04	0.20	0.51	2.91
Munsell							
Sony CRT	1.95	11.10	2.3E–03	1.8E–01	0.49	2.19	1.22
NEC DLP	1.94	10.61	8.5E–04	7.1E–02	0.30	0.62	0.00
Panasonic LCD	2.43	8.36	9.5E–04	1.2E–02	0.32	0.87	0.00
Rec2020 Laser	5.60	10.87	3.2E–01	2.6E+00	2.47	10.49	0.00
8-laser	–	–	–	–	–	–	–
RIT MPD	–	–	–	–	–	–	–

performance are noted, although the consumer P3 projector does fare better relative to the professional system than it did for average observer metamerism. The most telling trend for these results is the poor performance achieved by increasingly monochromatic primary sets. As such, enlarged chromaticity-area gamut is traded in these systems for a reduced observer metamerism.

Observer set variability, as modeled by color error ellipsoid volumes, tracks well with the trends in overall color difference magnitude. Again, the RIT MPD performs best and the eight-laser system worst. The variability index also proves to be much more sensitive to display change as there are roughly seven orders of magnitude in mean metamerism variability and maximum metamerism variability between the two. The CRT, DLP and LCD displays perform two orders of magnitude poorer than the RIT display, and the Rec. 2020 laser drops another two orders of magnitude from there. Figures. 11a–11f show the CIELAB error ellipsoids for the 24 MacBeth color checker patches illuminated by D65 for each of the simulated displays. Plots are presented with common scaling of axes to permit proper examination of the comparative variability. An interesting attribute of these figures is the lack of symmetry about the $\Delta L^*a^*b^*$ origin; metameric matches generated for the 1931 2° observer yield hue, saturation and lightness bias for the Sarkar/Fedutina observers.

Table II. Sarkar/Fedutina observer metamerism indices for various displays relative to test patchsets illuminated by CIE Illuminant A (1931 2° colorimetry match).

CIE IllumA	OM_s	$OM_{s,max}$	$OM_{s,var}$	$OM_{s,varmax}$	Mean RMSE	Mean peak err	Max DE00(31)
AMPAS190							
Sony CRT	4.35	42.56	4.3E–03	1.7E–01	0.58	2.62	17.97
NEC DLP	2.37	9.64	1.6E–03	6.2E–02	0.25	0.55	5.50
Panasonic LCD	2.20	8.48	2.3E–03	2.2E–01	0.27	0.70	4.18
Rec2020 Laser	5.38	12.63	1.9E–01	1.5E+00	1.77	7.79	0.00
8-laser	5.48	12.10	7.0E+00	1.1E+02	1.58	7.70	0.00
RIT MPD	0.46	1.81	4.8E–07	4.1E–05	0.21	0.49	0.00
MacBeth24							
Sony CRT	4.67	25.96	3.1E–03	2.1E–02	0.62	2.81	11.20
NEC DLP	2.28	7.12	8.5E–04	9.5E–03	0.26	0.54	0.00
Panasonic LCD	2.07	6.44	4.8E–04	6.2E–03	0.27	0.67	0.00
Rec2020 Laser	5.45	9.62	1.9E–01	1.5E+00	1.85	7.88	0.00
8-laser	6.01	12.12	8.9E+00	5.3E+01	1.68	8.01	0.00
RIT MPD	0.45	1.88	4.4E–06	1.0E–04	0.22	0.51	0.00
MacBeth DC							
Sony CRT	3.70	41.37	1.1E–02	1.2E+00	0.62	2.83	17.46
NEC DLP	2.48	12.28	5.4E–03	4.8E–01	0.28	0.58	4.84
Panasonic LCD	2.14	11.40	2.2E–03	1.2E–01	0.29	0.70	6.32
Rec2020 Laser	5.48	10.98	3.6E–01	2.2E+00	1.93	8.12	0.00
8-laser	5.89	11.40	5.6E+00	6.7E+01	1.70	7.78	0.00
RIT MPD	0.37	2.77	8.7E–07	1.6E–04	0.21	0.49	2.08
Big Metamers							
Sony CRT	7.89	44.91	3.2E–02	7.3E–01	0.49	2.11	17.97
NEC DLP	3.87	17.98	1.7E–02	4.2E–01	0.22	0.54	6.27
Panasonic LCD	3.57	15.49	1.3E–02	3.9E–01	0.25	0.73	6.08
Rec2020 Laser	4.87	12.63	9.6E–02	1.5E+00	1.39	6.57	1.30
8-laser	4.16	12.10	2.9E+00	6.0E+01	1.33	6.79	0.00
RIT MPD	0.69	9.03	1.0E–03	6.4E–02	0.17	0.45	0.00
Munsell							
Sony CRT	3.00	28.53	1.8E–03	1.0E–01	0.64	2.92	11.91
NEC DLP	2.25	8.28	5.7E–04	3.3E–02	0.29	0.59	0.65
Panasonic LCD	1.93	8.61	2.8E–04	2.0E–02	0.29	0.70	0.00
Rec2020 Laser	5.44	9.92	2.6E–01	1.8E+00	1.99	8.32	0.00
8-laser	–	–	–	–	–	–	–
RIT MPD	–	–	–	–	–	–	–

The most compelling conclusion from the RMSE and maximum spectral error ratios generated here is that none of these systems do a particularly good job at matching reference stimuli spectrally. The strongest average patch match from the best display still yields an RMSE of 25% across all visible wavelengths. The laser displays, not surprisingly, are significantly worse, as would be expected from attempted matches of continuous spectra with discrete monochromatic primaries. Still, the strong metamerism results achieved for some of these displays suggest that absolute spectral match might be an unnecessary objective for observer consistency in abridged multispectral system optimization.

Finally, maximum ΔE_{00} color differences for the 1931 2° observer show where not all of these displays are capable of rendering colorimetric matches for all of the patches in the stimuli set. The smaller gamut displays, CRT, DLP and LCD in particular, are consistently unable to produce exact matches according to traditional digital color management strategies.

Table V is an extension of Table I for D65-illuminated MacBeth patches and summarizes observer metamerism indices for the CIE 2006 and Heckaman CMF models. In general, the displays all perform in rank and relative magnitude similarly to the Sarkar/Fedutina results, although absolute numerical performance is worse for the CIE

Table III. Sarkar/Fedutina observer metamerism indices for various displays relative to test patchsets illuminated by HMI motion picture studio light (1931 2° colorimetry match).

HMI	OM_S	$OM_{S,max}$	$OM_{S,var}$	$OM_{S,varmax}$	Mean RMSE	Mean peak err	Max DE00(31)
AMPAS190							
Sony CRT	3.10	14.81	5.8E–3	2.2E–01	0.42	1.79	6.76
NEC DLP	2.83	9.53	7.8E–03	1.1E–01	0.25	0.54	2.75
Panasonic LCD	3.70	9.01	3.7E–02	6.1E–01	0.25	0.67	2.06
Rec2020 Laser	6.46	11.93	1.6E+00	1.5E+01	1.77	7.92	0.00
8-laser	11.04	26.01	2.9E+02	1.8E+03	1.59	8.34	0.00
RIT MPD	0.33	2.21	1.5E–07	4.4E–06	0.18	0.45	0.00
MacBeth24							
Sony CRT	2.82	7.25	4.7E–03	3.5E–02	0.42	1.82	0.33
NEC DLP	2.68	5.87	4.6E–03	2.1E–02	0.26	0.55	0.00
Panasonic LCD	3.74	7.33	3.1E–02	1.9E–01	0.25	0.67	0.00
Rec2020 Laser	6.62	10.65	1.5E+00	1.0E+01	1.86	8.13	0.00
8-laser	11.92	25.91	3.4E+02	1.9E+03	1.69	8.99	0.00
RIT MPD	0.32	1.19	4.6E–08	6.9E–07	0.17	0.44	0.00
MacBeth DC							
Sony CRT	3.35	24.19	8.4E–03	5.4E–01	0.45	1.92	11.25
NEC DLP	2.92	18.53	7.2E–03	7.3E–02	0.28	0.59	8.02
Panasonic LCD	3.72	18.09	2.2E–02	1.5E–01	0.27	0.71	8.07
Rec2020 Laser	6.58	12.96	2.1E+00	1.1E+01	1.92	7.90	0.00
8-laser	11.78	25.25	3.3E+02	1.6E+03	1.75	9.02	0.00
RIT MPD	0.36	5.43	9.5E–05	2.2E–02	0.18	0.46	4.03
Big Metamers							
Sony CRT	4.59	19.08	2.1E–02	3.4E–01	0.39	1.56	8.33
NEC DLP	4.01	16.29	1.4E–02	2.9E–01	0.22	0.53	5.47
Panasonic LCD	3.94	12.79	5.0E–02	1.8E+00	0.23	0.65	4.15
Rec2020 Laser	5.76	13.96	8.3E–01	1.2E+01	1.45	6.84	1.03
8-laser	8.47	26.10	1.1E+02	1.8E+03	1.32	6.95	0.00
RIT MPD	0.41	1.86	3.2E–06	1.2E–04	0.17	0.48	2.51
Munsell							
Sony CRT	2.94	9.13	3.7E–03	1.7E–01	0.46	2.00	1.86
NEC DLP	2.68	7.76	6.3E–03	6.5E–02	0.30	0.61	0.00
Panasonic LCD	3.71	7.71	2.2E–02	2.3E–01	0.28	0.73	0.00
Rec2020 Laser	6.71	10.72	1.6E+00	9.9E+00	2.01	8.27	0.00
8-laser	–	–	–	–	–	–	–
RIT MPD	–	–	–	–	–	–	–

2006 observers and then worse again for Heckaman's observers. As each represents an intentionally extreme array of potential observer response functions versus the Sarkar/Fedutina statistical CMF categories, these results are not surprising. Turning to observer variability ellipsoids, CIE 2006 actually predicts less disparity than Sarkar/Fedutina, although Heckaman again represents exaggerated differences considering his full observer set. What remains is to scale each model against real metamerism experiments to validate which correlates best with the degree of observer variability noted across an actual population of observers.

Data trends from models of CIE Illuminant A, HMI and fluorescent F2 sources reveal only a few notable differences from the D65 data. First, illuminant A offers significant gamut challenge to the Rec. 709 CRT and it thus performs quite poorly under this source across all patchsets and all indices. Also, under illuminant A, the eight-laser system fares a bit better than under D65, generating observer metamerism and observer variability levels more similar to the Rec. 2020 laser, but still worst among the candidate technologies. The RIT display improves its performance in tungsten light versus the D65 models by factors of near 2-to-1 and under HMI and fluorescent illumination by nearly 3-to-1 and 4-to-1 respectively. This advantages it consistently over the other investigated technologies. For the remaining displays,

Table IV. Sarkar/Fedutina observer metamerism indices for various displays relative to test patchsets illuminated by CIE F2 fluorescent (1931 2° colorimetry match).

F2	OM_s	$OM_{s,max}$	$OM_{s,var}$	$OM_{s,varmax}$	Mean RMSE	Mean peak err	Max DE00(31)
AMPAS190							
Sony CRT	5.61	26.90	1.0E–01	2.1E+00	0.49	2.15	11.47
NEC DLP	4.63	10.67	7.1E–02	4.9E–01	0.34	0.65	2.25
Panasonic LCD	5.13	9.83	1.7E–01	1.1E+00	0.29	0.71	1.66
Rec2020 Laser	7.63	13.04	4.5E+00	3.0E+01	1.76	7.71	0.00
8-laser	10.69	21.48	3.0E+02	2.3E+03	1.47	7.51	0.00
RIT MPD	0.21	0.93	9.9E–08	6.1E–06	0.13	0.36	0.00
MacBeth24							
Sony CRT	5.66	15.19	1.1E–01	8.0E–01	0.51	2.24	5.50
NEC DLP	4.90	8.79	6.6E–02	4.2E–01	0.35	0.68	0.00
Panasonic LCD	5.53	8.58	1.7E–01	1.1E+00	0.30	0.74	0.00
Rec2020 Laser	7.97	12.15	4.7E+00	3.6E+01	1.86	8.01	0.00
8-laser	11.47	21.05	3.2E+02	1.8E+03	1.57	7.95	0.00
RIT MPD	0.21	0.67	2.0E–07	4.6E–06	0.14	0.38	0.00
MacBeth DC							
Sony CRT	5.92	30.59	1.0E–01	2.3E+00	0.50	2.16	10.98
NEC DLP	4.91	11.05	9.5E–02	5.3E–01	0.36	0.69	2.90
Panasonic LCD	5.35	10.33	2.1E–01	1.1E+00	0.31	0.74	3.00
Rec2020 Laser	7.98	13.57	7.1E+00	3.9E+01	1.78	7.46	0.00
8-laser	11.11	20.26	4.0E+02	1.9E+03	1.46	7.00	0.00
RIT MPD	0.23	3.11	3.3E–06	7.7E–04	0.14	0.39	0.00
Big Metamers							
Sony CRT	5.12	31.75	1.7E–01	3.2E+00	0.45	1.93	12.61
NEC DLP	3.99	11.46	5.7E–02	7.5E–01	0.28	0.59	3.76
Panasonic LCD	3.95	10.70	1.1E–01	2.0E+00	0.25	0.67	3.35
Rec2020 Laser	6.02	14.12	1.4E+00	1.3E+01	1.60	7.53	0.00
8-laser	8.58	21.63	6.4E+01	6.9E+02	1.39	7.28	0.00
RIT MPD	0.31	4.96	5.6E–06	3.6E–04	0.14	0.37	0.00
Munsell							
Sony CRT	5.82	17.02	7.6E–02	1.5E+00	0.52	2.23	6.99
NEC DLP	4.96	10.36	7.5E–02	4.5E–01	0.37	0.71	0.00
Panasonic LCD	5.49	9.73	1.6E–01	9.7E–01	0.32	0.76	0.00
Rec2020 Laser	8.09	12.92	5.1E+00	3.1E+01	1.83	7.52	0.00
8-laser	–	–	–	–	–	–	–
RIT MPD	–	–	–	–	–	–	–

Table V. CIE 2006 and Heckaman et al. observer metamerism indices for various displays relative to MacBeth 24 test patches illuminated by CIE D65 (1931 2° colorimetry match).

CIE D65	OM_c	$OM_{c,max}$	$OM_{c,var}$	OM_h	$OM_{h,max}$	$OM_{h,var}$	Max DE00(31)
MacBeth24							
Sony CRT	2.81	9.47	4.5E–04	11.31	41.61	2.9E–02	0.44
NEC DLP	2.95	8.46	9.6E–05	11.00	41.93	2.4E–02	0.00
Panasonic LCD	3.31	5.74	4.0E–04	9.75	30.41	4.4E–03	0.00
Rec2020 Laser	12.84	20.61	1.9E–01	33.38	58.46	5.7E+00	0.00
8-laser	19.87	50.49	7.1E+00	43.29	75.42	2.3E+02	0.00
RIT MPD	2.67	6.35	3.1E–06	6.25	15.21	6.9E–04	0.00

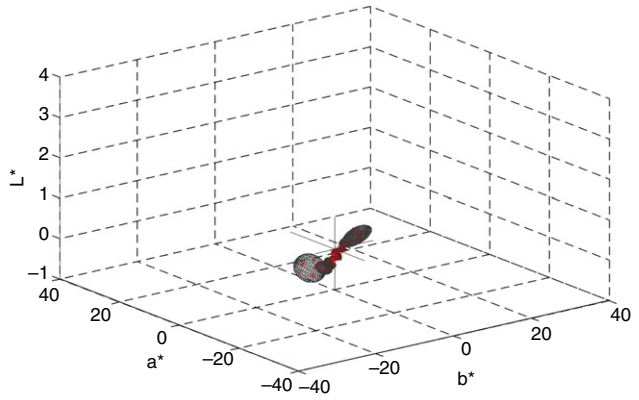

Figure 11a. Sony CRT observer variability ellipsoids based on reproduced 1931 2° colorimetry match to MacBeth 24 patches illuminated by CIE D65.

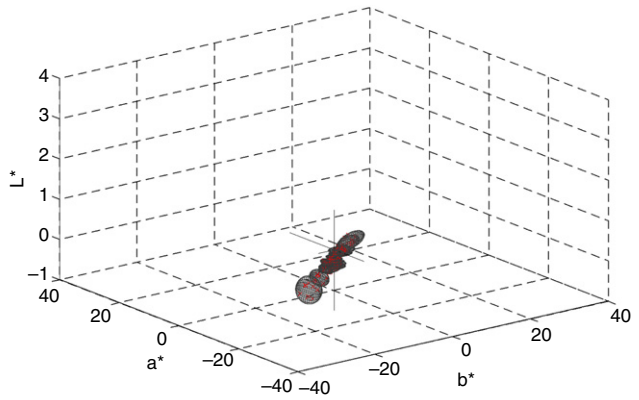

Figure 11b. NEC DLP observer variability ellipsoids based on reproduced 1931 2° colorimetry match to MacBeth 24 patches illuminated by CIE D65.

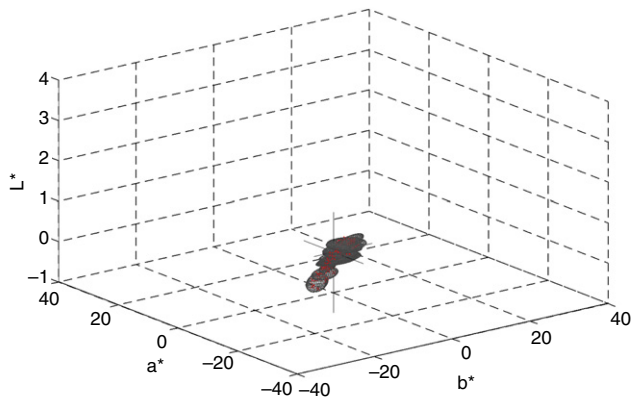

Figure 11c. Panasonic LCD observer variability ellipsoids based on reproduced 1931 2° colorimetry match to MacBeth 24 patches illuminated by CIE D65.

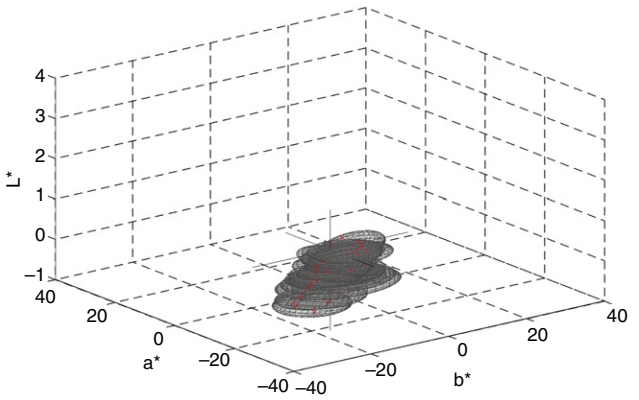

Figure 11d. Example ITU-R Rec. 2020 laser projector observer variability ellipsoids based on reproduced 1931 2° colorimetry match to MacBeth 24 patches illuminated by CIE D65.

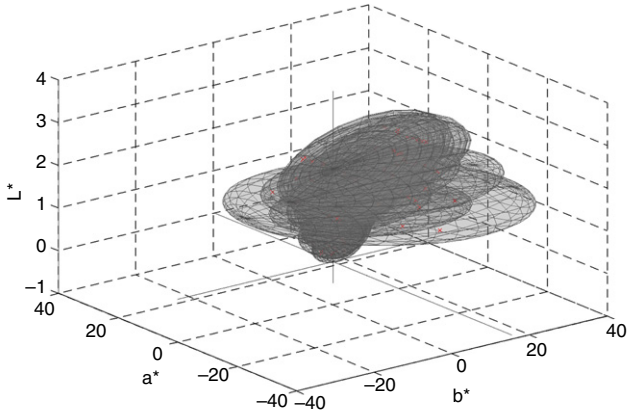

Figure 11e. Chromaticity-area-optimized eight-channel laser projector observer variability ellipsoids based on reproduced 1931 2° colorimetry match to MacBeth 24 patches illuminated by CIE D65.

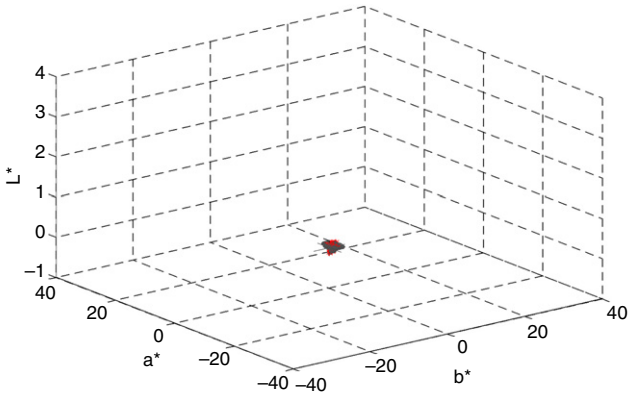

Figure 11f. RIT seven-channel projector observer variability ellipsoids based on reproduced 1931 2° colorimetry match to MacBeth 24 patches illuminated by CIE D65.

HMI and fluorescent lighting change their performance little versus under D65.

Particularly intriguing in these results overall is the disparity in observer metamerism and observer variability in the eight-laser system versus either a simpler Rec. 2020 three-channel laser display or the RIT optimized seven-channel display. Given its advantages of the greatest number of primary spectra, the greatest degrees-of-freedom for controlling metamerism (albeit with restriction to satisfy color matches for the 1931 observer), and the absolute largest overall chromaticity gamut area, this system

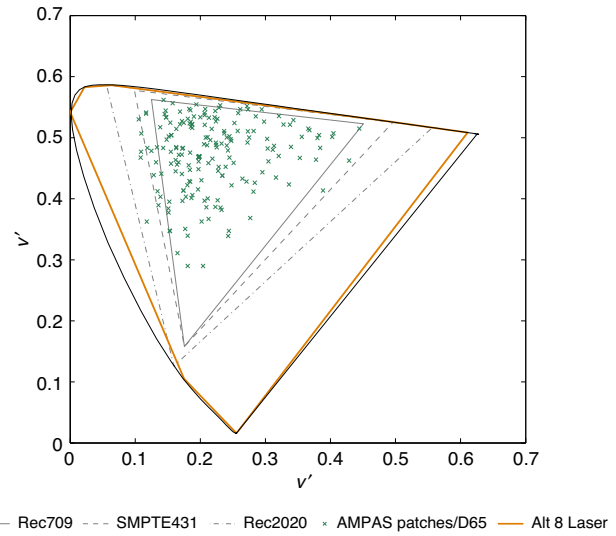

Figure 12. The alternate eight-primary laser projector chromaticity gamut; color points representing Kodak/AMPAS color patches illuminated by CIE D65 are also included (a) 1931 2° primary $u'v'$ chromaticity gamut; (b) 1931 2° primary chromaticity gamut.

underperformed considerably across the Sarkar/Fedutina observers. It is understandable that the RIT display had an advantage over this system since the primary spectra used to construct it were explicitly optimized to minimize observer metamerism against the eight Sarkar/Fedutina observers and specifically in consideration of the patchsets and illuminants represented in this test. However, the eight-laser system represents a common goal of multiple display manufacturers and technologists in the motion picture industry. It is capable of generating visible content across nearly the entire gamut of human color vision. The eight wavelengths were selected to produce the maximum geometric overlap with the 1931 chromaticity diagram yet yielded observer variability drastically higher than all of the smaller-gamut systems. The mathematical justification for this result probably stems from the alignment of the eight laser wavelengths with regimes of maximum CMF disparity amongst the eight observer categories.

To analyze this result further, an alternate eight-laser system was theorized and simulated. Given the benefit in observer metamerism for the three-laser Rec. 2020 system over the eight-laser display, three of the eight monochromatic primaries (485, 540 and 650 nm) were replaced by the Rec. 2020 wavelengths closest in chromaticity space, the idea being to take advantage of five additional degrees of freedom above the Rec. 2020 set. The resultant chromaticity gamut area was reduced only slightly from the ideal, but the metamerism results were significantly improved. Figure 12 shows the new $u'v'$ gamut. Table VI further summarizes the metamerism indices for the D65-illuminated MacBeth color checker. Although still not as good as the exemplary RIT MPD, the new eight-laser system yields much stronger metamerism and variability than either of the other laser systems and in fact exceeds the performance of the CRT, DLP and LCD displays. This

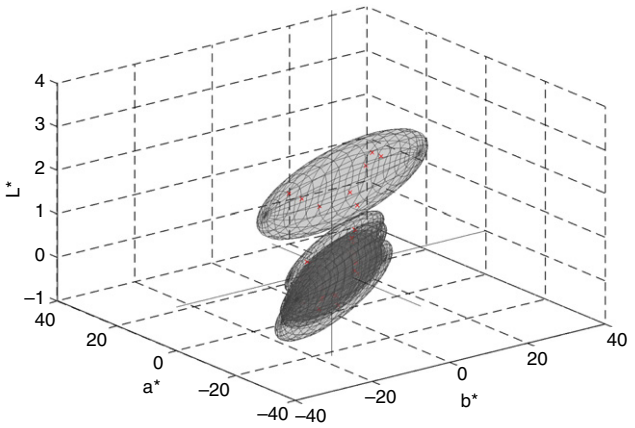

Figure 13a. Chromaticity-area-optimized eight-channel laser projector observer variability ellipsoids based on minimized observer metamerism for MacBeth 24 patches illuminated by CIE D65.

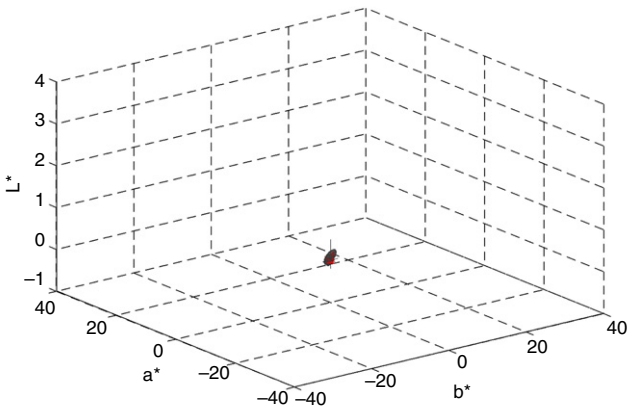

Figure 13b. RIT seven-channel projector observer variability ellipsoids based on minimized observer metamerism for MacBeth 24 patches illuminated by CIE D65.

solidifies the extreme sensitivity of observer metamerism and variability to tuned monochromatic primaries. Even small adjustments can generate large performance differences if the wavelengths chosen exacerbate physiological and psychophysical differences in response.

A final assessment was run to determine how the displays could perform if optimized for Sarkar/Fedutina observer metamerism magnitude, OM_s, rather than being forced to make metameric matches for the 1931 2° observer. Table VII summarizes results for the patches illuminated by D65. For all displays, the metamerism magnitude is notably improved, especially for the chromaticity-area-maximized eight-laser system, which proves to have been handicapped by its requirement to match the standard observer's response for each patch previously. In this new paradigm, it achieves results superior to the three-channel Rec. 2020 laser. Figures 13a and 13b further summarize variability ellipsoids for the eight-channel laser and seven-channel RIT display, validating that the RIT system still affords multiple orders of magnitude advantage. For both devices, the ellipsoid

Table VI. Sarkar/Fedutina observer metamerism indices for alternate laser displays relative to MacBeth 24 test patches illuminated by CIE D65 (1931 2° colorimetry match).

CIE D65	OM_s	$OM_{s,max}$	$OM_{s,var}$	$OM_{s,varmax}$	Mean RMSE	Mean peak err	Max DE00(31)
MacBeth24							
Rec2020 Laser	5.50	10.44	2.6E–01	1.3E+00	2.18	9.66	0.00
8-laser	11.61	27.31	3.1E+02	2.0E+03	2.08	11.01	0.00
8-laser + 2020	2.09	3.26	3.2E–03	2.8E–02	1.94	2.58	0.00
RIT MPD	0.78	2.43	6.2E–06	7.5E–05	0.31	0.66	0.00

Table VII. Sarkar/Fedutina observer metamerism indices for various displays relative to test patchsets illuminated by CIE D65 (optimized observer metamerism).

CIE D65	OM_s	$OM_{s,max}$	$OM_{s,var}$	$OM_{s,varmax}$	Mean RMSE	Mean peak err	Max DE00(31)
AMPAS190							
Sony CRT	1.74	16.20	4.8E–03	1.7E–01	0.45	1.97	8.53
NEC DLP	1.61	11.81	1.9E–03	1.2E–01	0.26	0.55	8.31
Panasonic LCD	2.01	7.96	2.6E–03	5.7E–02	0.28	0.78	6.04
Rec2020 Laser	4.84	7.52	3.8E–01	5.3E+00	2.10	9.49	3.70
8-laser	4.18	18.98	4.0E+01	1.7E+03	2.78	14.30	21.94
RIT MPD	0.15	0.95	2.3E–06	1.1E–04	0.25	0.60	3.27
MacBeth24							
Sony CRT	1.13	5.61	2.8E–03	4.9E–02	0.45	2.00	4.08
NEC DLP	1.37	2.44	3.0E–04	3.2E–03	0.26	0.52	3.93
Panasonic LCD	1.88	3.14	1.1E–03	6.1E–03	0.27	0.77	2.31
Rec2020 Laser	5.06	7.56	2.5E–01	1.2E+00	2.20	9.70	3.40
8-laser	4.69	18.50	8.0E+01	1.6E+03	3.03	15.96	19.95
RIT MPD	0.14	0.91	2.6E–06	4.9E–05	0.27	0.64	3.34
MacBeth DC							
Sony CRT	1.64	31.34	2.7E–02	2.9E+00	0.50	2.21	20.07
NEC DLP	1.65	23.85	8.7E–03	6.6E–01	0.30	0.60	15.24
Panasonic LCD	1.98	23.28	1.6E–03	8.3E–02	0.32	0.89	15.74
Rec2020 Laser	4.92	7.59	4.1E–01	3.0E+00	2.44	10.30	6.51
8-laser	3.99	17.81	8.1E+00	1.9E+02	3.18	16.78	24.47
RIT MPD	0.18	9.77	3.4E–04	8.1E–02	0.31	0.74	7.39
Big Metamers							
Sony CRT	4.51	23.12	4.7E–02	1.0E+00	0.42	1.70	11.79
NEC DLP	2.95	18.88	1.8E–02	2.0E–01	0.24	0.53	11.05
Panasonic LCD	2.87	14.33	9.2E–03	2.9E–01	0.26	0.76	9.28
Rec2020 Laser	4.26	8.24	3.2E–01	3.3E+00	1.61	7.60	8.01
8-laser	3.96	19.82	8.5E+01	1.6E+03	2.09	11.28	15.91
RIT MPD	0.27	2.75	7.8E–06	2.4E–04	0.18	0.47	2.91

errors are more symmetrically distributed about the CIELAB origin. The penalty for this strategy, though, lies with the standard observer color difference index, which is now appreciably higher for all of the displays. This result further confirms that the 2° CMFs are not statistically similar to any of the Sarkar/Fedutina observer categories in the context of this kind of analysis. Given that the Sarkar/Fedutina observers are derived from Stiles and Burch data focused on 10° field experiments, this may not be entirely surprising.

CONCLUSIONS

In designing color primaries for accurate color reproduction, spectral characteristics do carry significant importance. A move towards monochromatic color designs such as are found in laser displays adds significant chromaticity gamut area for users, but at the expense of observer metamerism and variability. Investigation of real displays designed around three current color standards for motion picture work reveals that the latest specification, ITU-R Rec. 2020, offers strong potential for viewer disparity when compared with

older broad-spectrum standards like ITU-R Rec. 709 and SMPTE 431. Expanding to more than three laser primaries can help, but only if those wavelengths are themselves optimized to the objective of improved observer consistency. Attempting, instead, to simply generate the largest color gamut possible from multiple laser wavelengths may actually exacerbate metamerism failure. Finally, it is possible to craft customized primary spectra with the intent of minimizing observer metamerism. A prototype seven-channel projection system at RIT has been constructed with modeled results significantly improved over any legacy three-color display.

REFERENCES

[1] E. Giorgianni and T. Madden, *Digital Color Management* (Addison-Wesley, 1998).

[2] R. S. Berns, *Principles of Color Technology*, 3rd ed.

[3] W. S. Stiles and J. M. Burch, "NPL color matching investigation: final report," Opt. Acta **6**, 1–26 (1959).

[4] N. I. Speranskaya, "Determination of spectrum colour coordinates for 27 normal observers," Opt. Spectrosc. **7**, 424–428 (1959).

[5] CIE, "Special metamerism index: change in observer," CIE Pub. No. 80, Vienna (1989).

[6] CIE, "Fundamental chromaticity diagram with physiological axes – Part 1," CIE Pub. 170-1:2006 (2006).

[7] A. Sarkar, L. Blonde, P. Le Callet, F. Autrusseau, P. Morvan, and J. Stauder, "Toward reducing observer metamerism in industrial applications: colorimetric observer categories and observer classification," *Proc. IS&T/SID Eighteenth Color and Imaging Conf.* (IS&T, Springfield, VA, 2010), pp. 307–313.

[8] A. Sarkar, F. Autrusseau, F. Vienot, P. LeCallet, and L. Blonde, "From CIE2006 model to improved age-dependent and average colorimetric observers," J. Opt. Soc. Am. **28**, no. 10, (2011).

[9] M. Fedutina, A. Sarkar, P. Urban, and P. Moran, "How do observer categories based on color matching functions affect the perception of small color differences?" *Proc. IS&T/SID Nineteenth Color and Imaging Conf.* (IS&T, Springfield, VA, 2011), pp. 2–7.

[10] P. Morvan, A. Sarkar, J. Stauder, L. Blondé, J. Kervec, and S. K. Hasan, "A handy calibrator for color vision of a human observer", *IEEE International Conference on Multimedia and Expo (ICME)* (Barcelona, 2011).

[11] R. L. Alfvin and M. D. Fairchild, "Observer variability in metameric color matches using color reproduction media," Col. Res. Appl. **22**, no. 3, 174–188.

[12] M. D. Fairchild and R. L. Heckaman, "Metameric observers: A Monte Carlo approach," *Proc. IS&T Twenty-first Color and Imaging Conf.* (IS&T, Springfield, VA, 2013), pp. 185–190.

[13] J. Pokorny, V. C. Smith, and M. Lutze, "Aging of the human lens," Appl. Opt. **26**, 1437–1440 (1987).

[14] J. Pokorny and V. C. Smith, "How much light reaches the retina?," *Colour Vision Deficiencies XIII*, edited by C. R. Cavonius, Documenta Opthalmologica Proceedings Series (1997), 59, pp. 491–511.

[15] J. Xu, J. Pokorny, and V. C. Smith, "Optical density of the human lens," J. Opt. Soc. Am. A **14**, 953–960 (1997).

[16] R. A. Bone, J. T. Landrum, and A. Cains, "Optical density spectra of the macular pigment in vivo and in vitro," Vis. Res. **32**, 105–110 (1992).

[17] T. T. Berendschot and D. van Norren, "Objective determination of the macular pigment optical density using fundus reflectance spectroscopy," Arch. Biochem. Biophys. **430**, 149–155 (2004).

[18] A. Stockman, L. T. Sharpe, and C. C. Fach, "The spectral sensitivity of the human short-wavelength cones," Vis. Res. **39**, 2901–2927 (2000).

[19] "Stockman and Sharpe Cone Fundamentals, 2-deg Fundamentals Based on Stiles and Burch 10-deg CMFs Adjusted to 2-deg," CVRL Functions, www.cvrl.org .

[20] M. Neitz and J. Neitz, "Molecular genetics of color vision and color vision defects," Arch. Ophthalmol. **118**, 691–700 (2000).

[21] K. Ohsawa, F. Koenig, M. Yamaguchi, and N. Ohyama, "Multiprimary display optimized for CIE1931 and CIE1964 color matching functions," *9th Congress Int'l Color Assn* (International Society for Optics and Photonics, 2002), pp. 939–942.

[22] M. Fairchild and D. Wyble, "Mean observer metamerism and the selection of display primaries," *Proc. IS&T/SID Fifteenth Color and Imaging Conf.* (IS&T, Springfield, VA, 2007), pp. 151–156.

[23] F. Koenig, K. Ohsawa, M. Yamaguchi, N. Ohyama, and B. Hill, "Multiprimary display: discounting observer metamerism," *9th Congress of the Int'l Color Assn* (International Society for Optics and Photonics, 2002), pp. 898–901.