Color Inconstancy Index: A Proposal and Case Study for Consumer Digital Cameras

Don Williams
Eastman Kodak Company
Rochester, New York USA

Abstract

The rationale, formulae, and examples for assessing the color inconstancy of consumer digital cameras are presented. Its use is appropriate for evaluating the color/neutral balance precision of consumer digital camera systems with respect to illumination quality; specifically for a given rendering intent such as sRGB or ROMM. Adopting the ΔE_{94}* color difference measure, an extended color inconstancy index (CII) is used to track the stability of color balancing algorithms across multiple illuminant types by effectively performing a variance analysis with respect to the illuminant variable for multiple color stimuli. Five differently branded 2 megapixel cameras were tested. The CII numerical results correlated well with visual engineering judgments. Results for both chromatic and nonchromatic samples are presented.

Background

In evaluating neutral or color balance algorithm performance one may be interested in not only the mean color position across illuminants, but also the constancy of that color position across illuminants. Color constancy is the tendency of samples to retain their color appearance in spite of changes in illumination.[1] Because performance in this area is quantified with a difference metric though, the aim is to minimize rather than maximize the metric. To reflect this aim, the appropriate measure is termed Color Inconstancy Index (CII); the lower the CII the better the performance.

Although a digital camera may not be required to capture identical images for the same scene with multiple illuminants, most consumer digital cameras try to avoid dramatic changes. Because the camera spectral sensitivity, determined by optics, CFA, and detector do not adapt to discount illumination quality changes, it falls to the signal processing to provide this characteristic. This is done in color/neutral balance type of operations. In evaluating digital cameras under various illuminants, therefore, a method for measuring the effective color inconstancy is developed that integrates multiple stimuli.

Minimizing color inconstancy can be important because it tracks the predictability, precision, and stability of color balancing algorithms. Pragmatically, this can be as important as mean color accuracy or preference because it aids in providing greater yields of good pictures instead of lower yields of excellent ones by minimizing color correction selections, and in turn, ambiguity. Wouldn't it be nice to know that if correction to a preferred color position had to be done (regardless of its mean position), a single profile choice would suffice instead of an illuminant dependent one? It is a fact that many Japanese photographic trade journals qualitatively, and exhaustively, test for this ability in evaluating digital camera performance.

To be thorough, it is emphasized that this index should be considered a color performance rather than a color quality or color appearance metric. Hubel[2] accurately points out that for certain scenes (such as sunsets), where a warm glow of an observer's adapted white point is desired in the rendered output, the proposed index would be an unacceptable indicator of quality because it would discount variations in adopted white points. To the extent that one is willing to consider such images as idiosyncratic (although frequent), then this metric would be suitable for the vast majority of color imaging situations. The author's stand is while there are wrong ways to perform this calculation, there are also numerous legitimate ways of doing so. What follows is a mathematical development of the index as an extension of past color difference formulations supported by experimental results.

Theory

The computational model for CII is rooted in the popularly accepted formula for calculating the classical values. This latter formula averages the squared differences of a population of N L*a*b* observations. These differences are calculated between observed values and reference or aim values for a number N, of intra-illuminant color samples. It is formulated according to Eq. 1.

$$\Delta E^*_{ab} = \frac{1}{N}\sum_{i=1}^{N}\sqrt{(\Delta L^*_i)^2 + (\Delta a^*_i)^2 + (\Delta b^*_i)^2} \qquad (1)$$

In Eq. 1, ΔE^*_{ab} can be considered the dispersion in CIELAB from an aim or mean, based on N different observations over a color volume. Typically, these observations correspond to the number of samples in the target being used. (e.g., 24 for a Gretag-Macbeth ColorChecker). Because the dispersion calculations of Eq. 1 are indexed over the number of samples, it can be considered an analysis of variance over the color volume under the conditions for which the data were captured; typically for a given illuminant.

To overcome reference chroma limitations of ΔE^*_{ab} the CIE adopted the ΔE^*_{94} color difference measure. Adopting Eq. 1 using the ΔE^*_{94} criteria yields the following.

$$\Delta E^*_{94} = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\left(\frac{\Delta L^*_i}{k_L S_L}\right)^2 + \left(\frac{(\Delta C^*_{ab})_i}{k_C S_C}\right)^2 + \left(\frac{(\Delta H^*_{ab})_i}{k_H S_H}\right)^2} \quad (2)$$

where

$$S_L = 1$$

$$S_C = 1 + 0.045\,\overline{C}^*_{ab}$$

$$S_H = 1 + 0.015\,\overline{C}^*_{ab}$$

$$k_L = k_C = 2$$

$$k_H = 1$$

and

$$\overline{C}^*_{ab} = \sqrt{(\Delta a^*)^2 + (\Delta b^*)^2}$$

$$h_{ab} = \tan^{-1}\left(\frac{b^*}{a^*}\right)$$

$$C^*_{ab} = \sqrt{a^{*2} + b^{*2}}$$

$$\Delta H^*_{ab} = 2\sqrt{C^*_{ab}\overline{C}^*_{ab}}\,\sin\left(\frac{\Delta h_{ab}}{2}\right)$$

The proposed CII calculation for digital cameras is similar in looks to Eq. 2, but is different in that the dispersion calculations are explicitly indexed across illuminant, M. While several color samples, N, over the color volume are used as replicates, the variance analysis is performed only with respect to the illuminant variable. For Eq. 1 or 2, the illuminant is often assumed fixed.

The proposed CII calculation for digital cameras expanded for M illuminants is formulated in Eq. 3 below.[3]

$$CII = \frac{1}{N}\sum_{ki=1}^{N}\sqrt{\left(\frac{\sigma_{L^*_k}}{2S_L}\right)^2 + \left(\frac{\sigma_{C^*_{ab_k}}}{2S_C}\right)^2 + \left(\frac{\Delta\overline{H}^*_{ab_k}}{S_H}\right)^2} \quad (3)$$

where

$$\sigma_{L^*_k} = \sqrt{\frac{1}{M-1}\sum_{i=1}^{M}\left[L^*_{i,k} - \overline{L}^*_k\right]^2}$$

$$\sigma_{C^*_{ab_k}} = \sqrt{\frac{1}{M-1}\sum_{i=1}^{M}\left[\left(C^*_{ab}\right)_{i,k} - \left(\overline{C}^*_{ab}\right)_k\right]^2}$$

$$\Delta\overline{H}^*_{ab_k} = \frac{1}{M}\sum_{i=1}^{M}\Delta H^*_{ab_{ik}}$$

and

$$\overline{L}^*_k = \frac{1}{M}\sum_{i=1}^{M}L^*_{i,k} \qquad \left(\overline{C}^*_{ab}\right)_k = \frac{1}{M}\sum_{i=1}^{M}\left(C^*_{ab}\right)_{i,k}$$

For this study, the dispersion across illuminants was calculated relative to mean $L^* C^*_{ab} h^*_{ab}$ values. Of course, depending on interpretive goals, the dispersion could be calculated relative to either absolute or preferred aims. These aims could even be weighted or constrained with respect to certain memory colors or exclusively neutrals. For example, studies[4,5] of the quantitative effects of blue sky, foliage, and flesh color on absolute quality have shown that deviations in flesh reproduction, as measured in L*a*b* values, produce approximately five times as much quality loss as equivalent errors in sky and foliage reproduction. Whatever the choice, the key is the variance is calculated across illuminants rather than within an illuminant. Naturally, the lower the CII, the better the performance.

Finally, whereas the calculations proposed here are very similar to those for classical ΔE_{94}* formulations, one may be able, with caution, to apply the same tolerance values on CII as those often adopted for ΔE_{94}* tolerancing. Not only are the formulations similar, but also the color space in which they are done.

Experimental

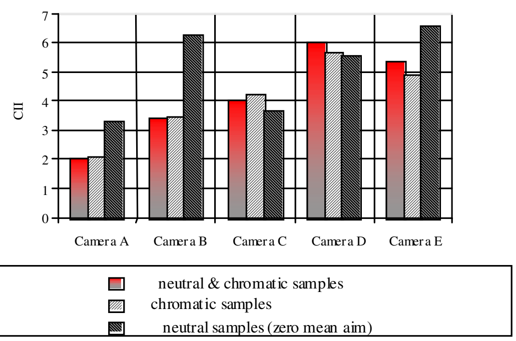
Single candidates of five differently branded two-megapixel digital cameras were evaluated.. For each of these cameras, a test target of a Gretag-Macbeth ColorChecker on an 18% gray background was photographed under five different studio illumination sources. The target itself filled approximately one-third of the areal field-of-view, and was centered within the frame. The color temperatures of each source were approximately 2400 K, 3200 K, 3600 K, 6500 K, and on-camera flash. It is across these five illuminants that the color variance analysis was performed. The on-camera flash illumination between cameras was treated as equivalent color temperatures. All cameras were placed in a center weighted, auto-white balance mode.

All of the calculations were based on full-resolution finished file captures that were assumed to be in an sRGB metric. Though technically this assumption is sometimes incorrect, digital camera finished files for consumer use are currently treated this way, for better or worse, in virtually all cases. After each image was converted to L*a*b* via Photoshop software, mean level statistics of the sample patches were calculated. These statistics were used for the computations in Eq. 3.

Results and Discussion

For demonstration, the CIIs for neutral, chromatic, and combined neutral/chromatic samples of the color target were calculated. These results are illustrated in Fig. 1 below.

Figure 1- CII results for consumer digital cameras



To aid in visualizing the inter-illuminant differences, a graphical color wheel was created for each camera. These are shown in Fig. 2 and are labeled for each camera, A through E. As the legend indicates, the different color patches of the target vary tangentially, while the illuminant results vary radially for each annulus, as labeled. Examining the differences within a given wedge indicates the variability in finished file color as a function of the five illuminants used. These wheels were created synthetically by extracting the finished file statistics from the captured chart images and simply filling the appropriate wheel ROI with the average value for the given patch-illuminant combination. Though rendered on a limited gamut hardcopy device, these wheels allow the reader to visually judge differences without the interaction of noise, artifacts, or sharpness.

The rendered color wheels of Fig. 2 correlate well with the values depicted in Fig. 1. Of the cameras tested, Camera A showed the lowest CII when evaluated across all 24 patches. This is consistent with performance reports in photographic trade journals as well as the color wheels. Camera A, in particular, showed half the CII of any competitor's camera. While Camera B also performed well in this regard, it would be difficult to tell by simply evaluating the neutral sample performance (upper-left corner of color wheels).

There was a noticeable color shift in the neutrals as a function of illuminant temperature for Camera B. This behavior was the reason for breaking out the neutrals alone in the CII calculations. It is emphasized here that the CII calculations for neutrals alone were performed relative to a zero-mean aim. Except for Cameras A and C, the other cameras did not handle the neutral CII for 2400 K and 3200 K illuminants very well.

Future Work and Conclusions

The rationale, formulae, and examples for an inter-illuminant imaging performance metric have been presented. This metric was designed primarily to quantify color/white balance performance from finished files of digital cameras and amounts to a one-way ANOVA across illuminants.

The execution of the CII as presented here is not meant to be flawless but rather an experimental basis from which to improve. With regard to illumination levels, number and quality of sample patches, and aims, more prudent choices may indeed be in order. Also, as one reviewer and other workers have pointed out[1] CIIs can be complex entities that may require evaluation in the context of chromatic adaption and color appearance.

References

1. Berns, R. Principles of Color Technology, Third Edition, 2000, pp. 129-130.
2. Hubel, Paul M., "Color Image Quality in Digital Cameras", *Proc. IS&T PICS* 1999, pp. 153-157.
3. See Appendix for alternate formulation.
4. Töpfer, K., Cookingham, R.,"The Quantitative Aspects of Color Rendering for Memory Colors", *Proc. IS&T PICS 2000*, pp. 94 – 98.
5. Töpfer, K., et al, "Preference in Image Quality Modeling", *Proc. IS&T PICS* 2002.

Biography

Don Williams received his M. Sc. Degree in Imaging Science from RIT, and works in Electronic Imaging Products, Research and Development at Kodak. His tasks focus on quantitative signal and noise performance metrics for digital capture imaging devices and imaging system simulations. He has been active in the development of imaging standards, and currently coleads the TC42/WG18 effort for print scanner resolution (ISO 16067-1), film scanner resolution (ISO 16067-2), and dynamic range measurement (ISO 21550) metrology. Mr. Williams is also a frequent contributor and advisor on digitization fidelity issues for the library and museum communities.

Appendix

There continues to be some confusion in the literature and in practice on the subtleties of pooling data for multiple difference estimates such as ΔE_{94}*. The approach taken for CII, in this paper, parallels past engineering practice by averaging the root-pooled variances as shown in Eq. 3.

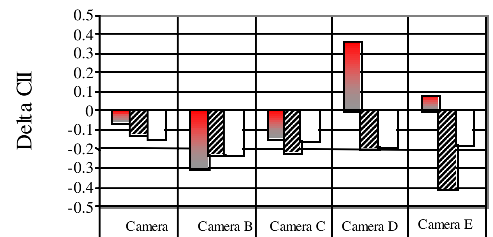
$$CII = \frac{1}{N}\sum_{ki=1}^{N}\sqrt{\left(\frac{\sigma_{L_k^*}}{2S_L}\right)^2 + \left(\frac{\sigma_{C_{ab_k}^*}}{2S_C}\right)^2 + \left(\frac{\Delta\overline{H}_{ab_k}^*}{S_H}\right)^2} \qquad (3)$$

An alternative formulation also frequently offered calculates the root averaged pooled variance. This latter approach is formulated according to Eq. 4.

$$CII \triangleq \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left[\left(\frac{\sigma_{L_k^*}}{2S_L}\right)^2 + \left(\frac{\sigma_{C_{ab_k}^*}}{2S_C}\right)^2 + \left(\frac{\Delta\overline{H}_{ab_k}^*}{S_H}\right)^2\right]} \qquad (4)$$

As illustrated in Fig. 3, the two approaches revealed small numerical differences, with a few exceptions. To a large extent these differences manifested themselves as a negative bias in Eq. 3.

Figure 3
(Eq. 3 - Eq. 4) CII values



	Camera	Camera B	Camera C	Camera D	Camera E
neutral & chromatic samples	-0.06	-0.3	-0.14	0.36	0.08
chromatic samples	-0.12	-0.23	-0.22	-0.2	-0.41
neutral samples (zero mean aim)	-0.14	-0.23	-0.15	-0.18	-0.17

neutral & chromatic samples; chromatic samples; neutral samples (zero mean aim)