Metamer Mismatching and its Consequences for Predicting How Colours are Affected by the Illuminant

Xiandou Zhang
Hangzhou Dianzi University
China
xiandouzhang@126.com

Brian Funt
Simon Fraser University
Canada
funt@sfu.ca

Hamidreza Mirzaei
Simon Fraser University
Canada
hmirzaei@sfu.ca

Abstract

Metamer mismatching (the phenomena that two objects matching in color under one illuminant may not match under a different illuminant) potentially has important consequences for color-based machine vision. Logvinenko et al. [1] show that in theory the extent of metamer mismatching can be very significant. This paper examines metamer mismatching in practice by computing empirical metamer mismatch volumes. A set of more than 20 million unique reflectance spectra is assembled using datasets from several sources. For a given color signal (i.e., RGB or CIE XYZ) recorded under a given first illuminant, its empirical metamer mismatch volume for a change to a second illuminant is computed as follows: the reflectances having the same color signal when lit by the first illuminant (i.e., that are metamers) are computationally relit by the second illuminant and the convex hull of the resulting color signals then defines the empirical metamer mismatch volume. The volume of these volumes is shown to vary systematically with Munsell value and chroma. The centroid of the empirical metamer mismatch volume is also tested as a predictor of what a given color signal might become under a specified illuminant.

1. Introduction

Metamer mismatching [2] refers to the fact that two objects reflecting metameric light under one illumination may reflect non-metameric light under a second; so two objects appearing as having the same colour under one illuminant may have different colours under a second. Metamer mismatching has important consequences for computer vision since the light illuminating an object is frequently changing, for example as it moves from direct sun to shadow, or when the lights are turned on in a room, or the image is taken at a different time of day, or the object is viewed under fluorescent light at one moment and tungsten light at another.

As Logvinenko et al. [3] argues, metamer mismatching imposes limits on color constancy since even when the full spectra of the two illuminants are known there is an inherent ambiguity in terms what a given colour signal (i.e., camera RGB or CIE XYZ coordinates) under a first illuminant will become under a second illuminant.

In the computer vision and color constancy fields, it is generally assumed that the color of an object is an intrinsic property of the object and hence the focus is on discounting the effects of the illuminant in order to recover the intrinsic color of the object. The intrinsic color is frequently expressed as the color signal that would be obtained from the object under some standard, "canonical" illuminant. However, Logvinenko [3] proves that color cannot be an intrinsic property of an object. His argument is straightforward: If two objects, A and B, are metameric matches (i.e., reflect light that generates an identical color signal) under the first illuminant, but do not match under the second illuminant then which is to be considered the carrier of the 'intrinsic' color? However, one might consider mapping the colors to some canonical coordinate system, a single color that becomes two different colors cannot possibly map to some unique 'intrinsic color' coordinate.

Metamer mismatching means that a color signal under a first light can become any one of an infinite convex set of different color signals under a second light. This convex set is usually called the *metamer mismatch volume*. Logvinenko et al. [1] provide an algorithm for computing the theoretical metamer mismatch volume of a given color signal and illuminant pair and show that these volumes can be surprisingly large.

Given a color signal arising from an object under a first illuminant, what can be said about its color signal under a second illuminant? Unfortunately, all that can be said definitively is that it could be any one of the color signals within the metamer mismatch volume. The metamer mismatch volume therefore represents the unpredictability we have in knowing what the color signal of an object will be under the second light. For computer vision applications it is important to know what the degree of this unpredictability is in practice. We will show that the degree of uncertainty varies with the type of color signal involved. In particular, the more a color signal differs from mid-grey, the smaller its metamer mismatch volume,

and hence, the less the uncertainty. This has important implications for color-based object recognition or any other color-based analysis (e.g., dermatological imaging, dentistry) whenever more than one illumination condition is expected.

The theoretical metamer mismatch volumes are based on the fact that the reflectances generating color signals on the boundary of the object color solid are special 2-transition reflectances [1]. Such 2-transition reflectances are either zero or one and make at most two transitions from zero to one or vice-versa across the visible spectrum. Clearly, we do not expect such reflectances to arise in practice, but there is no clear, non-arbitrary way to constrain the set of reflectances.

To establish the extent of metamer mismatching in practice, we examine empirically the metamer mismatch volumes arising under several typical illumination changes for a large set of reflectance spectra obtained from multispectral images and other data sets of reflectances. Foster et al. [4] performed a related analysis in which they evaluate the frequency with which metameric matches are likely to occur in a typical scene. They do not evaluate metamer mismatching or the consequences of metamer mismatching for color-based analysis.

Despite the unpredictability that results from metamer mismatching, it is nonetheless the case that we frequently wish to predict what the object's color signal is most likely to be under a second illuminant. Of course, any prediction can only be a guess since any of the color signals within the metamer mismatch volume is a plausible answer. However, when forced to choose, what is a good choice to make? We explore this issue by making predictions based on several different measures (e.g., mean, median, centroid) of the metamer mismatch volume and compare the mean prediction error to that obtained using the CAT02 [5] chromatic adaptation transform that underlies the CIECAM02 [6] color appearance model and to Mirzaei's [7] Gaussian Metamer method of color signal prediction.

2. Reflectance and Illuminant Spectra

In order to analyze the effects of metamer mismatching in practice, we construct a large dataset of reflectance spectra along with some illuminant spectra. The reflectance data are divided into disjoint training and test sets. Even though there is no machine learning involved, we use the term 'training set' since we will be predicting results for the test data based on a prior set of reflectance data.

2.1. Large Dataset of Training Reflectances

A large dataset of spectral reflectances was created by gathering spectra from various sources in order to create a representative dataset of the spectral reflectances of natural and man-made objects that are likely to occur in practice. All the spectral reflectances are sampled from 400nm to 700nm at a 10nm sampling interval.

The dataset is assembled from four main sources. The first group includes eight multispectral images [4] consisting of rocks, trees, leaves, grass, earth and urban scenes. The second group includes thirty-two multispectral images [8] containing scenes of faces, hair, paints, food, drinks and some other natural and man-made items. The third group includes thirteen multispectral images [9] containing scenes of people, houses, hands, fruits, flowers and other natural and man-made items. These three groups of images were all acquired with multispectral imaging systems. The fourth group mainly includes spectral reflectances of man-made, natural and industrial objects [10], which were measured by spectral photometers. The details of the fourth group are shown in Table 1.

Table 1: Spectral reflectances in group 4.

Dataset Label	No. of samples
DuPont spectra-master	672
Din	981
Sun Chemical	26784
Foliage	21
Calibration Data (ISO SOCS)	136
Skin (ISO SOCS)	8570
Flowers (ISO SOCS)	148
Graphics (ISO SOCS)	30624
Krinov (ISO SOCS)	346
Leaves (ISO SOCS)	92
Paint (ISO SOCS)	505
Photos (ISO SOCS)	2304
Printer (ISO SOCS)	7856
Textile (ISO SOCS)	2832
Oulu African	24
Oulu Caucasian	303
Oulu Oriental	30
RIT African	70
RIT Caucasian	78
RIT Oriental	112
RIT SubAsian	60
RIT Hispanic	20
Industry Cotton SPI	4028
Industry Plastic SPI	5338
Industry Pantone Polyester SPI	1925
Industry Pantone Cotton SPI	1925
PCC SPI	1063

Since many of the reflectance spectra are from multi-spectral images, it is likely that there will be many

extremely similar or duplicate spectra in the datasets. To eliminate these similar/duplicate spectra, the numerical precision of the spectral data is first reduced to integer values 0-50 (i.e., multiplied by 50 and rounded), and then any spectra that are identical at that level of precision are removed. This brings the initial set of 27,061,874 spectra down to 17,873,556 distinct spectra.

2.2. Dataset of Test Reflectances

For testing, a second, smaller set of reflectance spectra is created by combining the 1600 reflectances of the Munsell glossy edition [11] papers, the 1950 reflectances of the Natural Color System [12] samples, along with the 218 reflectances from the "Natural Colors" subset of the University of Eastern Finland's spectral database [13], and 1301 reflectances of natural objects in the ASTER Spectral Library from the Jet Propulsion Laboratory [14].

2.3. Chromaticities of the Reflectances

As an indicator of how complete the set of reflectances is we computed CIE1931XYZ values under CIE D65 (daylight) of all the spectral reflectances in the training and test sets and plotted them in xy-chromaticity (i.e., x = X/(X+Y+Z), y = Y/(X+Y+Z)) space as shown in Figure 1. The plot shows that the training set (black dots) covers most of the xy-chromaticity diagram.

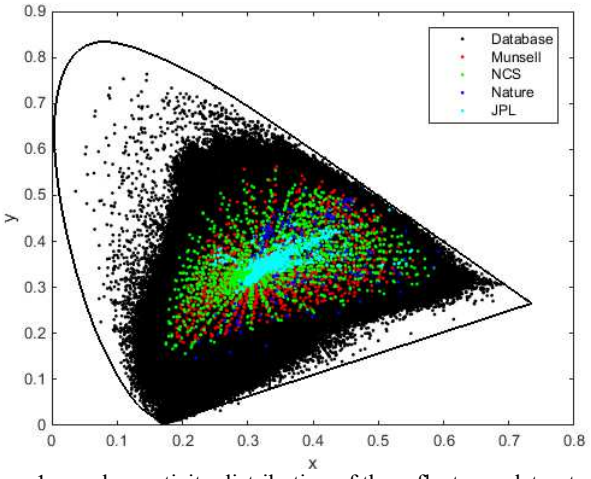

Figure 1: xy-chromaticity distribution of the reflectance datasets

Note that we will report results in CIE XYZ coordinates since they are well defined. The sRGB standard for color imaging devices formally defines a unique mapping from non-linear camera sRGB to CIE XYZ so our results apply equally well to an sRGB camera. The only difficulty is that most sRGB cameras conform only approximately to the sRGB standard.

2.4. Illuminant spectra

Seven illuminants, namely, the CIE standard illuminants A, D65 (6504K), D200 (20000K), F4, F8 and F11, along with a cell phone LED are used in evaluating the metamer mismatch volumes and color signal prediction results. They were chosen as a representative test set since A is a typical tungsten light bulb, D65 is typical daylight, F4, F8 and F11 are typical fluorescents with varying degrees of spikiness in their spectra. The spectral power distributions of these illuminants are shown in Figure 2.

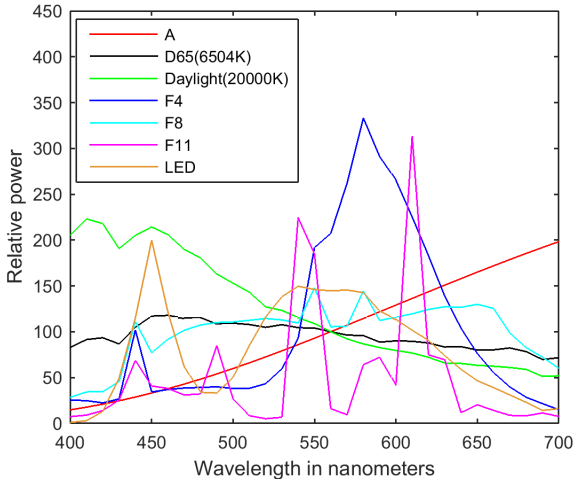

Figure 2: Spectral power distributions of the seven illuminants

3. Empirical Metamer Mismatch Volumes

While large, the training set of reflectances is still limited and hence it is unlikely that there will be many exact metameric matches to a given color signal. As a result, we consider any nearby color signal within a threshold distance T to be a metameric match. In other words, two color signals (X_c, Y_c, Z_c) and (X_i, Y_i, Z_i) will be consider metameric matches whenever

$$\sqrt{(X_i-X_c)^2+(Y_i-Y_c)^2+(Z_i-Z_c)^2} < T. \qquad (1)$$

All the results reported below are based on $T = 0.3$.

Using this definition of metameric matching, given a color signal (Xc, Y_c, Z_c), we find all the reflectances in the training set generating metameric color signals under the first illuminant. Using this set of reflectances, the *empirical metameric mismatch volume* is then determined as the convex hull of the colour signals generated by these reflectances under the second illuminant.

Since the range of possible color signals an initial color signal under the first illuminant can become under the second illuminant is only limited by the metamer mismatch volume, an interesting questions is: How does the volume of the empirical metameric mismatch volume vary with the initial color signal? To address this question, we computed the empirical metameric mismatch volume

for each of the 1600 reflectances from the Munsell color atlas for a change in illuminant from A to D65.

Figure 3 shows how the volume varies with the value and chroma of the Munsell samples. The red dots denote the volume averaged over all hues of samples having fixed value and chroma. The volume clearly peaks in the achromatic (chroma zero) samples and decreases with increasing chroma. In terms of value, it peaks at value 5 and decreases for values larger or smaller than 5. The relationship between volume and value is clearer in Figure 4 in which the volumes of the 37 Munsell neutral grays (chroma zero) are plotted as a function of their value.

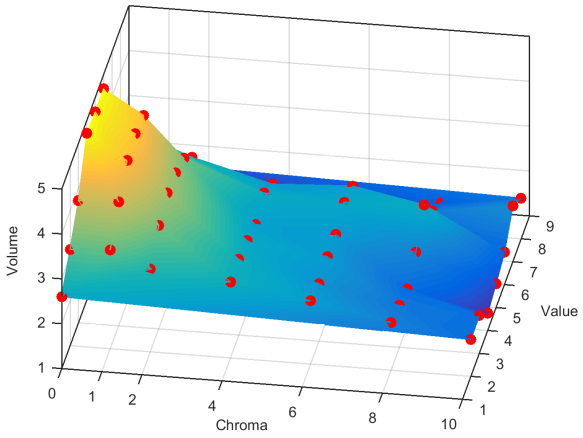

Figure 3: Volume (averaged across all Munsell hues) of the empirical metamer mismatch volume as a function of Munsell chroma and value for a change from A to D65. Red dots indicate the actual data points. The surface is interpolated through the data points to aid in visualization. The plots for the other illuminant pairs are qualitatively similar.

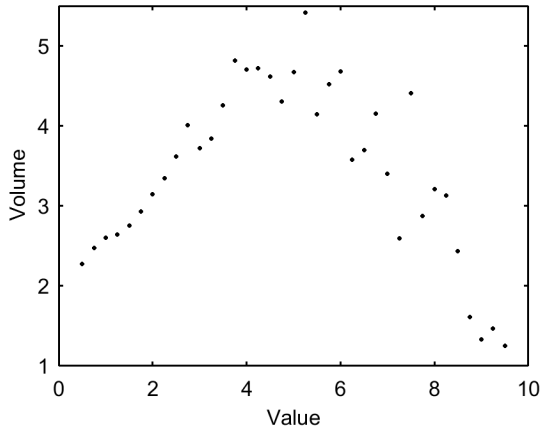

Figure 4: Volume of the metamer mismatch volumes of the 37 neutral gray Munsell papers as a function of Munsell value.

Logvinenko, loosely speaking, defines chromatic purity in his object color atlas [15] in terms of how far a color signal is from ideal grey. Figure 5 shows the relationship between average volume and Euclidean distance from mid-gray (chroma zero, value 5). The distance measure is based on assuming that all unit steps in either value or chroma in the Munsell color atlas are equal. Although this may be a questionable assumption, the qualitative trend in Figure 5 is clear—the further a sample is from mid-grey the smaller the volume of its empirical metamer mismatch volume.

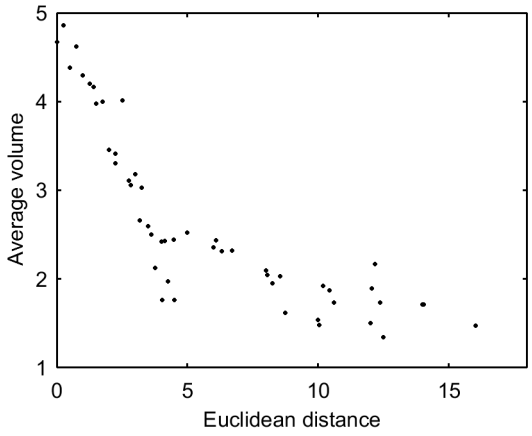

Figure 5: Average volume of the empirical metamer mismatch volume as a function of the Euclidean distance in Munsell space from the Munsell mid-grey neutral (chroma 0, value 5).

4. Color Signal Prediction Methods and Results

As mentioned above, given a color signal under one illuminant, all that can be definitively determined about what the color signal will become under a second illuminant is that it will lie within the theoretical metamer mismatch volume. Of course, if the reflectance that led to the given color signal is known then the new color signal can be simply calculated. However, in color imaging and human vision the reflectance is not available and any prediction must be made based on the color signal alone. We describe a new method of making such a prediction based on the properties of the empirical metamer mismatch volume and compare it to existing methods of color-signal prediction. For the tests described below, the training set described in Section 2.1 was augmented with reflectance spectra from two other multispectral image datasets [16, 17]. These datasets include nine hyperspectral images containing scenes of textile, wood, leaves, painting, paper and skin [16] and twenty-one multispectral images mainly composed of different man-made items [17].

4.1. Metamer-Based Prediction Method

Mirzaei et al. [7] propose a method of color-signal prediction based on relighting a ‘wraparound Gaussian metamer’. Given a color signal under a first illuminant, the idea is to find a Gaussian-like (the precise details are irrelevant for the present discussion) reflectance function

producing that same color signal under the first illuminant and then to calculate what that reflectance's color signal would be under the second illuminant. They report excellent results using this Gaussian Metamer (GM) method.

The GM method would appear to be limited in that the form of the metameric reflectance is fixed as something Gaussian-like. In comparison, the empirical metamer mismatch volume is based on relighting the many reflectances from the training set producing color signals that are approximately metameric to the given color signal under the first illuminant. The training set also contains only real reflectances, in other words, ones measured in practice rather than an idealized Gaussian-like reflectance function. We hypothesize that basing the color signal prediction on the properties of the empirical metamer mismatch volume rather than a single idealized reflectance will lead to better results.

The empirical metamer mismatch volume represents the range of possible color signals that might arise under the second illuminant so the centroid of the volume's convex hull is one method of predicting the color signal under the second illuminant. Other choices we investigate are the mean of the color signals under the second illuminant, or similarly, the median color signal. In all cases, we remove all color signals that are more than three standard deviations from the mean as being outliers.

4.2. Prediction Results

The spectral reflectances of Munsell, NCS, UEF Natural and JPL data sets described above are used for testing. These datasets are distinct from the training set. Predictions are made for a change from each of illuminants A, F4, F8, F11, D200 and LED to D65 as the 'canonical' illuminant. The centroid, mean and median methods were all tested. The results for all three are comparable, but the centroid method generally outperforms the others so only the results for it are reported here.

For comparison, the Gaussian metamer method and the von-Kries-based CAT02 chromatic adaptation transform are tested as well. CAT02 is the chromatic adaptation step underlying the CIECAM02 color appearance model. CAT02 includes a spectral sharpening transform [18]. Chong et al. [19] propose a tensor-based method of choosing the basis for the diagonal transform. The prediction error is measured using the CIEDE2000 [20] color difference measure. The results for each reflectance data set and illuminant are listed in Table 2. Table 3 lists the results averaged over all the test reflectances and the 6 illuminant pairings (A to D65, F4 to D65, F8 to D65, F11 to D65, D200 to D65, and LED to D65). Overall, the centroid method outperforms the other methods.

Table 2: CIEDE2000 prediction error statistics (mean, median, 95th percentile) for the different prediction methods on different reflectance test sets for a change from A, F4, F8, F11, D200, LED to D65.

		A			F4			F8			F11			D200			LED		
		avg	med	95th	avg	med	95th	avg	med	95th	avg	med	95th	avg	med	95th	avg	Med	95th
Centroid	Munsell	1.27	0.94	3.42	2.00	1.64	4.98	0.52	0.39	1.34	1.82	1.36	4.99	0.63	0.51	1.57	0.82	0.67	2.05
	NCS	1.08	0.90	2.64	1.67	1.32	4.20	0.43	0.35	1.06	1.69	1.29	4.69	0.61	0.51	1.45	0.66	0.55	1.55
	Nature	0.91	0.77	2.39	1.47	1.17	3.80	0.48	0.34	1.14	1.59	1.07	3.75	0.60	0.52	1.48	0.80	0.62	1.83
	JPL	0.77	0.51	2.31	1.13	0.76	3.36	0.52	0.33	1.37	1.00	0.68	2.83	0.54	0.37	1.49	0.78	0.50	1.86
GM	Munsell	1.37	1.02	3.74	2.05	1.59	5.28	0.65	0.44	2.03	1.88	1.56	4.68	0.81	0.55	2.42	1.22	1.06	2.68
	NCS	1.49	1.10	3.87	2.10	1.71	4.83	0.74	0.49	2.13	1.92	1.66	4.46	0.87	0.59	2.45	1.20	1.07	2.45
	Nature	2.01	1.48	4.70	2.60	2.54	5.15	1.00	0.88	2.40	2.25	2.07	4.22	1.22	1.17	2.67	1.66	1.76	2.76
	JPL	1.19	0.77	3.58	0.79	0.44	2.68	0.57	0.39	1.72	0.68	0.38	2.65	0.65	0.40	2.00	0.57	0.44	1.46
CAT02	Munsell	2.06	1.77	4.74	4.36	3.33	10.80	0.73	0.63	1.66	1.87	1.51	4.67	1.03	0.89	2.23	2.42	2.20	5.27
	NCS	2.16	1.82	4.95	4.65	3.71	11.46	0.80	0.69	1.73	2.08	1.63	5.60	1.09	0.94	2.40	2.54	2.29	5.29
	Nature	2.53	2.48	4.11	3.48	2.98	7.32	0.89	0.85	1.56	2.62	2.40	4.83	1.17	1.07	2.28	2.68	2.54	4.83
	JPL	1.04	0.79	2.75	1.68	0.92	6.07	0.57	0.50	1.33	0.49	0.28	1.60	0.84	0.75	1.92	1.09	0.69	3.31

Table 3: CIEDE2000 prediction error statistics (mean, median and 95th percentile) across the combined set of test reflectances and all 6 illuminant pairs

	mean	median	95th
Centroid	1.09	0.71	3.41
GM	1.21	0.82	3.59
CAT02	1.85	1.19	5.78

For the Munsell reflectances, the prediction error varies with the value and chroma of the Munsell sample. Figure 6 shows that the mean color difference generally is lowest for high chroma combined with low or high value — similar to the trend found in Figure 3 for the volume of the empirical metamer mismatch volumes. The trend is noisy since each point is based on a very limited number of samples so by chance some predictions may be significantly better or worse than average, but clear nonetheless.

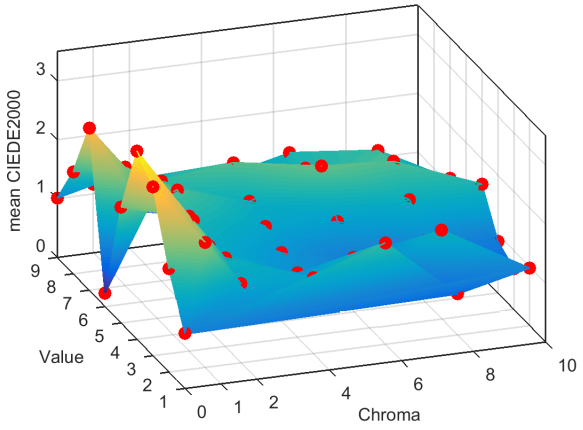

Figure 6 : Red dots represent the CIEDE2000 color difference at a given chroma and value averaged over all hues. The surface is interpolated through the data points to aid in visualization.

5. Conclusions

Metamer mismatching imposes a limit on the accuracy with which it is possible to predict the effect a change in illumination will have on a given color signal. Starting with a large set of unique reflectance spectra, the volume of the empirical metamer mismatch volumes is shown to decrease with increasing distanceof the color signal from mid-grey. This result means that for color-based recognition conducted under multiple illuminants these colors should generally be given priority. Studying how important this effect is in the context of object recognition is a topic of our ongoing research. In terms of predicting what a given color signal may become under a new illuminant, the centroid of the empirical metamer mismatch volume performs better overall than the Gaussian Metamer method, which in turn outperforms CAT02.

ACKNOWLEDGMENTS

The authors acknowledge the support of the National Natural Science Foundation of China (grant 61205168), National Science and Technology Support Program of China (grant 2012BAH91F03) and the Natural Sciences and Engineering Research Council of Canada.

References

[1] A. D. Logvinenko, B. Funt, and C. Godau. Metamer mismatching. IEEE Trans. on Image Processing, 23(1): 34-43, 2014.

[2] G. Wyszecki, and W. S. Stiles. Color science: concepts and methods, quantitative data and formulae. 2nd edition, John Wiley and Sons, New York, 1982.

[3] A. D. Logvinenko, B. Funt, H. Mirzaei, and R. Tokunaga. Rethinking colour constancy. PLOS ONE, 10(9): e0135029, 2015.

[4] D. H. Foster, K. Amano, S. M. C. Nascimento, and M. J. Foster. Frequency of metamerism in natural scenes. J. Opt. Soc. Am. A, 23(10):2359-2372, 2006.

[5] M. D. Fairchild. Color appearance models. 3rd edition, Wiley-IS&T, Chichester, UK, 197-198, 2013.

[6] CIE Publication 159. A Colour appearance model for colour management systems: CIECAM02. CIE Central Bureau, 2004

[7] H. Mirzaei, and B. Funt. Object-color-signal prediction using wraparound Gaussian metamers. J. Opt. Soc. Am. A, 31(7):1680-1687, 2014.

[8] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar. Generalized assorted pixel camera: post-capture control of resolution, dynamic range and spectrum. Technical Report, 2008.http://www.cs.columbia.edu/CAVE/databases/multi spectral/

[9] Joensuu Spectral Image Database. Spectral Color Research group, University of Eastern Finland. http://www.uef.fi/fi/ spectral/spectral-database.

[10] C. Li, M. R. Luo, M. R. Pointer, and P. Green. Comparison of real colour gamuts using a new reflectance database. Color Res Appl., 39(5): 442-451, 2014. (Data provided by R. Luo)

[11] A. H. Munsell. The Munsell book of color. Glossy edition, Grand Rapids, MI: X-rite Inc.

[12] A. Hard, and L. Sivik. NCS-Natural color system: A Swedishstandard for color notation. Color Res. Appl., 6(3): 129-138, 1981.
[13] J. Parkkinen, T. Jaaskelainen, and M. Kuittinen. Spectral representation of color images. ICPR, 2: 933-935, 1988. http://www2.uef.fi/fi/spectral/natural-colors
[14] A. M. Baldridge, S. J. Hook, C. I. Grove, and G. Rivera. The ASTER spectral library, Version 2.0. Remote Sensing of Environment, 113:711-715, 2009.
[15] A. D. Logvinenko. An object-colour space. J. Vis. 9(11): 1-23, 2009.
[16] S. Moan, S. George, M. Pedersen, J. Blahova, and J. Hardeberg. A database for spectral image quality. IS&T/SPIE, 9396, 2015. http://www.ansatt.hig.no/stevenl/sidq/
[17] S. Hordley, G. Finlayson, P. Morovic. A multi-spectral image database and an application to image rendering across illumination. Proceedings of Third international conference on image and graphics, Hong-Kong, 2004. http://www2.cmp.uea.ac.uk/Research/compvis/MultiSpectralDB.htm
[18] G. Finlayson, M. Drew, and B. Funt. Spectral sharpening: sensor transformations for improved color constancy. JOSA A, 11(5):1553-1563, 1994.
[19] H. Chong, S. Gortler, T. Zickler. The von Kries hypothesis and a basis for color constancy. ICCV, 1-8, 2007.
[20] CIE Publication 142. Improvement to industrial colour-difference evaluation. CIE Central Bureau, 2001.