Evaluating LED luminaires supporting colour critical assessment

Andreas Kraushaar; Fogra; Munich, Germany

Abstract

Colour critical assessment in the graphic arts industry is performed by using D50 simulators that obey rigorous criteria stipulated in ISO 3664 for years. Here multichannel LED based viewing cabinets are increasingly used. However the impact of phosphor converted LED based lighting for general lighting applications in lieu with colour assessment in prepress and press environments can be critical. Incorrect lighting might destroy ISO 3664 conformance and lead to colour differences that are larger than the domain of application of the advanced colour tolerances equations. A user friendly collection of metameric pairs was designed to solicit actual visual differences under 10 typical LED-based luminaires and six ISO 3664 compliant cabinets. These findings were correlated with modern colour rendering metrics including the new CIE fidelity index. Finally colour rendering tolerances will be proposed for general lighting that is used to support colour critical viewing and to certify LED based luminaires.

Introduction - Graphic arts colour appraisal

Ensuring that colour is evaluated under controlled lighting is a critical element of any colour-managed workflow. This step is often overlooked and is further complicated by a variety of lighting conditions in different parts of the world. Incorrect illumination leads to inferior or incorrect results and costly customer complaints throughout the production process. In the graphic arts rigorous aims for colour critical lighting conditions have been established since 1975 by ISO 3664. This standard was revised in 2009 and stipulates different requirements for the illuminated area of concern such as chromaticity tolerances of $\Delta u'v' <= 0{,}005$, strict tolerances for metamerism index in the visual (MI_{VIS}) and the ultraviolet spectral range (MI_{UV}) and last but not least minimum values for the general (Ra) and individual colour rendering indices. In all cases D50 is the reference illuminant for many good reasons, see [1].

In advent of advances in solid state lighting, LED light engines find their way into almost all kinds of lighting applications. The first multi-channel LED based viewing cabinet that complied to ISO 3664 criteria was introduced in 2009 [2]. However many print buyers, ad agencies, and even print shops don't always invest in controlled lighting conditions. Even in situations where ISO 3664 compliant viewing systems are in place the general lighting around the used cabinets are not always baffled from the plane of colour assessment. This might challenge the conformance to ISO 3664 by altering the stimulus hitting the hard or soft copies. Since general lighting is increasingly being replaced by means of retrofit or new LED based luminaires the colour rendering properties of typical tubular LED systems that happen to be mounted in prepress and press sites are subject for close inspection. Since the assessment should be conducted both in the lab and remotely in the field a handheld hawker's tray board was designed based on metameric pairs. In this paper we report the colorimetrical properties of that board, the conducted psychophysical experiment as well the correlation with established and modern colour rendering formulae.

Methodology

The colour rendering properties will be evaluated for 10 typical LED luminaires, five of which contain a DALI interface to check colour rendering properties for different dimming levels. Following the concept that a pair of specimen that exhibit no colour differences under a reference light will change it's appearance under another light source, metameric samples have been collected. It will be shown that the colour difference for the standard $r_1(\lambda)$ and the batch $r_2(\lambda)$ under the reference illuminant will not always be $\Delta E=0$ hence the samples are paramers and not metamers. Paramer definition - two reflectances $r_1(\lambda)$ and $r_2(\lambda)$ are called paramers for a specific illuminant $I(\lambda)$ if their colour difference for a given illuminant is smaller than a just noticeable difference [3]. In this paper different methods to correct this residual colour difference are used including no correction, which shows the remaining colour difference for the reference D50 illuminant.

In order to assess the perceived colour difference a non glossy grey scale was printed following the requirements of ISO 105-A02 [4]. This grey scale comprises 9 different colour differences from barely noticeable until clearly visible that allow for a gradual adjustment even for naïve observers, see Table 1.

Table 1: Colour difference (ΔE^*_{ab}) for the used grey scale

Scale ID	0	1	2	3+	3-	4+	4-	5+	5-
ΔE^*_{ab}	0	0,8	1,7	2,5	3,4	4,8	6,8	9,6	13,6

The scale was present in each assessment and allows for a stable identification of the perceived colour difference avoiding biases due ambiguous categorisations such as "acceptable" or "good". Nine painted pairs of plates from Clariant were used as parameric specimen. In cases where their colour under D50 illuminant is within a very large graphic arts proofing gamut, a third specimen was printed. This results in 8 colour centres including three pairs (out of gamut red and fuchsia as well as a paper pair) and 5 triplets, see Fig. 1. This results in 21 specimens and 3 (3*1) + 2 (2*1) + 15 (5*3) = 20 pair comparisons.

All but one pair do not show fluorescence. Since LED based lighting systems for general lighting applications do not contain any UV-amount, the rendering of substrates that using OBA (optical brightener agents) will be affected. In order to address this effect a paper with high amount of OBA was colorimetrically matched under D50 with an OBA-poor proofing paper.

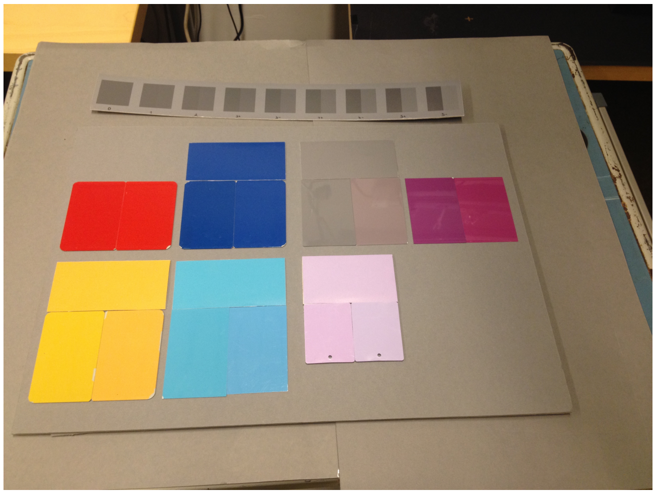

Figure 1: Photograph of the used hawker's tray under a typical office illumination hence exhibiting large colour differences. The used white colour pair is not shown here. Top: the used grey scale.

This results in a visual match under typical daylight including close D50 simulators (such as ISO 3664 compliant viewing cabinets) but will result in a mismatch for UV-free lighting situations. Since many substrates used in the publishing industry make use of OBA the correct, i. e. D50-like, rendering of prints on fluorescence paper is critical and is the reason for only one category for visual colour appraisal (as defined in ISO 3664:2009) without any gradual steps.

In order to judge the performance of the 10 LED luminaires used, seven ISO 3664-compliant viewing cabinets were added to the collection (including 3 LED-based) making altogether 17 IDs. LED T8 150 cm retrofits have also been tested in that project but they are not discussed in this paper.

Testing setup

The LED luminaires were mounted under the ceiling of a windowless laboratory. The modular height was used to adjust the illumination level of approximately E=2000 lx at the viewing position. The spectral power distribution has been measured with a CAS140 spectroradiometer from Instrument Systems measuring the incoming photons from 300 to 780 nm. The test setup is depicted in Fig. 2.

The CIE definition of the term "colour rendering" is:

"effect of an illuminant on the colour appearance of objects by conscious or subconscious comparison with their colour appearance under a reference illuminant" [5].

In recent scientific literature, the term "colour fidelity" has come into use with a very similar meaning; it has been chosen to help avoid a common error whereby some misinterpret the term "colour rendering" to have a broader meaning. Thus, the term "colour fidelity" is used in this paper in a manner that is consistent with the definition of "colour rendering" as above, especially to distinguish from other aspects of colour quality beyond colour fidelity such as colour saturation effects. This contrasts the evaluation of lighting systems in other applications where attributes such as preference, flattery, vividness, gamut enhancement etc. are of primary interest.

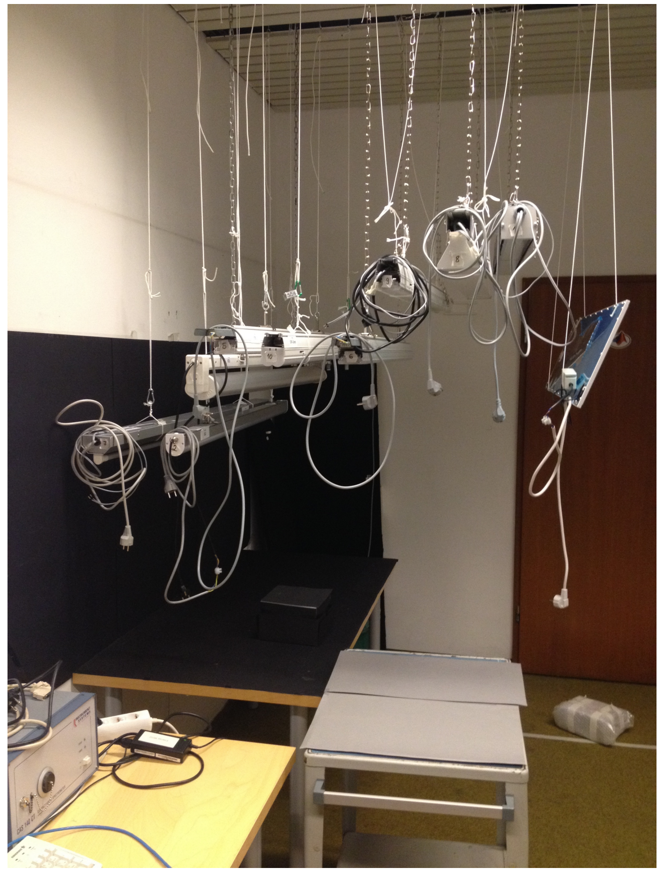

Figure 2: Mounted 10 LED luminaires.

The visual assessment is illustrated in Fig. 3.

Figure 3: A test person is judging the 9 different colour centres. For each subject a grey mask is used (on the left side, but not shown in action for a better overview).

In the graphic arts the paramount goal is to simulate D50 as closely as possible. It should be noted that even the tight tolerances of ISO 3664:2009 are reported to be too wide and after the phasing out of discharge lamps, stricter tolerances should find their way into the next revision of ISO 3664.

Colorimetrical properties

Based on the spectral measurements made at the centre of the plane of viewing, the correlated temperature (CCT), the general colour rendering index Ra, the ninth special colour rendering index R9 and the modern CIE fidelity index R_f [6] (which is based on IES TM-30-15 [7]) have been measured. The general colour fidelity index Rf is calculated as defined in equation 1:

$$R_f = 10ln\left(e^{100-c_f\overline{\Delta E}}/_{10} + 1\right) \quad (1)$$

whereas c_f=6,73 and ΔE the average of the 99 colour appearance differences. The special colour fidelity indices R_{fi}, also defined in [6] was not used in this study.

Limitations and inaccuracies of the CRI have been recently addressed [6], namely the outdated colour science (using 1964 U*V*W* and von Kries adaptation transform), the spectrally non-uniform distribution of the 13 test samples and insufficient information about colour shifts.

The spectral power distributions, normalized to a sum of 100, are plotted in Figure 4.

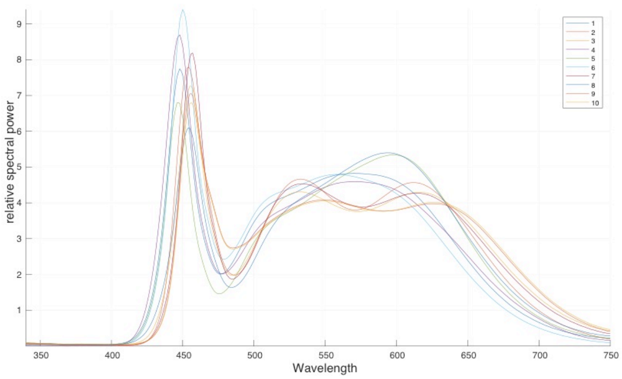

Figure 4: Relative Spectral power distribution of the 10 LED luminaries ID1, .., ID10. The spectral power distributions of the standard viewing cabinets (ID11 to ID17) are not shown here.

From Figure 4 it can be seen that most IDs are phosphor converted LEDs while 3 light sources have an additional LED installed. This can be seen in the additional bump in the spectral range between 500 and 600 nm. The colorimetrical results are listed in Table 2.

Table 2: Colorimetrical properties of the measured LED luminaires.

ID	CCT [K]	Ra	R9	Rf
1	5207	83,0	8,2	89,7
2	4997	95,3	79,7	90,2
3	4892	96,4	91,5	92,2
4	5604	83,6	18,1	81,2
5	4137	87,2	55,9	83,1
6	6116	77,3	-20,2	91,0
7	4972	94,4	78,6	93,3
8	4081	87,5	52,5	86,0
9	5004	95,6	90,7	94,7
10	4908	96,0	94,7	94,7

The listing in Table 2 shows that most light sources exhibit a colour temperature close to 5000 K whereas it was intentional to also include a warmer (ID 8) and a blue/cooler (ID6) light source. It can be seen that in this case the colour rendering of red colours (R9) is worse R9=-20,2 but there is also a case where CCT-differences to 5000 K do not correlate with better colour rendering figures. Contrary to CIE 15, D50 was fixed as the reference illuminant in all cases.

Metameric properties

The reflectance factors for the 21 specimens were measured according to ISO 13655:2009 M1 [8], i.e. using a measurement light source close to D50. It should be noted that the daylight simulation of the measurement light is actually defined by the same criteria as used in ISO 3664 [9] for colour viewing. That was the result of both standards being revised together in order to achieve the concept of “measure as you see”. The metamerism index, change in illuminant, was calculated for 27 illuminants, the typical standard illuminants namely D50, D75, D65, D55, A, F1, F2, F3, F4, F5, F6, F7, F8, F9, F10, F11, F12, and the ID1 to ID10. In all cases D50 served as the reference illuminant and all colour differences were expressed in CIEDE2000 (1;1;1). As mentioned before the colour difference is normally not ΔE=0 under D50. In this paper four different correction methods were used. First no correction was applied. This leads to a colour difference for the case of D50 for both reference and test illuminant. See figure 5 for the fuchsia coloured pair, where the residual colour difference is ΔE00 = 0.9, hence a paramer.

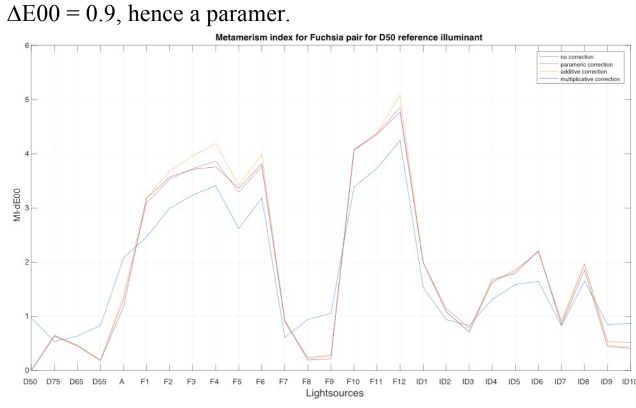

Figure 5: Metamerism index for a pair of fuchsia colour samples between D50 and 27 different light sources.

The second correction method is a parameric correction [10] showing a colour difference as expected of $\Delta E00 = MI_{D50\text{->}D50} = 0$. The third correction uses an additive and the fourth a multiplicative correction. Note the changes due to the different correction methods. These findings stress the importance of reporting the used correction method in any case. The results are also listed in Table 3.

Table 3: Metamerism index for a pair of fuchsia colored samples with D50 as the reference illuminant and 27 test illuminants by using four different correction methods to reduce the residual colour difference to zero (for the reference illuminant).

Test illuminant	Correction method			
	no	parameric	Add.	Mult.
D50	1,0	0,0	0,0	0,0
D75	0,5	0,6	0,7	0,6
D65	0,6	0,5	0,5	0,5
D55	0,8	0,2	0,2	0,2
A	2,1	1,3	1,4	1,2
F1	2,5	3,1	3,2	3,2
F2	3,0	3,5	3,7	3,6
F3	3,2	3,7	4,0	3,7
F4	3,4	3,9	4,2	3,8
F5	2,6	3,3	3,4	3,4
F6	3,2	3,8	4,0	3,8
F7	0,6	0,9	0,9	0,9
F8	0,9	0,2	0,2	0,2
F9	1,1	0,3	0,3	0,2
F10	3,4	4,1	4,0	4,1
F11	3,7	4,4	4,4	4,4
F12	4,2	4,8	5,1	4,8
ID01	1,5	2,0	2,0	2,0
ID02	0,9	1,1	1,1	1,1
ID03	0,8	0,8	0,7	0,7
ID04	1,3	1,6	1,6	1,7
ID05	1,6	1,9	1,8	1,8
ID06	1,6	2,2	2,2	2,2
ID07	0,8	0,9	0,9	0,8
ID08	1,7	2,0	2,0	1,9
ID09	0,8	0,5	0,5	0,4
ID10	0,9	0,5	0,4	0,4

The metameric properties of the other colour samples are of similar nature and confirm the usefulness for the following psychophysical experiment since it shows small and large colour differences for the test illuminants of interest and to show observers the influence of light – in a very practical and user friendly way.

Psychophysical experiment

The evaluation is focussed on the experiments conducted in the laboratory. The field tests are currently under way and aim to study the rendering of the built board under additional light sources measured at the same time on the exact position of the board.

18 observers, 5 female and 13 male, participated in the experiment. The comparisons were structured into P1, P2 and P3. P1 is the comparison between the sample pairs. As mentioned before, 5 colour centres have 3 samples (two metal samples and one print) hence there are 3 pair comparisons. P2 is denoted as the comparison between the first samples and the printed patch, and P3 is between the second metal sample and the printed patch. All persons were normal sighted. The task for all observers was:

"For each light source 21 colour pairs will be presented. Please judge the colour differences of all pairs by using the grey anchor pairs as the reference colour differences"

For all observers a mean opinion score MOS has been calculated by simply using the ratings of the observers based on the noted grey scale difference, see Table 1. Also the IDs 11 to 17, i.e. the standard viewing lights, have been added, see figure 6.

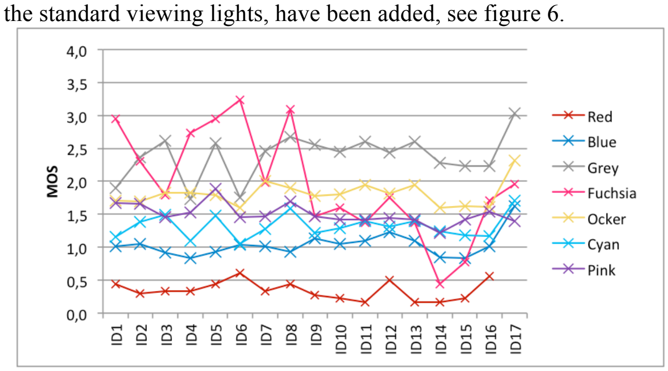

Figure 6: Mean opinion score for all colour centers (without white). The performance of the white samples are depicted in [13]

Averaging all colour centers, the statistical evaluation is depicted in figure 7. It is interesting to see that there is no significant difference between the tested LED luminaires and the used colour cabinets that comply with the aims and tolerances of ISO 3664.

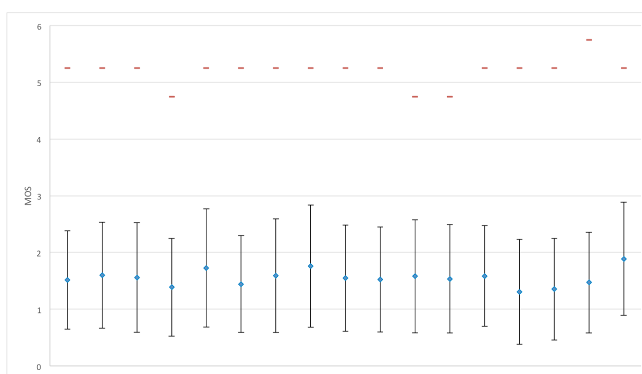

Figure 7: Mean, standard deviation and maximum MOS (mean opinion score) of 18 observers judging all colour centers (omitting the OBA paper pair) of the hawker's board under 10 LED test luminaries and 7 standard viewing cabinets.

The findings have been illustrated for two combinations that seem to be on the extremes when inspecting Table 2, namely ID6 ("worst") and ID9 ("best").

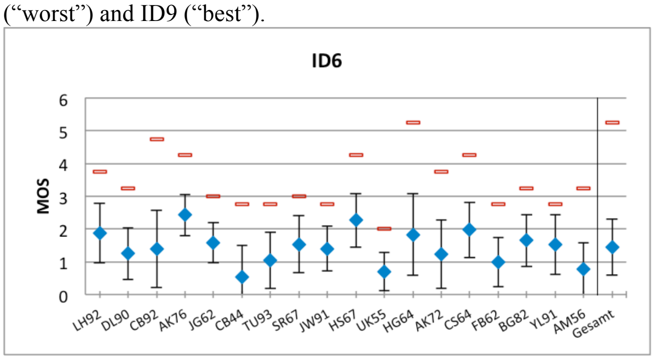

Figure 8: MOS of the 18 observers averaged over the 8 colour centers for light source ID6.

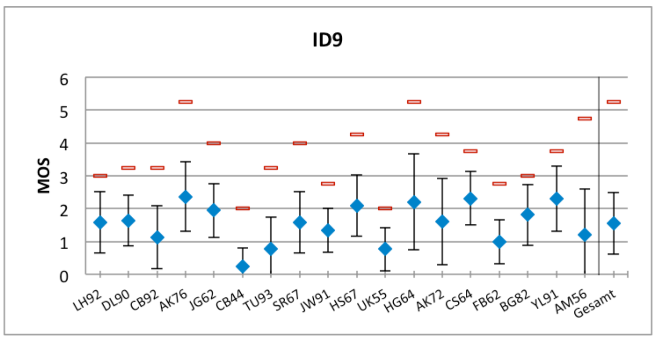

Figure 9: MOS of the 18 observers averaged over the 8 colour centers for light source ID9.

Inspecting the mean observer judgments of the grey scale for a particular colour center the good correlation can be confirmed. In this case the red colour pair (P1) was compared against R9 in figure 10.

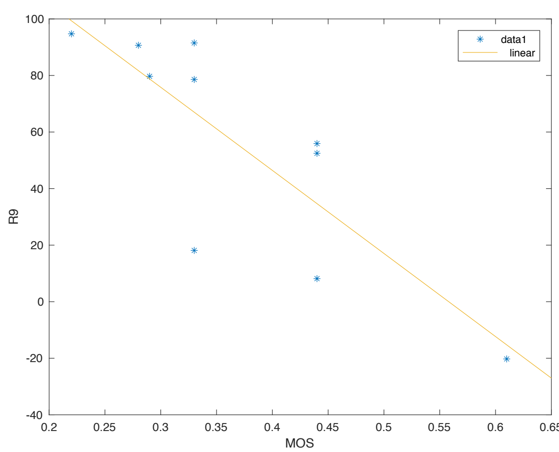

Figure 10: Correlation of mean grey scale rating (MOS) and R9 as well as a linear data fit with R^2=0,7 ("linear").

Comparing the average of all colours with typical colour rendering figures Ra, R9 and Rf in Figure 11, it can be seen that only Rf provides a good correlation.

The correlation of an average MOS over all patches with the colour rendering figures Ra, R9 and Rf is depicted in Figure 11.

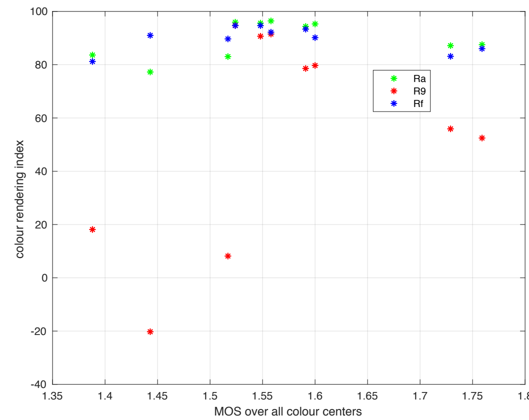

Figure 11: Correlation of average MOS (all patches) with Ra, R9 and Rf.

As expected R9 does not show a good correlation when being used to reflect the behavior of all 8 colour patches. Ra and Rf perform similarly and do not reveal much differentiation.

The picture looks different when adding the effect of fluorescence, an important feature of graphic arts products. Evaluating the tested LED luminaires with respect to the metamerism index in the UV region as stipulated in [9] and defined in [11] the following indexes can be computed, see Table 4. With respect to the five quality grades (A to E) defined in [11], with "A" being the best and "E" being the worst, defined by an average colour difference of MI_{UV} > 2, all luminaires fail the test as expected. In light of the newly developed metric to reflect the UV amount of paper like substrates with respect to D50 [12] all samples can be considered to exhibit a UV amount of 0.

Table 4: CIELAB colour difference for the three UV metameric pairs and it's average (MI_{UV}) for the tested 10 LED luminaries driven at 100% (no dimming).

ID	Pair 1	Pair 2	Pair 3	MIuv
1	3,2	4,4	6,0	4,5
2	2,7	3,7	5,3	3,9
3	2,7	3,7	5,2	3,9
4	3,5	4,6	6,4	4,8
5	3,4	4,5	6,2	4,7
6	3,1	4,2	5,9	4,4
7	2,6	3,5	5,1	3,7
8	2,9	3,9	5,5	4,1
9	2,5	3,4	4,9	3,6
10	2,5	3,4	4,9	3,6

The visual assessment for the paper samples, one OBA rich and the other one OBA poor with a colour shading that leads to a very close visual match under D50 (M1 as defined in [9]) was due to relocation activities not completed. However the intermediate findings show significant larger ($3 \leq MOS_{avg} \leq 4$) visual rankings for all ten IDs then for all used ISO 3664 complying viewing cabinets (IDs 11- 17) MOS. They comprise a $MI_{UV} < 1$, as required by ISO 3664. The complete findings will be made available on the project webpage [13].

Results

The designed and built hawker's board proved to be a very useful and convenient tool for practical colour rendering evaluation under both laboratory and field test environments. The grey scale being present in all observations served nicely as an anchor point for rating colour differences by their corresponding lightness difference.

Interestingly the difference between tested phosphor-converted "5000K-LEDs" for general lighting applications and high quality viewing cabinets is smaller than expected. For samples that do not fluoresce (e.g. without optical brighteners) the tested LED luminaires are not rated significantly worse than standard light cabinets. Even more interesting is the fact that a "4000 K" and a "6000 K" lamp render the selected metameric samples not significantly worse than the "5000 K" ones.

The paper samples, that show small visual differences under all standard viewing cabinets, are observed to show significantly larger colour differences for all (UV free) ID's. This confirms the need for using "high quality", i.e. ISO 3664 compliant lighting, for colour critical viewing since many substrates contain OBA's and the fluorescence effect is also visible in particular in lighter areas of the colour space.

However more research has to be done on the selection of metamers (physical samples or tabulated data) to check the present findings. This also refers to adding images in order to judge contrast (changes) instead of single uniform colour patches. Also more LED based luminaires have to be evaluated, e.g. no RGB-based LED lamps were evaluated.

The new colour fidelity index Rf shows no significant improvement over Ra. Based on the current findings both Ra and Rf values can be recommended for qualifying LED based luminaires for supporting critical colour appraisal. The exact set of tolerances needs more substantiation but will certainly be within the minimum Rf of 85 and 90. A tolerance in terms of CCT should not be used.

In terms of next steps, the framework presented here will be tested on a broader variety of metameric samples and test spectra and extended as necessary.

References

[1] A. Kraushaar, Why the printing industry is not using D65?, Munich Fogra Extra 18, URL: www.fogra.org/Extra18/, 2009.

[2] N. N., LED colour viewing light, URL: https://www.just-normlicht.de/media/files/JUST-4S-LED-Technology_INT.pdf, 2009.

[3] P. Urban and R. S. Berns, ``Paramer mismatch-based spectral gamut mapping," IEEE Trans. Image Process. 20, no. 6, 1599_1610 (2011).

.

[4] ISO 105-A02:1993 Preview, Textiles -- Tests for colour fastness -- Part A02: Grey scale for assessing change in colour, Beuth-Verlag, 10772 Berlin

[5] CIE. 2011. CIE S 017/E:2011, ILV: International Lighting Vocabulary. Vienna, Austria: CIE.

[6] CIE 224:2017 - Colour Fidelity Index for accurate scientific use.
Vienna, Austria: CIE. ISBN 978-3-902842-61-9

[7] IES. 2015. IES-TM-30-15: Method for Evaluating Light Source Colour Rendition. New York, NY: The Illuminating Engineering Society of North America.

[8] ISO 13655 Graphic technology -- Spectral measurement and colorimetric computation for graphic arts images, Beuth-Verlag, 10772 Berlin

[9] Norm ISO 3664 : 2017, Viewing conditions -- Graphic technology and photography, Beuth-Verlag, 10772 Berlin

[10] H.S. Fairman, "Metameric Correction Using Parameric Decomposition", Color Res. Appl. 12, 1987, pp.261-265.

[11] ISO 23603:2005, Standard method of assessing the spectral quality of daylight simulators for visual appraisal and measurement of colour, Beuth-Verlag, 10772 Berlin

[12] A. Kraushaar, "Development of methods to compensate the difference between proofing and production stock",
Fogra research project (No. 60.054), 2012

[13] K. Bieske; A. Kraushaar : 2017
"Evaluating criteria for LED based general lighting applications for print assessment", Research project 11.003, URL: https://www.fogra.org/CIC25/

ACKNOWLEDGMENTS

The project "Evaluating criteria for LED based general lighting applications for print assessment" [Fogra No. 11.003] was supported within the programme of the Federal Ministry of Economics and Technology for the support of "Industrielle Gemeinschaftsforschung [IGF]" via the German Federation of Industrial Research Association [AiF/No. 18340 BG]. Special thanks to Karin Bieske from TU Ilmenau (participating in that research project) for proof reading. We also thank Master student Caroline Breit for her great measurement support.

Author Biography

Andreas Kraushaar is head of the prepress division at Fogra, Munich, Germany. He joined Fogra in 2001. Within the prepress department he is responsible for research in the fields of colour management and digital printing. He holds the position of the convener of working group 3 (process control and related metrology) within ISO TC 130 (graphic technology) since 2005.

He received a master degree in media technology from the Technical University Ilmenau and his PhD in colour science from the RWTH university Aachen. He is a member of IS&T.