Out of the Wood

BY **MIKE WOOD**

Color rendering one more time. Could TM-30-15 be it?

I KNOW I'VE WRITTEN MANY COLUMNS ALREADY on the topic of color rendering but it's so important and generally misunderstood that I make no apologies for coming back to it again.

In the Winter and Spring 2010 issues of *Protocol*, I discussed CRI (Color Rendering Index) and its problems and a possible solution CQS (Color Quality Standard). If you recall, CRI was originally developed as a way for manufacturers to characterize white light from fluorescent lamps and is very limited. CRI behaves poorly with discontinuous spectra, such as those from RGB LED sources, and, because it uses a very limited set of color test samples, is easily optimized (or "gamed" depending on your point of view) to make a light source score better than it really is. Seemingly similar RGB LED-based luminaires can have CRIs that vary from a poor 40 up to a very respectable 80, but which, if any, of these values truly represents the color rendering of the light output?

> **"Can we assume that a good CRI means the source will get a good R_f, and vice-versa? The answer is an emphatic, no!"**

CRI wasn't designed as a measure for white light produced in the way that an RGB LED combination does it, and when CRI is applied to LEDs it can produce misleading results. For example, because the eight indices for the individual test colors are averaged together to produce the final CRI, a light source can score well even though it renders one or two colors poorly. With the large gap in the yellow wavelengths when using RGB LEDs, a luminaire can do a bad job of rendering yellows but still get a respectable CRI. Additionally, because the eight standard sample colors are all of fairly low saturation, the CRI tells you nothing about how a light source will perform when rendering deeply saturated colors. The widely separated and narrow peaks of the spectra from RGB LEDs can perform poorly when rendering saturated colors outside those peaks, but the current CRI definition doesn't pick up that deficiency. It also uses color science for its math that is long outdated.

In recognition of these and other problems, NIST (the National Institute of Standards and Technology) proposed a new metric called the Color Quality Scale (CQS). The goal was to keep the good points of CRI with its use of standard color chips and direct relation to the real-world, while addressing the short comings arising from the choice of standard colors and the math used to combine the results. CQS tried to address the major issues with CRI by using more color samples, 15 instead of eight, and choosing those as saturated colors evenly spaced across the entire visible spectrum.

This proposed metric met with a mixed reception. To some users, those making heavy use of RGB such as the entertainment lighting industry, it offered a huge improvement over CRI. In fact the PLASA Technical Standards Program recommended the use of CQS in its Photometric Standards as in, for example, *ANSI E1.41 – 2012, Recommendations for Measuring and Reporting Photometric Performance Data for Entertainment Luminaires Utilizing Solid State Light Sources*. However, to users who were interested in white light only, in particular phosphor converted white LEDs, the advantages were less clear.

Although CQS offered some advantages to CRI, it still didn't address the limitations of providing a single number as the metric. Yes, we know that a light with a CRI (or CQS) of 70 is relatively poor in its color rendering. However, we don't know from that single number where it is poor, is it in the reds, or the blues? Nor do we know whether it errs by under-saturating or over-saturating colors. A light source that over-saturated red could have the same CRI or CQS as one that under-saturated blue. They might both look acceptable on their own, but look horrible if used together.

The CIE (International Commission on Illumination) has been working on replacements for CRI for some time, it's gone through two or three committees, but so far they have failed to reach consensus and we have no official alternative.

I also have to say that some of the large light source manufacturers (You know who you are!) have been quite happy to allow the old and decrepit CRI to continue as the only official metric. It makes their lamps look good, and it's easy to make it say what you want. What's not to like?

To be effective, the push for a better metric has to come from the users and lighting practitioners as they are the ones this ultimately affects. To that end the IES (Illuminating Engineering Society) in North America charged its Color Committee with suggesting a solution and appointed a Color Rendering Task Group to develop and present a new metric, or set of metrics, to meet the disparate needs and address the issues of CRI and CQS. (*Disclosure: I am a member of the IES Color Committee and so have had knowledge of the progress of the proposals, but I was not a member of the Task Group developing them.*)

That proposal was published in August 2015 by the IES as a technical memorandum, TM-30-15, *IES Method for Evaluating Light Source Color Rendition*. TM-30 draws from a wide array of color perception research that has gone on in the last 30 years and seeks to answer the concerns with CRI while still maintaining a reference that is easily used and identifiable.

Some parts of TM-30 remain very familiar. It still uses color samples and compares the rendering of them with an ideal light source. However, instead of eight or 15 color samples it uses 99 samples that are uniformly spread across the visible color space and spectrum. This large number of samples means that optimizing or gaming the metric is much more difficult, if not impossible. TM-30 provides two main results rather than one. First a Fidelity Index, R_f, which is conceptually similar to the single value R_a provided by CRI. R_f is a number from 0 to 100 that indicates the fidelity with which a test light source renders colors to the human eye as compared with a reference white light source. The math is better than CRI, the number of samples is hugely increased, and the average value is more justifiable. (Although both R_a and R_f have similar scales, they are different and should not be directly compared or held to the same goals.)

The second metric provided by TM-30 is a Gamut Index, R_g. The Gamut Index provides a measure of the color gamut that the test light source provides relative to the test source. In other words, it gives you a metric for whether incorrectly rendered colors will be over-saturated or under-saturated. In all cases, an R_g of 100 shows that the source will render colors with the same average saturation level as the reference source, more than 100 indicates over-saturation, while an R_g less than 100 shows under-saturation. The range of R_g will depend on the Fidelity Index of the source. For example, a source with an R_f of 60 has a possible range for R_g of approximately 60 to 140. Depending on your usage, R_g values higher than 100 may or may not be desirable. To my mind, over saturating colors is often just as poor as under saturating them and can give objects a cartoonish appearance.

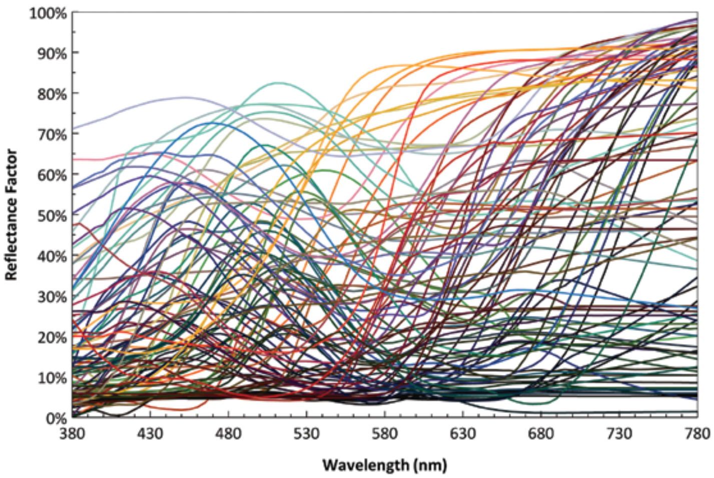

Figure 1 – 99 color samples of TM-30

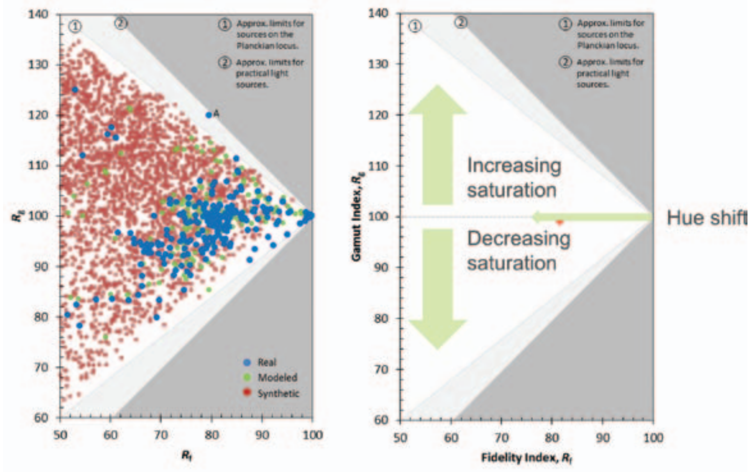

Figure 2 – Plot of R_f versus R_g

> . . . over saturating colors is often just as poor as under saturating them and can give objects a cartoonish appearance.

Figure 2 shows the range of possible values and thus the relationship between R_f and R_g by plotting real sources against both of them. Note that, the closer a light source gets to having a Fidelity Index, R_f, of 100, then the lower the range of Gamut Index, R_g, that is possible. A perfect light source is a perfect light source in both metrics. Conversely, the lower the Fidelity Index, R_f, then the greater the possible range for the Gamut Index. The lower the color fidelity of the light source then the more it can under- or over-saturate colors. Real world light sources tend to be clustered in an area with R_g less than 100, i.e. where colors are under-saturated, but there are many exceptions. If we take a closer look at a few real light sources we can see better what's going on.

Figure 3 shows a few familiar light sources plotted on a Gamut Index / Fidelity Index graph. Halogen is at the extreme right at the 100:100 point as our perfect source. Some possible RGBA sources are shown as the dark blue markers. They appear with a range of both R_f and R_g values. The RGBA source with an R_f of 90 might have excellent color fidelity, but notice how it over-saturates colors with

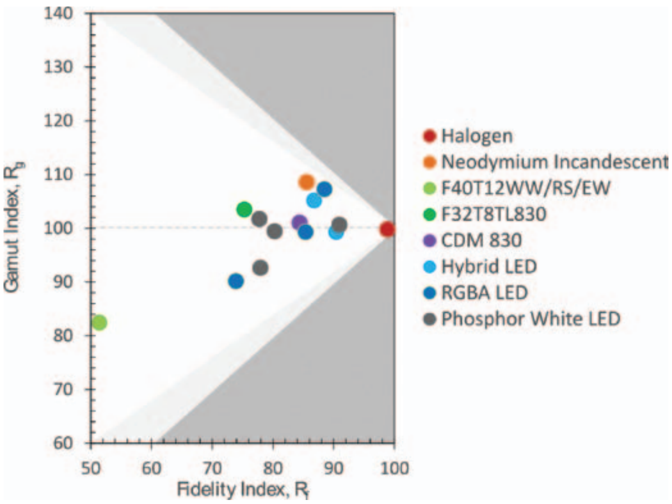

Figure 3 – Real light sources

an R_g approaching 110. That might be acceptable, but perhaps not. For many uses the RGBA with a slightly lower R_f of 85 but that has an R_g of 100 might actually be preferable. In an ideal world, perhaps sources used for entertainment lighting would be the 100:100 of an incandescent but, in reality, no real light source is, and it's up to us as users and designers to choose light sources that not only have acceptable color fidelity but also a reasonable gamut index. I suggest that, for theatrical lighting, most of the time we don't want to over-saturate colors. We prefer them to be natural. However, for a musical, a theme park, or a rock and roll show that may not be the case and oversaturation could be just fine. TM-30 gives us the information to make that judgment in a way CRI never could.

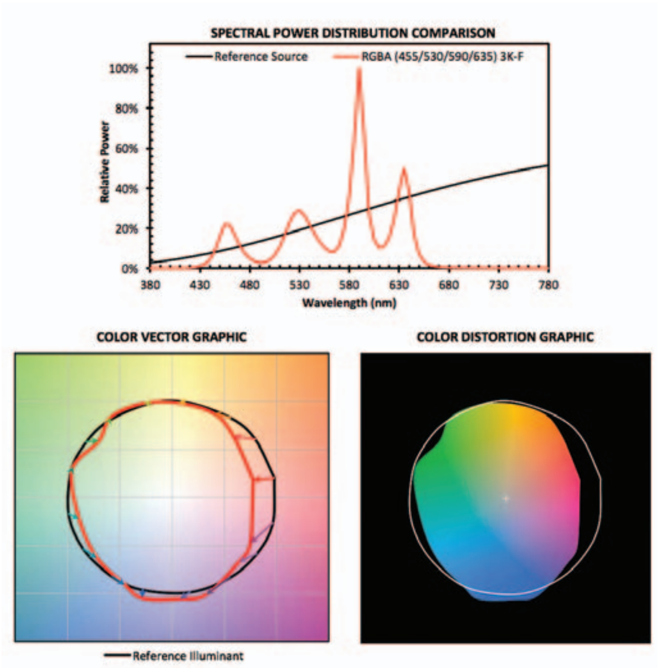

Figure 4 – Color vector graphics

To help the user understand what the two values R_f and R_g mean, TM-30 provides a couple of icon graphics that present the information on where and by how much color rendition is poor. For example, **Figure 4** shows the Color Vector and Color Distortion Graphics for an RGBA LED source.

This source has an R_f of 74 and R_g of 90. So we know it has some limitations in color fidelity, but the numbers alone don't tell us where. This is where the graphics really help, the icon on the left shows us where the light source under saturates, and where it over saturates. It also shows us, via the short vector arrows, how colors are distorted. You can see that greens tend to get pushed towards yellow for example. The second graphic only shows the under and over saturation, but in an icon that's clearer to print and, in my opinion, quicker to understand. You can instantly see that the maximum color distortion happens between orange, red, and magenta, while blues and yellows are not affected as much.

The graphics really become useful when you are looking at more than one light source and trying to figure out if they will match with each other. **Figure 5** shows two light sources which have identical R_f and R_g values, 82 and 95 respectively. From the numbers alone we might assume they were identical, but the spectra and graphic show that isn't the case.

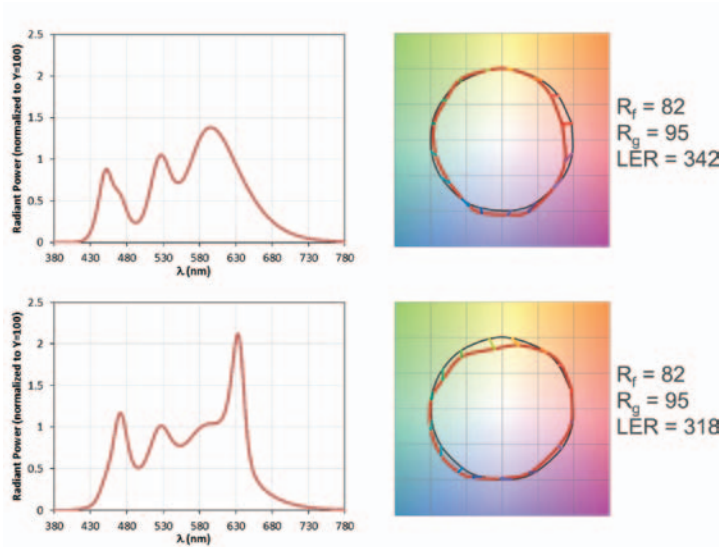

Figure 5 – Color vector graphic comparison

The top light source under-saturates reds and greens, while the lower one is almost perfect in those colors, and instead under-saturates yellows. This gives you a strong hint that, although these two sources may match very well on a white or pastel background, they could look very different when illuminating a saturated red object. Indeed, that large spike in the red at 630 nm on the lower source confirms that could be a problem.

There's a lot more to be learnt and discussed about the use of R_f

and R_g. Experience needs to be gained in what works for you and your clients and what doesn't, and I know I will be returning to this topic again. For now, I strongly encourage you to use TM-30's R_f and R_g and to ask your suppliers for these metrics as well as, or instead of, CRI on their light sources. It's easy enough to calculate, the IES provides Excel spreadsheets which do all the math for you and produce all the results and graphics I've shown here, plus more. All you need is a measured spectrum of the light source under test.

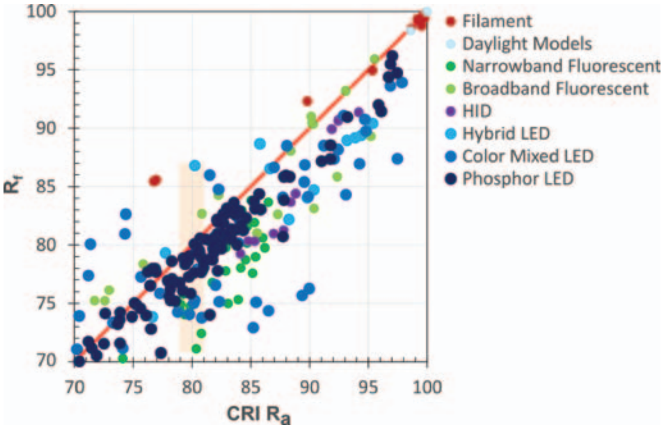

Figure 6 – Fidelity Index (R_f) versus CRI (R_a)

Let me finish this time with a chart (**Figure 6**) showing how the CRI (R_a) and R_f of real light sources compare. Can we assume that a good CRI means the source will get a good R_f, and vice-versa? The answer is an emphatic, no! The truth is all over the map, on average most light sources have a lower R_f value than CRI (i.e. that are below the red line on the chart), but there are exceptions. Within any CRI band we have a huge range of possible R_f values. For example, take a look at the shaded vertical band at CRI=80. Light sources that were rated with a CRI of 80 have R_f values that range from 71 all the way up to 87, almost 16 points of spread! Most sources are worse under R_f, but a couple of LEDs come out better. It should come as no surprise that the sources that come out the worst with the new metric are the very same narrow band fluorescent tubes that CRI was designed to make look good!

There's more about TM-30 I want to share so, to be continued . . .

Thanks to Michael Royer, Pacific Northwest National Laboratory, for use of some of the figures. ■

Mike Wood runs Mike Wood Consulting LLC, which provides consulting support to companies within the entertainment industry on product design, technology strategy, R&D, standards, and Intellectual Property. A 35-year veteran of the entertainment technology industry, Mike is the Immediate Past Chair of the PLASA Governing Body and Co-Chair of the Technical Standards Council. Mike can be reached at mike@mikewoodconsulting.com.