# COLOR FIDELITY OF CHROMATIC DISTRIBUTIONS BY TRIAD ILLUMINANT COMPARISON

*Marcel P. Lucassen, Theo Gevers, Arjan Gijsenij*

Intelligent Systems Lab Amsterdam, University of Amsterdam

## ABSTRACT

Performance measures for quantifying human color constancy and computational color constancy are very different. The former relate to measurements on individual object colors whereas the latter relate to the accuracy of the estimated illuminant. To bridge this gap, we propose a psychophysical method in which observers judge the global color fidelity of the visual scene rendered under different illuminants. In each experimental trial, the scene is rendered under three illuminants, two chromatic test illuminants and one neutral reference illuminant. Observers indicate which of the two test illuminants leads to better color fidelity in comparison to the reference illuminant. Here we study multicolor scenes with chromatic distributions that are differently oriented in color space, while having the same average chromaticity. We show that when these distributions are rendered under colored illumination they lead to different perceptual estimates of the color fidelity.

***Index Terms—*** Color constancy, color fidelity, triad comparison, chromatic distributions

## 1. INTRODUCTION

Color constancy is the ability of a visual system to maintain stable object color appearance despite substantial changes in the spectral power distribution of the illuminant. For quite some time, it has been recognized as one of the central themes in color research. Still, the methodological approaches in perception (human vision) studies and computational (computer vision) studies are very different. The background of this paper lies in our desire to bring the research areas on human color constancy and computational color constancy closer together.

The degree of constancy exhibited by human observers is often quantified by a color constancy index [1]. A common finding in the many psychophysical studies on color constancy is that human color constancy is not perfect. It depends on the experimental method employed and the observer's state of adaptation, among other things. The main techniques are color matching, color naming, nulling to maintain neutral appearance, discriminating between a change in illumination and a change in surface reflectance (operational approach), and identification of surfaces across illuminants [2]. What these techniques have in common is that each measurement (i.e. each observer response) relates to the appearance of a single object or patch in the scene.

In computer vision, the main approach to solving the color constancy problem is by estimating the unknown illuminant from the visual scene, after which reflectance may be recovered (e.g. [3], [4], [5]) or the color balance of images may be corrected for display or to support object recognition [6]. The performance of such color constancy algorithms is usually quantified by the angular error [7], a measure for the chromatic mismatch between the estimated illuminant and the true illuminant which is assumed to be known. So, the performance of computational color constancy is quantified by a number relating to a global illuminant.

This poses a problem for our comparison of color constancy performance measures: the psychophysical measurement relates to a single (local) object whereas the computational measure relates to a single (global) illuminant. To solve this, we here propose a novel method for assessing the degree of human color constancy, featuring two new methodological elements. First, the observers are asked to judge the color fidelity of the whole scene, instead of a single color or object. Second, in each experimental trial the observers deal with a scene rendered under three illuminants (one reference and two test illuminants) instead of the usual two (one reference and one test illuminant). We denote this by the term "triad illuminant comparison". Using this method, in this study we measure the color fidelity of chromatic distributions having the same average chromaticity, but different orientations in color space. Color constancy algorithms that are based on the scene averaged chromaticity would result in identical illuminant estimations for these distributions. We here show, however, that these distributions lead to different perceptual estimates of the color fidelity. This effect appears to be dependent on the alignment of the orientation of the chromatic distribution and the direction of the illuminant change. Such a perceptual effect might be incorporated into color constancy algorithms to improve the correspondence between human and computational color constancy.

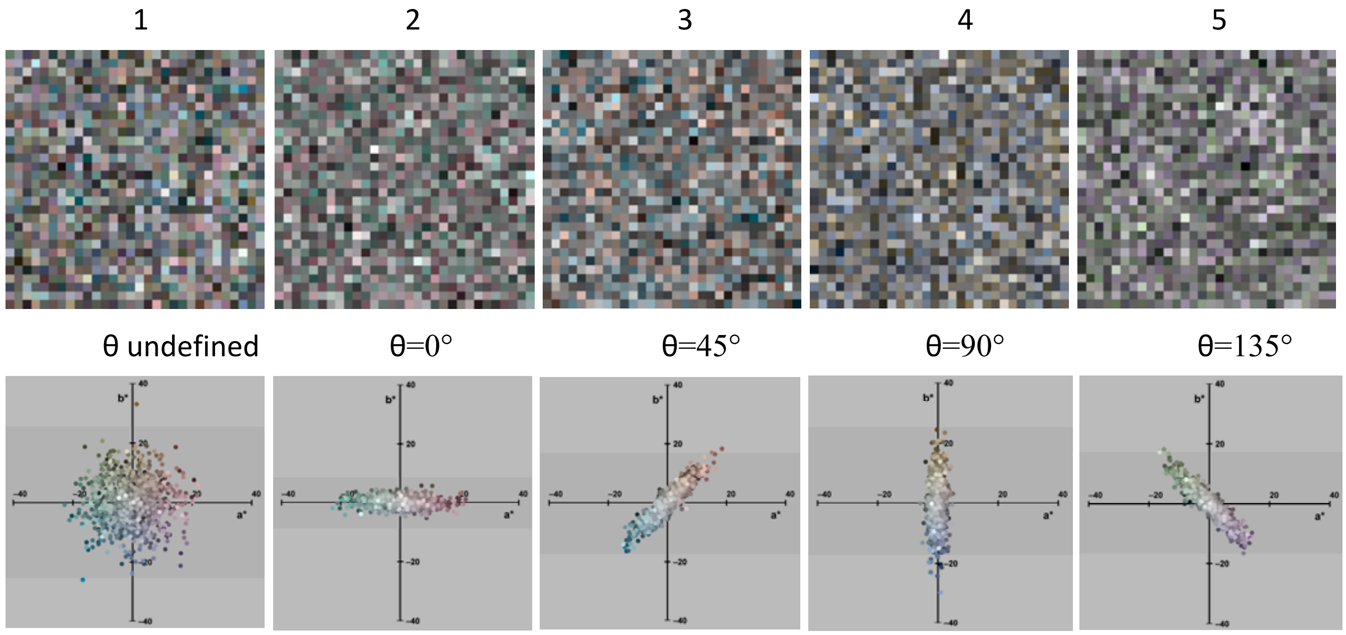


Figure 1: *Multicolor test scenes (top row) with their chromatic distributions under D65 reference illumination. The angle θ denotes the angle between the positive a* axis and the major axis of the distribution in the CIELAB a*b* plane (bottom row).*

## 2. METHODS

### 2.1. Multicolor test scenes

The test scenes used in this paper are multicolor images composed of 900 square color patches, arranged as a 30 x 30 matrix (see Figure 1). The distribution of colors in CIELAB color space was controlled. In our first distribution (labeled 1 in Figure 1), the 900 patches follow a Gaussian distribution, with standard deviations along the L*- axis twice the standard deviations along the chromatic a* and b* axes. This was done to match the distribution of RGB-derivatives of the Corel image database, which contains 40,000 images representative of the ‘real world’ [8]. So, our first distribution approximates the statistics of natural images. In addition we studied four chromatic distributions having a 5 to 1 ratio in the standard deviations along the major and minor axes of the distribution in the a*,b* plane. These four distributions (labeled 2 through 5 in Figure 1) differ in the elongation of the major axis, denoted by θ. This is the angle between the distributions’ major axes and the positive a*-axis in Figure 1. Although the 5 distributions have different orientations in chromaticity space, the average color remains centered on the neutral point.

### 2.2. Triad illuminant comparison

The triad illuminant comparison method involves a test scene rendered under three illuminants (Figure 2). One of these illuminants is the reference (R) while the other two are the test illuminants (T1 and T2). An observer has to indicate his/her preference for one of the test illuminants. They do this by visually comparing the colors of the test scene rendered under the test illuminants with the colors rendered under the reference illuminant. In essence the observer’s task is to judge the differences between two image pairs, the first pair of images (T1,R) being the scene rendered under test illuminant T1 and under the reference illuminant R, the second pair (T2,R) being the scene under test illuminant T2 and R.

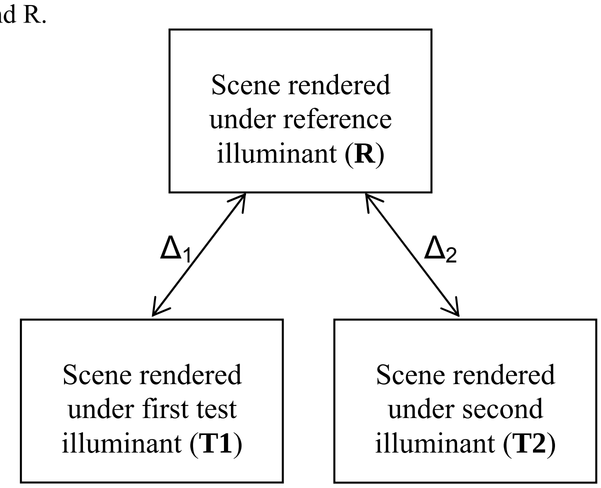


Figure 2: *Triad illuminant comparison. The relative color fidelity of the test scene under test illuminants T1 and T2 is measured by observers indicating which of the two differences ($\Delta_1$ or $\Delta_2$) appears smaller.*

### 2.3. Selection of illuminants

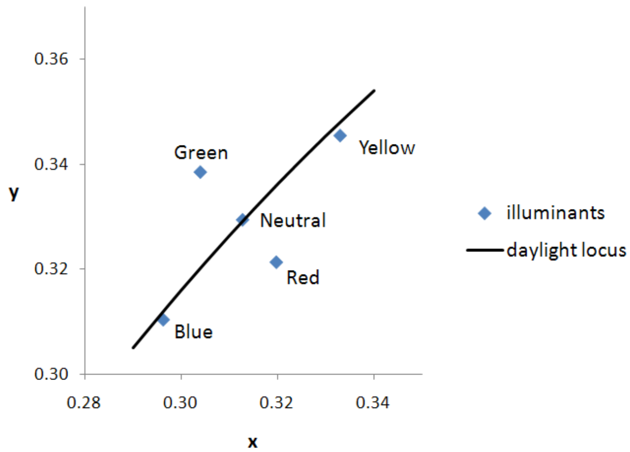


Figure 3: *CIE 1931 x,y chromaticity diagram showing the daylight locus and the positions of the illuminants (four chromatic and one neutral) used in this study. Note that the Blue and Yellow illuminants are located along the locus of natural daylight while Red and Green are perpendicular to it.*

In line with previous studies ([9], [10]) we select a neutral reference illuminant (D65) and four chromatic illuminants all composed with the CIE basis functions for spectral variations in natural daylight. In Figure 3 the positions of the illuminants are plotted in the CIE 1931 x,y chromaticity diagram. Delahunt & Brainard [9] investigated whether color constancy would be better for illuminant changes along the daylight locus than for illuminant changes perpendicular to it, but did not find experimental evidence that unambiguously supports this hypothesis. Here we use the same paradigm with the Yellow and Blue illuminants along the daylight locus, and another two (Red and Green) perpendicular to it, but without the constraint that they are perceptually equidistant from the neutral reference point. Perceptual equidistance of object colors rendered under these illuminants would only be guaranteed for spectrally non-selective (achromatic) samples seen in isolation, but it does not necessarily imply perceptual equidistance for individual chromatic samples or a distribution of chromatic samples.

Therefore, we adopt a different criterion for illuminant selection. We define a purely physical measure for the induced global difference in the reflected light signal when the illuminant is changed from neutral to one of the chromatic illuminants. Let the reflected light signal L be described as the wavelength-by-wavelength product of the illuminant spectral power distribution E(λ) and a sample's reflectance function ρ(λ):

$$L = \int_{\lambda} E(\lambda)\,\rho(\lambda)\,d\lambda \tag{1}$$

This signal is purely physical because it does not include the sensitivity of the visual system. When changing the illuminant from $E_1$ (neutral) to $E_2$ (chromatic), the associated change in the reflected light signal is given by

$$\Delta L = L_2 - L_1 = \int_{\lambda} \left[E_2(\lambda) - E_1(\lambda)\right]\rho(\lambda)\,d\lambda \tag{2}$$

We modify the purity of the chromatic illuminants (initially equidistant from the neutral point) such that more or less identical cumulative distributions of $\Delta L^2$ are obtained. So, a spectrally non-selective sensor receiving the image of the test scene would not notice the transition from the neutral illuminant to one of the chromatic illuminants. Modification of the purity of the chromatic illuminants is done by mixing the spectral power distribution of the chromatic illuminant with that of neutral illuminant D65. This process is described by

$$E'(\lambda) = \frac{E(\lambda) + x\,D65(\lambda)}{1+x} \tag{3}$$

in which E'(λ) represents the spectrum of the adjusted illuminant at wavelength λ, D65(λ) is the spectral power of illuminant D65 and x is the mixing factor that regulates the mixing of E(λ) and D65(λ). We derived mixing factors of 3.55, 2.54, 2.0 and 2.71 for the Red, Green, Yellow and Blue illuminants, respectively.

### 2.4. Simulation of colors under different illuminants

The usual way to simulate object colors under different illuminants is to calculate XYZ tristimulus values according to:

$$\begin{aligned} X &= k\int_{\lambda} E(\lambda)\,\rho(\lambda)\,\bar{x}(\lambda)\,d\lambda \\ Y &= k\int_{\lambda} E(\lambda)\,\rho(\lambda)\,\bar{y}(\lambda)\,d\lambda \\ Z &= k\int_{\lambda} E(\lambda)\,\rho(\lambda)\,\bar{z}(\lambda)\,d\lambda \end{aligned} \tag{4}$$

in which E(λ) represents the spectral power distribution of the illuminant, ρ(λ) is the spectral reflectance function of the object and $\bar{x}, \bar{y}, \bar{z}$ represent the 1931 color matching functions for the 2° standard observer. The factor *k* serves to normalize Y at 100 for a perfect white reflectance.

Our patches are defined in terms of CIELAB values rather than spectral reflectance functions. This poses a problem since an infinite number of reflectance functions may result in identical XYZ tristimulus values (and hence identical CIELAB values) under one illuminant. Therefore, a selection criterion for picking one reflectance function is needed. We applied van Trigt's method [11] to estimate the *smoothest* reflectance function (SRF) from a set of tristimulus values, where the smoothness measure is defined as the square of the derivative of the reflectance function with respect to wavelength, integrated over the visual range.

We thus assume that the CIELAB values of our color patches in the reference condition result from illumination of the estimated reflectance functions by D65, our reference illuminant. Conversion of the resulting XYZ values to RGB drive values for displaying the colors on the color monitor is done using the sRGB profile, to which our monitor was calibrated.

### 2.5. Procedure

Eight subjects (including the authors) participated in our experiment. They all have normal color vision according to the HRR color vision test and normal or corrected-to-normal visual acuity. The test scenes were rendered under the different (simulated) illuminants and shown on a calibrated color monitor (EIZO CG211). Using the triad illuminant comparison method as explained in section 2.2, in each trial the observers indicated which of the two test illuminants had better color fidelity. Since we are using four chromatic illuminants, six unique illuminant pairs are constructed. Each subject judged 5 chromatic distributions x 6 illuminant pairs = 30 trials. All trials were replicated, leading to 60 responses per subject.

## 3. RESULTS

In each experimental trial observers indicated a preference for one of the two competing test illuminants. The illuminant indicated as having better color fidelity received 1 point, the other received no points in that trial. In cases where the observer could not choose one or the other, both illuminants received 0.5 point. Given that each test illuminant appeared three times in competition against the other illuminants, the maximum score for an illuminant to obtain is 3. And since trials were replicated, the maximum final score was 6. In Figure 4 we show the visual score thus obtained, averaged over the 8 observers. Error bars denote the standard error. On average, in 73% of the trials the repetition resulted in the same response.

We here analyze the results by first discussing the data for the first chromatic contribution (having identical standard deviations in a*,b* color space). Then, the data for the other four chromatic distributions are discussed using Figure 5. This graph shows the same data as Figure 4, but the visual scores are shown relative to the visual scores of the first chromatic distribution.

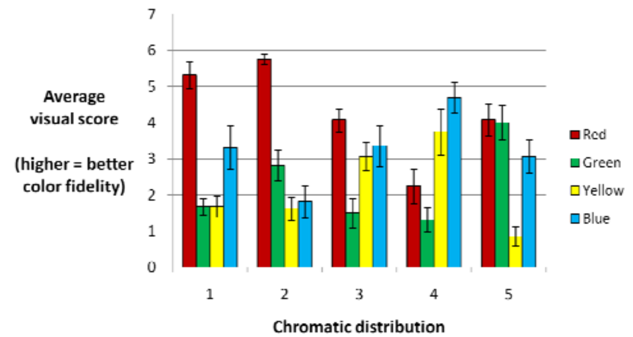


Figure 4: *Average visual scores per chromatic illuminant (Red, Green, Yellow, Blue) obtained for the five different chromatic distributions shown in Figure 1. A higher score means a better color fidelity. Error bars denote standard errors (n=8).*

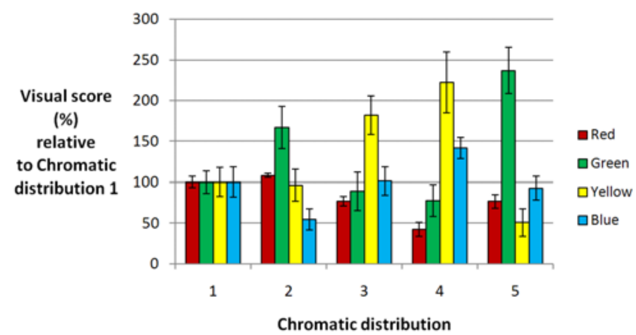


Figure 5: *Same data as in Figure 4, now plotted relative to the scores obtained for the first chromatic distribution (the 3D Gaussian distribution in CIELAB space).*

The data for the first chromatic distribution in Figure 4 clearly shows a large difference between the average visual score for the red illuminant and the other illuminants. Apparently, our multicolor test scene under the red illuminant results in higher color fidelity than the other illuminants. Second best is the Blue illuminant, and Green and Yellow have the lowest visual scores. Why do the Red and the Blue illuminant lead to higher color fidelity than Yellow and Green? We remind that the purity of the illuminants was adjusted to yield identical distributions of physical changes in the reflected light signal. The mixing factors for Red and Blue mentioned at the end of section 2.3 are larger than for Yellow and Green, resulting in lower purity of the first two illuminants. When comparing a scene rendered under a lower purity illuminant and a higher purity illuminant, the scene under the lower purity illuminant will have higher color fidelity. So, this (partially) explains the different visual scores for the first chromatic distribution.

The results for the other four chromatic distributions are best explained using Figure 5, which is a re-plot of the data of Figure 4 relative to the visual scores obtained for the first distribution. This allows an easy comparison of the changes that are measured due to changing the chromatic distribution. For the second chromatic distribution, having its major axis along the a*-axis of CIELAB color space (roughly the red-green axis), the color fidelity for the Red and Green illuminant increases, and decreases for Yellow and Blue. Likewise, for chromatic distribution number 4, having its major axis along the b* axis of CIELAB color space (roughly the yellow-blue axis), the color fidelity for the Yellow and Blue illuminants increases while it decreases for Red and Green. For chromatic distributions 3 and 5, having their major axes at a 45° angle with the a* and b* axes, color fidelity for Yellow and Green illumination is mainly affected. These findings are summarized as follows:

1) Chromatic distributions that differ in their direction in color space, lead to different color fidelity judgments under changing illumination.

2) When the direction of the illuminant change (from neutral to chromatic) is parallel to the major axis of the chromatic distribution, color fidelity of the rendered scene is better than when the direction of the illuminant change is orthogonal to the major axis.

## 4. DISCUSSION

We have introduced a psychophysical method (triad illuminant comparison) that quantifies the color fidelity of a visual scene under illuminant changes. More precisely, what is being measured is relative, since observers are asked to indicate which of the two test illuminants leads to the best color fidelity. The measurement does not specify the absolute difference between the rendered images and the reference. Still, by measuring all illuminant combinations, a relative ranking of these illuminants is possible. The advantage of the method is that it tells us something about the quality of the color image as a whole. Compared to the usual local measures for color constancy (on a single object), this global assessment is more in line with the performance measure of computational color constancy, being the angular error (e.g. [7]). This measure indicates the chromatic mismatch between the true illuminant and the estimated illuminant which is assumed to be spatially uniform. After illuminant estimation, usually a color correction is applied to an image. Any mismatch in the estimated illuminant results in a different global color correction, which would be measureable using our color fidelity paradigm.

If human color constancy would be perfect, the effects of changing illumination would not be perceived at all and the visual scores in Figures 4 and 5 should not change with chromatic distribution. However, we showed that the color fidelity of chromatic distributions under changing illumination depends on the alignment of the directions of the chromatic distributions and the illuminant change. When the directions of these two are parallel, color fidelity is best. In other words, when the direction of the illuminant change is in the direction of the major axis of the chromatic distribution, the color changes are least noticeable.

How may computational color constancy algorithms incorporate this perceptual ‘advantage’? Since many different algorithms are available, the locations of their estimated illuminants in relation to the major axis of the chromatic distribution (if present in the scene) may be used to select one or more algorithms. Our results predict that a mismatch in the estimated illuminant parallel to the chromatic distribution would be preferred over a mismatch perpendicular to it.

The current paper presents the work on synthetic images for which the chromatic distribution was under control. Future work will focus on image data sets with natural images and the prediction of data with different models relating to processing of visual information. This will allow us to test whether the preferences given by the observers are also given when presented with more natural scenes.

## 5. CONCLUSIONS

The triad illuminant comparison method results in useful data on color fidelity of chromatic distributions under changing illumination (the classical color constancy setting). For the synthetic test scenes studied here, illuminant changes parallel to the major axis of the chromatic distribution lead to higher color fidelity than illuminant changes perpendicular to it.

## 6. REFERENCES


[1] L. Arend, A. Reeves, J. Schirillo, and R. Goldstein, “Simultaneous color constancy: patterns with diverse Munsell values”, *Journal of the Optical Society of America A* 8, pp. 661–672, 1991.

[2] H.E. Smithson, “Sensory, computational and cognitive components of human colour constancy”, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360, pp. 1329–1346, 2005.

[3] L.T. Maloney and B.A. Wandell, "Color constancy: a method for recovering surface spectral reflectance," *Journal of the Optical Society of America A* 3, pp. 29-33, 1986.

[4] D.H. Brainard and W.T. Freeman, "Bayesian color constancy," *Journal of the Optical Society of America A* 14, pp. 1393-1411, 1997.

[5] G.D. Finlayson and G. Schaefer,” Solving for Colour Constancy using a Constrained Dichromatic Reflection Model”, *International Journal of Computer Vision*, 42 (3), pp. 127-144, 2001.

[6] B. Funt, K. Barnard, and L. Martin, “Is colour constancy good enough?”, *Proceedings of the 5th European Conference on Computer Vision*. Berlin: Springer-Verlag, pp. 445–459, 1998.

[7] S. Hordley and G. Finlayson, "Reevaluation of color constancy algorithm performance”, *Journal of the Optical Society of America A*, 23, pp. 1008-1020, 2006.

[8] J. van de Weijer, Th. Gevers and A. Gijsenij (2007), “Edge-Based Color Constancy”, *IEEE Transactions on Image Processing*, 16 (9), pp. 2207-2214, 2007.

[9] P. Delahunt and D. Brainard, “Does human color constancy incorporate the statistical regularity of natural daylight?” *Journal of Vision* 4, pp. 57–81, 2004.

[10] A. Gijsenij, T. Gevers, and M. P. Lucassen, "A perceptual analysis of distance measures for color constancy", *Journal of the Optical Society of America A*, 26 (10), pp. 2243-2256, 2009.

[11] C. van Trigt**, "**Smoothest reflectance functions. I. Definition and main results," *Journal of the Optical Society of America A,* 7, pp. 1891–1904 (1990).