Lighting Res. Technol. 2013; **45**: 689–709

CRI2012: A proposal for updating the CIE colour rendering index

KAG Smet PhD[a,b], **J Schanda** PhD[c], **L Whitehead** PhD[a] and **RM Luo** PhD[d,e]
[a]Sustainable Solutions Applied Physics Laboratory, University of British Columbia, Vancouver, Canada
[b]Light and Lighting Laboratory, Leuven University, Ghent, Belgium
[c]Virtual Environment and Imaging Technologies Laboratory, University of Pannonia, Veszprem, Hungary
[d]School of Design, University of Leeds, Leeds, UK
[e]Department of Optical Engineering, Zhejiang University, Hangzhou, China

Received 29 October 2012; Revised 6 February 2013; Accepted 12 February 2013

The CIE colour rendering index (CRI) has been criticized for its poor correlation with the visual colour rendering of many spiked or narrowband sources, its outdated colour space and chromatic adaptation transform and the use of a small number of non-optimal reflectance samples that have enabled lamp manufacturers to tune the spectrum of a light source to yield, in some cases, inappropriately high general CRI values. The CRI2012 metric proposed in this paper addresses these criticisms by combining the most state of the art colorimetric colour difference model, i.e. CAM02-UCS, with a mathematical reflectance set that exhibits a highly uniform spectral sensitivity. A set of 210 real reflectance samples has also been selected to provide additional information on the expected colour shifts when changing illumination.

1. Introduction

Traditionally, the influence of the light source on the colour appearance of objects – *colour rendition* – has been described using the colour rendering index (CRI)[1] developed by the International Commission on Illumination (CIE). However, during the past decades it has become clear that the CIE CRI fails to correlate well with the visual colour rendition of many spiked or narrow band sources such as tri-band fluorescent lamps and light-emitting-diodes (LED).[2–5] Over the years, several CIE Technical Committees, the most recent being TC1-69, have looked into the problems and criticisms of the current CRI recommendation and a number of different metrics have been proposed in an attempt to more accurately or fully describe the colour rendition of a white light source. Since papers by Judd[6] and Jerome,[7] a distinction is made between how closely test colours are rendered under the test source compared to the reference illuminant (nowadays often called *colour fidelity*), and how pleasing the rendered colours look. This latter effect was originally described by a *flattery index* – now termed a *colour preference metric*[8,9] – and has recently been revisited by the development of a memory colour quality index.[10,11] Some papers tried to bridge the gap between colour fidelity and preference by developing colour rendition calculation methods that do not penalize a lamp that distorts test sample colours if the distortion goes in a preferred direction.[12,13] Still others have quantified colour rendition

Address for correspondence: Kevin AG Smet, Light and Lighting Laboratory, Gebroeders Desmetstraat 1, B9000 Ghent, Belgium
E-mail: Kevin.Smet@kahosl.be

10.1177/1477153513481375

by looking at the change in gamut area of a number of test samples illuminated by test source compared to that under a reference illuminant,[14,15] by estimating the distortion in colour harmony caused by a change in illumination[16] or by counting the number of samples qualified as 'good' or 'erroneous' under the test source in comparison with the reference illuminant.[17]

Although most metrics mentioned above attempt to describe more subjective aspects of the colour rendition or colour quality of a light source (which definitely have their place in general lighting applications), for some professional applications, such as health care, colour reproduction, printing and quality control, colour fidelity has primary importance, because poor colour fidelity can cause substantial unwanted colour distortion and hence affect critical colour judgement. In addition, many regulating bodies require and many lamp manufacturers request an objective measure, like colour fidelity with regard to a standard reference illuminant, to quantify the influence of a light source on the colour appearance of objects. To address this particular need, this paper concerns only colour fidelity, and presents a proposed variant of the current CIE CRI that uses modern colorimetric models and an improved sample set. As such, the proposed index has the same purpose as the current CIE CRI, and is presented as a suggested improvement of it.

2. The CRI2012 colour fidelity index

Building on a large body of previous work by numerous groups during the past few decades, a suggested colour difference-based CRI, now named CRI2012 but previously called nCRI, has been developed by a subcommittee of TC1-69, the CIE technical committee dealing with the colour rendition of white light sources. Its workflow is given in Figure 1.

There are five major components in the CRI2012

(a) A reference illuminant with a well-defined spectral power distribution

(b) A set of colour test samples: a set of spectral reflectance functions, $R_i(\lambda)$

(c) A colour difference calculation algorithm

(d) An averaging algorithm for colour differences

(e) Scaling from colour differences to metric scores

used in six computational steps, which are the calculation of

1) the correlated colour temperature (CCT) of the test source,

2) the spectral power distribution (SPD) of the reference illuminant,

3) the tristimulus values of a set of colour test samples illuminated by the test source and the reference illuminant,

4) their CAM02-UCS coordinates, a_M, b_M and J,

5) the individual CAM02-UCS colour difference ΔE_i between the two illumination states and

6) the average colour difference ΔE_{rms} and scaling conversion to yield the final CRI2012 score.

Compared to the standard CIE CRI, all but the first components have been updated in the CRI2012. The reference illuminant (Component a and Step 2 in Figure 1) is kept the same as described in CIE 13.3.[1] The reference illuminant is calculated as either a blackbody radiator (CCT <5000 K) or a

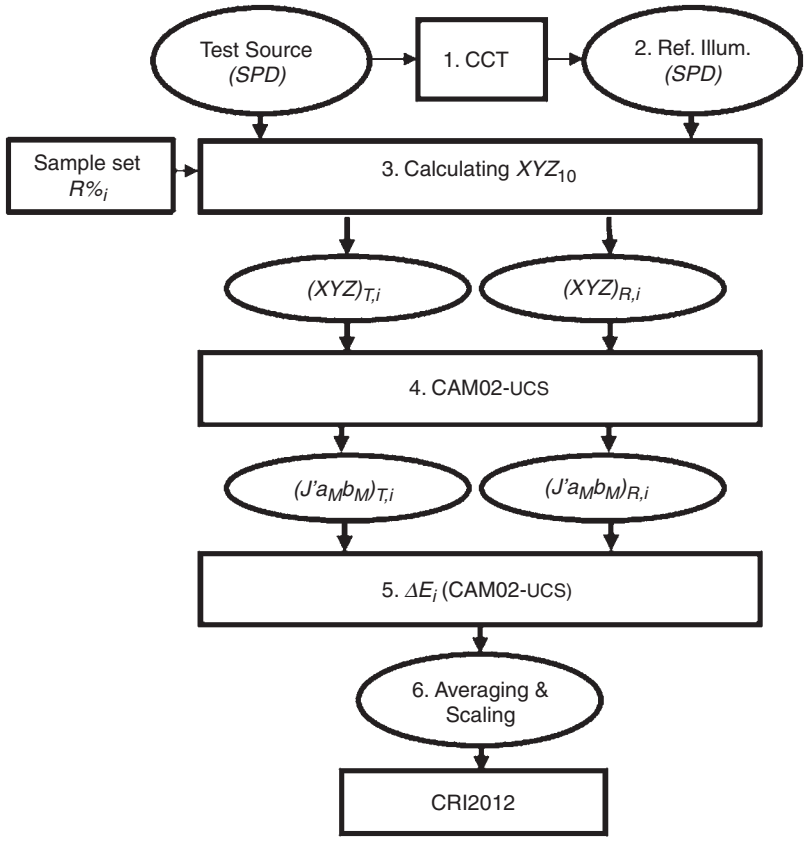

Figure 1 Workflow to calculate CRI2012

daylight phase (CCT $\geq$5000 K) of the same CCT as the test source. An important remark regarding the CIE reference illuminant of equal CCT is that it is not necessarily the most optimum source in terms of naturalness, preference or any other subjective aspects of colour quality.[4,11] Its only purpose is to serve as an objective standard.

Due to its critical nature, the selection of the test samples (Component b) is described in full detail in the Section 3 of this paper.

The colour difference calculation stage in the current CRI has long been in need of improvement, because of its use of the old $U^*V^*W^*$ colour space and the von Kries chromatic adaptation transform. Indeed, updating them to the state of the art in colorimetric models has been suggested several times during the past decades by a number of authors[12,18–25] and CIE Technical Committees on colour rendering.[26–28] Although colour differences calculated in CIECAM02, a CIE endorsed state-of-the-art colour appearance model, have been shown to be able to describe colour fidelity scores reasonably well,[3,17,29] the CIECAM02 space has been further elaborated by Luo and co-workers[29] to more accurately evaluate colour differences of various magnitudes. Because of its ability to handle small to large colour differences and to accurately predict corresponding colours, the CAM02-UCS[30] was

selected for this application (Component c and Steps 4 and 5 in Figure 1). The following parameters of CIECAM02/CAM02-UCS are recommended:

- Background $Y_b = 20$
- Average surround condition with F, N_c, c parameters of 1, 1, 0.69, respectively.
- Luminance of test adapting field: $L_A = 100\,cd/m^2$
- Adaptation factor: $D = 1$.

Another important change from the CIE CRI to CRI2012 is the use of the CIE 10° standard colorimetric observer for the calculation of the test sample chromaticities (Step 3 in Figure 1). It is well-known that the CIE 2° standard colorimetric observer is in error in the blue part of the spectrum. For modern SSL light sources – containing narrow band blue light – this results in a considerable difference between the instrumental and visual colour matches.[31,32] (Note that the CIE 2° standard observer is still used for the calculation of the CCT (Step 1 in Figure 1)).

Step 6 calculates the final CRI2012 index. It includes two parts: the averaging algorithm (Component d) and a new scaling function (Component e).

As suggested by several groups during the past decades,[9,12,23,26,28] the individual colour differences are averaged using a root-mean-square (RMS) average to ensure that the poor rendering of a few samples is adequately reflected in the general CRI score:

$$\Delta E_{rms} = \sqrt{\frac{\sum_{i=1}^{N} \Delta E_i^2}{N}} \qquad (1)$$

where E_i is the CAM02-UCS colour difference for sample number i illuminated by the test source and the reference illuminant; $i = 1$ to 17.

N is the number of test samples, $N = 17$.

Some users have expressed a preference for a CRI metric for which the colour rendering indices cannot go below zero. Although this is not needed in the sense that there is no fundamental meaning associated with the value zero in this case, a 0–100 scale would be less confusing.[12,20,23,28] Most ideas to accomplish this suggest some type of non-linear ΔE to R transformation.[12,17,20,28] In this metric proposal, a sigmoid-type function was adopted as this often better reflects human perceptual responses – which tend to saturate at both extremes of an intensity range – as compared to the typical linear rescaling used to rescale the colour differences in the CIE CRI.[11,33] Using E_i and E_{rms} the individual and average colour fidelity indices are calculated using the following equations:

General CRI:

$$R_{a,2012} = 100 \cdot \left(\frac{2}{e^{k \cdot |\Delta E_{rms}|^{1.5}} + 1} \right)^2 \qquad (2)$$

where $R_{a,2012}$ is the general colour fidelity index,

k is a constant that changes the sharpness of the *sigmoid* curve.

Specific colour fidelity indices:

$$R_{i,2012} = 100 \cdot \left(\frac{2}{e^{k \cdot |\Delta E_i|^{1.5}} + 1} \right)^2 \qquad (3)$$

where $R_{i,2012}$ is the ith individual colour fidelity index.

As for the constant k, lamp manufacturers had the request that their high colour rendering traditional lamps should retain – as far as possible – their current R_a value under the proposed CRI2012 metric. To this end, they were asked to supply their most favoured lamp spectra, so that the k value could be set such that the average $R_{a,2012}$ indices of these lamps equals the average of the current CIE CRI R_a indices of the same group of lamps.[26] For the time being – as not all the manufacturers have sent their spectra yet – the CIE F1 to F12 spectra (following the proposal in

Davis and Ohno[12]) of CIE publication 15[34] have been used to determine the k constant, yielding a value of 1/55.

3. The question of the test samples

The question of whether the spectra of the test samples influence the calculated colour rendering indices was investigated some 20 years ago[18,21,23] but did not result in the adoption of any new sample set in the calculation of the CIE CRI. Due to the small number and non-optimum selection of the test samples, with the introduction of three-band fluorescent lamps and, more recently, solid state lighting (SSL), manufacturers became able to tailor the spectral power distribution of the light source to yield high CRI values, even though for some such cases the visual impression of colour rendering was considered to be poor.[3,4]

The non-optimality of the test samples for the purpose of colour rendering evaluation is illustrated in Table 1, which shows how the R_i values change if different near metameric test samples are used. The metameric samples were selected from a database of over 100000 reflectance samples collected over many years at the University of Leeds. It consists of common everyday materials such as textiles, paints, print, photograph and plastic samples and also some natural colours such as skin, leaves and other types of foliage. As an example, consider the R_i values of CIE Test Sample #1 and a near metameric sample (sample No. 31 in the Leeds database). The CIELAB colour difference under D65 is shown in column ΔE^*_{ab}. Both Test Samples provides an $R_1 = 98$ for the '*4-LED-1 max Ra*' lamp, but for the '*4-LED no yellow*' source CIE Test sample and Sample #31 provide values of 89 and 95, respectively. Contrary to the six-point rise for the latter light source, the score drops from 91 to 76 for the '*3-LED-3*' source. As another example, consider the following. For the *'3-LED-3'* source a switch of CIE test sample #2 with near metameric sample #4 would drop the R_i value from 98 to 78, but for the '*4-LED with yellow*' source the score will go up from 89 to 92. Such differences are observable and should be reflected in the calculation of the CRI. It is apparent that the selection of the sample set is critical. Clearly, any selected set should not encourage 'gaming' of the spectral power distribution of newly developed light sources. One possible way to inhibit gaming could be to use a very large number of reflectance samples. However, this has the drawback of increasing the calculation time of the index. Although such calculations can be performed in a matter of seconds in the modern computer area, nonetheless this might present a problem for many lamp manufacturers who need to do such calculations for every white light source they produce. It is therefore desirable to reduce the number of samples to a minimum. In addition, even if the calculation time would not be an issue, a large number of samples does not necessarily imply that light

Table 1 Some examples of R_i values using near metameric test samples (underlined values are specifically discussed in the section 3)

		3-LED-1	3-LED-2	3-LED-3	3-LED model gamut	4-LED-1 max R_a	4-LED-2 max R(9-12)	4-LED no yellow	4-LED with yellow	Phosphor LED YAG
Samples	ΔE^*_{ab}					R_i				
CIE 1		93.5	77.9	90.5	79.4	97.6	89.7	88.6	98.7	83.2
#31	1.07	77.9	61.5	75.5	66.5	97.9	70.9	94.7	95.7	83.3
CIE 2		89.2	82.5	98.0	88.8	93.7	90.3	89.5	88.7	97.3
#4	2.20	78.5	71.2	77.8	78.0	85.9	84.9	70.6	91.6	91.3

sources cannot be gamed, as many reflectance samples are not independent due to the limited number of common dyes used to produce them. Thus, it would be important to select samples whose reflection spectra cover all real current and future pigment absorption spectra to preclude gaming. An ideal sample set would therefore contain only a limited number of samples, thereby reducing computation time, and would have its spectral features uniformly distributed across wavelength space. Such a sample set was developed for the calculation of the general CRI2012.

In addition, for applications that require more specific information, another, more extended set of real reflectance samples – with uniformly distributed colour coordinates – was also developed. This set enables a detailed visualization of the colour shifts produced by a change from the reference illuminant to the test source. Both sets, respectively named the *HL17* and *Real* set, will be introduced here.

3.1 HL17 sample development

3.1.1 Spectral Sensitivity

To better understand the importance of a sample set with uniformly distributed (non-localized) spectral features, it is helpful to consider the 'spectral sensitivity' of the CRI calculation – i.e. responsiveness of the CRI score to localized spectral features. The concept of spectral sensitivity was first introduced by Smet and Whitehead in a paper published in the proceedings of CIC19, but is discussed here as well for matters of completeness and clarity as it is key to development of the new HL17 sample set.[35]

Consider, for example, two illuminant spectra: (a) a 4500 K Planckian radiator and (b) the same function perturbed by adding a smooth, symmetrical function $p(\lambda - \delta)$ that integrates to 0 so it is zero outside the *perturbation zone* $\pm\Delta$, and has a maximum magnitude of ε; δ is an adjustable spectral offset value that determines the location of the perturbation zone. The two spectra are illustrated in Figure 2 (note that for the purpose of clarity the perturbation zone has been drawn with an exceedingly large width of approx. 100 nm, the actual spectral sensitivity calculations used perturbation zone values Δ of 10 nm and 20 nm).

When a coloured test sample of given spectral reflectance is illuminated by the two spectra – as is done in the CRI calculation – the

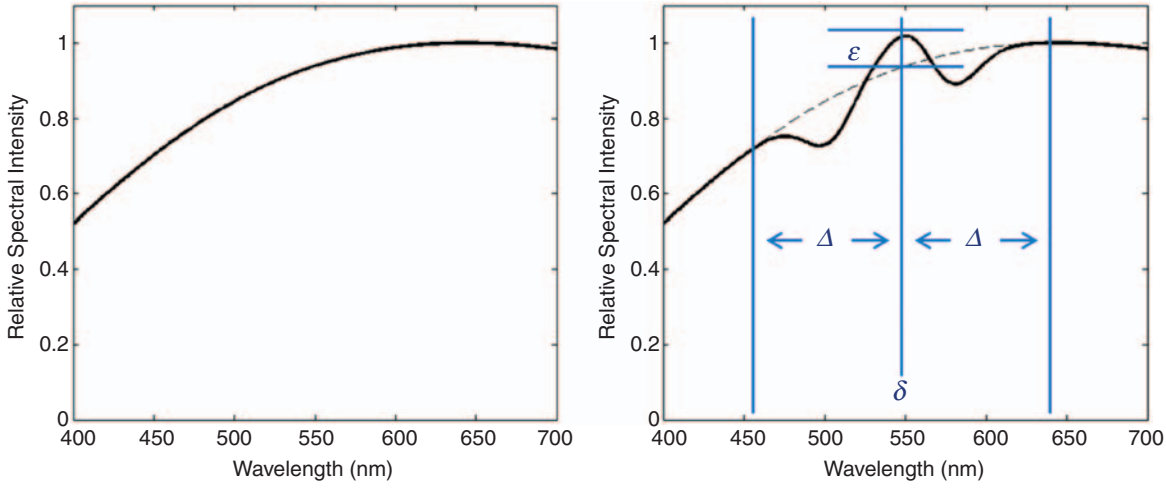

Figure 2 The (a) 4500 K Planckian spectrum and (b) its perturbed spectrum

colour difference of the sample under these spectra will be proportional to the size ε of the perturbation (for sufficiently small ε). This constant of proportionality, which in this paper is called the spectral sensitivity of the CRI, will generally be a function of the central wavelength δ the specific perturbation function $p(\lambda - \delta)$ and the underlying spectral response characteristics of human vision. In order to minimize the possibility of tailoring the spectral distribution in ways that improve the score but do not actually improve fidelity, the spectral sensitivity should be fairly smoothly distributed across the visible spectrum, while not critically depending on the specific details of the reflectance samples used.

To better understand how these goals constrain the selection of the reflectance sample set, it is helpful to consider how individual samples contribute to the overall spectral sensitivity. Obviously, for wavelengths outside the perturbation zone the reflectance values are irrelevant because at those wavelengths the two illuminants are identical. Samples with a uniform reflectance within the perturbation zone will also not contribute to the colour difference between the two illumination states because the perturbation function averages to zero. In fact, because of the symmetric shape of the perturbation function, even test samples with a linear gradient within the perturbation zone will yield little colour difference. Therefore, what is required to generate a significant colour difference between the two illuminants is curvature – a non-zero magnitude of the second derivative with respect to wavelength – in the sample spectral reflectance function $R(\lambda)$ As an example, consider the eight spectral reflectance functions used in the basic CIE CRI calculation.[1] Figure 3 shows the average magnitude of curvature of these 8 reflectance samples (which have been smoothed with a 20 nm moving average to reduce the influence of noise in the data). The average magnitude of curvature is very non-uniform, with the shape determined by the arbitrary spectral details of the individual reflectance samples. Since replacing each of the eight CIE CRI samples with a metameric sample would result in a quite different sensitivity curve, it is clear that this curve is

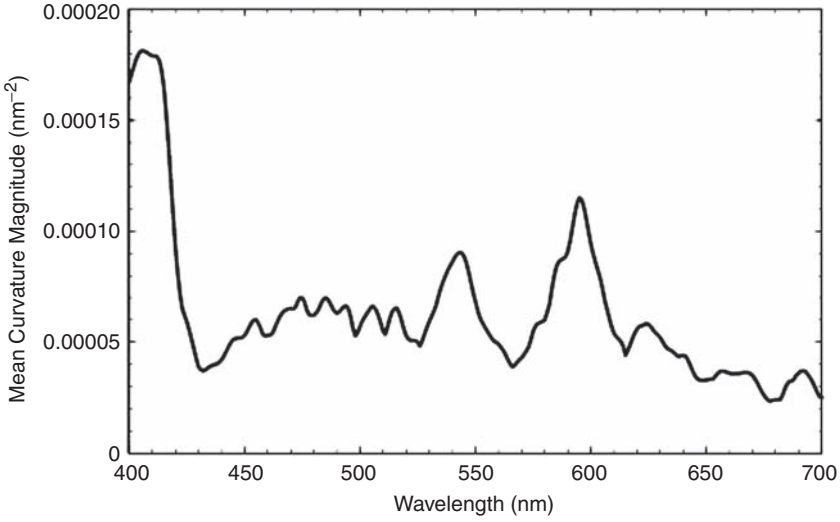

Figure 3 The average magnitude of curvature for the eight basic International Commission on Illumination (CIE) colour rendering index (CRI) reflectance samples

not purely a result of the characteristics of the human visual response, although it should be.

The likely error in the CIE CRI arising from the use of such a sample set can be estimated from several perspectives as described herein. It appears likely that this error is significant for some LED-based sources, some tri-band fluorescent lamps and possible future lamps featuring narrow-band spectral details[3] made possible by quantum dot technology. This observation agrees with anecdotal reports of dissatisfaction with the current CIE CRI when applied to such sources. Therefore, it seems there is an urgent need for an improved sample set that yields the correct spectral sensitivity characteristic for the CRI.

It is clear from the above that any selected sample set containing only a limited number of already-existing spectral reflectance functions is very unlikely to yield the appropriate smooth spectral sensitivity response. However, increasing the number of samples might reduce this problem, since features from individual spectra would tend to average out statistically.

3.1.2 Wavelength-shifted Leeds 1000 set

At the University of Leeds the reflectance spectra of 100000 real samples of all kinds of natural and artificial surfaces have been collected. The direct use of these as the CRI calculation sample set is impractical for two main reasons. First, it would be computationally challenging for reasons mentioned earlier. (Interestingly, from statistical considerations, only 1000 truly random samples ought to be enough to sufficiently smooth the spectral sensitivity.) Second, the samples contained in the Leeds 100000 set were not selected with the goal of uniformly representing the 3D-continuum of possible reflectance spectra. Because of the set's characteristics, a simple random sampling would not be an appropriate way to reduce the number of reflectance spectra to a more manageable number. A better approach would be to randomly sample the spectra in such a way that their colour coordinates (e.g. under the D65 or the equal energy white illuminants) would be fairly uniformly distributed in the $L^*a^*b^*$ colour volume. However, when such a sampling was tried, a more subtle problem became apparent: the non-uniformity of the mean curvature depicted in Figure 3 was not substantially reduced. If the samples were independent, one would expect the magnitude of the average of such a random selection to decrease with the inverse square root of the number of samples. The non-uniformities shown in Figure 3 should therefore have been reduced by a factor of about $\sqrt{8/1000} \cong \frac{1}{11}$ However, as can be seen by comparison of Figure 3 with Figure 4, this was not the case for the basic Leeds samples. This is most likely due to their mathematically non-independent nature. Most coloured surfaces are composed of various ratios of a reasonably small set of common dyes that have spectral features in fixed locations contributing to the curvature at the same wavelengths in the spectrum. Therefore, a large number of such dependent samples would not result in the degree of cancellation that would be expected statistically for independent samples.

An alternative would be to use a Monte Carlo approach for mathematically generating simulated spectra – one that did not differentiate between spectral locations. However, such generated spectra might not sufficiently resemble the shape of natural reflectance spectra.

Ideally, the reflectance spectra would contain only spectral features – and with approximately the same frequency distribution – as found in real samples. Such an ideal sample set can be created by adapting the Leeds spectra in a wavelength-shifted and composite manner as follows:

- First, a subset of 10000 spectra with an approximately uniform density in $L^*a^*b^*$ space – to reduce selection bias in sampling

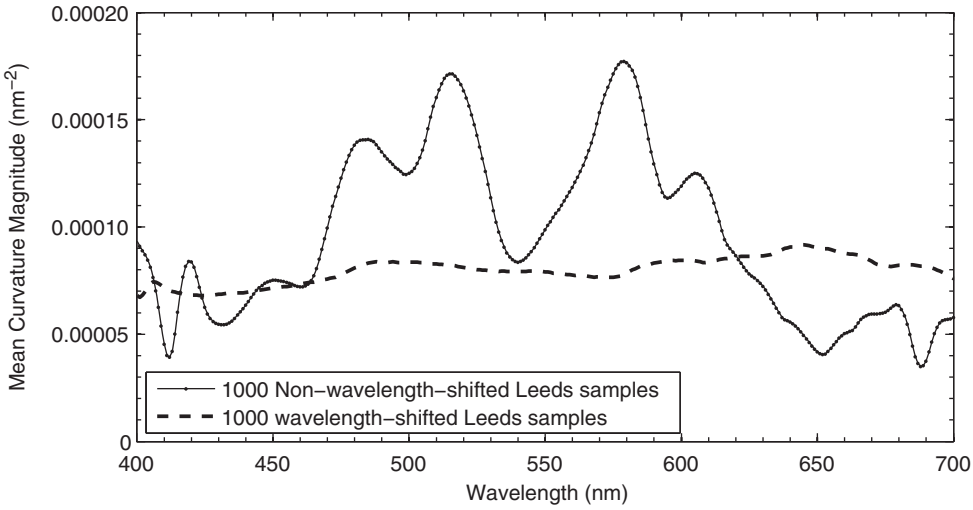

Figure 4 The average magnitude of curvature is not improved by using a random selection of 1000 Leeds samples but is improved by selecting wavelength-shifted Leeds samples

spectral features – is selected from the 100000 set.

- Second, these spectra were truncated to a 450–700 nm range as some artefacts were noted in some of the spectra outside this range.

- Using these 10000, the following steps were repeatedly carried out in 1250 independent cycles:

  1) First, a seed was generated by randomly selecting a spectrum from the 10000 subset.

  2) Next, a spectrum was selected that best matched the end slope and spectral reflectance of the seed, and was stitched onto its end. This provided a smooth, fairly continuous, natural-looking connection.

  3) This match/stitch process was repeated three times, producing a continuous spectrum long enough for the sampling operation described next.

  4) Selected spectral subsets were copied, to populate in 5-nm intervals, the range 380 to 780 nm (81 values) in an output file. This was done 80 times, and on each successive sampling, the sample interval was shifted by 5 nm. This ensured that every wavelength was treated the same as every other. In this way, 80 full spectra were created for each of the 1250 stitched spectra, yielding a total of 100000 final wavelength-shifted stitched spectra.

- Each of these 100000 spectra was slightly smoothed by 20 nm moving average, which did not change the spectral shapes but did remove any discontinuity in curvature at the connection points. The resulting spectra have *real* spectral features and look very natural.

- This final set of 100000 wavelength-shifted spectra was used for the selection of 16 subsets of 1000 spectra:

  1) For each of the 100000 spectra, a number was calculated for use in the next step

described below. The number is based on the $L^*a^*b^*$ values of the spectra. For each spectrum, the number is intended to represent the approximate local density of such spectral coordinates in $L^*a^*b^*$ space. It is approximately equal to the number of spectra whose $L^*a^*b^*$ coordinates lie within a 20-unit distance of the coordinates of the spectrum in question. For reasons of computational efficiency, the number was determined for each spectrum by studying the location of the $L^*a^*b^*$ values of a randomly selected subset of 100 of the other spectra. To minimize statistical noise in such a sample, the volume of interest was not represented as a discrete spherical volume centred on the $L^*a^*b^*$ value of the spectrum in question; instead, the contribution from each of the 100 spectra was weighted by a Gaussian decay function based on the distance from the centre, with the decay parameter selected to drop to 1/e at a distance of 20 units.

2) In order to achieve a uniform selection in $L^*a^*b^*$ space while ensuring reasonable diversity in multiple selections, a selection probability was calculated for each point, based on the density number determined above. By trial and error it was found that it is best to set the selection probability to 0.0038 divided by the above-calculated density number or to 0.05, whichever is smaller. This created a reasonably uniform distribution in $L^*a^*b^*$ space with no unusual discontinuities and good coverage of the colour volume.

3) A total of 16 random selections were carried out to select 16 different subsets of 1000 spectra.

- For each of the 16 subsets, the spectral sensitivity was calculated using two sets (with perturbation zones: $\Delta = 10$ nm and $\Delta = 20$ nm) of perturbed CIE reference illuminants of 3000 K, 4100 K and 6500 K. For each set the perturbation zone was shifted (using δ in 5-nm intervals.

- Finally, of the 16 subsets, the one that had a spectral sensitivity closest to the mean (of the 16 calculated spectral sensitivities) was selected as the final most representative 1000 wavelength-shifted spectrum set.

As expected, the curvature distribution – shown in Figure 4 – of the final wavelength-shifted 1000 Leeds set is much smoother than the non-wavelength-shifted set of 1000 randomly selected samples. Importantly, and not surprisingly, the stitched wavelength-shifted spectra look similar to the original Leeds spectra from which they are formed, as compared in Figure 5.

3.1.3 HL17 set

As explained before, it is important to reduce the size of the sample set as much as possible. A first simple approach would be to randomly sample the set of 1000 wavelength-shifted spectral reflectance functions. However, as expected and as can be seen in Figure 6, reducing the sample size leads to an increase in non-smoothness of the curvature. At sample sizes that are practical for colour rendering calculations from a manufacturers' point of view, the spectral uniformity is too erratic.

An alternate approach for obtaining a set of reflectance samples that produce a smooth mean curvature and undistorted spectral sensitivity would be to design a set of reflectance spectra with a fairly smooth spectral feature that shift through the spectrum, from one sample to the next. With such an approach the number of samples in such a set could potentially be small while yielding sufficient uniformity of spectral sensitivity. The first

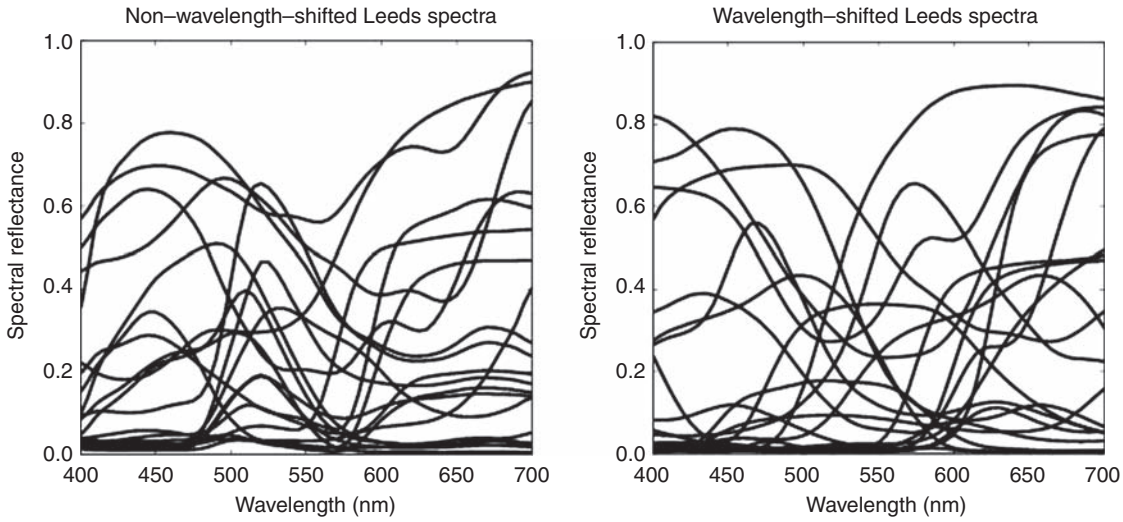

Figure 5 Comparing 20 randomly selected spectra from the original Leeds set and from the stitched wavelength-shifted spectra generated from the original set

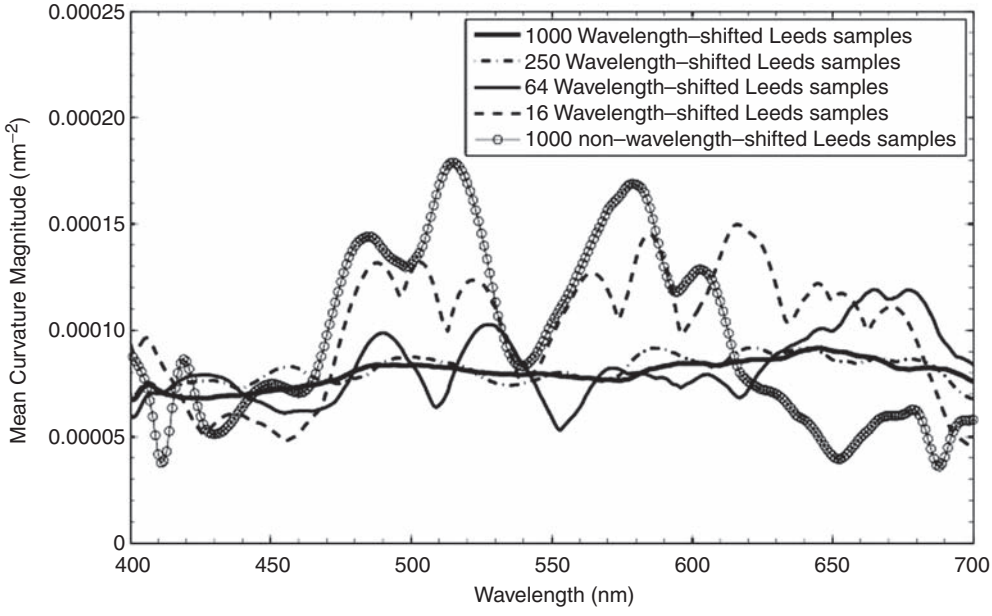

Figure 6 The influence of reducing the sample size of a set of wavelength-shifted reflectance samples on the smoothness of the mean magnitude of spectral curvature

shape that came to mind was a Gaussian function. However, due to the non-smooth characteristic of the absolute value of its second derivative – illustrated in Figure 7 – it is not possible to construct a spectrally uniform sample set composed of only a small number of (shifted) Gaussian reflectance spectra.

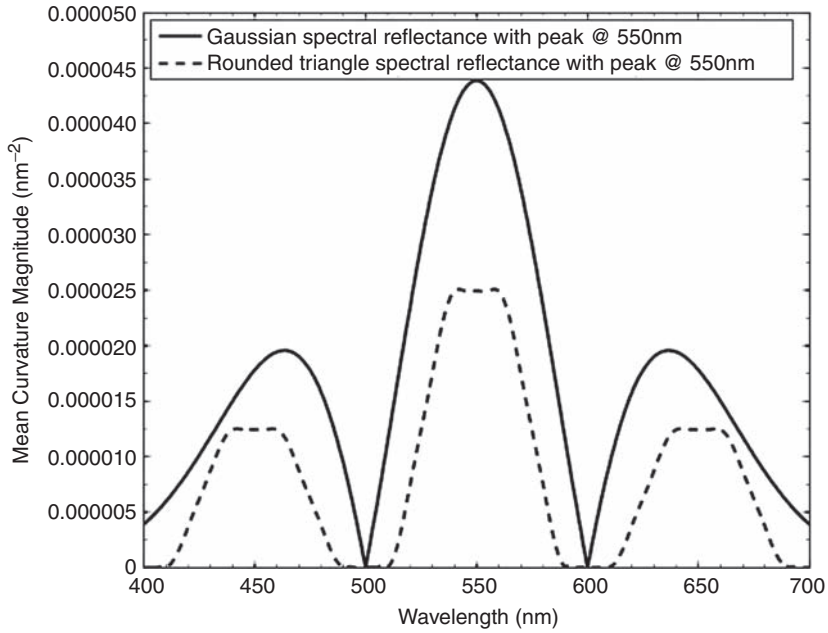

Figure 7 Mean curvature magnitude of a Gaussian-shaped spectral feature centred at 550 nm vs that of a rounded triangle centred at the same wavelength

A much better candidate for such an approach is a suitably rounded triangle. By suitably interleaving of the curvature features, shown for such a shape in Figure 7, such spectra can result in a more uniform spectral sensitivity. A model of the rounded triangle used in the optimization is shown in Figure 8. In total six parameters were optimized, five related to the (individual) rounded triangle – forward slope W_f, rear slope W_r, minimum R_{min} and maximum R_{max} reflectance and degree of rounding s – and one related to the sample set as a whole – the spacing of the reflectance samples across the spectrum, $\Delta\lambda_c$.

The goal of the optimization was to find a minimal simple sample set that emulates the 1000 wavelength-shifted samples set with sufficient accuracy in terms of match (evaluated as the root-mean-square-error (rmse)) in spectral sensitivity (as described in Section 3.1.1) and rmse-match in $\Delta E_{rms,cam02ucs}$ between 53 common commercially available lamps and their CIE reference illuminant. The Pearson correlation of the *CRI2012* scores with human observation of colour rendering differences was also considered. The model parameters were optimized by calculating the Pareto optimal boundary of the former three criteria using the NSGA-II genetic algorithm.[36] In the selection of the final solution, of the three criteria, the first one was considered the most stringent and the latter the least (only visual data from one fidelity experiment[5] was available).

Careful examination of some of the preliminary optimization results showed that the general model of an asymmetric rounded triangle could be constrained to the more simple symmetric model without a significant loss in the quality of the fit to the 1000 wavelength-shifted sample set.

In studying the wavelength-shifted 1000 set, it was found that the average spectral reflectance of the 1000 samples increases

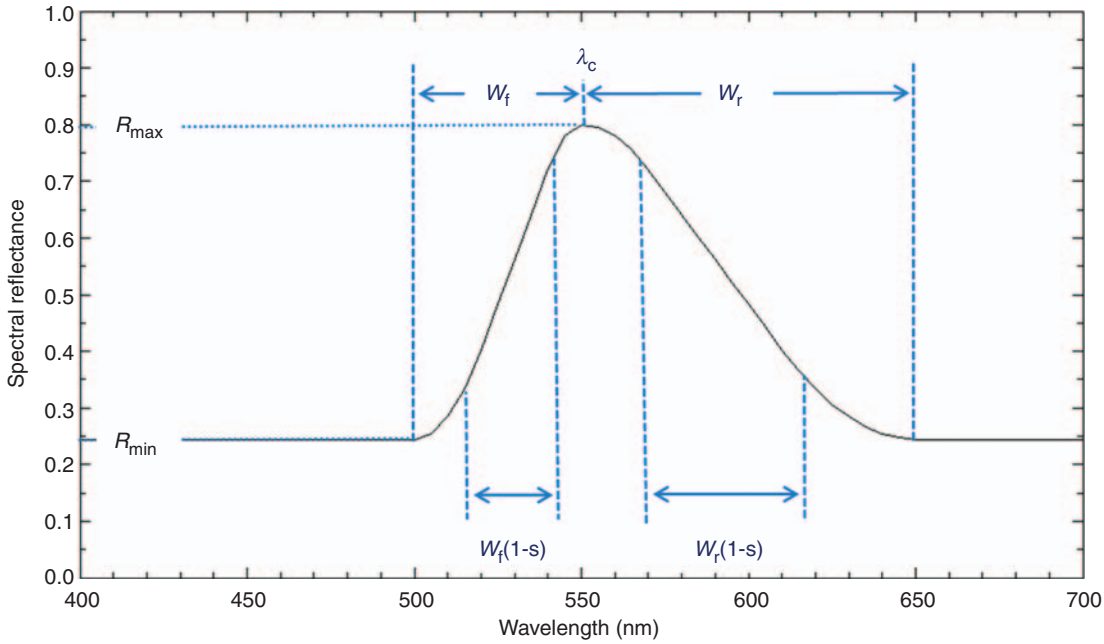

Figure 8 Rounded triangle model. The general form of the function is an asymmetrical rounded triangle with central wavelength λ_c, forward slope width W_f and rear slope width W_r, spectral reflectance values ranging from R_{min} to R_{max}, and with quadratic rounding applied at the corners leaving a fraction of the slope, 1-s, uncurved

gradually with wavelength throughout the spectrum – presumably due to the somewhat irregular shape of the colour volume in $L^*a^*b^*$ space. When the mathematically generated sample set was adjusted to correspondingly increase across the spectrum, the quality of the optimization increased from very good to excellent.

A natural extension of the rounded triangular sample set would be to include the complementary reflectance samples – i.e. the samples obtained by subtracting the spectral reflectance values of the regular set from unity. This would allow for bright, saturated purple samples, for example. Although intuitively this seemed like a good idea, this doubling of the number of samples did not offer any extra benefit with regard to the goal of the optimization. This is not surprising considering the nature of the mathematical calculations involved. Therefore, in recognition of the desire for simplicity, the complementary reflectance spectra are not included in the final mathematical sample set.

For the time being, the resultant set contains only 17 spectra and is called the '*HL17*' set to designate it has 17 spectra that are constructed to avoid gaming and are suitable for the theoretical testing of the colour rendering model. Figure 9 shows the plot of 'HL17' spectra. They can be labelled by an integer i ranging from 1 to 17. The central wavelengths vary from 350 nm to 750 nm in 16 intervals of 25 nm, according to $\lambda_{ci} = 550\,\text{nm} + 25 \cdot (i - 9)$ The spectral reflectance of the peak wavelength of each sample is given by $R_{\max,i} = 0.5 + 0.001 \cdot (\lambda_c - 550)$ with λ_c in nanometers. The forward slope W_f and rear slope W_r are identical – a symmetrical rounded triangle – and equal to 100 nm. The rounding, s, is set to 0.5. The formula for the

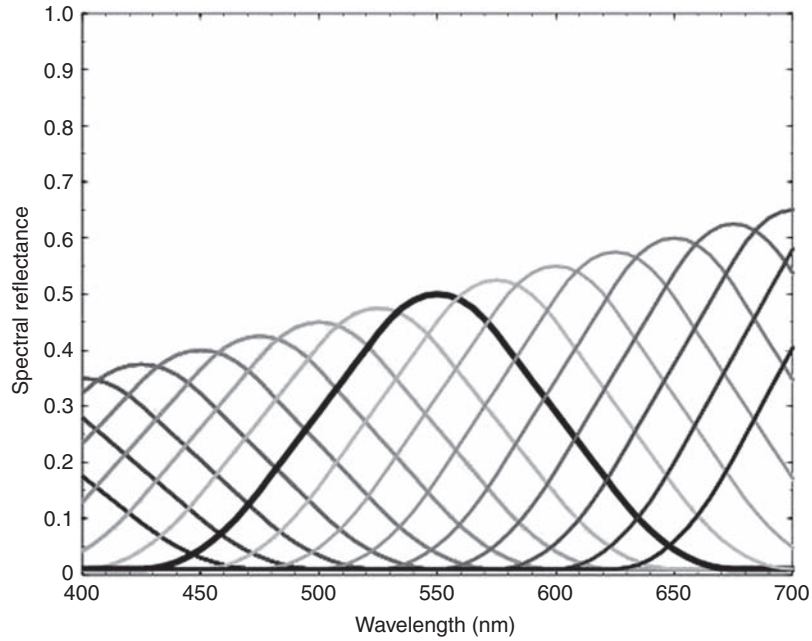

Figure 9 The HL17 set of spectra developed for the purposes of calculating an average colour rendering difference between two illuminant spectral power distributions that avoids gaming. As an example, one of the spectra is shown in a bold black solid line

reflectance values $R(\lambda)$ depends on the value of $|\lambda - \lambda_{ci}|$:

a) if $|\lambda - \lambda_{ci}| \geq 125\,\text{nm}$:

$$R(\lambda) = 0.01 \tag{4}$$

b) If $125\,\text{nm}|\lambda - \lambda_{ci}| \geq 75\,\text{nm}$:

$$R(\lambda) = 0.01 + (R_{\max,i} - 0.01) \cdot \left[\frac{1}{8750\,\text{nm}^2}\right] \cdot (125\,\text{nm} - |\lambda - \lambda_{ci}|)^2 \tag{5}$$

c) if $25\,\text{nm}|\lambda - \lambda_{ci}|$:

$$R(\lambda) = 0.01 + (R_{\max,i} - 0.01) \cdot \left[\frac{8}{7} - \frac{2}{175\,\text{nm}} \cdot |\lambda - \lambda_{ci}|\right] \tag{6}$$

d) If $25\,\text{nm}|\lambda - \lambda_{ci}|$:

$$R(\lambda) = 0.01 + (R_{\max,i} - 0.01) \cdot \left[1 - \frac{1}{4375\,\text{nm}^2} \cdot |\lambda - \lambda_{ci}|^2\right] \tag{7}$$

Because of the mathematical form of these spectra and the selected design parameter values, the average magnitude of curvature of these spectra is very uniform. This is a valuable characteristic of a sample set – for the purpose of calculating the colour rendering difference between two illuminant spectral power distributions that is not susceptible to gaming – that is not achievable with any small set of naturally occurring reflectance spectra.

The results of the optimization (using the symmetric triangular model with the $R_{\max}$ adjusted to increase across the spectrum) showed that it is possible to emulate the wavelength-shifted Leeds 1000 set – with essentially no loss of information – using a vastly smaller, but mathematically regular, spectral reflectance set:

The primary goal of the optimization was to match as closely as possible the spectral sensitivity of the wavelength-shifted Leeds 1000 set. The match, evaluated as a

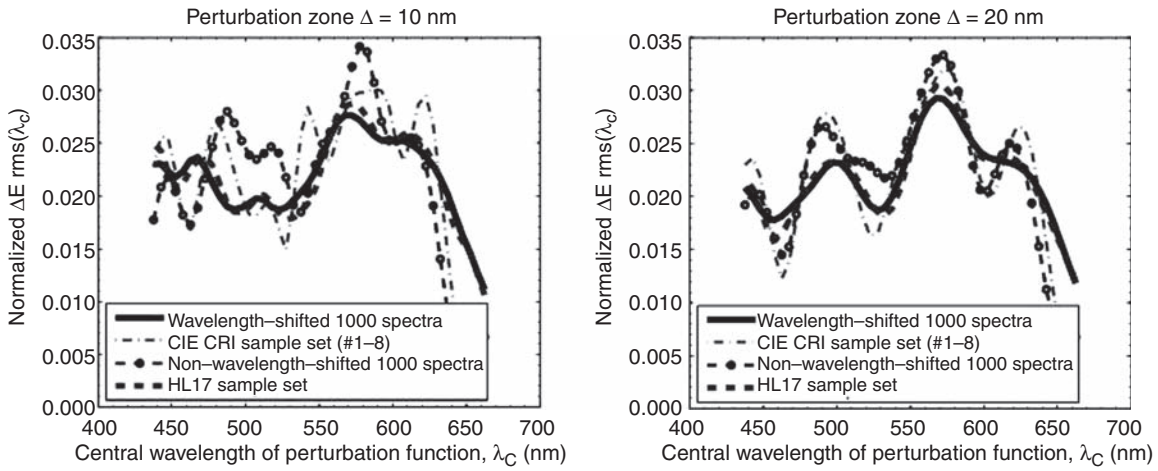

Figure 10 Plot of the normalized spectral sensitivity of different sample sets to small spectral perturbations, as a function of the central wavelength of the perturbation. The HL17 set matches the sensitivity of the wavelength-shifted 1000 set well. In contrast, the results for the CIE CRI sample set and a non-wavelength-shifted 1000 sample set match poorly

root-mean-square-error (normalized with the area under the spectral sensitivity curve) between the spectral sensitivity of the HL17 and wavelength-shifted Leeds 1000 sets, was found to be very small for the two selected perturbation functions: 0.0009 for $\Delta = 10$ nm and 0.0007 for $\Delta = 20$ nm. For comparison, the CIE CRI sample set has values of, respectively, 0.0046 and 0.0037 – which are about 5 times worse. The very good agreement between the spectral sensitivities of the wavelength-shifted 1000 Leeds and HL17 set can be clearly seen in Figure 10. The spectral sensitivity of the CIE CRI sample set and the non-wavelength-shifted 1000 set is also shown.

The second criterion adopted in the optimization process was the match in colour rendering predictions (between the HL17 and the wavelength-shifted 1000 samples) for a set of 53 common commercially available light sources (broad-band and narrow-band). The mean-absolute-error (MSE) of the ΔE_{rms} colour difference between the optimized HL17 set and the wavelength-shifted 1000 set was 0.36 units, which amounts to a $R_{\mathrm{a},2012}$ difference of about 0.86 units. The MSE of the $R_{\mathrm{a},2012}$ difference between the HL17 set (or the wavelength-shifted 1000 set) and the CIE CRI-8 test samples was 3.4 units. To put the scores of the proposed new CRI better in perspective, the analysis of the colour rendering predictions was extended to a total of 139 broad-band and narrow-band light sources. Respectively, for the ΔE_{rms} and the $R_{\mathrm{a},2012}$, an MSE of 0.37 and 1.2 was observed between the HL17 and the wavelength-shifted 1000 sets. The fact that the colour rendering assessments of the HL17 spectra closely match those of the wavelength-shifted 1000 samples for a large number of sources, not included in the optimization process, indicates that the optimized HL17 is robust and successfully emulates the behaviour of the larger set in general. In Figure 11, the $R_{\mathrm{a},2012}$ indices obtained with different spectral sets are compared to those obtained with the wavelength-shifted 1000 set. It is immediately clear that the size of the sample error is significantly reduced by using the HL17 set. The coefficient of determination, R^2, for the HL17 set was 0.99, while that of the CIE

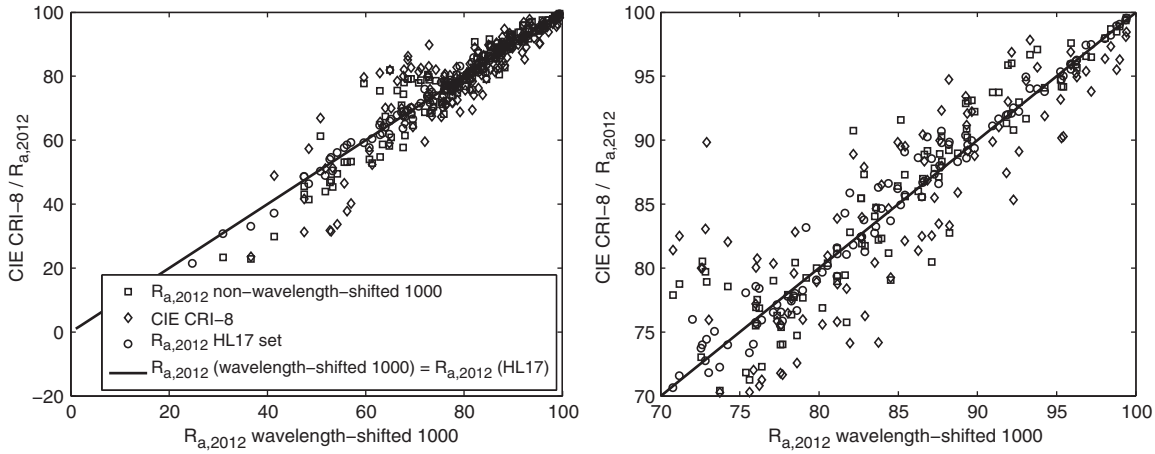

Figure 11 Two plots, at different scales, of $R_{a,2012}$ scores calculated using the new colour-rendering calculation engine with different spectral reflectance sets. The horizontal axis represents the value obtained using the wavelength-shifted Leeds 1000 set. On the vertical axis the results for the CIE CRI-8, the non-wavelength-shifted 1000 and the HL17 sample sets are plotted. The HL17 shows excellent correlation with the wavelength-shifted 1000 set

CRI-8 was 0.90. The non-wavelength-shifted 1000 set, sampled directly from the Leeds spectral data base, had an R^2 of 0.91.

Of particular interest to the lighting industry is how the proposed new CRI values compare to those of the CIE CRI. To this end, the MSE of the $R_{a,2012,HL17}$ with the CIE CRI was calculated and was found to be 5.9. This might seem fairly large, but it should be noted that just updating the colour space and chromatic adaptation used in the CIE CRI with the CAM02UCS colour space – while keeping the CIE samples #1–8 and the linear scaling function – introduced an MSE of 3.8 units. For most traditional (broad band emission spectrum) sources, the difference of CRI scores between the CIE CRI and the proposed CRI2012 with the above HL17 dataset is small. However, some sources do exhibit a significant discrepancy, showing larger CIE CRI values than those predicted by the CRI2012. A close examination of those sources, which have their CIE CRI values predominantly between 75 and 85, reveals that most of these sources are tri-band fluorescent sources with a CCT below 4000 K. The MSE between the CIE CRI and the $R_{a,2012,HL17}$ for these warm white tri-band sources was 10 units. In contrast, the MSE for tri-band sources with a CCT above 4000 K was only 2 units. Historically, anecdotal evidence from visual observations suggests that warm white tri-band fluorescent lamps do indeed deserve a lower CRI than they are currently awarded.

This brings us to the last criterion used in the optimization of the design parameters of the HL17 sample set, the correlation of the CRI values of the proposed new CRI, $R_{a,2012,HL17}$, with visual observation of colour rendering differences under several different light sources. Unfortunately, the data of only one *fidelity* study[5] was available. It should be noted that in any such study, necessarily only a small number of real samples are viewed, and therefore the same spectral sensitivity issues mentioned above will impair the accuracy of such limited visual experiments. Nevertheless, it is appropriate to consider what has been observed so far. In the visual experiments, colour rendering differences were obtained for sets of light sources

with five different CCTs: 2700 K, 3000 K, 4100 K, 5000 K and 6500 K. The average Pearson correlation, calculated according to the method of Hunter-Schmidt,[37] was found to be 0.90 ± 0.03. It is possible that most of the remaining deviation was a result of noise in the observations, a factor that would be consistent with the fact that similar scores are obtained with the CIE CRI metric. Obviously, more extensive fidelity studies are required to confirm these preliminary results as well as the anecdotal evidence regarding the lower $R_{a,\mathrm{nCRI}}$ values of the warm white tri-band fluorescent lamps.

Finally, since the HL17 sample set matches the behaviour of the wavelength-shifted 1000 set very well, it is interesting to consider the calculated CAM02-UCS $J'a_Mb_M$ values for these spectra under D65 illumination. They are shown in the three plots in Figure 12.

It is clear that the path demarked by the HL17 is reasonably representative of the larger wavelength-shifted 1000 set. Although there are no magenta samples in the HL17 set, this is not a requirement from the standpoint of calculating a *general* CRI. The spectral features present in any magenta test sample constructed from two rounded triangles, one red and one blue, would already have been sampled by those red and blue samples. The sole purpose of the HL17 is to be able to quickly assess the *general* colour rendering properties of a light source in a way that is insensitive to gaming – with smooth spectral sensitivity – and corresponds well with the colour rendering predicted by the wavelength-shifted 1000 set and human observation of colour rendering differences between two illuminants.

If one is also interested in more hue-specific measures of the colour rendering properties of a light source, the use of the real set discussed in the following section is suggested.

3.2 Development of a real set

The aim of the uniform *real set* is to provide more detailed (hue-specific) information on the colour shifts to be expected across colour space when changing the illumination of real world samples from one source (reference illuminant) to another.

Again, the set of 100000 reflectance samples, accumulated by the University of Leeds, was used as a starting point. First, the 100000 set was sampled in CAM02UCS space at a medium lightness ($L^* = 50$), high chroma ($C^*_{ab} \approx 30$) and hue spacing of 10 degrees for reflectance samples showing a high colour constancy (HCC)[38] under three illuminants (D65, F11 and A). These samples illuminated

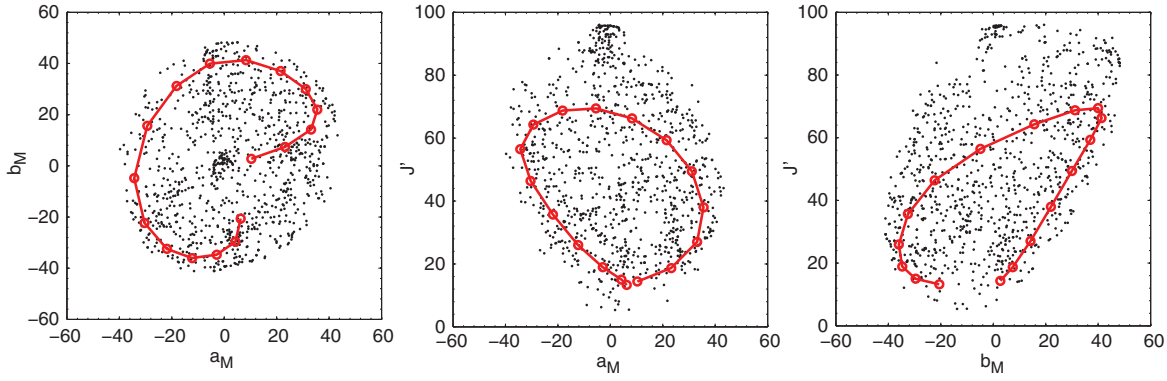

Figure 12 Distribution of the colour coordinates of the HL17 spectral set under D65 illumination in CAM02-UCS space (solid line). The colour coordinates under D65 of the wavelength-shifted Leeds 1000 set are also shown

by the CIE D65 illuminant are shown as diamonds in Figure 13. Next, the 100000 set was sampled for reflectance spectra that are near metameric to the first set under D65. In addition, the metameric samples were required to exhibit a significantly lesser degree of colour constancy (LCC) than the HCC under illuminants D65, F11 and A: the required minimum difference between the HCC and LCC CAM02UCS colour difference between CIE illuminant pairs D65-A and D65-F11 was set to 0.6 ΔE units (for larger values, the degree of metamerism substantially dropped). Colour constancy of a sample under illuminant A or F11 was assessed as the CAM02-UCS colour difference between the sample's colour appearance under the illuminant and under D65. The LCC reflectance samples are shown as squares in Figure 13. Third, the sampled colour volume was further filled by selecting HCC and near-metameric LCC samples with low chroma ($C_{ab}^* \approx 15$) at lightness values $L^* = 30$, 50 and 70 and a hue spacing of 20°. This gives a sample number of 180. The results in terms of achieved metamerism and colour constancy of the HCC and LCC sample sets are presented in Table 2. An additional analysis with the Mann-Whitney U

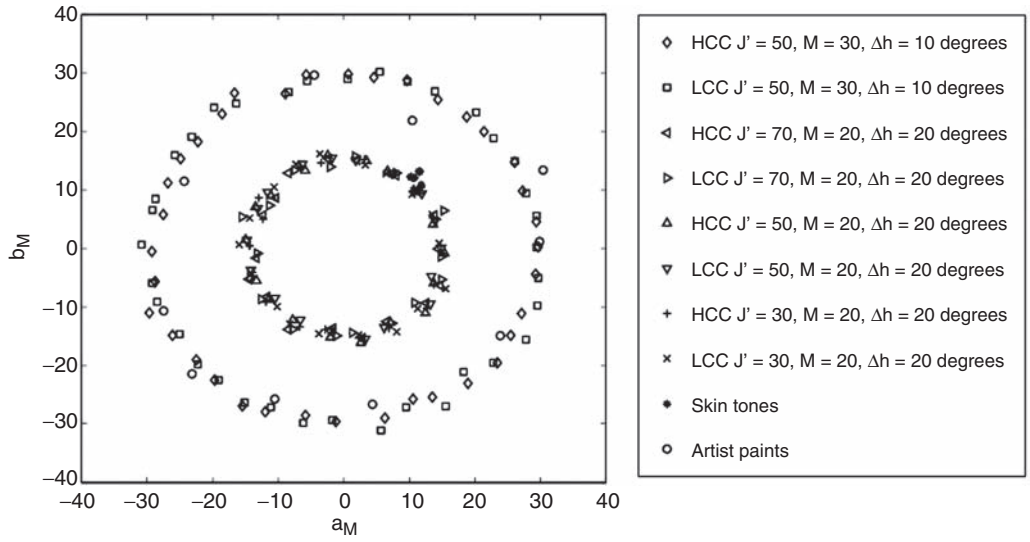

Figure 13 CAM02-UCS coordinates a_M, b_M of the 90 high colour constancy (HCC) and 90 low colour constancy (LCC) metameric samples, and 4 × 5 skin tone and 10 artist paint samples

Table 2 Final achieved performance of the 90 high colour constancy (HCC) and 90 low colour constancy (LCC) reflectance samples in terms of D65 metamerism and colour constancy selection criteria

	Degree of metamerism, $\Delta E_{\text{HCC-LCC}}$			Degree of colour constancy, $\Delta E_{\text{D65-Illuminant X}}$				HCC-LCC difference in degree of colour constancy, $\Delta E_{\text{HCC,D65-Illuminant X}} - \Delta E_{\text{LCC,D65-Illuminant X}}$	
	D65	A	F11	HCC, A	LCC, A	HCC, F11	LCC, F11	A	F11
Average	1.62	2.92	3.26	3.27	4.72	2.14	4.34	1.44	2.20
Max	4.33	8.50	7.71	7.23	7.89	5.58	11.02	4.80	7.10
Min	0.28	0.51	0.42	0.70	1.88	0.39	1.10	0.61	0.69
Median	1.55	2.87	2.98	3.10	4.48	1.86	3.95	1.19	1.84

test showed that for both illuminant A and F11, the ΔEs for the HCC set differed significantly (two tailed $p<0.0001$) from those of the LCC set, confirming the higher degree of colour constancy of the HCC set compared to the LCC set.

Additionally, five representative skin tones of major ethnic groups (African, Caucasian, Hispanic, Oriental and South Asian) were selected. To give a somewhat higher weight to these complexion colours, they were added to the sample set four times. In addition, ten reflectance spectra representing typical fine art paint were added as well; bringing the total number of samples to 210. The chromaticities are shown in Figure 13. This final set of real samples, which includes both natural and artificial materials, gives a good uniform coverage in lightness, chroma and hue range, and can provide more detailed information on colour shifts required for some applications. Note that this set is not intended to be used for the calculation of a *general CRI* as it exhibits a substantially non-uniform spectral sensitivity (samples are non independent due to the limited amount of commonly used dyes).

4. Conclusions

An update to the CIE CRI is proposed. First, the outdated colour space and chromatic adaptation transform would be replaced by the CAM02-UCS colour appearance space. Second, the arithmetic averaging and linear rescaling of the colour differences would be replaced by an RMS average and a sigmoid function, respectively. To enable a smooth transition from the CIE CRI to the suggested CRI2012, the scale of the proposed new system is set such that for CIE illuminants F1 to F12 the average R_a and $R_{a,2012}$ are the same. Third, the previous CRI set of 14 sample reflectance functions would be changed in two ways. The first 8 functions, which are used to calculate the general index R_a, would be replaced by a mathematically derived set containing 17 samples. For calculation of the special indices, instead of using the original 14 CRI functions, we propose to use a new set of 210 real reflectance functions, having high and low colour constancy, to provide more detailed hue-specific information. The 17-sample mathematical set has been shown to yield a spectral sensitivity function for the general index that is smoothly varying (which is important to preclude gaming of light source spectra), and matches that of a larger set of 1000 wavelength-shifted samples constructed from the Leeds data base of 100000 spectral reflectance functions and selected to uniformly represent the entire colour volume. The difference in colour rendering scores – calculated using the updated formulas – between the HL17 and the wavelength-shifted 1000 sets is negligible (average r.m.s. difference less than 1 unit). However, the differences with the CIE CRI scores are not negligible, suggesting that the proposed improvement is indeed needed. Fortunately, for traditional, broad band lamps the changes are modest and non-disruptive. Generally, most of the change would arise not from improving the test samples but from updating the colour difference calculation from $U^*V^*W^*$ to CAM02-UCS. However, in the specific case of low CCT (warm white) tri-band fluorescent sources, substantially lower CRI2012 scores occur. This is in agreement with the anecdotal reports of visual experience, suggesting these sources may indeed be overrated by the current CIE CRI. Furthermore, the proposed new CRI2012 correlates well ($r=0.90$) with visual observations of colour rendering differences between a number of light sources. For all these reasons, the CRI2012 is proposed as an update to the CIE CRI, in order to improve the accuracy of assessment of colour fidelity for all light sources, including those with highly non-uniform spectral power distributions.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

1 Commission Internationale de l'Eclairage, CIE13.3-1995. *Method of Measuring and Specifying Colour Rendering Properties of Light Sources*. Vienna: CIE, 1995.
2 Bodrogi P, Csuti P, Hotváth P, Schanda J. 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 Photobiologica Assessment*, Tokyo, Japan: June 7–8: 2004: 24–27.
3 Sándor N, Schanda J. Visual colour renderin based on colour difference evaluation. *Lighting Research and Technology* 2006; 38: 225–239.
4 Smet K, Ryckaert WR, Pointer MR, Deconinck G, Hanselaer P. Correlation between color quality metric predictions and visual appreciation of light sources. *Optics Express* 2011; 19: 8151–8166.
5 Szabó F, Schanda J, Bodrogi P, Radkov E. A comparative study of new solid state light sources: *Proceedings of the CIE Session 2007,* CIE, Beijing: July 4–11: 2007: D1-18–D1-21.
6 Judd DB. A flattery index for artificial illuminants. *Illuminating Engineering* 1967; 62: 593–598.
7 Jerome CW. Flattery versus rendition. *Journal of the Illuminating Engineering Society* 1972; 1: 208–211.
8 Halstead MB. *Colour Rendering: Past, Present and Future*. Proceedings of 3e AIC conference AIC 77, Bristol: 1977: 97–127.
9 Jerome CW. The flattery index. *Journal of the Illuminating Engineering Society* 1973; 2: 351–354.
10 Smet K, Ryckaert WR, Pointer MR, Deconinck G, Hanselaer P. Colour appearance rating of familiar real objects. *Color Research and Application* 2011; 36(3): 192–200.
11 Smet KAG, Ryckaert WR, Pointer MR, Deconinck G, Hanselaer P. A memory colour quality metric for white light sources. *Energy and Buildings* 2012; 49: 216–225.
12 Davis W, Ohno Y. Color quality scale. *Optical Engineering* 2010; 49: 033602–033616.
13 Schanda J. A combined colour preference - colour rendering index. *Lighting Research and Technology* 1985; 17: 31–34.
14 Hashimoto K, Yano T, Shimizu M, Nayatani Y. New method for specifying color-rendering properties of light sources based on feeling of contrast. *Color Research and Application* 2007; 32: 361–371.
15 Rea MS, Freyssinier-Nova JP. Color rendering: A tale of two metrics. *Color Research and Application* 2008; 33: 192–202.
16 Szabó F, Bodrogi P, Schanda J. A colour harmony rendering index based on predictions of colour harmony impression. *Lighting Research and Technology* 2009; 41: 165–182.
17 Bodrogi P, Brückner S, Khanh TQ. Ordinal scale based description of colour rendering. *Color Research and Application* 2011; 36: 272–285.
18 Opstelten J. The dependence of the general colour rendering index on the set of test colours; the standard observer and the colour-difference formula. *Lighting Research and Technology* 1980; 12: 186–194.
19 Opstelten JJ. The establishment of a representative set of test colors for the specification of the color-rendering properties of light sources: *Proceedings of the CIE-20th session*, Amsterdam: August 31–September 8: 1983: D112/1-4.
20 Schanda J. Chromatic adaptation and colour rendering. *CIE Journal* 1982; 1: 30–38.
21 Schanda J. The influence of the test samples used for calculating colour rendering indices: *Proceedings of the AIC Congress COLOUR,* Monte Carlo, Monaco: June 16–22: 1985: 6.
22 Schanda J. The concept of colour rendering revisited: *Proceedings of the First European Conference on Color in Graphics Imaging and Vision*, Poitiers, France, April 3–5: 2002: 37–41.
23 Seim T. In search of an improved method for assessing the colour rendering properties of light sources. *Lighting Research and Technology* 1985; 17: 12–22.

24 van der Burgt P, van Kemenade J. About color rendition of light sources: The balance between simplicity and accuracy. *Color Research and Application* 2010; 35: 85–93.
25 van Kemenade JTC, van der Burgt PJM. Light sources and colour rendering: additional information to the Ra index *Proceedings of the CIBSE National Lighting Conference*, Cambridge, 1988: 133–143.
26 Commission Internationale de l'Eclairage, CIE135-1999. Colour Rendering *(TC1-33 closing remarks)*. Vienna: CIE, 1999.
27 Commission Internationale de l'Eclairage, CIE177-2007. *Colour Rendering of White LED Light Sources*. Vienna: CIE, 2007.
28 Commission Internationale de l'Eclairage, TC-E3.2. Colour rendering (Minutes from the 25th meeting, Paris, (Nov. 1980), the 26th meeting, Berlin (Sep. 1981) and the 27th meeting, Oslo (Sept. 1982)): 1980–1982.
29 Li C, Luo MR, Cui GH, Li CJ. Evaluation of the CIE colour rendering index. *Coloration Technology* 2011; 127: 129–135.
30 Luo MR. The quality of light sources. *Coloration Technology* 2011; 127: 75–87.
31 Csuti P, Schanda J. Colour matching experiments with RGB-LEDs. *Color Research and Application* 2008; 33: 108–112.
32 Csuti P, Schanda J. A better description of metameric experience of LED clusters. *Light and Engineering* 2010; 18: 50.
33 Bodrogi P, Krause N, Brückner S, Khanh TQ, Winkler H. Psychological relationship between colour difference scales and colour rendering scales. *Proceedings of the AIC 2011 midterm meeting, Interaction of colour and light in the arts and sciences,* AIC, Zurich, June 7–10: 2011: 210–212.
34 Commission Internationale de l'Eclairage, CIE15-2004. *Colorimetry*. Vienna: CIE, 2004.
35 Smet K, Whitehead L. *Meta-Standards for Color Rendering Metrics and Implications for Sample Spectral Sets: 19th Color Imaging Conference*, San Jose, USA, IS&T, November 7–11: 2011.
36 Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 2002; 6: 182–197.
37 Hunter JE, Schmidt FL. *Methods of Meta-Analysis:Correcting Error and Bias in Research Findings*. 2nd Edition, Newbury Park, CA: Sage, 2004.
38 Chou YF, Luo MR, Schanda J, Csuti P, Szabo F, Sárvári G. Recent developments in colour rendering indices and their impacts in viewing graphic printed materials: *19th Color Imaging Conference*, San Jose, USA, AS&T, November 7–11: 2011: 61–65.