Psychophysical optimization of lighting spectra for naturalness, preference, and chromatic diversity

Sérgio M. C. Nascimento* and Osamu Masuda
Centre of Physics, Campus de Gualtar, University of Minho, 4710-057 Braga, Portugal
*Corresponding author: smcn@fisica.uminho.pt

Received September 1, 2011; revised November 5, 2011; accepted November 28, 2011; posted November 28, 2011 (Doc. ID 153878); published January 25, 2012

The optimal spectral profiles of lighting for naturalness, individual preference, and chromatic diversity were estimated with psychophysical experiments in which observers selected illuminants from a set of metamers of D65 to render outdoor and indoor scenes. For naturalness, the illuminant selected was more spectrally structured than daylight and had a low color rendering index. For preference, the illuminant was similar but produced colors a little more saturated. For chromatic diversity, the spectrum was much more structured, with clearly defined peaks in the blue, green, and red spectral regions. These experiments show that light sources with specific structured spectra may have important visual applications in lighting.

1. INTRODUCTION

The color appearance of the objects depends on physical factors, like the reflecting properties of surfaces, the spectral distribution of the light source, the mutual reflections between the objects, and on visual factors, like the adaptation state of the visual system [1,2]. The illumination has an important role. In spite of visual mechanisms compensating for the effects of varying lighting conditions, normally considered in the general context of color constancy [3], the chromatic appearance of complex scenes can be considerably different under different natural or artificial illumination, as suggested by the moderate levels of color constancy often estimated in psychophysical experiments [4].

The complex effects of the illumination on the global appearance of scenes pose difficult challenges to the lighting industry which aims at developing light sources not only energetically efficient and physically safe, but also visually adequate for specific applications. Museum lighting, for example, requires not only protection of the works of art from radiation [5] but also lighting that is visually adequate [6]. Street lighting and lighting retail food counters are other examples of important specific visually demanding applications.

In practice, several colorimetric aspects of lighting are evaluated by rendering indices [7,8]. The color rendering index (CRI) [9,10] compares the chromatic effects of the test illuminant with those produced by a natural illuminant like daylight or the blackbody radiation. Its maximum value is 100. It measures, indirectly, how natural the colors appear when rendered by a specific light source. Its limitations are well documented and, among others, are the need for a reference illuminant and a set of test surfaces [11], ignoring the memory of familiar colors as potential internal reference [12] and the failures with spectrally structure light sources like LEDs [13].

An effect that may influence subjective evaluation of color rendering is the fact that colors of familiar objects seem to be remembered more colorful than the real ones [14,15]. In an experiment using a variable LED light source and real objects observers rate the similarity of the variable illuminated object to their notion of what the objects look like in reality [12]. It was found that the highest rates correspond to illuminations producing higher chroma in comparison with D65. Thus, naturalness condition is not maximized for D65 but for an illuminant that produces more saturated colors. There are also experimental evidences suggesting that the known color of the objects affects color appearance [16,17]. All these cognitive aspects of memory colors are likely to influence in a complex way the perception of naturalness from illumination.

Preference is typically an aspect often considered in image quality in the context of color reproduction [15,18] and it has been demonstrated to be related to naturalness [15]. In the context of art illumination, preference has been studied mainly with daylight [6,19].

Another important aspect of color rendering is the color rendering capacity of a light, that is, the potential for producing many different colors or chromatic diversity [20]. Several quantities have been suggested to estimate this property, like the volume of the object-color solid [21], the extension of the color gamut generated defined as the gamut area index (GAI) [22], and the color volume enclosed by the Munsell set, referred to as the chromatic diversity index (CDI) [23].

Several other indices have been suggested but it is generally accepted that none capture all aspects of rendering [2]. Color rendering can also be addressed in a more empirical way by optimizing illumination spectra for specific conditions or tasks. This approach avoids theoretical assumptions and is adequate in a world where lighting can be produced with almost unrestricted spectral composition like in LED lighting [24].

The goal of the present work was to investigate psychophysically what are the optimal spectral profiles to render complex indoor and outdoor scenes with specific visual properties. In the experiments observers could select illuminants from a set of metamers of D65 to optimize different visual aspects of complex outdoor and indoor scenes. The observers were instructed to select the illumination...
maximizing naturalness, individual preference and perceived chromatic diversity. The stimuli for the psychophysical experiments were derived from hyperspectral imaging of outdoor urban and rural scenes and indoor scenes and objects.

2. METHODS

A. Spectral Reflectance Data from Complex Scenes

Figure 1 represents colored pictures of the scenes tested in the experiments. The stimuli were simulations of these 12 scenes rendered with illuminants of variable spectral composition. The simulations were generated from the spectral reflectance of each pixel estimated from hyperspectral data. The data for the first and third scene of the first row were obtained from the publicly available hyperspectral image database from David Brainard's laboratory (http://color.psych.upenn.edu/hyperspectral/index.html). Data for the outdoor scenes were from a previously digitized set [25] and data for the other indoor scenes were obtained in our laboratory with a hyperspectral system similar to the one used for outdoor scenes. Brainard's data were acquired from 400 to 700 nm in 10 nm steps using narrowband interference filters and a monochromatic CCD camera with a spatial resolution of 2000 × 2000 pixels and 12 bit output [26]. The data for the other scenes were acquired over the range 400–720 nm at 10 nm intervals using a fast-tunable liquid-crystal filter and a low-noise Peltier-cooled digital camera with a spatial resolution of 1344 × 1024 pixels and 12 bit output.

The spectral reflectance of each pixel was estimated from the hyperspectral data by dividing the image data for each wavelength by the spectrum of the illuminant obtained at a given reference location for the same wavelength. In Brainard's hyperspectral images the spectral reflectance of each pixel was obtained by dividing the raw data by the illuminant spectrum of the scene obtained at a given reference location with a tele-spectroradiometer. In these estimations the illumination is assumed to be uniform over the scene. The images acquired with the tuneable filter were compensated for spatial nonuniformities of the optical system.

Scenes were selected to represent a diversity of conditions but the selection was constrained by the requirement that at least 95% pixels were within the gamut of the display device in all conditions of the experiment (see details of the colorimetric accuracy in the Sec. C, Stimuli). The scenes on the first row of Fig. 1 represent indoor scenarios with a mixture of colored objects. The first two contain fruits with familiar colors. The scenes in the second and third rows represent outdoor scenarios, the former containing more rural elements and the latter more urban elements. In all outdoor scenes there are elements with colors that can be familiar to the observers, like the vegetation and the roofs' colors, for example.

B. Illuminants

The intention was to test illuminants with variable spectral composition. For that purpose a large number of metamer of the standard illuminant D65 were generated. Metamers are spectra with the same color but different spectral compositions [27]. More specifically, a set of 1000 metamers of D65 were generated by the Schmitt's simple elements approach [28]. This methodolody generated spectral distributions with variable number of spectral maxima and variable degree of spectral smoothness. A metamer set of real positive functions $F$ can be described by a convex hyperpolyhedron volume in an $M$-dimensional space, where $M$ is the number of spectral bands considered. The spectral resolution considered was that of the hyperspectral data and therefore $M = 33$. The apexes of that hyperpolyhedron $S_j$ are functions that have at most three nonzero coordinates, that is, no more than three spectral bands. These are the simple elements and can be...
computed by deriving the combinations of three spectral bands that produce the same chromaticity. Any element $f_i$ of the set can be written as a positive barycentric combination of simple elements, i.e., for any $f_i \in F$ there is at least one set of $N \leq M$ positive numbers $\alpha_j$ such that:

$$f_i = \sum_{j=1}^{N} \alpha_j S_j,$$

where

$$\sum_{j=1}^{N} \alpha_j = 1.$$  \hspace{1cm} (2)

To generate the set of metamers of D65 the simple elements of D65 were first calculated, i.e., all combinations of sets of three spectral bands producing the chromaticity of D65. Then, by combining a variable number of simple elements metamers with different number of spectral bands were obtained. The spectral smoothness is determined by the number of simple elements and their weights used in the synthesis. The spectral smoothness was quantified in relation to the D65 by the absolute spectral difference $\delta_i$ between each $f_i$ and D65:

$$\delta_i = \sum_{k=1}^{23} |D65(\lambda_k) - f_i(\lambda_k)|.$$ \hspace{1cm} (3)

This parameterization of the metamers is instrumental and allowed the selection of metamers with different chromatic effects on the scenes. The metamers were generated for the spectral range 400 nm–720 nm, with 10 nm resolution, were normalized in energy, and were generated to be approximate uniformly distributed for $\delta$ in the range 0–2.0. Figure 2 shows some examples of metamers with different values of $\delta$. D65 is the spectrum with $\delta = 0$. To illustrate the colorimetric effects of metamers Fig. 2 also represents the same scene rendered with the metamers exemplified. The effects can be seen more clearly in the fruits and in the elements of the color chart.

For each metamer the CRI and CDI were computed. The CRI was computed accordingly to CIE and the CDI was computed by calculating the CIELAB color volume occupied by the set of 1269 samples from the Munsell Book of Color [29] when rendered by each metamer. The spectral reflectances were used as tabulated by the University of Joensuu Color Group University of Joensuu (http://spectral.joensuu.fu) and the volume was computed using a three-dimensional convex hull routine.

C. Stimuli

Stimuli were displayed by front projection of a Christie Mirage S + 4K 3 × DLP projector with a refresh rate of 70 Hz controlled by a dedicated and programmable hardware and software graphics system providing 24 bits per pixel in true-color mode (ViSaGe Visual Stimulus Generator; Cambridge Research Systems, Rochester, UK). The projector was calibrated in color and luminance with a telespectroradiometer (SpectraScan Colorimeter, PR-655; Photo Research Inc., Chatsworth, California). The projection area in the screen was 3 m (H) by 2 m (V) and at 3 m observation the images subtended horizontally 45 deg visual angle. The colorimetric performance of the DLP projector was similar to that of a good CRT. Checks of the quality of calibration were carried out during the period of the experiments to ensure the stability of the color reproduction.

Images were displayed with the resolution of 1024 × 768 obtained by cutting from the image with the original resolution. The average luminance of the displayed images was 20 cd/m². The colorimetric accuracy of the display device for all the images selected was such that at least 95% of the pixels brighter than 10 cd/m² in each image in all conditions of the experiment were within the gamut of the projector. When the color of the pixels was outside the gamut the colors were clipped to the closest RGB value. The chromatic effects resulting from this clipping were such that the average chromatic error across all possible 11815 images expressed as $\Delta E_{ab}$ in CIELAB space was 0.69 with a standard deviation of 1.7 and only 8% of all images tested exceeded the error value of 2.2, the detectable magnitude in complex images [30,31]. Visual inspection of the location of the larger chromatic errors

![Fig. 2](image-url) (Color online) Spectral power of metamers of D65 with different values of $\delta$, the absolute spectral difference relative to D65, and their colored effects for one of the scenes tested. D65 is the spectrum for $\delta = 0$. For illustration and better visualization each spectrum was normalized to unity at maximum power.
revealed that these were scattered across the image and did not concentrate on specific objects and therefore it was implausible that they had influence on the results reported here.

D. Procedure
Each session started with 3 min adaptation in which the observer looked at a uniform screen displaying the color of D65 with 20 cd/m². In each subsequent trial the projector displayed the simulation of each scene rendered under elements of the metamer set. In the first image of the trial each scene was displayed illuminated by a metamer selected at random from the set by the software. The observer could then select other rendering metamer by actuating in a miniature keyboard. The metamers were ordered accordingly to a specific criterion and actuating in the right or left key changed rapidly the metamer in the direction of increasing or decreasing order. To illustrate the setup of the experiment Fig. 3 shows the displayed screen and one observer carrying out the experiment.

Three conditions were tested. In one, the naturalness condition, the observers were instructed to select the illuminant on the scenes such that the colors of the objects appear as natural as possible. In other, the preference condition, observers were instructed to select the illuminant such that the appearance of the scenes was the most pleasant. In other, the chromatic diversity condition, observers were instructed to adjust the illuminant such that the number of the colors perceived in the scenes was as large as possible. The instructions were written and showed to the observers in the beginning of each experimental session.

The order of the metamers in the main experiment was different for each condition of the experiment. For the naturalness condition they were ordered linearly as a function of CRI, for the chromatic diversity condition they were ordered linearly as a function of CDI, and for the preference condition they were ordered linearly as a function of the product of CRI and CDI. In all cases the size of the step in each parameter was 5% of the possible range. Thus, in each trial only a subset of the metamers was available for the observers but because the initial metamer was selected at random from the complete set a large number was tested during the experimental sessions. This choice of parameterization was adopted to make the variation in appearance smoother and more directly related with the property under study in each condition. To study how much the parameterization can influence the outcome of the experiment all conditions of the experiments were also tested with the metamers ordered linearly as a function of δ also in steps of 5%. This was the control experiment and was carried out by a different set of observers.

The experiments took place in a dedicated laboratory located in the cave of the building and therefore without windows. The laboratory was completely in the dark except for the light for the DLP projector. Each observer did two sessions and in each the three experimental conditions were tested. In each session conditions were blocked and images and conditions were tested in random order. Each image was tested six times per condition per observer.

E. Observers
For the main experiments there were six observers. Five were naïve and one was one of the authors of this work. For the control experiments there were four different naïve observers. Each observer had normal or corrected acuity and normal color vision accessed by Rayleigh anomaloscopy and Ishihara plates. In the main experiment, two were female, aged 24 and 29 years and four were male aged 22, 25, 27, and 40 years old. In the control experiment three observers were male, aged 33, 16, and 20 years old, and one was female, aged 29 years.

3. RESULTS
A. Spectral Structure
Figure 4 shows the values of δ averaged across scenes and observers for the three experimental conditions. Data for the main experiment is represented by white bars and data for the control experiment is represented by listed bars. Error bars represent the standard error of the mean across scenes averaged for all observers. Figure 5 shows the average illuminant spectra selected by the observers for the three condition of the main experiment. On the left are represented the spectra for naturalness and for preference and on the right are represent the spectra for naturalness and for maximum...
chromatic diversity. For comparison purposes the spectra of D65 is represented in both graphs by the dotted line.

The average $\delta$ varies with the experimental condition both for the main and control experiments. The data for the control experiment show the same structure as for the main experiment but are systematically a little lower. This effect of the parameterization will be considered below.

For the main experiment, the average $\delta$ for the naturalness condition was about 1.0. If D65 was selected as the best illuminant for naturalness $\delta$ should be zero. The graphs in Fig. 5 show clearly that the average illuminant selected in this condition was spectrally more structured than D65 and had three local maxima around 400 nm, 490 nm, and 700 nm, that is, in the blue, green, and red regions of the visible spectrum. To estimate how chromatically different are the images of the scenes under the optimal illuminant and under D65, a pixel by pixel comparison was carried out and expressed as a color difference in the CIELAB color space. It was found that the average color difference across images and observers was 1.9 and the average chroma difference was $0.2$, indicating that observers select an illuminant producing smaller saturation than D65. This result is consistent with the fact that the average color volume for D65 is $2.9 \times 10^5$ whereas the average color volume for the optimal spectrum is $2.8 \times 10^5$, that is, a little smaller.

For the selected condition the average $\delta$ was about 1.25 indicating that observers prefer an illuminant with a little more structure than for naturalness. The spectral profile shown in Fig. 5 shows that the spectrum is indeed similar to the one selected for naturalness but the green maxima is a little sharper and shifted to lower wavelengths by about 10 nm. In this case the pixel by pixel comparison resulted in a larger color difference of 2.8 and a chroma difference of 0.31 indicating that the illuminant selection was to produce more colorful and chromatically diverse images.

For the maximum chromatic diversity condition the average $\delta$ was about 1.5. As shown in Fig. 5 this corresponds to a much more structured spectrum with well defined local maxima around 400 nm, 525 nm, and several for wavelengths longer than 625 nm. The differences between D65 and the optimal illuminant were clearly higher resulting in a pixel by pixel comparison color difference of 4.4 and chroma difference of 1.6.

The fact that there was a small but systematic decrease of $\delta$ across conditions when the metamer was ordered linearly as a function of $\delta$ (control experiment) may be due to the fact that for large values of $\delta$ the color appearance of the images can change dramatically for small increments or decrements. This arises because very different spectra may have similar absolute spectral differences to D65. In Fig. 6, which represents CRI and CDI expressed as a function of $\delta$ for the set of metamer tested, it can be seen that small variations in the horizontal axis produce large variation in CRI or CDI. This effect may compels the observers to scan more the regions of low $\delta$ thereby producing systematically smaller values. This small influence of the experimental parameterization does not affect the analysis and conclusions of this work.

### B. Rendering Indices

Figure 7 shows the CRI and CDI averaged across observers and scenes for all conditions of the experiment. Error bars represent the standard error of the mean across scenes averaged for all observers. For the naturalness condition the average CRI for the main experiment was about 60, much lower than the maximum value of 100. Thus, as expected from the spectral data, large CRI do not correspond to the most natural colors. Rather, naturalness is obtained with illuminants with relatively low CRI, confirming that this index is not a good descriptor for naturalness. Also, consistently with the spectral data, CRI for the preference condition is lower, about 40, and for the maximum chromatic diversity is about 17.

Maximum chromatic diversity is not obtained for the maximum value of CDI, $5.0 \times 10^5$, but for a value about 25% lower, suggesting that this quantity does not describe fully observers’ performance. CDI for the naturalness and preference conditions are considerably lower, being the latter a little higher than the former.

In general, these graphs show that these indices are correlated with human performance but they do not describe the data satisfactorily. The comparison between the main and
the control experiment show a systematically higher value of CRI and a systematically lower value of the CDI, consistently with the consideration made above.

C. Influence of Scene Content

To investigate the influence of the scene content on observers' performance the data was analyzed for individual scenes. Figure 8 shows the average values of δ, CRI, and CDI across observers expressed as a function of the scene number. There is a variation across scenes for all parameters and conditions of the experiment. This variation, however, does not seem to be systematically correlated with the scene content. The scenes with objects with more obvious diagnostic colors are the first two as they contain familiar fruits and the scenes with less chromatic references are the second two. It would be reasonable to expect that the presence of the familiar objects differentiates observers' preference but that is not observed. The comparison between indoor and outdoor scenes also does not show an obvious effect.

4. DISCUSSION

This study shows that the optimal spectral structure with CCT of 6500 K to render natural colors in complex outdoor or indoor scenes is not D65 but a less spectrally smooth illuminant with local maxima in the blue, green, and red regions of the visible spectral range (see Fig. 5). Consistently, the corresponding color rendering index is low, around 60. The average difference between the effects of D65 and of the optimal illuminant are, in absolute terms, small. However, the optimal illuminant produces colors a little less saturated than D65. The fact that subjective naturalness is different from objective naturalness is consistent with the results of experiments reported elsewhere [15]. In these experiments, however, observers preferred more saturated colors, an effect that is normally explained by the fact that the memory color of familiar objects is more saturated than the real colors [14].

In the experiments reported here the observers adjusted the colors of the scenes by selecting the illuminant from a large set of illuminants metamers of D65. The display screen was much larger than those normally used in this type of experiment, allowing the stimulation of about 45 deg visual angle. These viewing conditions improved the extent of adaptation in relation to the typical viewing angles of a few degrees. The methodology was different from the studies on memory color which were based on rating and therefore no action from the observers in the scenes was required.
Rather, this methodology is closer to the methods employed in the experiments about how the color of the object affects color appearance [16,17]. In these experiments observers adjusted the color of pictures of natural fruit objects until they appeared achromatic. It was found that the fruits were perceived as achromatic away from the physical achromatic point in a direction opposite of the typical color of the fruit. The conclusion is that objects with familiar colors are perceived chromatically different from objects of the same colors but that are not familiar. In our experiments all the scenes except one had some familiar object: fruit, grass, roofs, among others, and therefore these effects are probably present. How they influence observers’ settings is not clear.

The results for the preference condition show that observers prefer a similar illuminant to that for naturalness but producing colors in the scene a little more saturated. This result is consistent with image quality studies where it was found that quality is closely related with naturalness [15]. The visual factors determining preference for specific spectral power distribution are unknown. One possibility could be that the preferred spectral power distribution maximizes visual comfort. The relationships between visual discomfort and the properties of images have been studied psychophysically [32,33] in relation to the amplitude spectra obtained from a Fourier analysis of the images. Natural images show a characteristic variation of the amplitude spectra with the spatial frequency $f$, being the amplitude proportional to $1/f^\alpha$ with $\alpha$ roughly equal to one [34–36]. It was suggested that deviations from these natural statistics may produce discomfort [32]. To explore if these effects were present in the experiments reported here we compared the values of $\alpha$ obtained from a luminance and chromatic contrast analysis across the three conditions of the experiment and did not find any effect of the condition. The visual effects present in our experiments are probably of a different nature as our stimuli did not look uncomfortable or disagreeable.

For the maximum chromatic diversity condition the observers select an illuminant with very marked spectral structure and well defined local maxima in the blue, green-yellow, and red spectral regions, with a spectral structure that resembles the hypothetical lamp derived by Thornton for maximum chromatic discrimination [37]. It may be argued that in this condition of the experiment the observers are just adjusting the illumination to produce more saturated colors. These two factors are indeed difficult to disambiguate and for that reason during the instruction phase it was explained to the observers that more saturated colors does not mean necessarily more colors. The observers understood well this difference.

The spectral structure of the metamers was quantified with the absolute spectral difference $\delta$, which has the advantage of simplicity over more complex quadratic or multidimensional parameters. As the function of the parameter is to order the metamers in a way roughly correlated with the spectral structure, its relevance for the final spectra obtained from the experimental data is probably small. On the other hand, comparison between the main and control experiment shows that the influence of the parameter is real but minor.

Spectral optimization of the illuminant to produced maximum naturalness, preference, and chromatic diversity showed that specific spectrally structured illuminants may have important visual applications in lighting. The ideal illuminants to use to make products look natural or pleasant need not to be daylight illuminants; rather, spectra with spikes in the right spectral positions are the best ones for those purposes.

ACKNOWLEDGMENTS

This work was supported by the Fundação para a Ciência e a Tecnologia (grant PTDC/EEA-EEL/098572/2008). We are grateful to Kinjiro Amano, David H. Foster, and João M. M. Linhares for the use of the hyperspectral images.
REFERENCES