Evaluation of the IES Method for Evaluating Light Source Color Rendition in terms of Metamer Mismatching

Brian Funt[1], Ben Hull[1] and Xiandou Zhang[2], [1]School of Computing Science, Simon Fraser University, Vancouver, B.C., V5A1S6 Canada; [2]School of Media & Design, Hangzhou Dianzi University, Hangzhou 310018, China

Abstract

The Illumination Engineering Society's Rf color rendering index [IES TM-30-15, 201] is compared to the MMCRI [Metamer Mismatching as a Measure of the Color Rendering of Lights, Mirzaei & Funt, Proc. AIC 2015]. IES Rf is based on color differences using a special set of 99 surface reflectances; while, in contrast, MMCRI is based on all theoretically possible reflectances. The two indices evaluate many lights similarly, but the MMCRI ranks some lights—especially those having strong peaks and wavelength regions of minimal power—lower than does R_f*. Is this difference in rating simply due to the fact that MMCRI uses all theoretically possible reflectances including step functions? A 'practical' version of MMCRI based on a set of 41 million real, measured spectral reflectances, rather than all theoretically possible reflectances, turns out to concur with the original MMCRI and shows that the disagreement between* R_f *and MMCRI is more fundamental. Overall, the present study suggests that* R_f *may overrate the color rendering properties of some lights; and, at the very least, indicate the type of lights upon which future psychophysical testing should concentrate.*

Introduction

The Illumination Engineering Society (IES) recently approved IES TM-30-15 "Method for Evaluating Light Source Color Rendition" May, 18, 2015 [4]. The method improves upon the previous CIE CRI [1] (color rendering index) in several ways, but is similar to it in that it is based on measuring the color differences arising for a change from a reference to a test illuminant, both at the same correlated color temperature, averaged over a specific set of reflectances. The most significant difference between the CIE CRI and the R_f index defined in IES TM-30-15 is the set of reflectances used for evaluation. The CIE index R_a is based on 8 Munsell samples, whereas the IES R_f is based on 99 reflectances derived from a much larger initial set of approximately 105,000 reflectance spectra. The 99 were specifically chosen [3] to be representative of the larger set, while at the same time covering the visible spectrum uniformly. Other differences between R_a and R_f are described below in the Background Section.

In a very different approach, Mirzaei et al. [7] propose evaluating the color rendering properties of lights based on the amount of metamer mismatching they induce. They report that their metamer mismatch CRI (MMCRI) corresponds well to the CIE R_a, but with some notable exceptions. The exceptions were lights such as CEI fluorescent illuminants F11 and F12, which because of their distinct spectral spikes, may well have poorer color rendering properties than R_a predicts. Since then the IES has published TM-30-15 with an accompanying software tool and data that are available from http://www.ies.org/redirect/tm-30/. In particular, the data includes the 99 reflectance functions and the spectral power distributions of 318 illuminants.

Using this data, we compare the MMCRI to the R_f for these 318 illuminants as shown in Figure 1. The plot appears to have two distinct branches. The illuminants falling on the lower branch have comparable Rf and MMCRI ratings. For the illuminants on the upper branch, the Rf rating is higher than the MMCRI rating. The natural question arises: Is the MMCRI underestimating the rendering properties of these lights or is Rf possibly overestimating them? Since there is limited psychophysical data concerning the rendering properties of these 318 lights a completely definitive answer is not possible at this time, but we investigate the differences further and argue that the MMCRI is the more accurate of the two measures. Our investigation is based both upon an examination of the particular illuminant spectra involved and a test of the metamer mismatching method based on a data set of 41 million reflectance spectra collected by Zhang et al. [10] as part of an investigation that compared theoretical metamer mismatch volumes to experimentally-determined metamer mismatch volumes.

Note that 'experimentally' in the term 'experimentally-determined' is simply intended to make the distinction between the metamer mismatch volumes calculated based on the 41 measured reflectance spectra and those calculated based on all theoretically possible spectra. We are not conducting new experiments when computing the experimentally-determined volumes.

Background

The MMCRI introduced by Mirzaei et al. [7] measures the degree of metamer mismatching using a variant of Logvinenko's [5] metamer mismatch volume index (MMVI). Metamer mismatching refers to the fact that two different reflectances that match in the sense that they yield identical color signals (i.e., XYZs) under one illuminant (i.e., are metamers) may no longer match under a second illuminant. For a given color signal obtained under a given first illuminant, the metamer mismatch volume is the set of all possible color signals that could theoretically arise under a given second illuminant. To provide a measure of the degree of metamer mismatching that is independent of any linear transformation of the color coordinate space, Logvinenko's MMVI is based on normalizing the metamer mismatch volume by the volume of the object color solid [9] (i.e., the set of color signals produced by all the reflecting objects under a given illuminant). Informally (see [5] Eq. 15 for a formal definition), the MMVI is defined as:

$$MMVI = \frac{volume\ of\ metamer\ mismatch\ volume\ for\ a\ given\ illuminant\ pair}{volume\ of\ the\ object\ color\ solid\ under\ the\ second\ illuminant}$$

One important feature of the MMVI is that it is independent of any linear transformation of the sensor space. In other words, it does not matter whether the metamer mismatch volumes are computed in XYZ or an LMS space derived as a linear transform of XYZ.

The MMCRI [7] is evaluated using only the MMVI for 'flat grey' (i.e., $reflectance(\lambda) = 0.5\ for\ 380nm \leq \lambda \leq 780nm$) and is defined as:

$$MMCRI = (1 - \sqrt[3]{MMVI}) \times 100$$

The argument for using only the volume of the metamer mismatch volume of flat grey is that the volumes for all other colors are roughly proportional and so provide no additional information. The primary advantage of the MMCRI as a measure of color rendering properties is that it is based on the set of all theoretically possible reflectances metameric to flat grey and hence not subject to the unavoidable bias any finite sample of reflectances such as those used for R_a and R_f necessarily entails.

Both R_a and R_f are based on measuring color differences resulting for a change from a reference to a test illuminant. The reference illuminant is chosen to have the same correlated color temperature as the test illuminant and is either a Planckian radiator, a CIE D-series daylight or, in some cases for R_f, a linear combination of the two. For comparison to R_f, the MMCRI used here is based on the R_f definition of reference illuminant.

Beyond the key difference of the use of 99 specially selected reflectances instead of only 8, and the use of slightly different reference illuminants, R_f differs from R_a in several other aspects. In particular, it is based on the 10-degree rather than the 2-degree standard observer; it uses a different chromatic adaptation transform; and it evaluates color differences in the CAM02-UCS [6] which is based on the CIECAM02 color appearance model [2] rather than CIE LUV.

Method

Figure 1 compares the R_f and MMCRI values. If the plot showed just a cloud of points and no correlation between the MMCRI and R_f values one might argue that the MMCRI was simply wrong; however, since there are two distinct linear clusters it seems possible that the MMCRI is spotting some key difference between the illuminants that the R_f is missing. MMCRI and R_f essentially agree on the lower branch, but disagree on the upper branch. Inspection of the cases on the upper branch reveals that it primarily represents illuminants with 'spiky' spectra such as the example shown in Figure 2. Such spiky spectra are more likely to be given a low MMCRI, potentially significantly lower than the R_f. Spectral spikes often lead to significant metamer mismatching. Assuming, for the moment, that spiky spectra are the key difference, is it the case that R_f is systematically overrating the rendering properties of spiky illuminants or rather that MMCRI is underrating them?

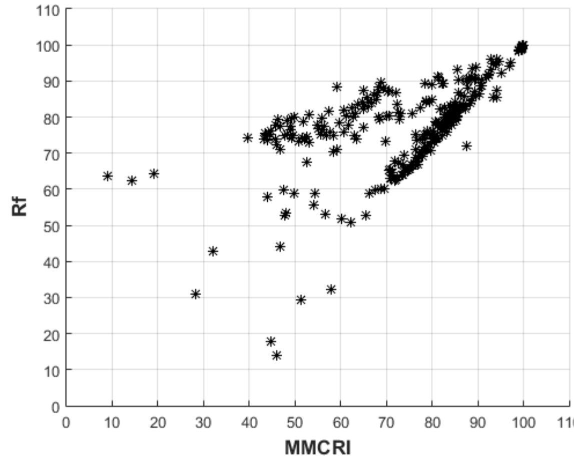

Figure 1. Comparison of the IES Rf color rendering index to the metamer mismatch rendering index MMCRI.

The computations for object color solids and metamer mismatch volumes are both based on 'optimal reflectances' [9]. Optimal reflectances are rectangular looking functions involving two step-function transitions with the rectangle appearing either as a pulse or a valley. Such idealized reflectance functions certainly will not arise in practice. Do these optimal reflectances create a problem? As Lorne Whitehead [8] points out this is a possibility: "…MMCRI effectively assesses the maximum possible color difference, under illumination from a test source, for spectral reflectance functions that are metameric under illumination from a reference source. Presumably that includes all theoretically possible spectral reflectance functions. In contrast, the R_f calculation uses a carefully selected set of ***real*** spectral reflectance functions. These do not include reflectance functions that contain step function changes. Such sharp changes will be particularly sensitive to the location of narrow spectral features. So I would expect the metric to be much more sensitive to sharp spectral features than is R_f. (emphasis Whitehead's)"

Whitehead's observation/criticism that the MMCRI is based on all theoretically possible reflectances is completely reasonable, but are these theoretical reflectances potentially revealing important details that the finite set of 99 reflectances is missing? To address this question we turn to a recent study in which Zhang et al. [10] showed that the volumes of experimentally-determined metamer mismatch volumes correlated well with the volumes of theoretically-determined metamer mismatch volumes. They report that the experimental volumes are substantially smaller than the theoretical volumes, but that 'diameters' of the theoretical and experimental volumes are linearly correlated. To the extent that this correlation holds then the theoretical MMCRI should, in fact, be representative of the metamer mismatching of real reflectances as well.

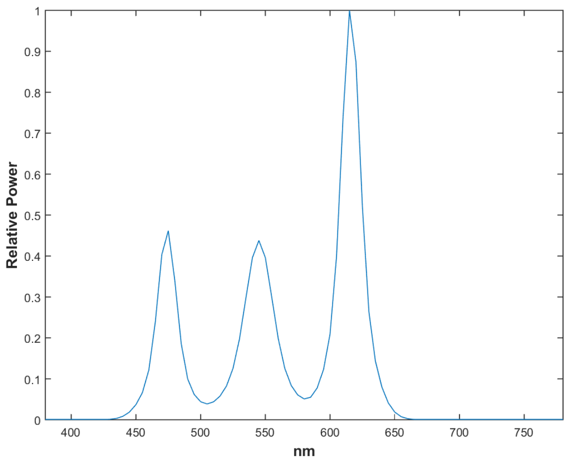

Figure 2. Spectral power distribution of test illuminant number #134 (RGB (474/545/616) Duv=-0.006) from the IES dataset having Rf = 74.2 and MMCRI = 40.8. The spectrum has three distinct peaks and low valleys.

To determine whether a 'practical' MMCRI (PMMCRI) based on real reflectances rather than theoretical ones would correlate more closely with R_f or with MMCRI, we obtained a copy of the 41 million real reflectance spectra used by Zhang et al. [10] in their experiment and followed their method of generating metamer mismatch volumes. In particular, their method is based on finding all reflectances in the database that are metameric under the reference illuminant within in a specific threshold and then relighting them under the test illuminant. The experimentally-determined metamer mismatch volume is then computed as the convex hull of the color signals obtained under the test (second) illuminant. The volume of the metamer mismatch volume is the volume of the convex hull of the resulting color signals. Analogous to the MMCRI, PMMCRI is defined in terms of a 'practical' MMVI (PMMVI), with PMMVI defined as:

$$PMMVI = \frac{volume\ of\ the\ experimental\ metamer\ mismatch\ volume}{volume\ of\ hull\ of\ the\ 41\ million\ color\ signals\ under\ test\ illuminant}$$

Note that the denominator is the volume of the convex hull of the gamut of all 41 million color signals obtained from the database of reflectances under the test illuminant. As such, it is the experimental equivalent of the volume of the theoretical object color solid. The experimentally-determined index PMMCRI is then defined as:

$$\mathrm{PMMCRI} = (1 - \sqrt[3]{\mathrm{PMMVI}}) \times 100$$

Results and Analysis

Figure 3 compares the PMMCRI (i.e., metamer mismatching based CRI calculated using 41 million reflectances) to R_f. It is clear that this plot is qualitatively similar to the plot for the theoretical case shown in Figure 1 of R_f versus MMCRI in that it also has a distinct lower branch illustrating agreement between the two indices and

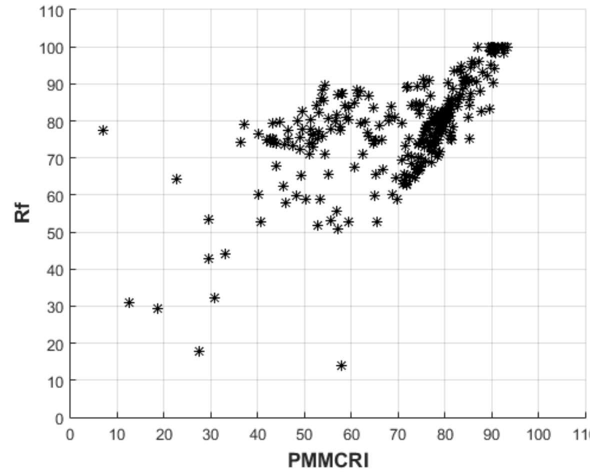

Figure 3. Comparison of the IES R_f color rendering index to PMMCRI.

an upper region (less clearly a branch in this case) in which the R_f ratings are consistently higher than those of PMMCRI. In other words, it is not the case that the differences between the MMCRI and R_f can all be attributed entirely to the fact that the theoretical calculations involve optimal reflectance functions since we now see that the same pattern arises for results based on real, measured reflectances as well.

Figure 4 compares PMMCRI with MMCRI and shows there is a strong correlation between the two, as was anticipated given the findings of Zhang et al. [10] concerning the correlation between theoretical and practical metamer mismatch volumes. Since the computation required to compute PMMCRI using the 41 million reflectances is high, it is significant that the MMCRI and PMMCRI results are correlated and that both are similar relative to R_f since it implies that MMCRI can be used instead of PMMCRI. Not only is MMCRI faster to compute, it is potentially more accurate and less prone to sampling bias.

The fact that the theoretical and experimental findings are similar still leaves the question as to why R_f and MMCRI agree much of the time, but not all of the time. Visual inspection of the illuminant spectra for which R_f>MMCRI reveals that they are exclusively functions one might describe as spiky, usually with three distinct, tall narrow peaks in their spectral power distribution functions, and generally having regions of very low power between the peaks. Figure 2 shows one example. Figure 5 shows another more extreme case in which R_f is low and MMCRI is very low. Cases such as illuminant #317 in Figure 5 may not be of too much concern since both metrics are clearly indicating the illuminant is of low color rendering. Figure 6, however, reveals a possibly more significant difference. For illuminant #37 R_f is relatively high at 80.6, whereas the MMCRI is only 53.0. Only future psychophysical experiments will tell which is the more accurate, but intuitively one might expect—given the spikiness of its spectrum accompanied with the fact that there are wavelength ranges over which its power is less than 5%—that there are least some real surfaces that illuminant #37 would not render very effectively.

In terms of the illuminants on which R_f and MMCRI agree, Figures 7 and 8 show two typical examples. Although illuminant #221 in Figure 7 has a spike, it is small relative to the overall peak and there is no wavelength region where its power is low. Illuminant #28 in Figure 8 is an interesting example in that it contains some distinct spikes, but both R_f and MMCRI are high at 85.8 and 87.0, respectively. In other words, it is clear in this case, at least, that MMCRI is not simply measuring the spikiness of the spectra or being overly influenced by the present of sharp peaks.

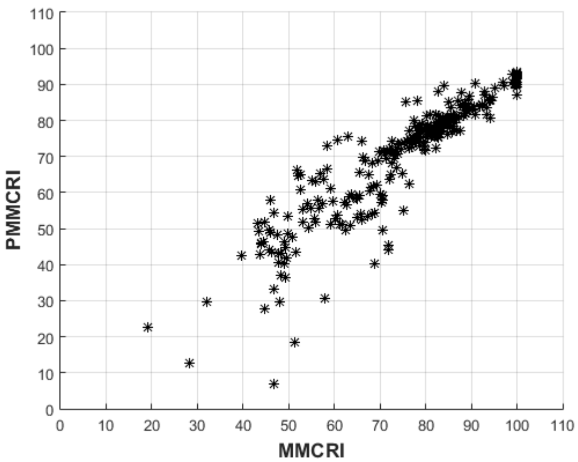

Figure 4 Plot of PMMCRI versus MMCRI showing a clear correlation between the experimental and theoretical measures of color rendering. Lighter dot shade corresponds to higher correlated color temperature.

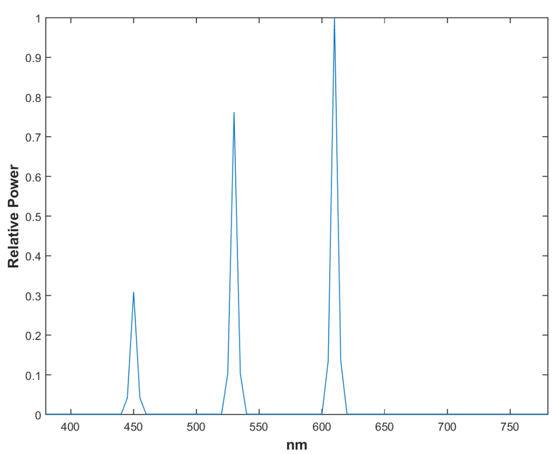

Figure 5 Spectral power distribution of test illuminant number 317 (Tri-band Gaussian) from the IES dataset for which R_f and MMCRI disagree. The spectrum contains three distinct spiky, very narrowband peaks with effectively zero power between the peaks. R_f =63.7 and MMCRI = 9.0

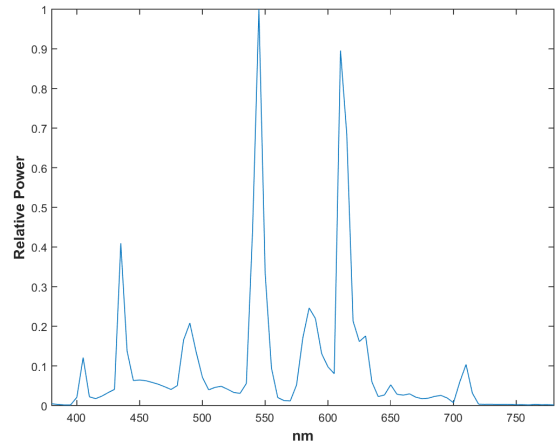

Figure 6 Spectral power distribution of test illuminant number 37 (F32T8/TL835 (1)) from the IES dataset for which R_f and MMCRI disagree and for which R_f is relatively high. The spectrum has numerous narrowband peaks with regions of very low power between them. R_f =80.6 and MMCRI = 53.0.

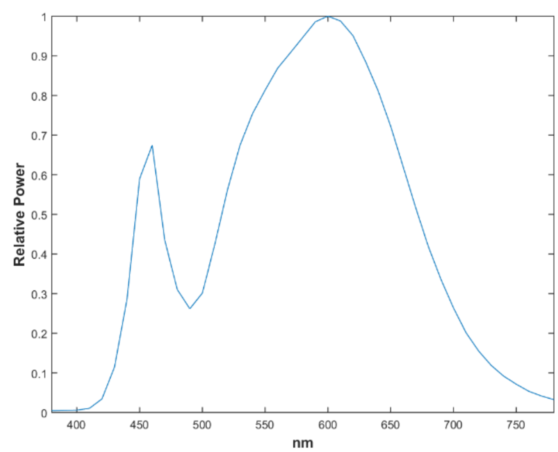

Figure 7 Spectral power distribution of test illuminant number 221 (LED Phosphor Blue Pump (53)) from the IES dataset for which R_f and MMCRI agree. The spectrum is smooth in comparison to those shown in Figures 5 and 6. It also contains no low values anywhere except at the very end of the visible spectrum. R_f =82.6 and MMCRI 85.1.

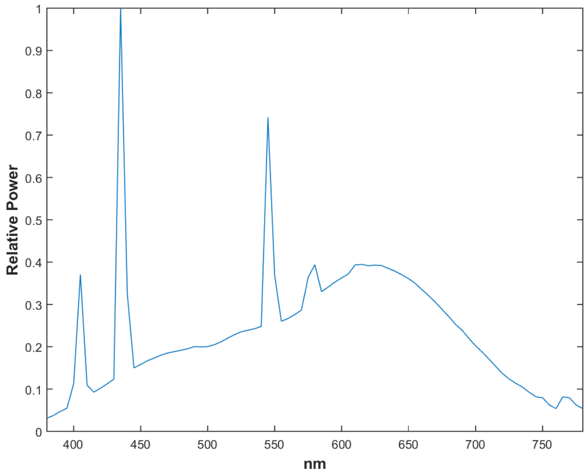

Figure 8 Spectral power distribution of test illuminant number 28 (F40T12/N (1)) from the IES dataset for which R_f and MMCRI agree. Although the spectrum has clear spikes, the tallest one is in a region of relatively low cone sensitivity and there are no near-zero values anywhere. R_f =85.8 and MMCRI= 87.0.

Conclusion

The MMCRI (metamer mismatching color rendering index) proposed by Mirzaei et al. [7] is compared to the newly IES-approved R_f color rendering index [4]. Tests on the 318 illuminant spectra provided by IES show agreement between MMCRI and R_f for the majority of spectra, but significant disagreement for many others. Since it has been conjectured that the differences may arise from the fact that the MMCRI computation involves the use of optimal reflectance spectra that do not occur in practice—with the implication being that such spectra make the index perhaps too sensitive to any sharp, spiky peaks in the illuminant spectra, thereby underrating their color rendering properties—we compare the MMCRI to a similar metamer mismatching index but based on using 41 million real reflectance spectra rather than optimal reflectance spectra. Interestingly, the results of the theoretically-determined and experimentally-determined metamer mismatch indices are comparable. Since MMCRI and PMMCRI yield similar results, in the future we can rely simply on the MMCRI. This is important because computing the PMMCRI is time consuming since it involves such a large dataset of reflectances. As well, even though the set is large, it still will reflect the bias of being a finite sample. In contrast, the MMCRI is based on all theoretically possible reflectances.

The illuminants for which MMCRI and R_f agree tend to have relatively smooth spectra without any wavelength regions of low power. The illuminants for which MMCRI and R_f disagree tend to include both strong peaks and wavelength regions of very low power. Regions of low power tend to increase dramatically the degree of metamer mismatching and, as a result, MMCRI gives them a lower color rendering score. The R_f, based as it is on an average of color differences, appears not to be particularly sensitive to spectra of that type. In terms of its use as a color rendering index, R_f may work well on average, but it does not guarantee that it will render the colors of all reflectances well. In comparison, MMCRI provides a measure of what the worst-case rendering might be for any surface under the given light. Especially in some color-critical industrial situations, it may be more important to know what can possibly go wrong than what will go right on average.

Unfortunately, there is insufficient psychophysical data yet available to establish which index might be better and under which circumstances. However, because the two indices are based on such different principles, the cases where R_f and MMCRI differ suggest the type of illuminants on which to focus future psychophysical experiments on color rendering.

Acknowledgements

Funding provided by Natural Sciences and Engineering Research Council of Canada and National Natural Science Foundation of China (NSFC) (61675060).

References

[1] Commission Internationale de l'Eclairage. "Method of measuring and specifying colour rendering properties of light sources," 3rd ed. Vienna (Austria): CIE. Publication No. 13.3:1995, 1995

[2] M. Fairchild, Color Appearance Models, 3rd Edition. Wiley Inc., Hoboken, New Jersey, 2013.

[3] A. David, P. Fini, K. Houser, Y. Ohno, M. Royer, K. Smet, M. Wei and L. Whitehead, "Development of the IES Method for Evaluating the Color Rendition of Light Sources," Optics Express, vol. 23, no. 12, 15888–15906, 2015.

[4] IES TM-30-15, "IES Method for Evaluating Light Source Color Rendition," Illuminating Engineering Society of North America, New York, 2015

[5] A. Logvinenko, Funt, B. and Godau, C., "Metamer Mismatching," IEEE Trans on Image Processing, vol. 23, pp. 34-43. 2014

[6] M. R. Luo, G. Cui, C. Li, "Uniform Colour Spaces Based on CIECAM02 Color Appearance Model," Color Research and Application, vol. 31, no. 4, pp. 320–330, 2006.

[7] H. Mirzaei and B. Funt, "Metamer Mismatching as a Measure of the Color Rendering of Lights," in Proc. AIC2015 International Colour Association Conference, Tokyo, 2015.

[8] L. Whitehead, Lorne, Personal Communication, March 3, 2016

[9] G. Wyszecki and W.S. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd Edition, Wiley Inc., Hoboken, New Jersey, 2000.

[10] X. Zhang, B. Funt and H. Mirzaei, "Metamer Mismatching in Practice versus Theory," Journal of the Optical Society of America A, Vol. 33, No. 3, pp. A238-A247, 2016.

Author Biographies

Brian Funt is Professor of Computing Science at Simon Fraser University where he has been since 1980. He obtained his Ph.D. from the University of British Columbia in 1976. His research focus is on computational approaches to modeling and understanding color."

Ben Hull obtained his M.Sc. in Computing Science from Simon Fraser University. He is currently a research engineer at Point Grey Research Inc. in Vancouver.

Xiandou Zhang is Associate Professor of digital media technology at Hangzhou Dianzi University where he has been since 2013. He obtained his Ph.D. from Zhejiang University in 2010. His research focus is on color science and imaging technology.