Merit function for the evaluation of color uniformity in the far field of LED spot lights

Anne Teupner*[a], Krister Bergenek[b], Ralph Wirth[b], Juan C. Miñano[a], Pablo Benítez[a]
[a]Technical University of Madrid, CeDInt, Campus de Montegancedo, 28223 Madrid, Spain; [b]Osram GmbH, Wernerwerkstraße 2, 93049 Regensburg, Germany

ABSTRACT

The scope of the present paper is the derivation of a merit function which predicts the visual perception of LED spot lights. The color uniformity level Usl is described by a linear regression function of the spatial color distribution in the far field. Hereby, the function is derived from four basic functions. They describe the color uniformity of spot lights through different features. The result is a reliable prediction for the perceived color uniformity in spot lights. A human factor experiment was performed to evaluate the visual preferences for colors and patterns. A perceived rank order was derived from the subjects' answers and compared with the four basic functions. The correlation between the perceived rank order and the basic functions was calculated resulting in the definition of the merit function U_{sl}. The application of this function is shown by a comparison of visual evaluations and measurements of LED retrofit spot lamps. The results enable a prediction of color uniformity levels of simulations and measurements concerning the visual perception. The function provides a possibility to evaluate the far field of spot lights without individual subjective judgment.

Keywords: LED, color uniformity, color blending, optical systems, color perception

1. INTRODUCTION

A light source for high performance illumination requires a high color rendering index (CRI) and high efficiency. Spot lights are predominantly used in shops and hospitality to illuminate and emphasize major objects or areas. A defined illumination distribution and a uniform color are the basis for comfortable perception. A variety of LED chips in combination with collimators offers the possibility to fulfill quality and efficiency requirements. The optical system for spot lights consists of two main parts, the light source itself and the optical element for light collimation and far field design. Due to several methods for generating white light like phosphor converting methods or RGB LEDs (Figure 1), there is a wide range of light distributions. A high quality spot light comprises a wide continuous spectral power distribution (SPD) and a uniform spatial color distribution. Both properties are mainly influenced by the types of light source and optical element.

Figure 1. One example of a multicolored LED light source with a board (1), diffuser cast (5) and several red, green and blue LED chips (2, 3 and 4), (a) is the cross section and (b) the top view of the light source.

The optical system has to collimate and mix the light in the far field up to a sufficient color blending level. These optical elements are mainly reflectors or total internal reflection lenses (TIR) but are not limited on these elements (Figure 2). Due to plenty of possible combinations, the appearance of spot lights is varying[1]. The color uniformity in the far field can be improved by additional elements like facets, diffuser elements or scattering layers. The color blending can be highly improved but often then the efficiency is affected and decreases.

Linked with the visual perception of the human eye, the color uniformity has to be as good as the just noticeable difference (JND). There is no need for a better color mixing because it could not be recognized by the human eye. The JND for the comparison of two colored sheets was firstly defined by MacAdam[2]. A general definition of the perceivable color difference is not available because the perception depends on many factors. On the one hand environmental factors

are important. These are the visual field, stimulus size[3], observer distance, luminance[4], background and expose duration[5]. On the other hand the appearance of the far field itself determines the threshold[6], depending on the color and the pattern in the spot light and subject's abilities for color perception and color memory[7]. To study the perception of color appearances in the far field of the spot lights a human factor experiment was performed while the environmental conditions were fixed. The scope of the experiment was to find preferences for perceived evenness of colors and patterns in the far field of spot lights.
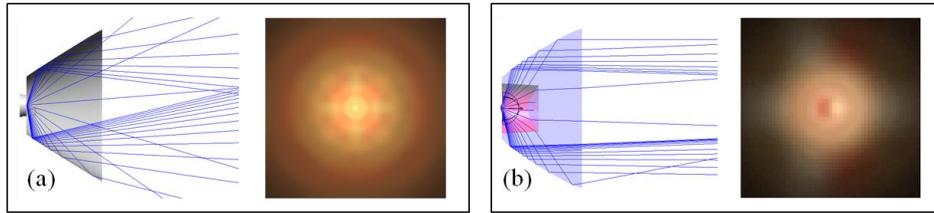

Figure 2. Two optical systems with multicolored LEDs and the spatial color distribution of the far field; the optical element for collimation is a reflector (a) and a TIR lens (b).

2. METHODOLOGY

2.1 Definition of spot lights

The visual appearance of spot lights is affected by the two main properties of color and pattern. The color is described by the attributes hue, chroma and luminance. The pattern can be described by gradient and symmetry features. The far field of the spot lights is divided into the spatial luminance distribution and its spatial color distribution. The luminance distribution has to correspond with the required gradient. It is not included in the following considerations because of the preexisting luminance gradient from the lighter center to darker edges. The focus is on the spatial color distribution.

Figure 3. 5 spot lights out of 46 which were used in the human factor experiments; the reference spot (a) and four examples with different color and pattern combinations (b-e).

The color distribution and design of the spot lights are based on optical simulations. They were performed with multicolored LED light sources in combination with different types of reflectors and TIR lenses. Figure 3 shows the reference spot light without color features and four other examples.

2.2 Human factor experiment

A human factor experiment was executed to evaluate visual preferences and to get a perceived rank order of the spots. The spot lights vary only in their spatial color distribution. The luminance level, spot size and background properties are almost identical in each spot. The maximum luminance in the center of the spot lights was 300 cd/m². The test method was an alternative forced choice test (2AFC)[8]. Figure 4 shows the experimental setup. The experiment was done with 72 subjects. The 46 different spot lights were combined with each other pairwise. Two spot lights were shown successively. Afterwards the subject had to select the more even spot light regarding its color distribution. The order of pairs was randomized to avoid biased errors. During the session of one subject, each pair was shown twice, the second time in reverse order to eliminate time order errors.

The answers of the experiment were analyzed after the method of Thurstone[9] to receive a perceived rank order. It signifies the order from the best perceived spot (always chosen by subjects) to the worst perceived spot (never selected by subjects). The perceived rank order is applied to detect correlations between the basic functions.
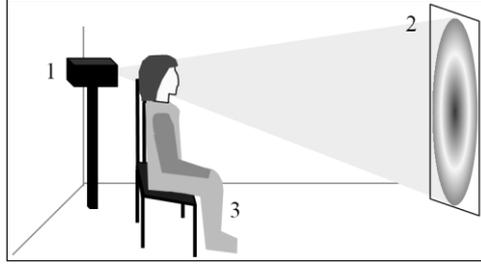
Figure 4: Experimental setup for the human factor experiments with a projector (1), the illumination plane of the spot light far field (2) and the subject (3).

2.3 Mathematical evaluation

Several functions describe different color features and aspects of spot lights. The aim is to derive a merit function from the visual perception to evaluate spot lights regarding human vision. An objective evaluation criterion independent from subjective opinions is required. There are many influence factors for the color and shape. The features are hue, chroma and lightness for color and contrast, shape and symmetry properties for patterns. All factors have to be included in an evaluation method to reach an adequate prediction[10]. The existing mathematical descriptions are adapted to spot lights. The *CIE L*a*b** [11] is used due to applicability for surface colors and for small color differences near the white point. The same perceived color difference are represented by approximately the same Euclidean distances in *CIE L*a*b**. *L* is the lightness, *a* is the red-green color opponent dimension and *b* is the yellow-blue one.

By default, the evaluation of spatial color differences is calculated with ΔE. For spatial color differences, the average color deviation Δab_{mean} is used. It is calculated as follows:

$$\Delta ab_{mean} = \frac{1}{k} \cdot \sum_{i=1}^{k} \sqrt{\left(a_i^* - a_{Ref}^*\right)^2 + \left(b_i^* - b_{Ref}^*\right)^2} \,, \tag{1}$$

k is the total number of pixels. The reference color a_{Ref}^* and b_{Ref}^* are:

$$a_{Ref}^* = \frac{1}{k} \sum_{i=1}^{k} a_i^* \,, \text{ and } b_{Ref}^* = \frac{1}{k} \sum_{i=1}^{k} b_i^* \,. \tag{2}$$

Δab_{mean} involves explicitly only the color differences. A similar function, the root-mean square color difference is used by Sun et al. [12] for the evaluation of RGB-LED spot lights. The color deviation is used to specify the color uniformity within the spot lights. Another possibility holds the gradient $Grad_{ab}$. It describes the change of color in the spot light from one point to another point. It is calculated by:

$$Grad_{ab} = \frac{1}{m \cdot n} \sum_{i=1}^{m} \sum_{j=1}^{n} \left(\left(a_{i+1,j}^* - a_{i,j}^*\right)^2 + \left(b_{i+1,j}^* - b_{i,j}^*\right)^2 + \left(a_{i,j+1}^* - a_{i,j}^*\right)^2 + \left(b_{i,j+1}^* - b_{i,j}^*\right)^2 \right), \tag{3}$$

m and n are the number of pixels in x- and y- direction.

The visual perception is also sensitive to symmetrical objects in the visual field[13]. The smoothness of the radial component S_{rad} and linear component S_{lin} are calculated to differ between spots with circular symmetry and others. It is the root-mean square value of the color variation for radii and angles. It is calculated as follows:

$$S_{rad} = \sqrt{\frac{1}{n} \sum_{\varphi=0}^{2\pi} \left(\left(a_\varphi^* - \overline{a_r^*}\right)^2 + \left(b_\varphi^* - \overline{b_r^*}\right)^2 \right)} \,, \tag{4}$$

$$S_{lin} = \sqrt{\frac{1}{n} \sum_{r=0}^{r\,max} \left(\left(a_r^* - \overline{a_\varphi^*}\right)^2 + \left(b_r^* - \overline{b_\varphi^*}\right)^2 \right)} \,. \tag{5}$$

The reference colors for each radius $\overline{a_r^*}$, $\overline{b_r^*}$, and for each angle $\overline{a_\varphi^*}$, $\overline{b_\varphi^*}$ are:

$$\overline{a_r^*} = \frac{1}{n} \sum_{\varphi=0}^{2\pi} a_\varphi^* , \quad \overline{b_r^*} = \frac{1}{n} \sum_{\varphi=0}^{2\pi} b_\varphi^* \quad \text{and} \quad \overline{a_\varphi^*} = \frac{1}{n} \sum_{r=0}^{r\,max} a_r^* , \quad \overline{b_\varphi^*} = \frac{1}{n} \sum_{r=0}^{r\,max} b_r^* . \tag{6}$$

The maximum value of S_{rad} of all angles and of S_{lin} of all radii is used as one evaluation value for the spot lights.

The four basic functions 1, 3, 4 and 5 focus on different visual influence factors. The dependency between visual influence factors and the functions is shown in Table 1.

Table 1. Each of the visual influence factors is represented by another basic function.

Visual influence factors	Δab_{mean}	$Grad_{ab}$	S_{rad}	S_{lin}
Hue	X			
Chroma	X			
Contrast		X		
Shape		X		
Circular symmetry			X	
Linear smoothness				X

The analysis of the spot lights according to each of those functions was performed with data from measurements with a luminance meter. The spot lights were projected onto a white screen, as in the experimental setup (Figure 4). The values of the four functions were calculated for each spot. The characteristics of each function were analyzed and the correlation between the functions and the perceived rank order was compared.

3. RESULTS

3.1 Results of the human factor experiment

The results of the human factor experiment show preferences for spatial color distributions. The analysis of the answers indicates that more than 88% of the spot lights could be distinguished by subjects. They selected the same spot, independent from the position of presentation. In other cases the spot lights could not be kept apart from each other. Either the difference between the two spots was too low to remember a difference or both spots were colorful or had clear patterns. In both cases the subjects had no preference and made an arbitrary choice for one spot. This resulted in inconsistent answers of the subjects.

The high number of same selections by several subjects for same pairs implies a homogeneous opinion about the evenness of spot light. The experimental setup as a 2AFC test involves time order errors for successive shown pairs[14]. In many cases subjects tend to select more often the first or the second spot. This was contrasted with double shown pairs. Consequently the tendency of one subject is equal to all spots but the number of unclear pairs is increased. Furthermore the reference spot was not identified as the most uniform spot light but the spot from Figure 3(b) was chosen mostly. The worst spot lights were those with high hue or saturation and dot patterns (Figure 3(e)) as well as spots with sectioned rings at the edge (Figure 3(d)).

A Wilcoxon signed rank test ($\alpha=0.001$) showed a significant difference between the preference of chroma, pattern and symmetry. Spot lights with less chroma were preferred over others. Ring patterns were more frequently chosen than dots. The circular symmetry was clearly preferred over other. Even perceived spots have less chroma, rings instead of dots and need circular symmetry for patterns.

3.2 Evaluation of basic functions

The correlations of the perceived rank order and the basic function values were analyzed. The correlations are listed in Table 2. The Pearson's correlation coefficient r_p reflects the linear correlation between the perceived rank order and the basic function. The Spearman's rank correlation coefficient r_s refers to a non linear monotonic correlation.

Table 2. Correlation coefficients for the perceived rank order and the four basic functions.

	Δab_{mean}	$Grad_{ab}$	S_{rad}	S_{lin}
r_p	0.75	0.71	0.66	0.72
r_s	0.77	0.83	0.67	0.73

There is a correlation of all basic functions with the perceived rank order. The nonlinear correlation coefficients r_s are higher than r_p, especially for $Grad_{ab}$. A nonlinear correlation is expected because of the randomly chosen spot lights and no previous knowledge about the perceived differences between single spot lights. The correlation of all functions shows systematical aberrations for the prediction of the visual perception. For several pattern and color features there is a general over- or underestimation. Each function considers only few visual influence factors. For one single function as the merit function a large amount of wrong ratings would be the consequence. The percentage of clearly wrong predicted spots is lowest for $Grad_{ab}$ thus the correlation is highest. For example, Δab_{mean} overestimates single ring patterns with high chroma. The subject rated these spots better than the prediction of the function. These spots do not disturb subjects so much. Spot light patterns with dots are underestimated and subjects perceive a lower evenness than the Δab_{mean} function value states. The spots are perceived more uneven than the estimation of the function. The underestimated features for S_{rad} are patterns with two or more rings in different colors because it is not sensitive to axial color changes. The function is not sensitive for different colors at different radii. In contrast S_{lin} underestimates patterns with one uniform or segmented ring. The function tends to overestimate one ring patterns because it is not sensitive to radial color changes.

3.3 Merit function U_{sl}

The systematical aberrations to different color features of each basic functions resulted in the derivation of the merit functions as a linear regression. All four basic functions are included in the merit function to avoid systematical aberrations and to get a higher correlation between the perceived rank order and the merit function. The linear regression function is described by:

$$U_{sl} = 4 \cdot \Delta ab_{mean} + 15 \cdot Grad_{ab} + 12 \cdot S_{rad} + 7 \cdot S_{lin}. \tag{7}$$

The merit function reaches a correlation coefficient of $r_s = 0.91$. The non linear coefficient of U_{sl} is clearly higher than the values of the basic functions. The combination of the four basic functions enables a prediction which avoids most systematical aberrations of the basic functions. All visual influence factors are combined and all have now an effect on U_{sl}. Figure 5 shows the four basic functions and the derived merit function U_{sl}.
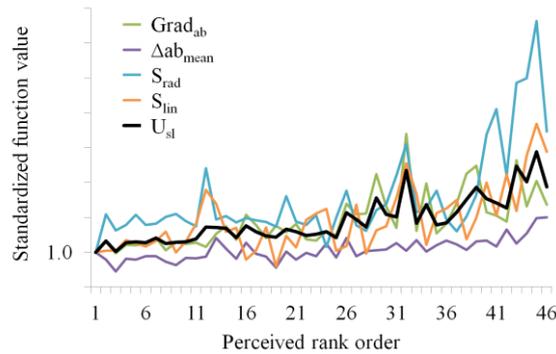

Figure 5. Standardized comparison of the four basic functions and merit function U_{sl}.

Compared to the basic functions, U_{sl} is smoothed. There is a continuous increasing of the functions. Higher function values indicate lower perceived color uniformity level. There are fewer outliers. The outliers can be explained by the combination of the basic functions in one function. Extreme values for the same spot in several functions add up in the combined merit function. The sensitivity is higher for $Grad_{ab}$ and S_{rad} because they have higher coefficients.

3.4 Validity of the merit function

The function is derived from the evaluation of spot lights used in the human factor experiment. It is optimized to fit the visual perceived evenness of the presented spot lights. Based on the experimental setup conditions, a high correlation is

reached. Many studies and experiments examined the effect of single or more influence factors but rarely to complex scenes for spot lights. These results can be adapted to specify the validation of the merit function. The color perception is similar in a wide luminance range and do not change significantly[15]. Thus the function is not limited to a luminance of 300cd/m² but could be applicable in a range from 50cd/m² to 1000 cd/m². Also the single colors used in the experiment like red, green, yellow and blue can be expanded. The change of LED chips in the light source would cause different colors in the far field. But for colors in the same color category, the visual perception is similar. The evenness of the spot light would be perceived equal for similar color combinations[16].

The application of U_{sl} is approved for same measurement setup. Hereby, the absolute values are comparable with each other. This includes mainly the spot size and resolution of measurement (pixel/cm²). The merit function U_{sl} is sensitive for the visual perceived color uniformity in range of about 60 and up to 130. Function values under 60 imply uniform color appearance and for values higher than 130 the spot lights appear very uneven and colorful. There is an uncertainty of the values for all calculation. It results from measurement uncertainty and the uncertainty of the perceived rank order. Furthermore the resolution is relevant for the measurement because of the noise in the measurement, especially $Grad_{ab}$ is sensitive to noise. This is countered by standardization to the number of pixels. Also optical simulations are concerned. A low noise simulation requires a high number of rays but need a lot of time. Fast simulations with fewer rays contain more noise. A balance for the signal-to-noise ratio should be defined.

4. APPLICATION

4.1 LED spot light retrofits

LED retrofit lamps are replacements for traditional lamps like reflector lamps. There are advantages for energy saving, higher efficiency and tunable color and CCT. A comparison of a traditional halogen lamp in Figure 6 (a) and LED replacement (b) was performed to define color uniformity level of both technologies. The scope of the comparison is the evaluation of the color uniformity U_{sl} in the spot lights and to illustrate differences and similarities in the light quality regarding color homogeneity. Therefore the central area of the spot light is considered down to 25% of the maximum luminance.
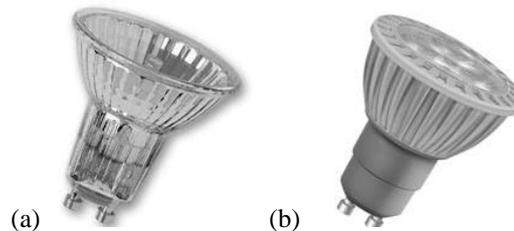

(a)   (b)

Figure 6. Halogen PAR 16 lamp (a) and LED PAR 16 lamp (b), both with GU10 lamp socket.

4.2 Measurement and visual judgment

In this example we analyzed the color distribution of 35 W replacements with a CCT of 2700 K and a full width half maximum (FWHM) angle of 36°. The LED lamps generate warm white light by phosphor conversion. They are compared with the corresponding halogen lamp as the standardized reference. Two LED replacements were selected to show the appreciable quality differences in the spatial color distribution of the LED technology. The lamps were mounted in a fixture and illuminated the measurement screen. A luminance meter measured the chromaticity values as a function of the coordinates in the illumination plane of the far field.

The results of the measurements are presented in Figure 7. The spatial color distribution is represented in the CIE xyY color space. The range of the scale is equal for all CIE x diagrams as well as for the CIE y diagrams. The traditional halogen lamp (a) performs best. The two LED replacements have very different color uniformities. LED lamp I (b) has quite good color uniformity. In contrast, the color uniformity of LED lamp II (c) is very poor. The color of the spot is very yellowish in the center and the outer ring. For the other two spots there is just a color gradient towards the edge because of the decreasing luminance in the spot light. This shows how big the differences in light quality of recent LED retrofit lamps still are.

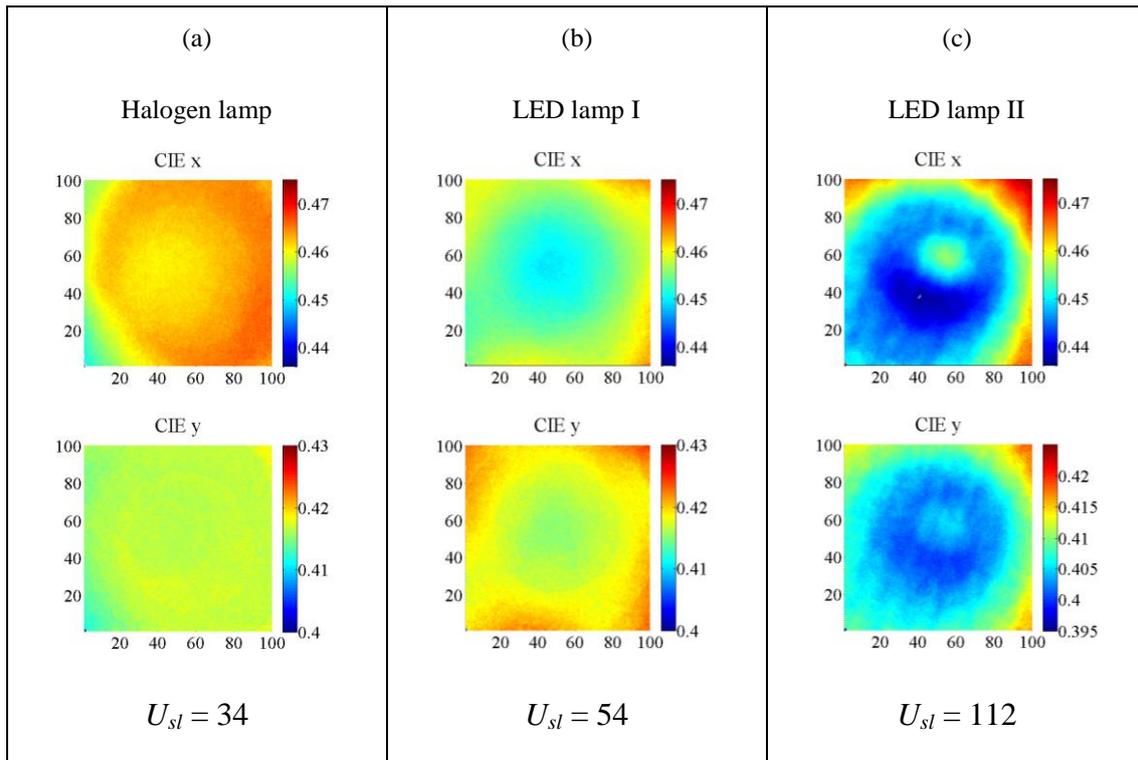|	(a)	(b)	(c)
	Halogen lamp	LED lamp I	LED lamp II
	$U_{sl} = 34$	$U_{sl} = 54$	$U_{sl} = 112$

Figure 7. False color images of the measurements of the halogen lamp (a), and two different LED lamps (b) and (c) in the CIE xyY color space at a 1 m² wall. The upper line contains all diagrams for CIE x chromaticity coordinates and lower line for CIE y chromaticity coordinates dependent on the spatial position.

The analysis of the measurements is verified by visual judgment. The lamp (c) showed clear colors inside the spot and was not accepted by subjects. It was perceived as a very nonuniform spot light. The other lamps (a) and (b) performed much better. There was no definite preference for one of the lamps (a) or (b). These spots showed no or very few colors and were perceived very even regarding the color distribution. The perceived evenness of the spot lights is consistent with the U_{sl} values. The difference of the values between (a) and (b) is difficult to perceive. The value of (b) implies already a very good color uniformity. The color gradient due to luminance decrease is not perceived as color difference. The difference to (c) is enormous in the visual judgment and the measurements. The merit function value U_{sl} reflects the visual judgment reliable.

5. CONCLUSION

5.1 Summary

A new metric is provided to predict the visual perception of color blending level in the far field of spot lights. It enables the classification of spot lights after the perceived evenness. The relationship between the color uniformity level and the visual perception of human is established by the human factor experiment. On the basis of a wide range of spatial color distributions, subjects had to compare various spot lights with each other in a 2AFC experiment. It was the basis for the evaluation of several basic functions describing far fields of spot lights. The correlation was not sufficient for single functions because systematical aberrations occurred. There are typical features in several spots which are over- or underestimated by the single basic function. An accurate prediction is not possible. A linear regression of four basic functions combines the main visual influence factors. The merit function U_{sl} avoids over- and underestimations for most features. The function provides a possibility to evaluate data of spot lights from measurements and simulations. The function values can be compared for the same measurement setup. There are efforts to make the absolute function values comparable to each other and to expand the application field. A first application of U_{sl} was the comparison of the

calculated and perceived color uniformity of a traditional halogen lamp and LED replacements. In accordance with the perception the far fields of these lamps were evaluated by the merit function. The results demonstrate the possibility to evaluate optical designs prior to manufacturing on an objective background.

REFERENCES

[1] Moreno, I. and Contreras, U., "Color distribution from multicolor LED array," Opt Express 15 (6), 3607-3618 (2007).

[2] MacAdam, D.L., "Visual Sensitivities to color differences in Daylight," J Opt Soc Am 32 (5), 247-274 (1942).

[3] Abramov, I., Gordon, J. and Chan, H., "Color appearance in the peripheral retina: effects of stimulus size," J Opt Soc Am A 8(2), 404-410 (1991).

[4] Carreño, F. and Zoido, J. M., "The influence of luminance on color-difference thresholds," Color Res Appl 26(5), 362-368 (2001).

[5] Hita, E., Romero, J., del Barco, L. J., and Martínez, R., "Temporal aspects of color discrimination," J Opt Soc Am 72(5), 578-582 (1982).

[6] Krauskopf, J. and Gegenfurtner, K., "Color discrimination and adaptation," Vision Res 32(11), 2165-2175 (1992).

[7] Jin, E. W., and Shevell, S. K., "Color memory and color constancy," J Opt Soc Am A 13(10), 1981-1990 (1996).

[8] Katkov, M., Tsodyks, M. and Sagi, D., "Analysis of a two-alternative force-choice signal detection theory model," J Math Psychol 50(4), 411-420 (2006).

[9] Thurstone, L. L., "A law of comparative judgment," Psychol Rev 34(4), 273-286 (1927).

[10] Moreno, I., "Illumination uniformity assessment based on human vision," Opt Lett 35(23), 4030-4032 (2010).

[11] Melgosa, M., "Testing cielab-based color-difference formulas," Color Res Appl 25(1), 49-55 (2000).

[12] Sun, C.C., Moreno, I., Lo, Y.C., Chiu, B.C. and Chien, W.T., "Collimating lamp with well color mixing of red/green/blue LEDs," Opt Express 20(S1), A75-A84 (2012).

[13] Wagemans, J., "Detection of visual symmetries," Spatial Vision 9(1), 9-32 (1995).

[14] Hellström, Å., "Comparison is not just subtraction: effects of time- and space-order on subjective stimulus difference," Percept Psychophys 65(7), 1161-1177 (2003).

[15] Pridmore, R.W. and Melgosa, M., "Effect of luminance of samples on color discrimination ellipses: analysis and prediction of data," Color Res Appl 30(3), 186-197 (2005).

[16] Uchikawa, K. and Shinoda, H., "Influence of basic color categories on color memory discrimination," Color Res Appl 21(6), 430-439 (1996).