Towards the definition of new visual color quality representations

Josep Carreras[1], Jesús Quintero[1] and Charles E. Hunt[1,2]

IREC, Institute for Energy Research of Catalunya, Barcelona, Spain

CLTC, California Lighting Technology Center, University of California, Davis CA, USA

1. - Introduction

Luminous efficacy of radiation (LER) is one of the multiplicative factors affecting the wall-plug efficiency of a light source. In order to maximize it, it is important to keep the spectral content of the lamp in a small region around 555 nm. However, doing so limits the color span that the source can correctly reproduce. Obviously, whether one determines to prioritize energy efficiency over color reproduction depends ultimately upon the particular application under consideration. Thus, it is important both to understand theoretically the functional relationship between these inversely correlated variables (design) and at the same time being able to control and tune it in real lighting devices (implementation). The simple computational example shown in Figure 1 serves as an illustration of the expected functional dependence, where the efficacy and general CRI (Ra) are plotted for a family of Gaussian spectra with standard deviations ranging from 10 to 200 nm.

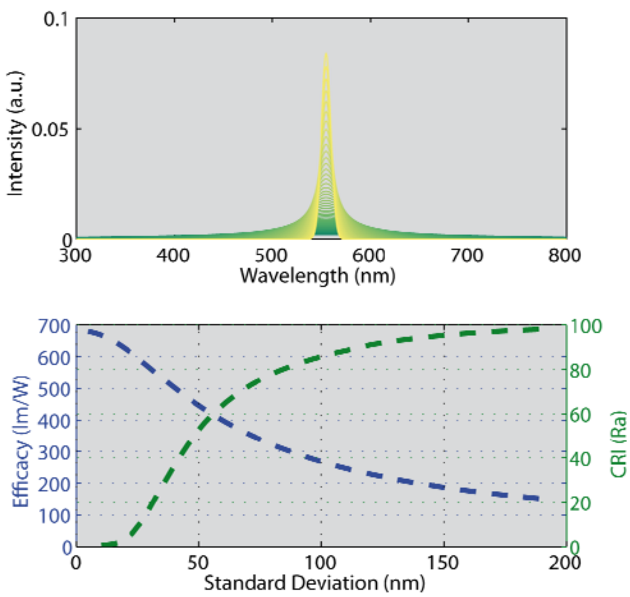


Figure 1.- Trade-off relationship between CRI (Ra) and efficacy

Although there is a simple and well-determined procedure for measuring the luminous efficacy of a light source, at present time there is still controversy and lack of agreement in the elements that should be taken into account in the definition of a new scale for color assessment. The CIE Color Rendering Index

(CRI) [1] is widely used and is the only internationally accepted metric for assessing the color rendering performance of light sources.

However, the CRI has a number of shortcomings and problems[2]: (i) uniform color space used to calculate color differences is outdated and no longer recommended for use, (ii) the Von Kries chromatic adaptation transform used by the CRI is also considered obsolete, (iii) the CRI method specifies that the CCT of the reference illuminant be matched to that of the test source, which assumes complete chromatic adaptation to any light source CCT (this assumption fails at extreme CCTs), and (iv) none of the eight reflective samples used in the computation of R_a is highly saturated.

The problems with the CRI become more pronounced in narrowband solid-state lighting solutions. However, since no agreement has been found among lighting experts on the definition of the right index to treat such sources, the whole industry behind continues to try to optimize production parameters to maximize CRI, the only metric they are knowledgeable of and able to trust. This implies that, assuming CRI is not telling the whole story, a wrong color quality assessment is at risk of inhibiting market acceptance of LEDs or solid-state luminaires.

2. - Extended visual representation of a color rendering metric

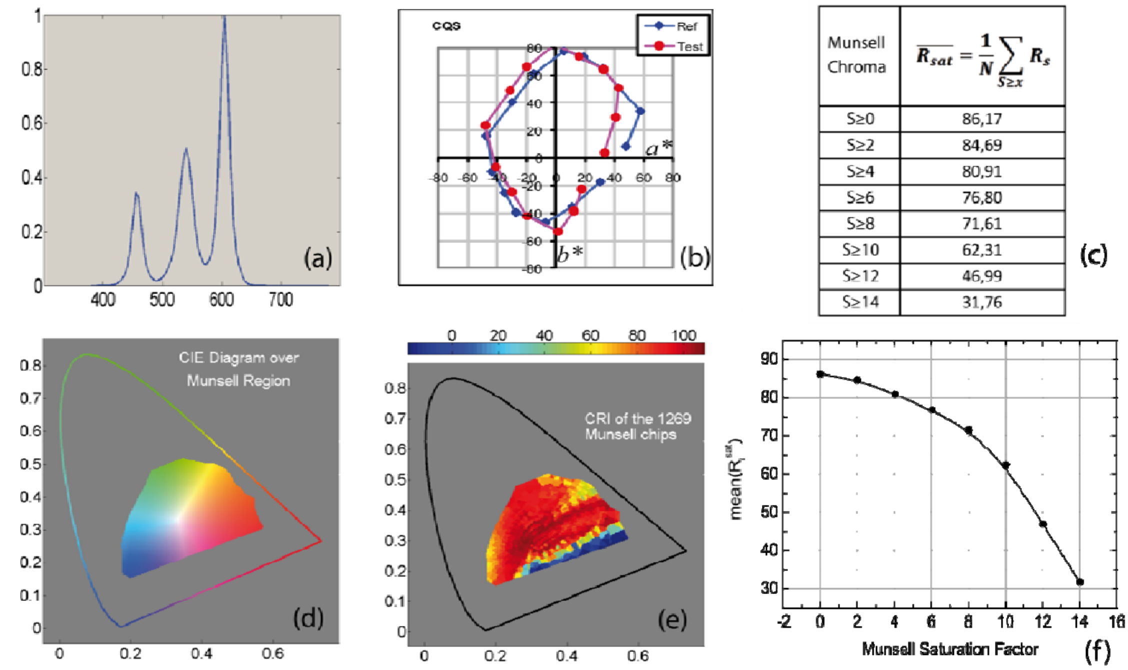
There are more profound problems in the definition of the CRI and other metrics that base their output in a single number. The purpose of a color quality metric is to condense information into something manageable, and that users having very limited knowledge of color science can read the right message from that particular index. However, in this process where a single number is obtained to represent all the subtleties from the spectrum, the chances that users get poor or wrong knowledge of whether a light source is suitable or not for a particular implementation requiring to render well certain colors, are really high.

Current definition of CRI (based on color fidelity to blackbody or to a daylight phase), as well as the improved CQS (an enhanced version with additional elements correcting for color preference), utilizes only a reduced number (15) of representative reflective samples, so their output are 15 scores, one for each of the reference samples. Since this information is still considered too much to handle by the general user, an averaging over this set is performed to get the general version of the index. The 15 reference samples are Munsell chips, a database composed of 1269 reflectance spectra. This huge amount of reflective spectra could be used to create a dense set of scores in a color space, which could be plotted in a more convenient graphical representation over a large gamut area, suitable for fast visual inspection over the complete Munsell set, which would open up the possibility to identify important color rendering weaknesses for certain colors that might compromise the application under consideration.

The determination of using a single index based on a single number like CRI or CQS is adopted for the sake of simplicity, preventing the general user from having to process excessive information. However, the visual system can process a huge amount of information per time unit (of the order of Gb/s), being able to filter superfluous information and to find relevant details in a fraction of a second. Thus, it would be desirable to have a color rendering metric that, not only does not require any expertise in color science, but also is based on a well-designed visual score mapping itself.

As an example, consider the RGB LED spectrum already discussed in reference [2], shown here in Figure 2(a). The spectral content of this particular RGB LED cannot render saturated reds and purples at all. Despite the low appealing that such a light source may have for some applications (i.e. fruit or meat illumination), the general CRI for this source is 80, which would be considered as a good light source for the general user. The CQS index does a much better job rating this LED source, giving a score of 73, partly due to the highly saturated reference Munsell chips used and to their combination through a RMS (root mean square) to get the final CQS value.

However, none of the two scales tells anything about the reason why they are rating a light spectrum in a particular manner. CQS is the only scale this far that could be considered as backward compatible with CRI but able to penalize not properly rendered saturated objects (not penalizing over saturated ones on behalf of color preference). But again, from a scalar number is not possible to infer which particular colors are pushing the score down the table. For this reason, a visual mapping of the chosen rating scale into a human-readable color space is required and should arrive to the final user.



Munsell Chroma	$\overline{R_{sat}} = \frac{1}{N}\sum_{S\geq x} R_s$
S≥0	86,17
S≥2	84,69
S≥4	80,91
S≥6	76,80
S≥8	71,61
S≥10	62,31
S≥12	46,99
S≥14	31,76

Figure 2.- (a) RGB 3-LED spectrum with good R_a value. (b) Color differences in an *a*b plot between the LED in (a) and the reference samples used in the Color Quality Scale (CQS). (c) Table showing the results of a CRI-like metric using a complete Munsell set and averaged over the saturation indexes. (d) CIE diagram. (e) Visual representation of CRI over the dense Munsell set. (f) Plot of the data in table (c).**

As an example, Figure 2(b) shows the coordinates of the reference and test spectra in a^*b^* space. The inexperienced user may find difficult to interpret this plot and may wonder (i) how the test spectrum renders other colors different than those chosen as references, (ii) how color differences in an a^*b^*

space translate into final scores, and even (iii) how can be known the color that a particular point in the

a^*b^* space is representing.

For the sake of illustration, let us assume that CRI represents a good indicator of color fidelity and that the most known color space to the general user is the original CIE 1931 [3]. Then, by analyzing a test spectrum against the 1269 Munsell reflective samples and following the CRI procedure for each of them, a dense set of scores (R_a in this example) can be generated over a large area of the CIE diagram around the CIE white point, as illustrated in Figure 2(e). The mesh created by these scores are evaluations over a set of reflective samples that represent a subset of the most significant saturation levels and hues, as many as defined in the Munsell color system. Within the framework of the proposed visual representation, the identification of the color that a particular score is representing, is straightforwardly found by comparison with the standard CIE plot in Figure 2(d). Thus, by direct inspection one would conclude that this LED disastrously (negative CRIs) reproduce all the region along the line of purples, saturated reds, and fails to correctly reproduce saturated blue and yellow/orange (CRI of 40-60). Since there is not a clear picture of what a negative CRI means, all CRI values have been transformed through a suitable expression[2] to lie between [0-100], not affecting R_a greater than 20.

Although this intuitive scheme should contains enough information for the most demanding user, this representation could be further condensed by collapsing the hue information, and averaging over the saturation index *S* of the Munsell notation. This would enable a set of eight indexes $R_{S}\geq x$, splitting the extended CRI in 8 levels of saturation, as shown in the table of Figure 2(c). Every particular application may choose up to what extent is desirable to optimize the full set of saturation indexes, but the majority of applications would probably require setting up a cut off value, i.e. $R_{S}\geq 6$. Finally, let us revisit the RGB LED spectrum of Figure 2(a). This spectrum gives an R_a of 80, nearly the same value as $R_{S}\geq 4$, which is a clear indication that R_a only deals with slightly saturated samples. However, in the plot of Figure 2(f), it can be seen how rapidly the color rendering of this spectrum decreases with saturation levels, making it unsuitable for certain applications where saturated objects or color preference schemes are required.

3. - Conclusions

A new graphical approach for color quality assessment that does not relies in a scalar number as a final score has been presented. This representation contains information over a large set of Munsell reflective spectra, not pretending to condense information into an irreversible manner, but instead utilizes a colorful representation that can be effectively and visually processed. Our visual approach is compatible with the most accepted indexes like CRI and CQS, and extends its validity to objects with all levels of saturation and hues.

It should be noted that the election of CRI as a quality scale as well as the CIE 1931 color diagram as the space in which to draw this representation is somewhat arbitrary, and similar results could have been

obtained by utilizing the CQS and the a^*b^* coordinate system. This implies that our visual index would minimize the difference between CRI and CQS, and possibly among other forthcoming metrics, mainly because it eliminates the need for an averaging to obtain a single number and considers separately the whole range of levels of saturation/hues in its representation.

References

[1]. CIE (1995), *Method of Measuring and Specifying Colour Rendering Properties of Light Sources*, Publication 13.3, Vienna: Commission Internationale de l'Eclairage, ISBN 978-3900734572

[2]. Y. Wendy, Y. Ohno, *Color Quality Scale*, Optical Engineering, **49(**3), 033602 (2010)

[3]. CIE (1932), Commission Internationale de l'Eclairage proceedings, 1931. Cambridge University Press, Cambridge.