EVALUATION OF COLOR DIFFERENCE FORMULAE FOR COLOR RENDERING METRICS

Davis, Wendy and Ohno, Yoshi
National Institute of Standards and Technology

ABSTRACT

Solid-state lighting is providing strong incentive to develop a new color rendering or color quality metric for sources of general illumination. Differences in perceived color between reflective samples illuminated by the test lamp and a reference source will likely be computed in an object color space, as is done in the Color Rendering Index (CRI). Using CIE 1976 L*a*b* color space, the results of using of ΔE_{00} and ΔE^*_{ab} to calculate color differences were examined when implemented in the Color Quality Scale (CQS), a proposed new color quality metric.

Multiple difficulties with implementing ΔE_{00} in a color rendering metric are discussed. In addition to being substantially more mathematically complex, it cannot parse apart the effects of hue and chroma shifts as simply as ΔE^*_{ab}. Further, it is only to be used when ΔE^*_{ab} is less than five, necessitating the use of both methods in typical color rendering applications. Longevity is also a concern, as ΔE_{00} was introduced only a few years ago.

The CQS scores of a number of traditional lamps, as well red-green-blue (RGB) and red-green-blue-yellow (RGBY) light-emitting diodes (LEDs), are compared when color differences were calculated with ΔE_{00} and ΔE^*_{ab} in CIELAB. Changes are minor and the relative scaling of lamps is well preserved.

Keywords: color rendering, color differences, color quality, solid-state lighting

1. INTRODUCTION

The CIE has been interested in updating or replacing the metric used to assess the color rendering properties of lamps for a number of years [1]. The current color rendering index (CRI) [2] has a number of problems, particularly when applied to solid-state lighting (SSL) [3]. One of the several problems with the CRI is that it relies on 1964 W*U*V* uniform color space to calculate sample color differences when illuminated by the reference and test sources. This color space is outdated, no longer recommended for use, and is very non-uniform, particularly in the red region. Instead, the CIE currently recommends the CIE 1976 L*a*b* (CIELAB) and the CIE 1976 L*u*v* (CIELUV) [4] for the determination of color differences. CIELAB is widely used in many applications, and so it seems an obvious choice for use in a new color rendering metric. The traditional method for calculating color differences within CIELAB, ΔE^*_{ab}, is simply the Euclidian distance between points within the space.

CIELAB is not without its own problems, however. A number of researchers have conducted vision experiments that show that perceived color differences are not uniform across CIELAB [e.g., 5,6]. Attempting to correct for the known non-uniformities, different ways of calculating color differences have been proposed, most recently with the introduction of ΔE_{00} [7].

There are several reasons to hesitate in implementing ΔE_{00} for the calculation of color differences in a new color rendering metric. The calculation of ΔE_{00} is considerably more complicated than ΔE^*_{ab}. Further, the use of ΔE_{00} is only recommended when ΔE^*_{ab} is less than five; for the range of color differences typical in color rendering, the use of both methods would then be required. The added complexity of the new calculation method goes beyond determining color differences. There is momentum to adopt a color quality metric, which does not strictly adhere to the definition of color rendering, but rather assesses overall color quality of objects under the lamp. In that case, not all color shifts would be treated equally. There is some evidence that increases in object chroma are preferred [8], whereas it is generally agreed that hue shifts and decreases in chroma are disfavored. One proposed new metric, the Color Quality Scale [9] uses a "Saturation Factor," so that lamps' scores are not penalized by increases in object chroma under the test source (relative to the reference source). This is implemented by calculating ΔE^*_{ab} as usual, and later subtracting the contribution of ΔE^*_{ab} that was caused by increases in object chroma under the test lamp. Since ΔE_{00} includes an interactive term between chroma and hue differences, the best way to implement an analogous Saturation Factor would need to be carefully considered.

Incorporating the use of ΔE_{00} into a new color rendering metric also raises longevity concerns. Care should be taken to minimize the likelihood any one component of the new color rendering metric will become obsolete well before the others. Some argue that ΔE_{00} is merely a milestone and will not likely become the final recommendation [10]. If ΔE_{00} is, however, one step closer to the final recommendation, one might conclude that it is still better to implement than not.

2. THE COLOR QUALITY SCALE (CQS)

The goal of these computations is to develop a sense of whether the implementation of ΔE_{00} into a new color rendering metric would result in meaningful changes in the outputs of the new metric. Since the authors are currently developing a new metric for assessing the color quality of light sources, the Color Quality Scale (CQS) [9], the effect of implementing ΔE_{00} within this metric was examined.

A brief explanation of the CQS is warranted here, though interested readers are directed to more detailed publications [9]. The CQS takes inspiration from CRI and uses the same basic procedure: the perceived color differences are calculated in a uniform object color space for a set of reflective samples illuminated by a test lamp and by a reference source. As with the CRI, the reference source is matched in correlated color temperature (CCT) to the test lamp and is either a blackbody (if CCT is below 5000K) or a phase of daylight illuminant (if CCT is equal to or above 5000K).

Several details differ from the CRI, however. The uniform object color space used in the CRI (CIE 1964 W*U*V*) is outdated, so the CQS uses CIELAB instead. The set of reflective samples is also different. The eight samples used in the calculation of R_a are all of a low to moderate chroma. Simulations have shown that some sources perform well with low chroma samples, but perform poorly with high chroma samples [3]. In order to detect those instances, the 15 Munsell samples used in the CQS are all of high chroma and distributed throughout the hue circle.

Based on the observation that increases in object chroma are generally desirable [8], the CQS does not penalize a lamp that causes such increases. Though color differences are calculated in a typical fashion (ΔE^*_{ab}), the amount of color difference that is attributed to increases in object chroma is excluded from the calculated color differences via the “Saturation Factor” of the CQS. In this way, the CQS deviates from the definition of color rendering and accounts for general observer preferences (though not fully) as they relate to light quality.

Though the CQS matches the CCT of the test and reference sources, it acknowledges that this assumption of complete chromatic adaptation breaks down at extreme CCTs. As a tentative solution, the “CCT Factor” scales the scores of lamps at various CCTs by the ratio of the gamut area of the 15 reflective samples when illuminated by the reference source, normalized by that at CCT=6500 K. A few CCTs actually have a larger gamut area than 6500K, but those scaling factors are truncated to one.

Because some sources render only one or two samples very poorly, the color differences in the CQS are combined with root-mean-square (RMS) rather than simple averaging as is done in the CRI. This accounts appropriately for the effect of large individual color differences.

To eliminate the confusion of negative scores that the CRI occasionally produces, the CQS is converted to a 0-100 scale. This conversion does not effect scores greater than 30. Finally, a new scaling factor was implemented in order to maintain consistency with the CRI (in light of the other differences between the metrics) for traditional lamps. The scaling factor is determined so that the average score of the CQS for the CIE standard fluorescent lamp spectra (F1 through F12 [4]) is equal to the average score of the current CRI R_a (=75.1) for these sources.

Other computational changes are being integrated into the CQS. For instance, CMCCAT (as discussed in [11]) has been tested to replace the von Kries chromatic adaptation correction. But, only those computations completed and described above are used in the calculations presented here.

3. IMPLEMENTING ΔE_{00} IN THE CQS

A few issues must be considered before replacing the use of ΔE^*_{ab} with ΔE_{00} in the CQS. As mentioned earlier, ΔE_{00} is only to be used when the color difference calculated by ΔE^*_{ab} is less than five. Given the wide range of color differences observed in color rendering calculations, ΔE^*_{ab} must be calculated for every color difference first. If it is greater than or equal to five, it should be used as the color difference. If, however, ΔE^*_{ab} is found to be less than five, then ΔE_{00} must be calculated and used as the color difference.

The Saturation Factor poses another complication. As discussed earlier, this is a simple calculation when ΔE^*_{ab} is used. However, ΔE_{00} includes an interaction between chroma and hue, and it is not obvious how one would parse the influence of chroma out of the total color difference. Researchers from the University of Leeds have developed a computational method for calculating ΔE_{00} involving only three final terms: ΔL_{00}, ΔC_{00}, and ΔH_{00}, corresponding to the differences in lightness, chroma, and hue respectively [12]. These terms are combined simply as the square root of the sum of each term squared. A spreadsheet of these calculations was graciously provided to the authors by Guihua Cui [13]. With this method, ΔE_{00} is calculated and, if there is an increase of an object chroma (compared to that of reference source), the ΔC_{00} term is set to zero during the calculation of ΔE_{00}.

This implementation also requires adjustment of the scaling factor. Since some lamps' scores are changed with the implementation of ΔE_{00}, the scaling factor is modified to maintain consistency with the CRI (and the CQS when ΔE^*_{ab} is used) for the set of standard fluorescent lamps.

The scores from two groups of test sources will be presented here: traditional lamps, including fluorescent and filament lamps, and LED models, including red-green-blue (RGB) and red-green-blue-yellow (RGBY) chip combinations. The scores for the first group, traditional lamps, are shown in Figure 1. The top panel shows three types of fluorescent lamps: cool white (4290K), daylight (6500K), and tri-phosphor (3380K). The bottom panel shows the results for metal halide (4280K), super high-pressure sodium (2530K), incandescent (2810K), and neodymium (2760K) lamps.

There is little difference between CRI and CQS scores for these discharge lamps. This is to be expected, since the CQS is scaled to be congruent with the CRI for fluorescent lamps. Further, the effects of implementing ΔE_{00} are minimal, which is also expected since it is similarly scaled.

Figure 2 shows results for LED models. The top panel shows scores for four RGB LEDs. The peak wavelengths for these white lights are: 1) 457, 540, and 605 nm 2) 473, 545, and 616 nm 3) 465, 546, and 614 nm 4) 455, 547, and 623 nm. The bottom panel shows results for four RGBY LED models. The peak wavelengths for these sources are: 1) 461, 526, 576, and 624 nm 2) 447, 512, 573, and 627 nm 3) 445, 495, 555, and 617 nm 4) 450, 525, 580, and 625 nm.

The differences between CRI and CQS scores for RGB LEDs are important and the intentional consequences of differences between these metrics. The implementation of ΔE_{00} makes little difference in the CQS scores for RGB LEDs and the relative scaling of these sources is maintained.

The variation between CQS scores when calculated with ΔE_{00} and ΔE^*_{ab} is most pronounced for RGBY LEDs (due to smaller color differences caused by these sources, for which ΔE_{00} is implemented), but score differences are less than three points.

4. CONCLUDING REMARKS

Though only a small sampling of data is presented here, the effect of implementing ΔE_{00} in the calculation of color differences does not appear to make a substantial difference in CQS scores for LEDs or

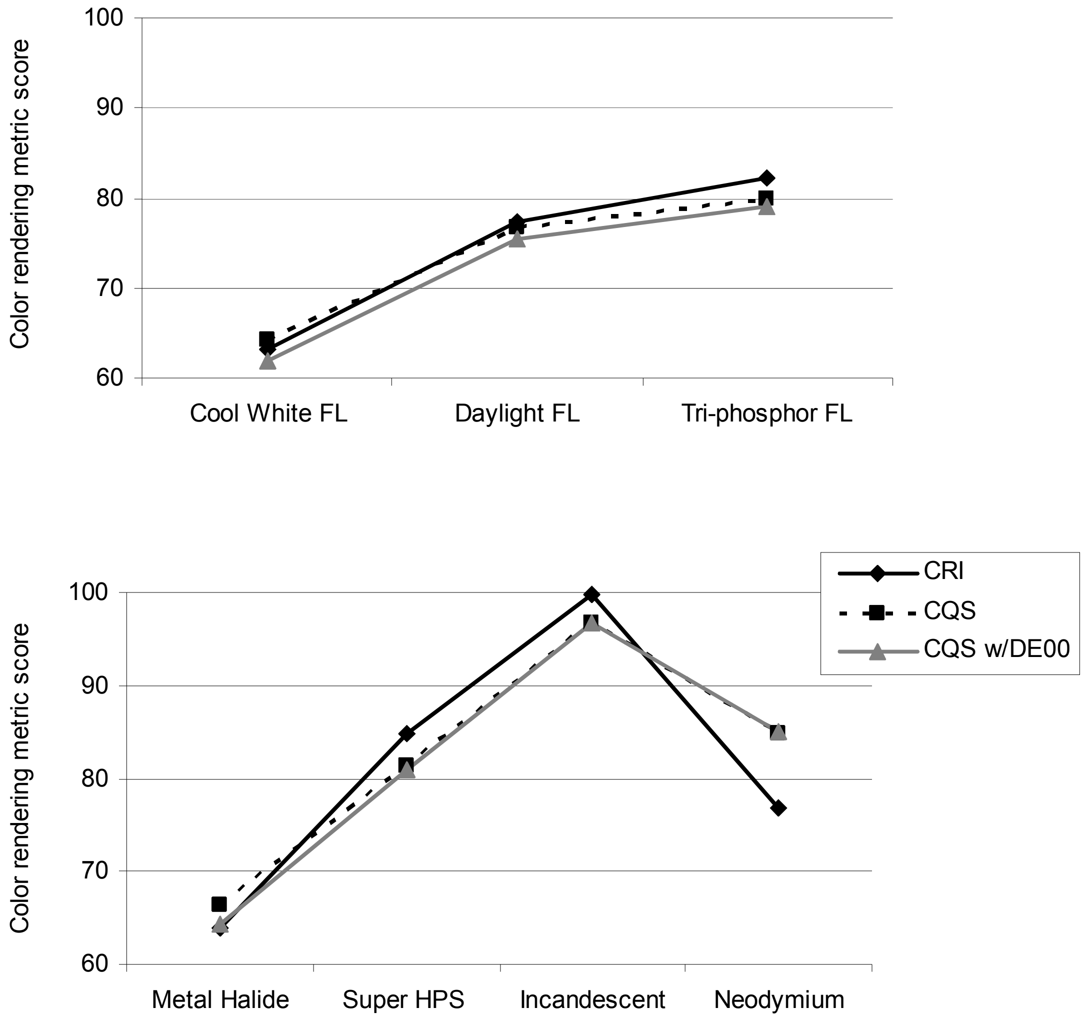

Figure 1. The CRI and CQS scores for three fluorescent lamp types (top panel) and four additional traditional lamps (bottom panel). The black diamonds show the CRI scores; black squares show the CQS scores (ΔE^*_{ab} used to calculate color differences), and the grey triangles show the CQS scores when ΔE_{00} is implemented.

traditional lamps. Very close correspondence is found between CQS scores when color differences are calculated with either ΔE_{00} and ΔE^*_{ab}. Clearly this analysis will need to be expanded to include many more lamps. Major differences between scores when the different methods for calculating color differences are used should be examined for meaningfulness when the new metric is undergoing testing and validation by visual experiments. Such experiments are planned to be carried out at NIST shortly.

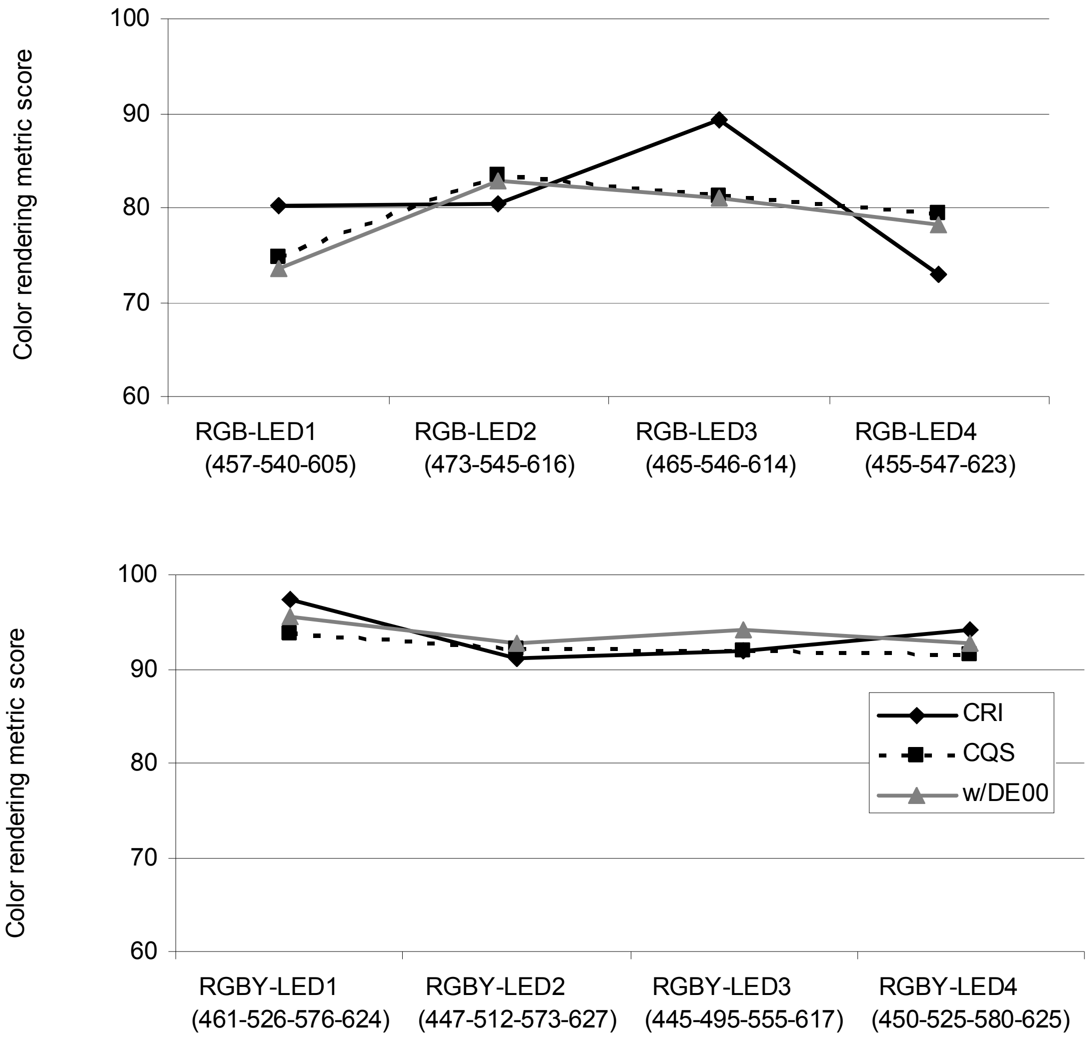

Figure 2. The CRI and CQS scores for four RGB LED models (top panel) and four RGBY LED models (bottom panel). The black diamonds show the CRI scores; black squares show the CQS scores (ΔE^*_{ab} used to calculate color differences), and the grey triangles show the CQS scores when ΔE_{00} is implemented. The peak wavelengths of the LEDs are indicated in parentheses.

REFERENCES

1. CIE 135/3-1999. *Colour rendering, closing remarks.*

2. CIE 13.3-1995. *Method of Measuring and Specifying Colour Rendering Properties of Light Sources.*

3. OHNO, Y. Spectral Design Considerations for Color Rendering of White LED Light Sources, *Opt. Eng.* **44**, 111302, 2005.

4. CIE 15-2004. *Colorimetry, 3rd ed.*

5. LUO, M.R., RIGG, B. Chromaticity-discrimination ellipses for surface colours. *Color Res Appl*. **11**, 25-42, 1986.

6. BERNS, R.S., ALMAN, D.H., RENIFF, L., SNYDER, G.D., BALONON-ROSEN, M.R. Visual determination of suprathreshold color-difference tolerances using probit analysis. *Color Res Appl*. **16**, 297-316, 1991.

7. CIE 142-2001. *Improvement to Industrial Colour-Difference Evaluation.*

8. HASHIMOTO, K, NAYATANI Y. Visual Clarity and Feeling of Contrast. *Color Res Appl*. **19**, 171-185, 1994.

9. DAVIS, W, OHNO, Y. Toward an Improved Color Rendering Metric. *Proc. SPIE*. **5941**, 59411G, 2005.

10. KUEHNI, R.G. CIEDE2000, Milestone or final answer? *Color Res Appl*. **27**, 126-127, 2002.

11. CIE 160-2004. *A Review of Chromatic Adaptation Transforms.*

12. NOBBS, JH. A lightness, chroma, and hue splitting approach to DEIDE2000 colour differences. *Adv. Col. Sci. Technol*. **5**, 46-53, 2002.

13. CUI, G. CIEDE2000 Excel calculation spreadsheet: Three terms, private communication, 2006.

ACKNOWLEDGEMENTS

The authors are grateful to Guihua Cui and Ronnier Luo for their generous assistance in implementing the three-term method of calculating ΔE_{00}.

Authors:

Wendy Davis
National Institute of Standards and Technology
100 Bureau Dr., MS 8442
Gaithersburg, MD 20899-8442
Phone: 301-975-6963
Fax: 301-840-8551
e-mail: wendy.davis@nist.gov

Yoshi Ohno
National Institute of Standards and Technology
100 Bureau Dr., MS 8442
Gaithersburg, MD 20899-8442
Phone: 301-975-2321
Fax: 301-840-8551
e-mail: ohno@nist.gov