The Effect of Display Gamut Volume on Image Preference

Rodney L. Heckaman[2], Masato Sakurai[1], Mark D. Fairchild[2], Takehiro Nakatsue[1], and Yoshihide Shimpuku[1]
1: Sony Corporation, Tokyo, Japan
2: Munsell Color Science Laboratory, Rochester Institute of Technology, Rochester, New York, USA

Abstract

Similar psychophysical experiments were performed at the Munsell Color Science Laboratory (MCSL), Rochester Institute of Technology, and at the Sony Corporation, Tokyo, in collaboration to determine observer preference as a function of color gamut volume as simulated using a Sony prototype, extended gamut, LED backlit, LCD video display. In both experiments, the overall results indicate preference was virtually indistinguishable above a color gamut volume ratio 0.8 times that of the full, extended gamut of the display. However, when the coarser MCSL data was subjected to a cluster analysis, sizable scene dependencies were revealed, and while the highly colorful scenes followed the overall result, preference for those scenes containing flesh tone clearly benefited from higher color gamut volume as did certain outdoor scenes. Because of these dependencies, finding an optimal color gamut that fits all may not be advisable. Instead, an extended color gamut might be better utilized by rendering object color to less than a full gamut leaving room for producing those portions of the scene that extend beyond object color offering a richer video experience.

Introduction

Recently, large format, wide color gamut displays have been introduced into the market by several manufactures so that consumers can enjoy rich color in their home - particularly in the home theater, media experience as the gamut of cinema film is wider than that of conventional video signal (BT.709[1]) used in television today. [2] Further to this end, multi-primary color displays with four to six primaries have been developed that also extend color gamut[3],[4]. Hence, the trend for producing wider color gamut displays will continue. Yet, while these displays express vivid, clear colors colorimetrically and photometrically and exclusive of a paper presented in paper dated 1997[5] Fedorovskaya, *et al*, there is no current research on how humans perceive these extended gamut displays when varying the extent of their gamut. Clearly, such research is important to the design of such displays.

In their paper, Fedorovskaya, *et al*, investigated the effect of variations in chroma on the perceived quality of natural scenes. While their work did not cover the ull region of extended gamut displays found today, they did find similar results to those reported here and a prior experiment - that " … colorfulness is the main perceptual attribute underlying image quality when chroma varies". Furthermore, they report that "the perceptual quality of images … [is] closely related to the naturalness of the images". In making these latter assertions, the authors made two assumptions – that only global changes need be considered and that the "optimum image equals the original (real life) scene". Hence, their assumptions almost beg the question as to whether naturalness is the sole factor of the quality of an image reproduction – particularly in digital cinema and video media [and certainly film based media] where it is the creative intent of the cinematographer and the director that is of paramount importance and where it is then the job of the media itself to both carry through on this intent and, over the long term, continue to expand the available palette to them.

Purpose and Methodology

A psychophysical experiment was performed to determine observer preference as a function of color gamut volume for each of ten representative scenes as displayed on a Sony prototype, 40 inch, LED backlit, LCD display with an expanded color gamut in its three primaries.

Observers

At the Munsell Color Science Laboratory (MCSL), twenty observers were chosen from a range of demographics that included close to an even split of males and females and a fairly wide distribution of age from young adults to the elderly – both expert and non-expert – and ethnic background.

Scenes

A set of ten scenes (see Appendix A) including a color chart was chosen as representing degrees of both lightness contrast and colorfulness. The Flowers, Lake, Color Chart, Grand Tetons, and Barn scenes were chosen for a high degree of colorfulness over a full range of hues, the Pastel and Fog scenes for reduced colorfulness or saturation and contrast, the Musician for Caucasian, Black, and Asian flesh tones, the Lady for reduced contrast flesh tones, and the Coast image for its high contrast.

Viewing Conditions

In MCSL, observers were sat approximately two meters from the display with their field of view perpendicular to the center of the display screen. A uniform gray wall within a meter and a half of the back of the display in a darkened room and within the field of view of the observer was illuminated uniformly by two Buhlite 150 watt, diffuse studio lamps. The lamps were placed behind the display to eliminate or at least minimize viewing flare. The illumination off the wall was measured to be 94 $\mathrm{cd/m^2}$ at a color temperature of $3150\,^{o}K$ by a Spectrascan PR650 spectrophotometer.

Display Characterization

The display was characterized to within an average DE94 of 1.0 and a standard deviation of 0.67 to the CIE color matching functions for the 1931 observer and illuminant D65 using a LMT C 1210 Colorimeter. A series of 10 step, RGB ramps were measured for lookup table (LUT) generation that convert RGB linear scalars to RGB counts assuming the three display primaries are linearly independent [i.e. strictly additive]. Figure 1 illustrates the degree in which this assumption holds true. A random sampling of 100 RGB values was measured for obtaining a least squared solution to the CIEXYZ values of the display's three primaries given in Table 1.

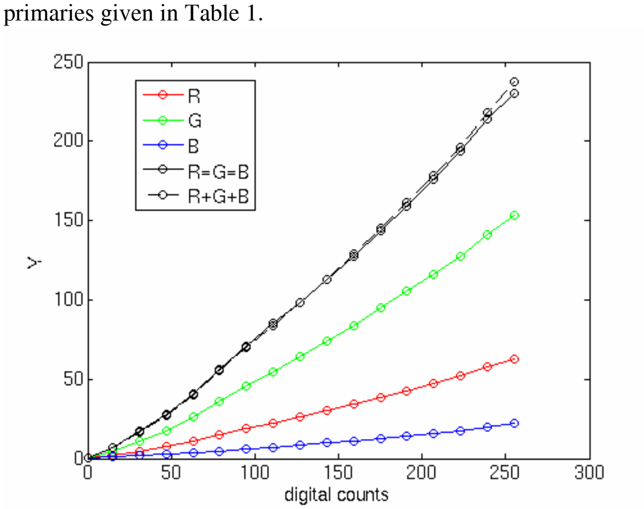


Table 1: CIE chromaticities and luminance of the measured display primaries.

	x	y	Y(cd/m^2)
R	0.7038	0.2953	123
G	0.2505	0.7107	295
B	0.1351	0.0786	33
White Point	0.3521	0.3137	450
Black Point	0.3305	0.3136	0.65

Figure 1: CIEXYZ Y versus their respective digital counts for each to the display's primaries and their additivity (R+G+B) versus measured grayscale (R=G=B)

Stimulus Preparation

According to the methodology given in Appendix B, successive versions of images from each scene were obtained with successive color gamut volumes of 1.00, 0.80, 0.60, and 0.40 times the display's actual color gamut volume (k-factor of 1.00, 0.89, 0.77, and 0.63 per Appendix C) in CIELAB. By this methodology, the perceptually uniform reduction in color gamut was achieved by a simulated set of display primaries (Figure 2) derived from the display's actual primaries and constrained to maintain both hue and the white point of the display. Hence, the luminance or lightness contrast of each image version was the same as was hue to within the ability of CIELAB to maintain perceptual hue.

The images sourced in RGB digital counts were then converted to CIEXYZ using these simulated primaries, not sRGB primaries, to insure a full range of colors within the simulated gamut (see Figure 3 as an example). Hence, they are scaled, not clipped, to each of the reduced gamut of the display. Finally, these CIEXYZ values were converted to RGB digital counts for display using the display's actual primaries.

Psychophysical Testing and Analysis

Preference was determined using the method of paired comparison[6]. All six possible pairs of the versions of each of the images were displayed. The observers were then asked to pick his or her preference via the following instructions.

"You will be shown pairs of different versions of a number of scenes. For each pair, I am asking you to simply select which image you prefer."

Thurstone's Law of Comparative Judgments, Case V, was assumed to hold *a priori*. That is, each stimulus gives rise to a discrimination process whose result is a value on a continuum of values and whose statistics are described by a normal distribution.

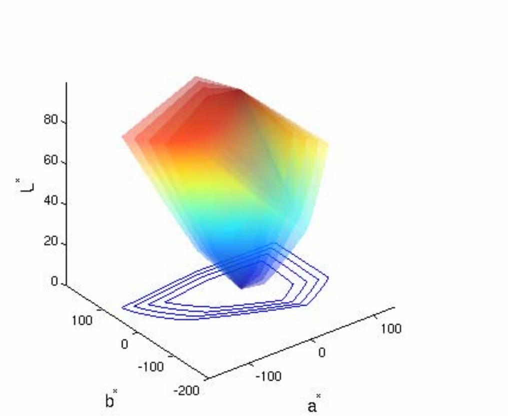


Figure 2: Actual perceptual gamut of the display in CIELAB with successive color gamut volume reductions of 1.0, 0.8, 0.6, and 0.4 times the full, extended gamut of the display for the MCSL experiment

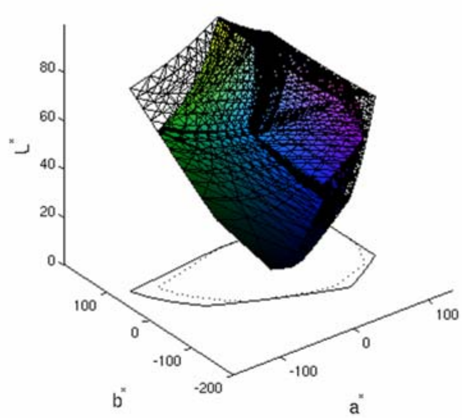


Figure 3: Perceptual gamut in CIELAB of the Flowers image (colored solid) compared to the full display gamut (wire frame)

Further by the Case V assumption, dispersions in discrimination across all observations are assumed equal. Under this assumption, the paired comparison data was analyzed by observer and by scene using a Z-Scores or interval scale methodology that first computes the average proportion each image version was preferred over all comparisons, then their respective standard normal deviates z or Z-Score from the tables for a normal distribution.

Finally, a cluster analysis was performed to determine scene or observer dependencies on the results. To this end, the set of interval scale values (Z Scores) for each version of an image was represented as a linear combination of orthogonal vectors (PCA analysis) which are, in turn, assumed to be normally distributed across observers and scenes. For four versions of each scene, there are then a 4-plex of vectors with their respective coefficients each accounting for a certain proportion of the variance in the interval scale. The average of those coefficients whose combination with the 4-plex of vectors account for the bulk of variance are then subjected to a nearest neighbor cluster analysis to find groupings of like results across scenes and across observers.

Test Results and Discussion

Figure 4 plots the average Z-score or interval scale of preference versus the simulated fractional volume of the full, extended gamut of the display in CIELAB over all MCSL scenes and observers. Additionally, the corresponding 95% confidence intervals are shown as computed according to the method prescribed by Montag[7]. As shown, observers exhibited a significant overall preference for gamut volumes of beyond 0.8 times the display's full gamut with some, but not significant so, maximum preference at 0.8 times the volume.

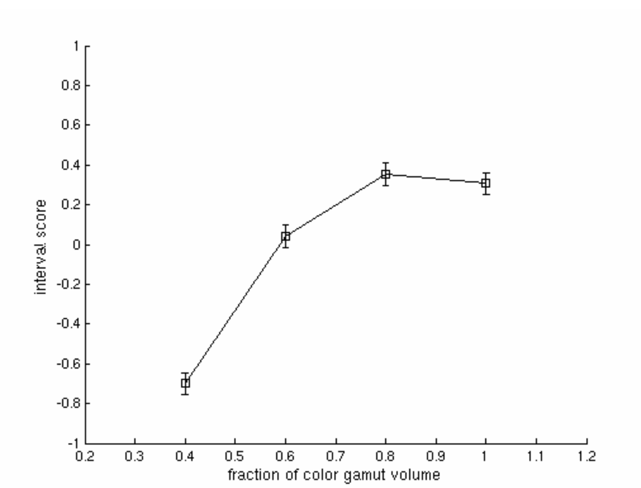


Figure 4: Overall MCSL results for preference in terms of interval score averaged across all 20 observers and 10 scenes with 95% confidence interval shown

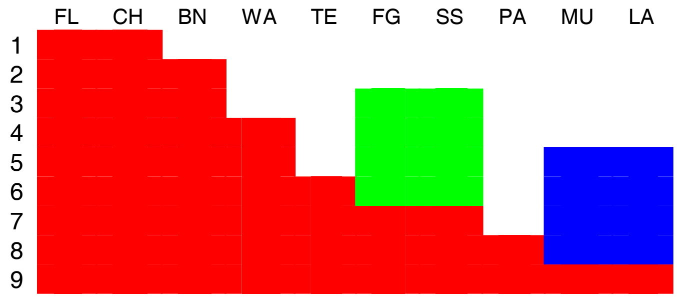
Observer Dependencies

A nearest neighbor, cluster analysis performed on the observer-by-observer results across all images revealed that all the observers' judgments essentially were in concert with each other. I.e., no consistent clusters developed out of the observer group.

Image Dependencies

A cluster analysis performed on the image-by-image,results averaged over all observers provides interesting insight. Table 2 presents the results. Across the top are each of the ten images and on the side, level number in the hierarchy of the clustering. At each succeeding level, either one image is added to an existing group or another group formed.

Table 2: Image-by-image preference cluster hierarchy



The first group (Group I in red) at level 1 consists of the Flowers and the Color Chart images. At level 2, the Barn image was added to the Group. At level 3, the Fog and Sunset image were joined in a second group (Group 2 in green) and, at level 5, a third group (Group 3 in blue), the Musician and Lady images, were formed. Of course, ultimately at level 9, all the images are formed into one group. For this analysis, the cluster results were taken at level 5 where three distinct groups are formed. At this level, the Grand Teton and the Pastel images are not members of any group.

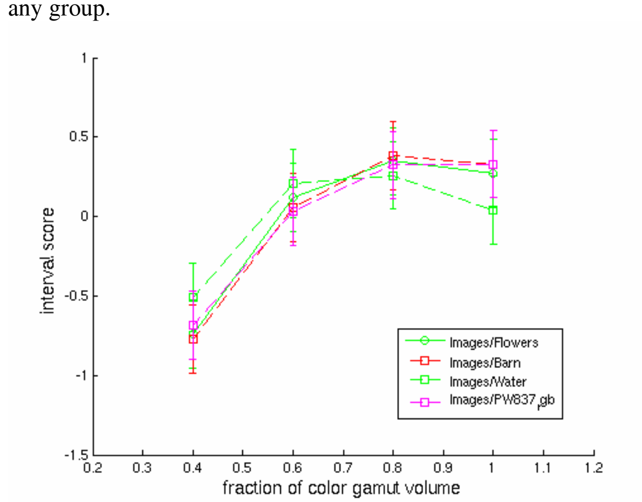


Figure 5: Preference results for Group 1 – highly colorful images

The interval scale or Z-Score results for the Group I images are shown in Figure 5. It is noted the Group I images are distinguished by their high degree of colorfulness, yet their ratings at a color gamut volume fractions of 0.6, 0.8, and 1.0 are virtually indistinguishable with overlapping confidence intervals. However, there is the notion that the rendering of these images would better serve preference at less than full gamut – particularly the Flower image which could be said to appears unnatural at full gamut.

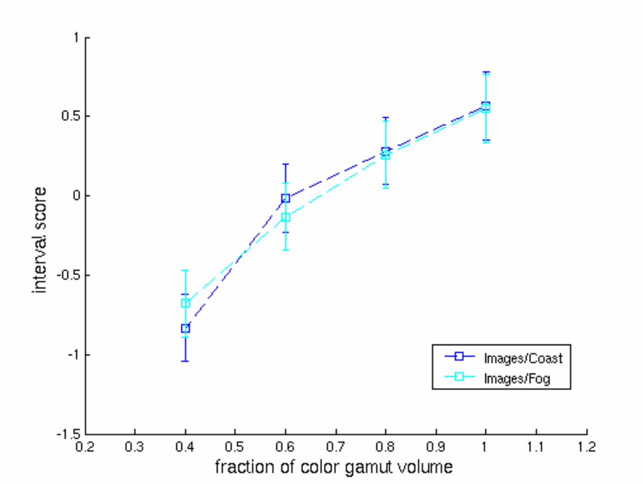


Figure 6: Preference results for Group II – Scenic images

The Group II images consist of two outdoor images, one of a sunset over water and the other, a foggy, lakeside scene in pastels. Their interval scale results are shown in Figure 6, and unlike Group I, their preference increases monotonically with increasing gamut volume. Hence, unlike the Group I scenes, their preference benefits from ever increasing color gamut. The sunset image at least seems intuitive as the experience of an actual sunset is extreme in colorfulness – certainly beyond object color perception. Yet, why the foggy image is rated similarly is not so intuitive.

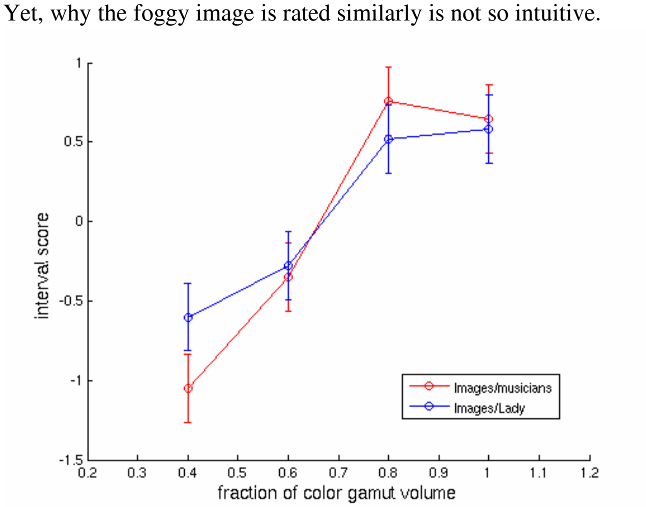


Figure 7: Preference results for Group III – flesh tones

Group III, consisting of the Musicians and Lady images, are clearly distinguished as representing flesh tones, and the preference scale results shown in Figure 7 indicate a statistically significant preference for larger gamut volumes. And when viewing these images as rendered in the smaller of gamut volumes, the perception of grayness becomes apparent in the flesh tones – obviously not considered a desirable trait. Hence, like Group II, these images benefit from ever increasing gamut at least within the scope of this experiment.

Of the remaining images, Grand Tetons and Pastel, preference for the Pastel image is indistinguishable across the range of volumes tested – an expected result as the de-saturated colors are proportionately less affected by a shrinking gamut. Unexpectedly, as noted in the above, the Fog image does not exhibit the same result.

The results for the Grand Teton image are also unexpected. Such an image naturally occurs on dark cloudy days when rays of sunlight illuminate only a stand of colorful, fall foliaged trees. The resulting intense color sense they provoke is quite compelling when experienced. Yet, the observers' preferences are in conflict with this experience as they tended to prefer a much less intensely colored rendition. One possible explanation is that this image was segmented and composited to create this effect which may have lent an artificial look to it most apparent at higher gamut volumes.

Sony Results

The Sony results were taken over a finer grained set of color gamut volume factors between 0.9 and 1.0 and are based on eight expert and non-expert judges and a similar set of images and are shown in Figure 8. Similar to the coarser overall Munsell results, these finer grained results show that preference is virtually indistinguishable between a gamut volume factor of approximately 0.95 and 1.0.

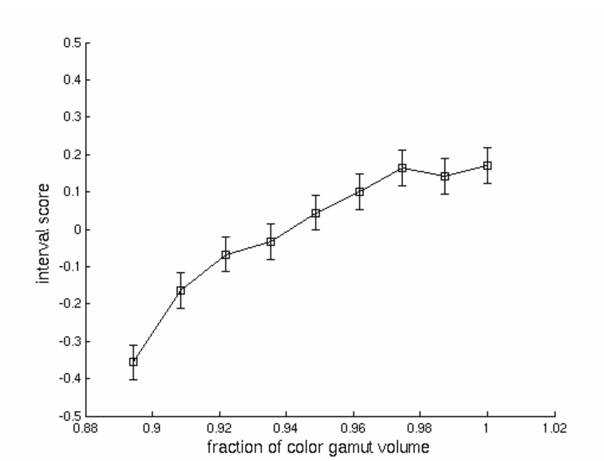


Figure 16: Overall Sony results for preference in terms of interval score averaged across all eight observers and 10 scenes with 95% confidence interval shown

Because of the finer grained volume differences, cluster analysis similar to that performed on the RIT results made no distinction between groups of images. As in the Munsell data, no distinction was found between groups of observers.

Conclusions

These results show that image preference as a function of color gamut volume is scene dependent. At reasonable color gamut volumes of 0.8 to 1.0 times that of the Sony extended gamut display, the perception of highly colorful scenes are less sensitive to reductions in gamut than certain outdoor scenes (e.g. sunsets) and scenes with a sizable portion of flesh tone. Hence, while the overall preference results would indicate an optimal color gamut volume, such a conclusion would only produce an average result. Scenes that are already quite colorful would be unaffected, yet the opportunity for rendering really compelling outdoor scenes and flesh tones squandered away.

The result is, in many ways, analogous to similar results obtained from high dynamic range (HDR) display media[8,9]. In typical display media – particularly digital video, even those with fairly high dynamic range, the tendency is to set the white point at the maximum output of the display producing a brighter display for point of sale considerations. In this scenario, object color is rendered fully in the gamut, but anything beyond object color – like a sunset – is clipped.

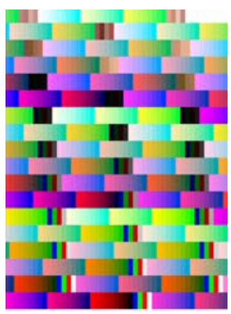
A more satisfactory alternative may be to render object color at less than the full gamut of the display thus leaving room for those colors beyond that can make for a truly compelling image. However, such an alternative cannot just be relegated to the display media alone but requires the cooperation of the media system as a whole – from the production of its content through to its rendering and actual display.

Appendix A: MCSL Images

Barn

Sunset

Fog



Color Chart

Musicians

Lady

Flowers

Pastel

Grand Tetons

Lake

Appendix B: Simulation of Primaries

The following methodology is given for simulating the display of an RGB image in a reduced color gamut while maintaining the white point of the display $\mathbf{XYZ}_{\max}$, keeping hue constant, and using the Sony LCD display primaries with characteristic matrix $\mathbf{M}_{Sony}$, lookup table LUT_{Sony} of scalar RGB values versus digital counts, and black point $\mathbf{XYZ}_{\min}$.

The CIEXYZ values for the display's primaries are:

$$\mathbf{XYZ}_{RGB} = \mathbf{M}_{Sony} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{1}$$

Transforming to CIELAB at the display's white point, a set of simulated primaries are obtained by reducing a^* and b^* each by a factor k to maintain constant hue and by optimizing their relative strengths $\mathbf{l} = [l_R \quad l_G \quad l_B]'$ in order to maintain L^* or the white point:

$$\mathbf{Lab}_{RGB,k} = [\mathbf{1} \quad \mathbf{k} \quad \mathbf{k}]\mathbf{Lab}_{RGB} \tag{2}$$

The characteristic matrix M_k of the simulated gamut is then:

$$\mathbf{M}_k = \begin{bmatrix} X_{R,\max}-X_{\min} & X_{G,\max}-X_{\min} & X_{B,\max}-X_{\min} & X_{\min} \\ Y_{R,\max}-Y_{\min} & Y_{G,\max}-Y_{\min} & Y_{B,\max}-Y_{\min} & Y_{\min} \\ Z_{R,\max}-Z_{\min} & Z_{G,\max}-Z_{\min} & Z_{B,\max}-Z_{\min} & Z_{\min} \end{bmatrix}_k \tag{3}$$

where $\mathbf{XYZ}_{RGB,\max}$ are the transformed CIEXYZ values of the simulated RGB primaries at the display's white point as before, and where $\mathbf{XYZ}_{\min}$ are the display's black point as before.

Now, for an RGB image of size N, the scalar RGB values are first obtained by linear interpolation of the inverse of the lookup table LUT_{Sony} from the image's digital counts. The CIEXYZ values for the image in the simulated gamut are then given by:

$$\mathbf{XYZ}_{image,k} = \mathbf{M}_k \begin{bmatrix} \mathbf{RGB}_{image} \\ ones(1,N) \end{bmatrix} \tag{4}$$

To keep L^* of the original image (k = 1), L^* of the image is converted tp CIELAB according to:

$$\mathbf{Lab}_{image,k}(1,:) = \mathbf{Lab}_{image,1}(1,:) \tag{6}$$

$$\mathbf{Lab}_{image,k} \Rightarrow \mathbf{XYZ'}_{image,k} \tag{7}$$

then converted to scalar RGB values with corresponding scalar RGB values for displaying the image in the reduced gamut on the Sony display:

$$\mathbf{RGB}_{image,k} = \mathbf{M}_{Sony}^{-1}\left[\mathbf{XYZ'}_{image,k} - repmat(\mathbf{XYZ}_{\min}, N)\right] \tag{5}$$

Finally, RGB digital counts for the image are obtained by linear interpolation of the lookup table LUT_{Sony} from the RGB scalars values.

Acknowledgements

The authors would like to thank the Sony Corporation for supporting this work though funding and those who participated as observers, both at Sony in Sony and the Munsell Color Science Laboratory in Rochester, New York.

References

[1] Recommendation ITU-R BT.709-5 Parameter values for the HDTV standards for production and international programme exchange (2002).

[2] T. Matsumoto et al., Color conversion from film to xvYCC video signal, SID 07 (in press).

[3] H. Sugiura et al., Improved six-primary-color 23-in. WXGA LCD using six-color LEDs, SID 06 Digest, 19.1 (2006).

[4] J. S. M. de Vaan, Competing display technologies for the best image performance, CIC 14, 274-279 (2006).

[5] Fedorovskaya, EA, de Ridder, H, and Blommaert, FJJ, Chroma variations and perceived quality of color images of natural scenes, Color Research and Application, John Wiley & Sons, Inc. [1997]

[6] Bartleson, CJ, Measuring differences, Chapter 8, Optical Radiation Measurements, Volume 5, Chapter 8, Academic Press, 1984

[7] Montag, ED, Empirical formula for creating error bars for the method of paired comparison, *Journal of Electronic Imaging*, **15**, 1, (2006)

[8] R.L. Heckaman and M.D. Fairchild, "Expanding display color gamut beyond the spectrum locus," *Color Research and Application, 31* 475-482 (2006)

[9] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs, High dynamic range display systems (ACM Transactions on Graphics, 23(3), 2004) pgs. 760-768.

Biography

Rodney L. Heckaman is a MacBeth/Engel Fellow and PhD candidate for a degree in Imaging Science at the Rochester Institute of Technology. Mark Fairchild is his advisor.

Masato Sakurai received his BS, MS, and PhD degrees in engineering from the Utsunomiya University, Japan in 1998, 2000, and 2003, respectively. After that, he worked as a post-doctoral fellow in McGill University, Canada for two years. Since 2006, he has been working at Sony Corporation, Japan. His work has focused on the development of color reproduction and color management system from the point of view in human color perception

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