Taylor & Francis
Taylor & Francis Group

LEUKOS

The journal of the Illuminating Engineering Society of North America

ISSN: 1550-2724 (Print) 1550-2716 (Online) Journal homepage: http://www.tandfonline.com/loi/ulks20

Why Color Space Uniformity and Sample Set Spectral Uniformity Are Essential for Color Rendering Measures

Kevin A. G. Smet, Aurelien David & Lorne Whitehead

LEUKOS, 00:1–12, 2015

ISSN: 1550-2724 print / 1550-2716 online
DOI: 10.1080/15502724.2015.1091356

Taylor & Francis
Taylor & Francis Group

Why Color Space Uniformity and Sample Set Spectral Uniformity Are Essential for Color Rendering Measures

Kevin A. G. Smet[1],
Aurelien David[2],
and Lorne Whitehead[3]
[1]KU Leuven, ESAT/Light & Lighting Laboratory, Ghent, Belgium
[2]Soraa Inc., Fremont, California, USA
[3]University of British Columbia, Department of Physics and Astronomy, Vancouver, British Columbia, Canada

ABSTRACT Recently, a group including the present authors developed a new color rendering (fidelity) measure, approved by the IES and henceforth referred to as IES R_f, that has two major updates with respect to the general color rendering index R_a of the CIE. First, it proposes an update to the more perceptually uniform CAM02-UCS color space. Secondly, instead of using only a small number (eight) of moderately saturated samples to determine a general color fidelity index, the IES method proposes using a set of 99 samples uniformly distributed in color space. In addition to the latter, the sample set has one other important property: spectral uniformity. This ensures that the sample set is not wavelength biased: all wavelengths contribute equivalently to the general color fidelity score. This article explores the importance of these two updates for color fidelity evaluation. It shows how the color space update results in a substantial spread of the new color fidelity scores relative to the old CIE R_a values and how the sample set update results in an overall reduction of color fidelity scores for most light source spectra with high CIE R_a ($\geq$80) and high luminous efficacy of radiation (LER) values, especially those with narrowband or spiked spectral features. It is shown that calculated color fidelity scores are affected by the degree of sample set spectral uniformity and that, in the absence of sample set spectral uniformity, light source spectra can be tuned to yield anomalously high CIE R_a values without necessarily yielding an increase in actual perceived color fidelity.

KEYWORDS color fidelity, color rendering, color space uniformity, spectral uniformity, sample set, CIE R_a, IES R_f

Received 11 August 2015; revised 3 September 2015; accepted 3 September 2015.

Address correspondence to Kevin A. G. Smet, ESAT/Light & Lighting Laboratory, KU Leuven, Gebroeders Desmetstraat 1, Ghent B9000, Belgium. E-mail: Kevin.Smet@kuleuven.be

1. INTRODUCTION

Color rendering, defined by the CIE [2011] as the "effect of an illuminant on the colour appearance of objects by conscious or subconscious comparison with their colour appearance under a reference illuminant," has been assessed by the CIE color rendering index (CRI) [CIE 1995] for decades. Since its standardization in 1974, the CIE CRI has remained virtually unchanged, despite repeated attempts, starting in the early 1980s, by several CIE technical committees to address some problems that had become apparent early on. One of these problems was that the CIE CRI scores tend to be unrepresentative of the visual color rendering of certain types of fluorescent sources [CIE 1980–1982]. Over the years, the body of literature and anecdotal evidence of the failure of the CIE CRI for spiked or narrowband sources, such as light emitting diodes (LEDs), has continued to grow [Bodrogi and others 2004; CIE 2007; Sándor and Schanda 2006; Smet and others 2011; Szabó and others 2007]. Especially with the advent of solid state lighting, the matter has become pressing, because the inability to correctly predict the color rendering of lamps based on narrowband LEDs can result in unexpected lighting quality effects and dissatisfaction of the end user.

Over the years, several groups have proposed possible updates to the CIE CRI. Some of those proposed measures focused on color rendering in CIE-defined sense—that is, on the color fidelity of a set of test samples illuminated by test sources as compared to those same samples illuminated by a reference illuminant [for example, David and others 2015; Davis and Ohno 2010; Smet and others 2013; van der Burgt and van Kemenade 2010]—whereas others attempt to provide predictions for subjective illumination color quality; that is, on aspects of color rendition related to attributes such as preference, naturalness, and harmony [for example, Davis and Ohno 2010; Freyssinier-Nova and Rea 2010; Hashimoto and others 2007; Smet and others 2012; Yano and Hashimoto 1998] or even tried combining the two into a single measure [for example, Davis and Ohno 2010]. Each of those proposed measures acknowledged, among others, the importance of updating both the color space and the sample set. The only exception was the Gamut Area Index [Rea and Freyssinier-Nova 2008], which, for reasons of complementarity, continued the use of the eight Munsell samples defined in the CIE CRI. The reasons for updating the color space are simply that the CIE CRI $U^*V^*W^*$ has long been known to have poorer perceptual uniformity than more modern color spaces and chromaticity diagrams [Luo and others 2006], such as CIELAB, CIE 1976 $u'v'$, IPT, and CAM02-UCS [Fairchild 2005]. Reasons for updating the sample set were more diverse, although all agreed that a uniform distribution of the samples around the hue circle is a desirable feature. For example, the authors of the Color Quality Scale (CQS) claimed that more saturated samples were required for accurate color rendering evaluation [Davis and Ohno 2010], whereas the choice of sample set in the memory color rendition index (MCRI) was of a practical nature; that is, the samples had to be familiar objects [Smet and others 2012]. The CRI2012 sample set was guided by the requirement of spectral uniformity, which ensures that each wavelength contributes equivalently to the total color rendering score [Smet and others 2013], to ensure that it is not possible to selectively optimize a lamp's spectral power distribution (SPD) in ways that improve the CRI value but not the actual color fidelity. Most of these proposed sample sets have, however, a relatively small number of samples. It has been proposed by van Kemenade and van der Burgt [2010] that larger sample sets are more appropriate because they allow a better coverage of the color space and therefore provide more detailed information, especially when plotted in color vector icons.

The recently recommended color rendering framework of the IES [David and others 2015; IES 2015], based in part on the work of the present authors, has attempted to utilize the best of other proposed measures: it uses the most state-of-the-art color (difference) space, CAM02-UCS [Luo and others 2006], as well as a larger number of samples uniformly distributed across color space while exhibiting excellent spectral uniformity [Smet and others 2013].

This article presents an investigation of the impact of the choice of color space and sample set properties on color fidelity predictions, in particular that of the IES fidelity metric R_f, to address some of the issues recently raised in CIE technical committee TC1-90: for example, despite the generally high correlation between the CIE R_a scores and the IES R_f scores, systematic differences are apparent for certain types of light sources. The article shows why both color space uniformity and sample set spectral uniformity are essential for a fidelity metric and how these systematic differences in scores can be traced back to a failure of the CIE CRI to comply, in contrast with the new IES measure, with these requirements.

2. SYSTEMATIC DIFFERENCES BETWEEN THE CIE R_a AND IES R_f SCORES

The existence of systematic differences for certain types of sources is illustrated in Fig. 1, where the IES R_f scores of a set of 139 light sources are plotted as a function of the CIE R_a values. The set of 139 light sources was composed of broadband fluorescent (halophosphate, multiband), broadband LEDs (phosphor white LEDs, phosphor + multichromatic LEDs), and an assortment of narrowband LEDs, fluorescent, and other discharge lamps.

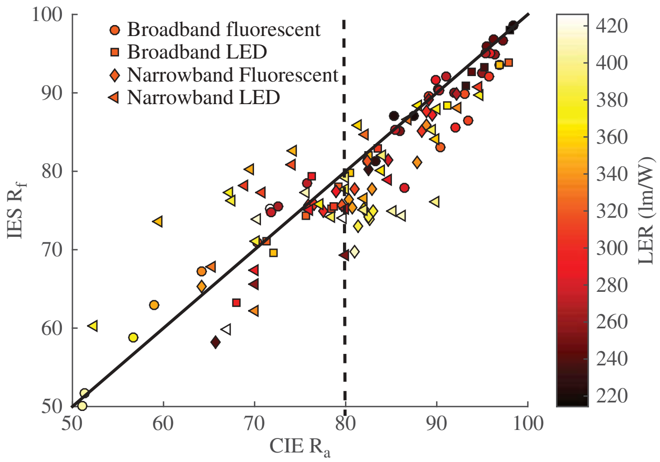


Fig. 1 **IES R_f scores versus the CIE R_a scores for a set of 139 broadband and narrowband sources.**

From Fig. 1 it is clear that, not surprisingly, there is a strong correlation between the two metrics ($R^2 = 0.86$). However, there is a substantial absolute difference in score values, which is more pronounced for narrowband sources, especially for those with greater values of luminous efficacy of radiation (LER). Though there appears to be a rather symmetrical spread around the 1:1 line, most sources with $R_a \geq 80$ and higher LER values tend to have lower R_f values, suggesting that these sources might have been optimized by taking advantage of some of the inaccuracies in the CIE CRI. Note that a CIE R_a value of 80 is in many norms and guidelines the lower limit for acceptable color rendering for general lighting. The mean absolute deviations (MADs) for broadband and narrowband sources are respectively 2.1 and 5.0. For the narrowband sources with high R_a ($\geq$80) and high LER, the MAD values can be as high as 10–14 units.

In the following sections these discrepancies with the CIE CRI will be explored and it will be shown how they impact the resulting color fidelity scores.

3. COLOR SPACE

For years, it has been known that the color space adopted in the CIE CRI (that is, CIE1964 $U^*V^*W^*$) lacks sufficient perceptual uniformity,[1] which is why most other color rendition proposals have chosen to update it to more modern and more perceptually uniform spaces, like CIE1976 $u'v'$ (GAI [Rea 2010]), CIELAB (CQS [Davis and Ohno 2010], FCI [Hashimoto and others 2007], Philips CRI [van der Burgt and van Kemenade 2010]), IPT (MCRI [Smet and others 2012]), CAM02-UCS (CRI2012 [Smet and others 2013], and IES R_f [David and others 2015; IES 2015]).

The IES R_f metric has chosen the CAM02-UCS color space, which has been shown to have a good perceptual uniformity [Luo and others 2006]. Additionally, it implicitly includes a better chromatic adaptation transform (CAT)—that is, CAT02 [CIE 2004]—than the outdated von Kries CAT adopted by the CIE CRI. The CAM02-UCS adopted white point is the equi-energy white.

This built-in chromatic adaptation transform and the implicitly adopted white point in CAM02-UCS have the advantage that fidelity calculations are essentially independent of the correlated color temperature (CCT) of the test light source: prior to calculating the color differences, the color coordinates of the reflectance samples under both the test source and the reference source are first transformed to their corresponding colors under the equi-energy white. In contrast, in the CIE CRI the color coordinates of the samples under the test source are transformed to the reference illuminant, thereby making the color difference calculation dependent on the shape of the color space at that CCT.

This can be seen in Fig. 2, which shows the shape spanned by the 1269 samples of the Munsell color order system in the $U^*V^*W^*$ and CAM02-UCS color space for four different CCTs.

It is clear that the $U^*V^*W^*$ color space tends to increasingly compress along the yellow–blue axis as the CCT

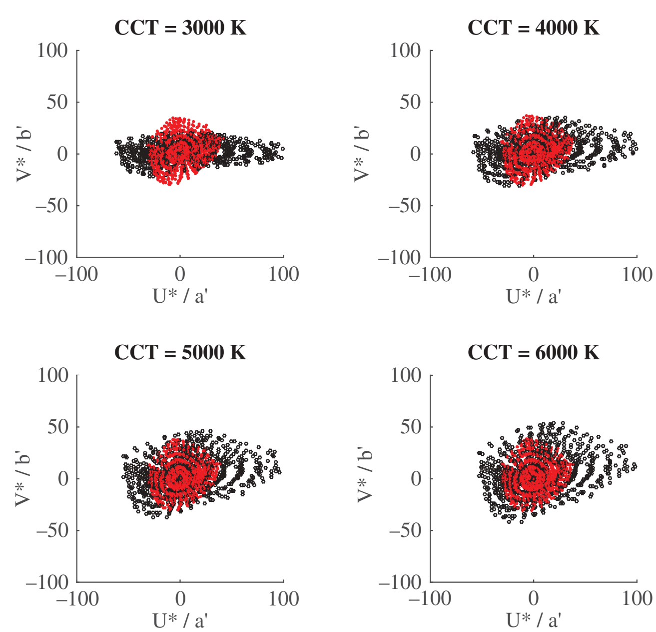


Fig. 2 **Shape of $U^*V^*W^*$ (black open circles) and CAM02-UCS (red dots) color spaces for four different CCTs as illustrated by the 1269 Munsell samples illuminated by a blackbody radiator.**

decreases. Therefore, in $U^*V^*W^*$ color space the weight of each dimension in the color difference calculation will change as the CCT changes. The CAM02-UCS space, on the other hand, is quite stable across changes in CCT. The relative size of the effect was quantified as follows: For each CCT the average color difference was calculated between the Munsell set illuminated by a blackbody radiator of the specified CCT and that same set perturbed by four different types of random Gaussian error: three-dimensional errors along CAM02-UCS dimensions J', a', b', two-dimensional errors along a', b', and one-dimensional errors along a' and b'. For each error type, the magnitudes of the errors were set such that the average was approximately one unit in the uniform CAM02-UCS color space. The median results of 100 repetitions are shown in Fig. 3. First, it is clear that $U^*V^*W^*$ is not a uniform color space, even at higher CCTs, because errors generated along different directions do not result in equally large color differences; that is, the lines for the different error types do not coincide for $U^*V^*W^*$. Note that the $U^*V^*W^*$ color differences are generally larger than those calculated in CAM02-UCS. This discrepancy can be adjusted for by rescaling the color differences to a color fidelity score using an appropriate set of lights sources (for example, the CIE illuminants F1–F12 as proposed by Davis and Ohno [2010]). Second, it is clear that relative to the uniform CAM02-UCS, the color differences in $U^*V^*W^*$ are indeed CCT dependent, especially when color differences present themselves along a single dimension (for example, a' and b' errors). For low CCT values, the $U^*V^*W^*$ color differences corresponding to the a' and b' errors tend to be respectively over- and underestimated compared to the unit CAM02-UCS color differences, with the color differences related to the a' errors being almost triple those of the b' errors. Such substantial CCT-dependent differences in the sensitivity of the color difference calculation to the prevalent direction of the color error are undesirable in a fidelity metric.

As a side note regarding the impact of CCT on fidelity scores, some have also argued that fidelity measures that use the CIE-defined reference illuminants are unnecessarily restricting the development of new, better low-CCT light sources. They fear that creating whiter looking light sources, by having the chromaticity slightly below the blackbody locus [Rea and Freyssinier 2013], will have a substantial negative effect on the color fidelity scores. However, such fears are unwarranted; in reality, high color fidelity scores can still be generated for off-Planckian sources. To illustrate this, several SPDs, each composed of five Gaussian primaries, were selected (in a simulation) to yield various levels of Duv [Ohno 2014] at a CCT of approximately 3000 K, while optimizing IES R_f. The results are shown in Fig. 4.

Finally, the effect of updating the CIE R_a calculation with the CAM02-UCS color space while holding constant all other components (eight Munsell samples, CIE 1931 2° observer, et cetera) is illustrated in Fig. 5. It is clear that using a more uniform color space results in a spread around the 1:1 (diagonal) line. Unlike the spread shown in Fig. 1, the spread in Fig. 4 is more symmetrical, with only a tiny tendency toward lower scores for the updated CIE R_a (using CAM02-UCS). The MAD values for broadband and narrowband sources are respectively 1.4 and 3.0 units, with values of up to 5–6 units for high LER–high R_a ($\geq$80) sources. As an example, CIE R_a and updated CIE R_a (using CAM02-UCS) values for a selection of broad-/narrowband fluorescent and LED light sources are listed in the table in Appendix A.

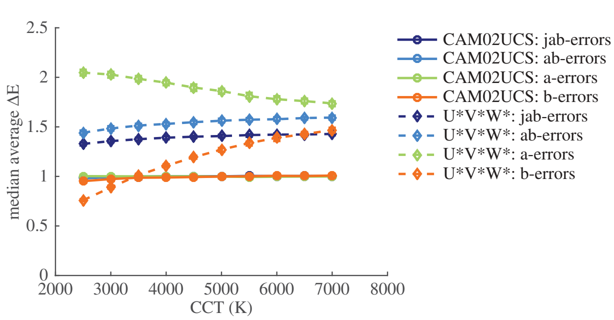


Fig. 3 **CCT dependence of color error in $U^*V^*W^*$ relative to a unit CAM02-UCS color error. CAM02-UCS color errors are represented by filled circles (solid lines); $U^*V^*W^*$ color errors space are shown as diamond shaped symbols (dashed lines).**

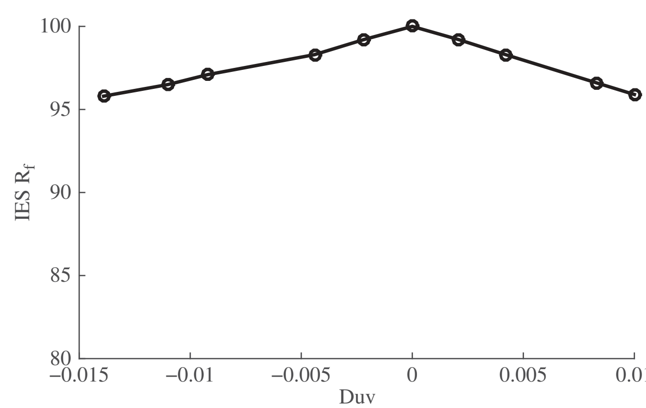


Fig. 4 **Tradeoff between Duv and R_f for 3000 K simulated SPDs composed of five Gaussian primaries.**

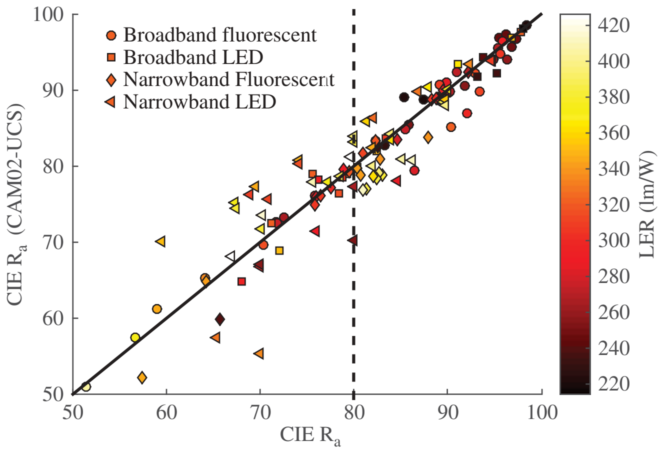


Fig. 5 **CIE R_a scores calculated with CAM02-UCS versus the CIE R_a scores for a set of 139 broadband and narrowband sources.**

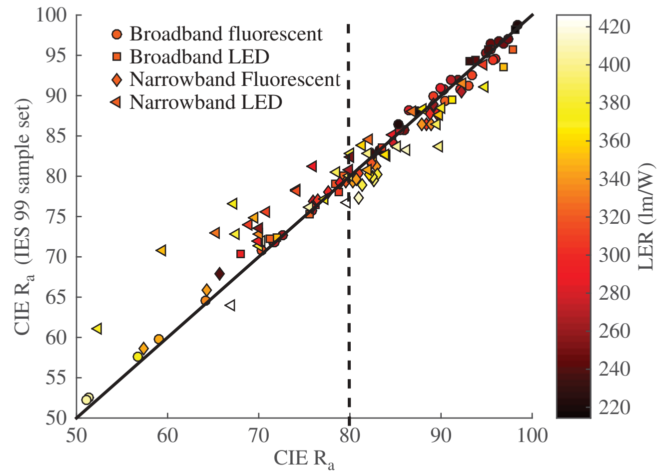


Fig. 6 **CIE R_a scores calculated with the IES 99 sample set versus the CIE R_a scores for a set of 139 broadband and narrowband sources.**

4. SAMPLE SET PROPERTIES

Previously in this article, it was suggested that the overall decrease in score level could be attributed to selective optimization of the spectral power distributions of various sources to take advantage of the weaknesses of the CIE CRI. Having shown that the color difference calculation in the nonuniform $U^*V^*W^*$ was one such weakness, let us now consider errors arising from the CIE CRI sample set.

A comparison of the MAD values between the CIE R_a and the IES R_f values with the MAD values between the CIE R_a and the updated CIE R_a (using CAM02-UCS) values suggests that the effect of the sample set improvement is about three times stronger for narrowband sources (difference ≈ 2.0) than for broadband sources (difference ≈ 0.7 units). This is confirmed by a direct, color space–unbiased comparison between the CIE and IES sample sets: Introducing the IES 99 sample set into the CIE CRI calculation engine (in $U^*V^*W^*$) results in MAD values for broadband and narrowband sources of respectively 0.8 and 2.4 units. However, CIE R_a and IES R_f values can differ by as much as 4 to 6 units for high LER and high R_a sources. The results of this comparison are illustrated in Fig. 6. As an example, the CIE R_a (using the IES 99 set) for a select number of broad-/narrowband fluorescent and LED light sources are listed in the table in Appendix A.

As was reported in David and others [2015] and Smet and others [2013, 2015], spectral or wavelength uniformity is an important property of sample sets used for color rendering evaluation. In the following subsections, the impact of spectral uniformity, or spectral flatness, on color fidelity scores will be investigated in more detail.

4.1. Sample Set Size

As previously mentioned, the CIE CRI uses a limited number of samples to evaluate the color rendering properties of a light source. Eight samples, all of low to moderate saturation, do not provide adequate coverage of the color space to provide sufficiently detailed information on the specifics of the color rendering shifts generated by the test sources in comparison with the reference illuminant. It has been proposed to provide this additional information to the user in the form of color rendering vector plots [van der Burgt and van Kemenade 2010]. The IES metric provides this information in the form of a similarly motivated color distortion graphic [David and others 2015; IES 2015]. The advantage of using a larger sample set is not only the more detailed information about color shifts but, importantly, a more statistically accurate estimate of the overall color rendition properties: a single extreme sample will only have a limited impact on the overall color rendering score. This is an important property to deal with possible metameric samples. For example, Smet and others [2013] have shown that metameric samples could easily differ by 10 CIELAB color difference units or more depending on the test light source.

To additionally illustrate the effect of sample size on the statistical precision of the IES R_f scores, 100 random, but uniformly distributed, sample sets were generated of sizes $N = 8$, 16, 32, 64, 125, 250, 500, 1000, and 2000, and the IES R_f values were calculated for each of the 139 light source spectra mentioned earlier. For each SPD and sample size the standard deviation of the R_f scores was calculated over all 100 repetitions. Those standard deviation values

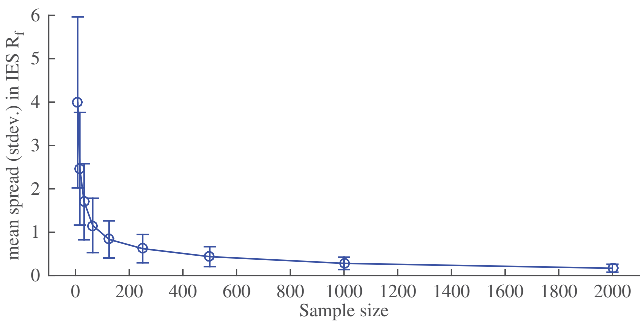


Fig. 7 **Effect of sample size on statistical precision of the IES R_f scores. The standard deviation bars around the mean spread are also shown.**

were then averaged over all sources as a general estimate of the effect of sample size on statistical precision. The results are shown in Fig. 7. It is clear that as sample size gets larger the precision increases. For a sample set size of 100 the statistical precision is about 1 R_f unit, whereas it is approximately 4 units for a sample set size of eight but can be as high as 6 units.

Color fidelity precision values for a sample size of eight are listed for a selection of broad-/narrowband fluorescent and LED light sources in the table in Appendix A.

4.2. Wavelength Equivalency

Although advantageous from an information point of view, having a large number of uniformly distributed samples is not sufficient to ensure a good fidelity sample set. Even with many thousands of samples there remains a high likelihood of wavelength bias in the color fidelity calculation, because many samples from available reflectance databases are based on a limited number of underlying dyes. This can result in a substantial wavelength bias, whereby specific spectral features in the sample set tend to be located at specific wavelengths, causing those wavelengths to contribute more than they should to the overall color rendering score. An important additional property of the sample set is therefore good sample set spectral uniformity [David and others 2015; Smet and others 2013, 2015].

As an extreme example of poor spectral uniformity, consider a sample set of 99 samples, with the same chromaticity values as those of the IES 99 sample set but generated using only three dyes: magenta, cyan, and yellow. As can be seen in Fig. 8 (middle graph), almost all of the spectral variation is located around only two wavelengths, and such a sample set would therefore be unsuitable for color fidelity evaluation. This bias in wavelength sensitivity is clearly reflected in the spectral uniformity (Fig. 8, right-hand graph), defined here as the mean square slope $(r')^2$ [see also David and others, 2015], which also shows large peaks at those two wavelengths. For comparison, the figure also plots the spectral uniformity of three other sample sets—the IES 99 set, a set of 99 samples drawn at random from the 1269 Munsell set, and the CIE CRI test samples. It is clear that for all three sample sets, all composed of real objects, the degree of spectral uniformity is better; that is, their $(r')^2$ show a much smoother variation. The IES 99 sample set has a very good spectral uniformity because it was specifically designed to be spectrally uniform. In contrast, the other two sets still have a substantial degree of spectral disuniformity, with clear peaks and valleys that can be taken advantage of in the spectral optimization of light sources. As is clear from the 99 sample Munsell set, solely increasing the sample size from eight to 99 did not improve spectral uniformity.

Though it is possible that certain wavelengths are indeed more privileged due to the specific materials and dyes prevalent in specific lighting applications and situations, a spectrally uniform set could be argued to be the most reasonable choice ensuring the most general validity of the color rendering measure in which it is adopted. In addition, unknown future advances or changes in coloring agents might affect the current distribution of privileged wavelengths, further supporting the IES approach of using a spectrally uniform sample set in the color fidelity calculation because it ensures not only implicational but also temporal, general validity. Uniformity is a desirable property in cases where general validity is deemed important and is commonly agreed upon. This is consistent with the general agreement that it is good to have samples uniformly distributed around color space for color fidelity measures, even though the frequency distribution of colors actually viewed in various settings may differ considerably.

4.2.1. Spectral Nonuniformity and Selective Optimization of Light Source Spectra

To illustrate selective spectral optimization of the spectrally nonuniform CIE CRI sample set, a number of laser-line spectra with a CCT of 3000 K (CIE x, y = 0.4369, 0.4041) were simulated. The full-width–half-maximum spectral width of each laser line was set to 1 nm. The peak wavelengths were chosen as follows: blue at 450 nm, amber at 590 nm, and green and red covaried between respectively 510 to 540 nm and 590 to 670 nm. For each simulated spectrum, four different fidelity indices were calculated: the CIE R_a, the color space updated CIE R_a (using

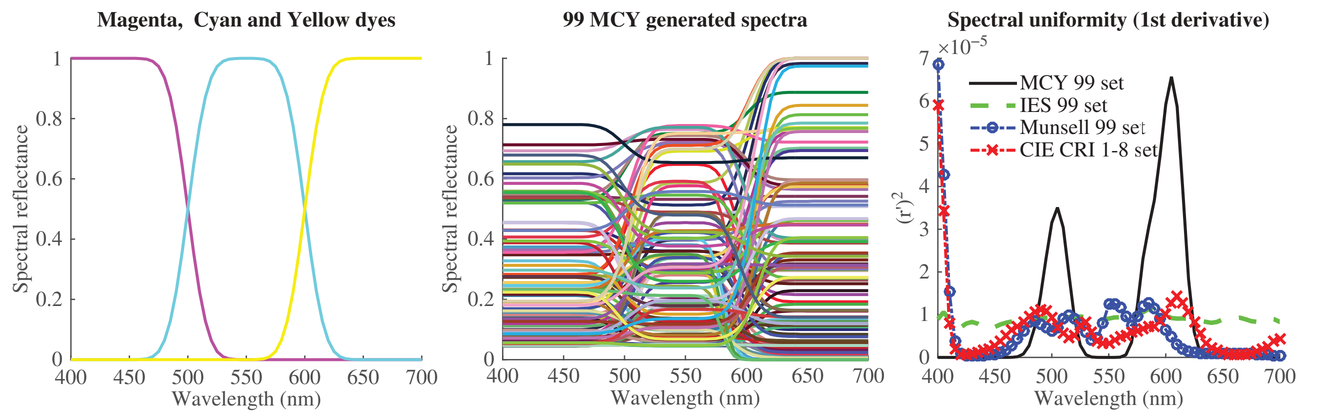


Fig. 8 **Extreme example of spectral nonuniformity for a three-dye sample set. Left: Spectral reflectance functions of the three dyes. Middle: Spectral reflectance functions of a set of 99 samples composed of the three Magenta, Cyan and Yellow (MCY) dyes that have the same chromaticity as the IES 99 set. Right: Spectral uniformity of the MCY 99 sample set, of the IES 99 set, of a set of 99 randomly selected Munsell samples, and of CIE CRI sample set (1–8).**

CAM02-UCS), the IES R_f (99 samples), and the IES R_f (4900 Refset). Note that the latter, the IES 4900 Refset, was the original, spectrally uniform set from which the IES 99 set was down-sampled to achieve a match of about one color fidelity unit [David and others 2015]. The results are plotted in Fig. 9.

The fine structure observable in the surfaces corresponding to the nonuniform CIE CRI set (Fig. 9, top row)

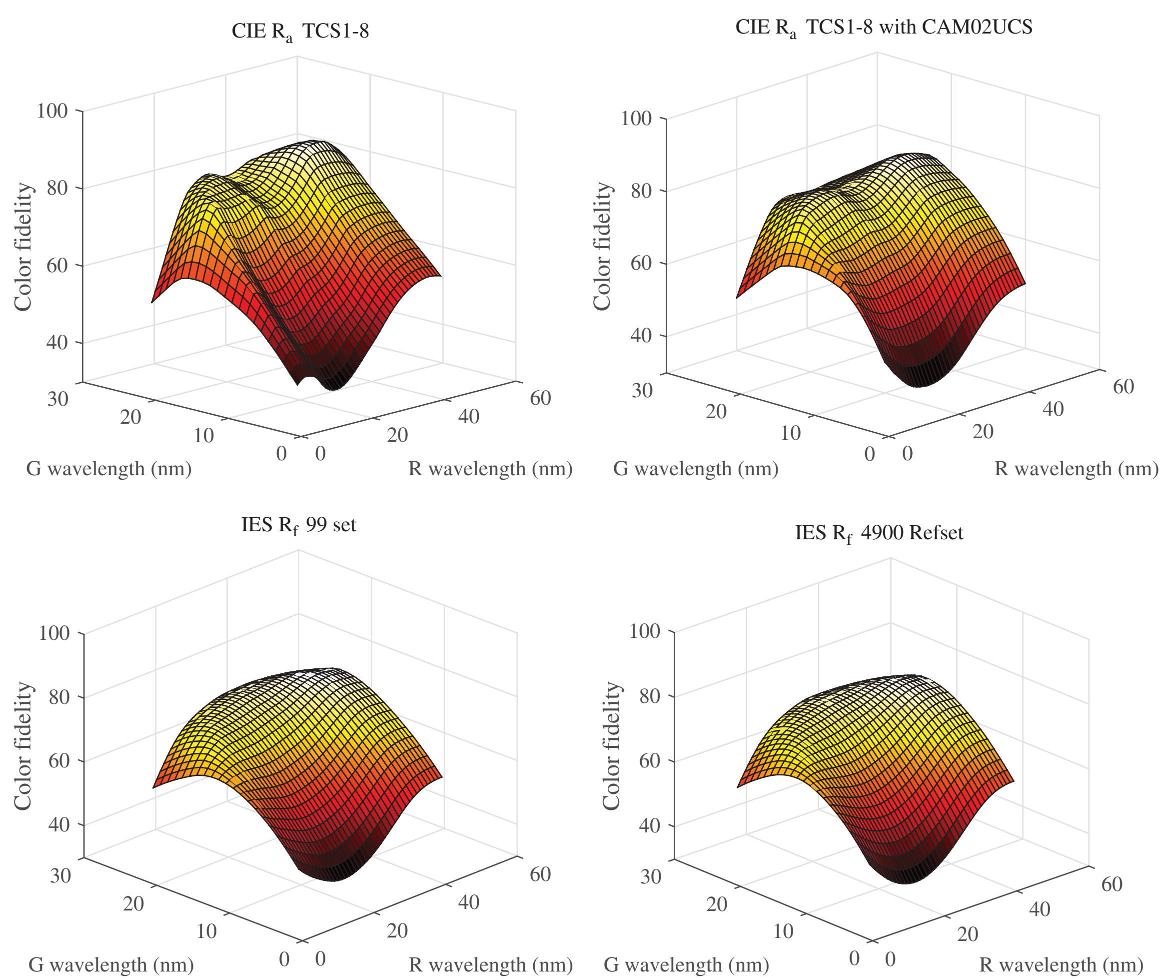


Fig. 9 **Spectral (non)uniformity of the sample set and selective spectral optimization using laser-line spectra.**

provides the possibility of gaming a light source's spectral power distribution for higher color rendering scores by slight tweaks in its shape: shifting the peak position of a laser line results in substantial difference in color fidelity score. In contrast, no such fine structure is observable for the spectrally uniform IES 99 and 4900 sets (Fig. 9, bottom row).

To make this distinction clearer, cross sections of the graphs in Fig. 9 for a green peak wavelength of 520 nm have been plotted in Fig. 10. From the right-hand graph of Fig. 10 it is obvious that the spectrally nonuniform CIE CRI sample set has certain optimal wavelength locations at which color fidelity scores are higher, effectively enabling gaming of the light source spectrum. For example, choosing a peak wavelength of 620 nm instead of 630 nm results in a about a 5-unit color fidelity difference. On the other hand, and as intended, the IES 99 and 4900 set do not exhibit this behavior. In addition, from Fig. 10 it is clear that, as was intended, both sets agree very well with one another. (It should be noted that the plots for the IES 99 set and IES 4900 Refset are not flat; their shape is determined by the response characteristics of the human visual system as modeled in the color difference calculation. The key difference is that, unlike the case with the CRI sample set, the response characteristics are not distorted with fine features arising from sample set spectral nonuniformity that have no perceptual meaning.)

4.2.2. Impact of Degree of Uniformity on Color Fidelity Scores

In the previous subsection it was shown that spectral nonuniformity of the sample set has an impact on color fidelity scores and that that impact can be used to selectively optimize the spectral power distribution of a light source. Earlier it had also already been suggested that some of the existing light sources might already have been gamed for high CIE R_a values by taking advantage of the wavelength bias exhibited by the CIE CRI sample set. In this subsection, the impact of the degree of spectral uniformity, or spectral flatness, on color fidelity scores will be investigated.

To this end, several sample sets with varying degrees of spectral uniformity were created by repeatedly selecting samples from a spectrally uniform set (IES 4900 Refset) and from a spectrally nonuniform set (105,000 reflectance function from David and others [2015]). By changing the contribution of each set to the total size of the final sample set spectral flatness can be varied. Each final set contained 2000 samples uniformly distributed in color space. Sample set spectral uniformity was then assessed using the $(r')^2$ function as defined in David and others [2015]. The overall degree of sample set spectral uniformity, henceforth referred to as "flatness," was assessed as the average magnitude of the first derivative of the mean spectral uniformity function $(r')^2$. An example of the variation in spectral uniformity for such a set of sample sets is shown in Fig. 11. In the left graph, the spectral uniformity of each individual sample set is shown, and in the right graph the flatness of the spectral uniformity of each sample set has been plotted.

To observe the effect of changing the sample set spectral flatness on the color fidelity calculation, the aforementioned sample sets were sequentially applied to evaluate the color fidelity (R_f) of each of the previously mentioned 139 light source spectral power distributions. The results are plotted in Fig. 12.

In the left graph in Fig. 12, the color code for the individual plotted points is the same as that in Fig. 11. It is

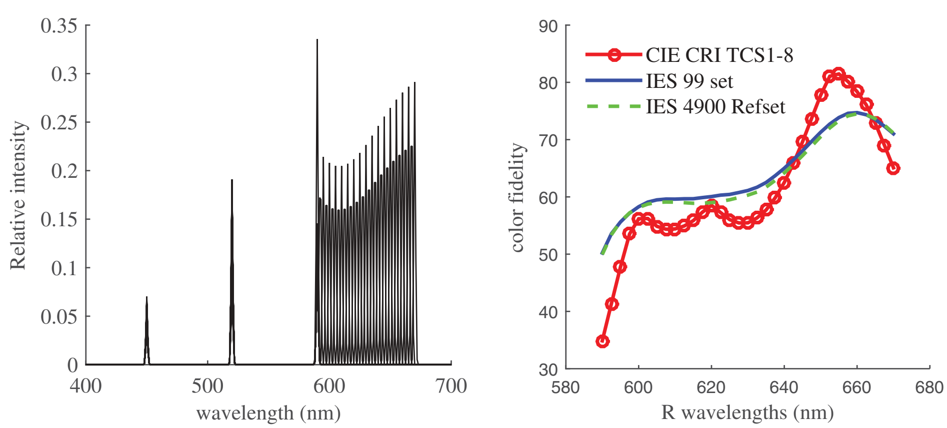


Fig. 10 **Spectral (dis)uniformity of the sample set and selective spectral optimization. Left: Laser-line spectra. Right: Color fidelity values for various peak wavelength locations of the red laser line.**

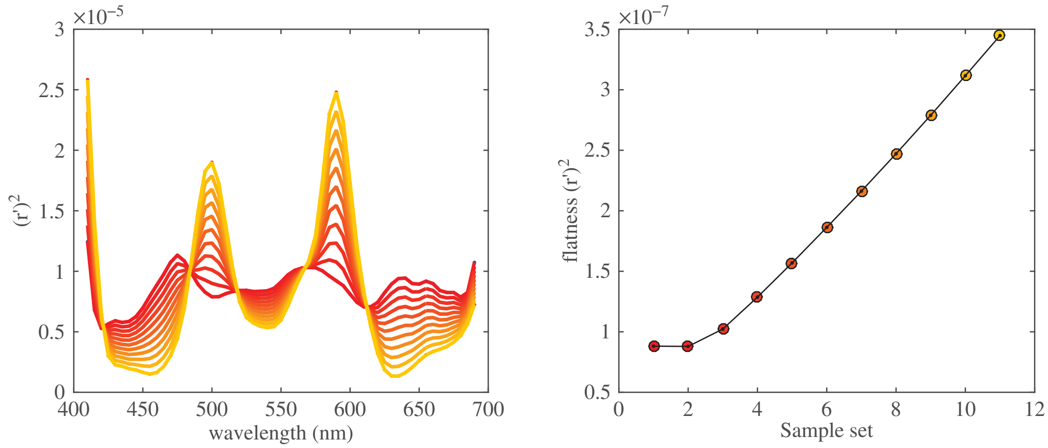


Fig. 11 **Left: Spectral uniformity, mean of $(r')^2$ of sample sets of varying spectral flatness. Right: Nonflatness of the spectral uniformity of each sample set.**

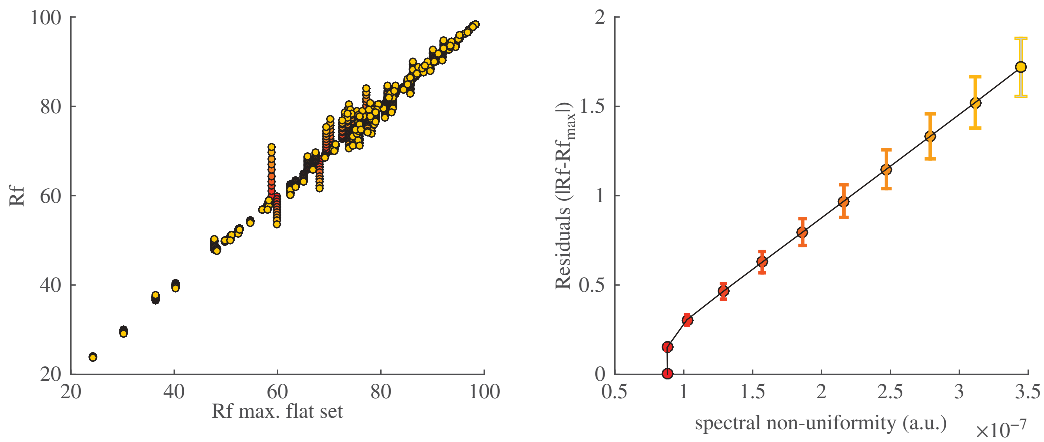


Fig. 12 **Database of 139 light sources. Left: Color fidelity scores for sample sets of varying degrees of spectral uniformity. Right: Overall average difference in color fidelity scores between the various sample sets and the flattest set as a function of the degree of spectral uniformity (flatness) of each sample set.**

clear that for many sources the impact of sample set spectral uniformity is limited. However, for some there is a substantial and systematic shift in which the fidelity score deviates from the base value by an amount that increases as sample set flatness decreases. The overall color fidelity difference therefore tends to increase with decreasing spectral uniformity of the sample set, as shown in the right graph of Fig. 12. On average, for this database of 139 SPDs, the difference is about two color fidelity units. However, the more extreme scores tend to be in the 5- to 12-unit range, indicating a substantial impact of spectral uniformity on color fidelity scores.

This can also be seen by observing, for example, the spectra of two light sources that have approximately the same color fidelity value, $R_f = 58$ and $R_f = 59$, for the maximum flat sample set but with very different values for the sample set with minimum flatness: $R_f = 58$ and $R_f = 71$, respectively. In Fig. 13, it can be seen that the SPD with the largest shift in R_f values (12 units) has a substantially larger overlap with the peak in the $(r')^2$ spectral uniformity measure in the long wavelength range—an area sensitive to perturbations of the SPD—than the one with the lowest shift in R_f values (one unit): the larger overlap causes that SPD to respond more to changes in the degree of spectral uniformity of the sample set, illustrating the importance of spectral flatness.

The color fidelity values for the sets with maximum and minimum flatness are reported in Table A1 in Appendix A for a select number of broad-/narrowband fluorescent and LED light sources.

As a further example of some light source spectra resulting in extreme differences, the same analysis was performed for a database of commercially available triband fluorescent sources. For these sources there are anecdotal reports suggestive of the CIE CRI overestimating the fidelity with which they render the color of many real objects, especially

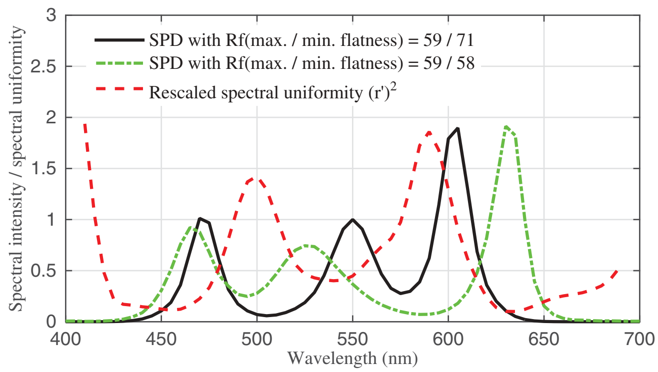


Fig. 13 **Example illustrating how spikes in an SPD in a region of high $(r')^2$ can substantially affect its color fidelity scores.**

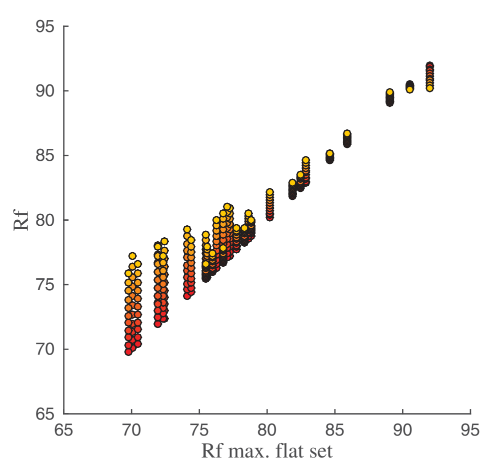
for warm white sources with CCTs below 4000 K. The results of the analysis are illustrated in Fig. 14.

The results are similar to those with the broader range of light sources but more pronounced. The color fidelity difference with the most spectrally uniform set increases with increasing spectral nonuniformity. Significantly, the shift in color fidelity difference is very predominantly toward higher fidelity scores, which lends substantial credibility to the hypothesis that such sources tend to be overestimated by a spectrally nonuniform sample set such as the one used in the CIE CRI.

From the above analysis and discussion it is clear that spectral uniformity is an important property of a sample set for use in color fidelity calculations. Without this feature, color fidelity scores are highly dependent on the wavelength bias present in the sample set, which lamp manufacturers could use (and likely have been using) to boost a light source's color fidelity score.

5. CONCLUSIONS

It has been known for many years that the CIE CRI is insufficiently accurate to assess some light source spectra, especially those with narrowband and/or spiked features. In the literature, there have been quite a few proposals to update the CIE CRI calculation method. Almost all agreed on updating both the color space and the sample set. One of those proposed updates was formulated by a group including the present authors, in the context of an IES working group. The IES color fidelity measure updates the color space to the state-of-the-art CAM02-UCS and updates the sample set to a set of 99 samples uniformly distributed in color space. In addition to uniform distribution in color space, the proposed sample set has one other important property; that is, good spectral uniformity. In this article, it has been shown that both properties are important updates that have a substantial impact on the color fidelity scores. It was found that the combined effects were, on average, about 2 and 5 units for broadband and narrowband sources respectively, but could be as high as 10–14 units for high LER–high R_a sources. Both properties—that is, color space uniformity and sample set spectral uniformity—were also analyzed separately. It was found that the effect of the update of color space is mainly to eliminate directional biases in color space with regard to the color errors, which introduces an approximately symmetrical spread of approximately 1.5 and 3.0 units around the CIE R_a values for broad- and narrowband sources, respectively. Maximum values could be as high as 5–6 units. The sample set update was shown to be an important remedy against possible selective optimization of light source spectra by taking advantage of the



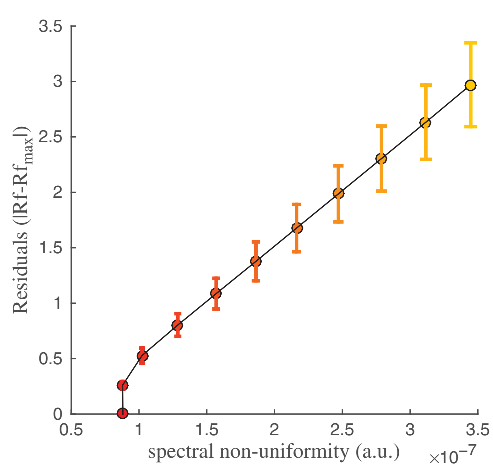


Fig. 14 **Database of 31 triband fluorescent sources. Left: Color fidelity scores for sample sets of varying degrees of spectral uniformity. Right: Overall average difference in color fidelity scores between the various sample sets and the flattest set as a function of the degree of spectral uniformity of each sample set.**

wavelength bias present in the CIE CRI sample set. In general, the effect of the introduction of a spectrally uniform sample set was a downward shift of color fidelity scores compared to the CIE R_a scores. On average, the magnitude of the shift for broadband and narrowband sources was found to be respectively 0.8 and 2.4 units but could be as high as 4–6 units. Importantly, this downward shift appears to be limited to high CIE R_a ($\geq$80) and higher LER light sources and to be more extensive for narrowband and spiked spectra. The fact that the impact is mostly of importance for this limited subset of light sources suggests that these sources may have previously been tuned to take advantage of the errors arising from the previous CIE sample set. Finally, the increased sample size of the IES 99 set was shown to substantially increase the statistical precision of the calculated color fidelity scores: Though a set of 99 samples tends to have a precision of about one color fidelity unit, a set of only eight samples (like the CIE CRI set) has a precision of about 4 units but can be as high as six color fidelity units.

To conclude, the results of this article indicate that the IES R_f is a more appropriate color fidelity measure than the CIE R_a, because of the following properties: (1) it has a more perceptually uniform color space, (2) its color difference engine has no built-in CCT dependence, (3) it has better statistical precision due to its increased sample size, (4) it is capable of providing much more detailed information due to the larger number of samples, (5) it is much harder to selectively optimize due to sample set uniformity, and (6) it can be argued that spectral uniformity is a more reasonable assumption for general validity of the color fidelity measure than a wavelength-biased sample set that might only have validity for a specific lighting application or situation and for current coloring agents.

ORCID

Kevin A. G. Smet http://orcid.org/0000-0003-3825-6274

FUNDING

Kevin Smet was supported by a Postdoctoral Fellowship of the Research Foundation Flanders (12B4913N and 12B4916N).

NOTE

1. A color space is perceptually uniform if the distance between any two colors is proportional to their perceived dissimilarity.

REFERENCES

Bodrogi P, Csuti P, Hotváth P, Schanda J. 2004. Why does the CIE colour rendering index fail for white RGB LED light sources? CIE Expert Symposium on LED Light Sources: Physical Measurement and Visual and Photobiological Assessment. Tokyo, Japan. p. 24–27.

[CIE] Commission International de l'Éclairage. 1980–1982. Colour rendering (Minutes from the 25th meeting, Paris (Nov. 1980), the 26th meeting, Berlin (Sep. 1981), and the 27th meeting, Oslo (Sept. 1982)). Publication No. CIE TC-E3.2.

[CIE] Commission International de l'Éclairage. 1995. Method of measuring and specifying colour rendering properties of light sources. Vienna (Austria): CIE. Publication No. CIE13.3-1995.

[CIE] Commission International de l'Éclairage. 2004. A review of chromatic adaptation transforms. Vienna (Austria): CIE. Publication No. CIE16x-2004.

[CIE] Commission International de l'Éclairage. 2007. TC 1-62: Colour rendering of white LED light sources. Vienna (Austria): CIE. Publication No. CIE177-2007.

[CIE] Commission International de l'Éclairage. 2011. ILV: CIE international lighting vocabulary. Vienna (Austria): CIE. Publication No. CIE S 017/E:2011.

David A, Fini PT, Houser KW, Ohno Y, Royer MP, Smet KAG, Wei M, Whitehead L. 2015. Development of the IES method for evaluating the color rendition of light sources. Opt Express. 23(12):15888–15906.

Davis W, Ohno Y. 2010. Color quality scale. Opt Eng. 49(3):033602-033602-16.

Fairchild MD. 2005. Color appearance models. Chichester (UK): John Wiley & Sons. 408 p.

Freyssinier-Nova JP, Rea MS. A two-metric proposal to specify the color-rendering properties of light sources for retail lighting. In: Tenth International Conference of Solid-State Lighting; Proceedings of SPIE; 2010 Aug 1–5; San Diego, CA. p. 77840V.

Hashimoto K, Yano T, Shimizu M, Nayatani Y. 2007. New method for specifying color-rendering properties of light sources based on feeling of contrast. Color Res Appl. 32(5): 361–371.

[IES] Illuminating Engineering Society of North America. 2015. Method for evaluating light source color rendition. New York (NY): The Illuminating Engineering Society of North America. Publication No. IES-TM-30-15.

Luo MR, Cui G, Li C. 2006. Uniform colour spaces based on CIECAM02 colour appearance model. Color Res Appl. 31(4):320–330.

Ohno Y. 2014. Practical use and calculation of CCT and Duv. LEUKOS. 10(1):47–55.

Rea MS. 2010. A practical and predictive two-metric system for characterizing the color rendering properties of light sources used for architectural applications. In: International Optical Design Conference 2010; Jackson Hole, WY. p. 765206-765206-7.

Rea MS, Freyssinier JP. 2013. White lighting. Color Res Appl. 38(2):82–92.

Rea MS, Freyssinier-Nova JP. 2008. Color rendering: a tale of two metrics. Color Res Appl. 33(3):192–202.

Sándor N, Schanda J. 2006. Visual colour rendering based on colour difference evaluation. Lighting Res Technol. 38(3):225–239.

Smet KAG, Ryckaert WR, Pointer MR, Deconinck G, Hanselaer P. 2011. Correlation between color quality metric predictions and visual appreciation of light sources. Opt Express. 19(9): 8151–8166.

Smet KAG, Ryckaert WR, Pointer MR, Deconinck G, Hanselaer P. 2012. A memory colour quality metric for white light sources. Energ Buildings. 49:216–225.

Smet KAG, Schanda J, Whitehead L, Luo RM. 2013. CRI2012: a proposal for updating the CIE colour rendering index. Lighting Res Technol. 45(6):689–709.

Smet KAG, Whitehead L, Schanda J, Luo RM. 2015. Toward a replacement of the CIE color rendering index for white light sources. LEUKOS. doi:10.1080/15502724.2014.994747

Szabó F, Schanda J, Bodrogi P, Radkov E. 2007. A comparative study of new solid state light sources. In: CIE Session 2007. Proceedings of the 26th Session of the CIE. Beijing, China. CIE. p. D1-18-D1-21.

van der Burgt P, van Kemenade J. 2010. About color rendition of light sources: the balance between simplicity and accuracy. Color Res Appl. 35(2):85–93.

Yano T, Hashimoto K. 1998. Preference index for Japanese complexion color under illumination. Journal of Light and Visual Environment. 22(2):269–274.

APPENDIX A

TABLE A1 Color fidelity values of a selection of broad-/narrowband fluorescent and LED light sources calculated for several color fidelity measures: CIE R_a, IES R_f (99 set), IES R_f (4900 Refset), updated CIE R_a (using CAM02UCS), updated CIE R_a (using IES 99 set), IES R_f (max. flat set), and IES R_f (min. flat set). The mean precision of the IES R_f measure for a sample set size of eight is also given

Light source	CCT (K)	LER (lm/W)	CIE R_a	R_f (99 set)	R_f (4900 set)	CIE R_a (CAM02-UCS)	CIE R_a (IES 99 set)	R_f (max. flat set)	R_f (min. flat set)	R_f precision for sample size eight
Fluorescent sources										
CIE F1	6428	271	76	78	78	76	76	78	77	2.5
CIE F2	4225	334	64	67	67	65	65	67	66	2.0
CIE F3	3446	373	57	59	59	58	58	58	58	2.0
CIE F4	2938	404	51	52	52	51	52	51	51	2.8
CIE F5	6345	274	72	75	75	73	72	75	74	2.8
CIE F6	4149	344	59	63	63	61	60	62	61	2.3
CIE F7	6495	241	90	90	90	90	91	90	90	0.6
CIE F8	4997	244	96	96	96	97	96	96	96	0.8
CIE F9	4149	270	90	90	90	90	91	90	90	0.9
CIE F10	4998	313	81	78	78	82	80	79	80	3.8
CIE F11	3999	340	83	78	78	81	81	78	79	3.7
CIE F12	3000	382	83	75	75	79	80	75	79	4.1
CIE FL3.4	2904	282	87	78	78	79	88	78	81	5.3
CIE FL3.7	2979	384	81	73	73	77	79	74	77	4.5
Triband FL	3000	384	83	74	74	78	80	78	74	4.7
Triband FL	4042	342	81	75	75	79	80	51	50	4.0
Triband FL	6480	290	79	77	77	80	79	93	94	3.5
LED light sources										
LED, phosphor white + R	2775	365	91	88	88	93	89	89	89	2.9
LED, phosphor white + R	4905	235	95	93	93	92	96	92	93	1.6
LED, RGB	3300	398	90	76	76	88	84	86	90	6.2
LED, RGB	3391	300	38	58	58	51	48	57	57	5.3
LED, RGB	5014	271	40	58	58	51	47	58	57	5.5
LED, RGB	6500	265	80	75	76	77	81	76	75	5.5
LED, RGBA	2962	389	90	85	85	89	87	48	50	4.0
LED, RGBA	4000	284	85	79	78	78	85	79	79	5.0
LED, RGBA	6500	253	80	69	70	70	82	69	75	7.4

R, G, B, and A refer to respectively to single color Red, Green, Blue and Amber LEDs.