Metamer Mismatching and Its Implications

by

Emitis Roshan

M.Sc., Sharif University of Technology, 2014
B.Sc., Isfahan University of Technology, 2011

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
School of Computing Science
Faculty of Applied Sciences

Declaration of Committee

Name: **Emitis Roshan**

Degree: **Doctor of Philosophy**

Thesis title: **Metamer Mismatching and Its Implications**

Committee: **Chair:**   Ali Mahdavi-Amiri
                                      Assistant Professor, Computing Science

                         Brian Funt
                         Supervisor
                         Professor Emiritus, Computing Science

                         Yagiz Aksoy
                         Committee Member
                         Assistant Professor, Computing Science

                         Kathleen Akins
                         Examiner
                         Professor, Philosophy

                         Anya Hurlbert
                         External Examiner
                         Professor
                         Centre for Transformative Neuroscience
                         Biosciences Institute
                         Newcastle University

Abstract

Due to the phenomenon of metameric mismatching, Logvinenko [PLoS One, 2015 Sept. 10, 10(9)] has pointed out that color is not an intrinsic property of an object. Not only can a change in the illumination lead to a substantial change in the measured or perceived color of an object, two objects that are metamers (i.e., their color signals match under one illuminant) may not match under a different illuminant. This phenomenon is referred to as metamer mismatching. Given a surface reflectance illuminated by a given light, there can be many other surface reflectances for which the eye provides an identical LMS cone-triple response or an identical RGB response in the case of a camera. The tristimulus color values of these metamer reflectances can disperse into many different tristimulus values under a different illuminant. The set of all such possible tristimulus values defines a convex hull known as a Metamer Mismatch Body (MMB). Metamer mismatching poses several challenges in color-based machine vision, such as color prediction, color discrimination variability and the color accuracy of a digital camera. In this work, a variety of approaches to measuring the extent of metamer mismatching and partially mitigating its effects are considered.

We begin by investigating the performance of existing color prediction methods in predicting the results of asymmetric color matching experiments (i.e., finding the least-dissimilar matching pairs of colored papers under different illuminants). Because of the possibility of metamer mismatching, it is a mistake to interpret one answer as the 'correct' answer. However, we demonstrate that all the computational methods studied are not capturing some important aspects of the observers' least-dissimilar matching strategy. MMBs are 3D volumes that tend to be quite wing-like in shape. By modeling the MMB by its equivalent (in terms of its inertial moments) ellipsoid, a new metric for evaluating the colorimetric accuracy of digital color cameras is proposed. The advantage of the new metric is that it is based on a theoretical principle rather than simply computing the average error over a chosen set of representative test reflectances. Furthermore, a theory about the inverse relationship between color discrimination ellipsoids and the extent of metamer mismatching is proposed. Statistical analysis over the existing datasets provides evidence that metamer mismatching can possibly explain why color discrimination varies throughout color space as it does. It is shown that the proposed theory paves the way to predict the Just Noticeable Differences (JNDs) in different regions of a color space.

Dedication

To my parents for their unconditional love and unwavering faith in my abilities
To my beloved husband for his constant encouragement, patience and support

Acknowledgements

I would like to express my deep and sincere gratitude to my senior supervisor, Professor Brian Funt, for his invaluable guidance, continuous support, and patience throughout my doctoral studies. I consider myself extremely fortunate to have had the privilege of being his student. His mentorship goes beyond this dissertation and extends to countless real-life lessons that I have learned from him. Furthermore, I would like to thank the members of my examination committee, Dr. Yagiz Aksoy, Dr. Kathleen Akins, Dr. Anya Hurlburt, and the defense chair, Dr. Ali Mahdavi-Amiri for their time and insightful comments on the thesis.

I would like to extend my heartfelt appreciation to my dear parents, Mohammad and Khorshid, and my lovely sister, Fatemeh, from the bottom of my heart, for their unwavering support along the way, and their constant motivation that kept me going.

Finally, I would like to express my deepest gratitude to my beloved husband, Hamid, for being my rock, and providing me with limitless love and companionship. This thesis would have been impossible without him being by my side at each and every step of this journey. I am forever grateful for his presence in my life.

Table of Contents

List of Tables

List of Figures

Chapter 1

Color Vision

Color vision provides humans with important sensory information that gives the ability to see the physical environment. We all are intuitively aware of what color is but studying color theory, its development, the state-of-the-art algorithms and their wide applications in different fields provide us the material required to overcome some of the current existing problems.

The biological process of vision uses light reflected from the surrounding world as the input to interpret the environment. Our knowledge about the very complex process of human color vision is used to model its behavior which can serve as an insight about how machine vision can be solved. There are three types of cone cells with different spectral sensitivities in our major sensory organ (eye) that enable us to see colors. Digital cameras, monitors, printers and scanners are all attempts to imitate the characteristics of human's trichromatic color vision. Trichromacy assumes that only the three types of cones provide a perception of color and two lights will be indiscriminable when they produce equal rates of photopigment absorption for each cone type. However, the human color perception is much more complicated and is only initiated by these receptors in the retina. Although it fails to explain many phenomena such as visual aftereffects, simultaneous contrast and surface-based color perception, trichromacy and the opponent process theory are the explanatory bases that help to describe how our color vision system works.

In this chapter we briefly talk about color image formation, the metamer mismatching concept and its effect on the measured color under different viewing conditions. The material presented here provides some fundamental background knowledge required for this research. In the next chapter, some of the state-of-the-art color prediction methods are reviewed and their performance in predicting the asymmetric color matching experiment results is analyzed. In the third chapter, the existing sensor quality metrics for digital cameras are studied. Also, a novel method for evaluating the colorimetric accuracy of digital color cameras is proposed. The main advantage of the proposed metric is that it determines how much color error can arise in principle rather than measuring the error over a set of sample reflectances. Color discrimination experiments and their applications in the test and
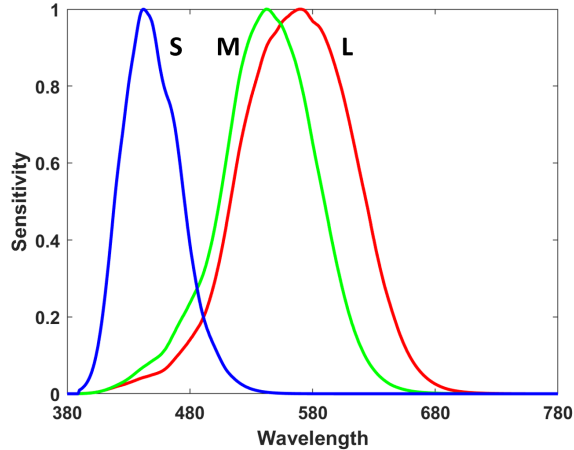
Figure 1.1: The Normalized spectral sensitivities of human cones, S, M and L types.

the development of new color spaces and color difference formulas are discussed in the fourth chapter. In addition, a theory of color discrimination based on the uncertainties reflected in the extent of metamer mismatching is proposed and investigated. The proposed theory paves the way for predicting the Just Noticeable Differences (JNDs) in color in chapter 5.

1.1 Color Image Acquisition

For many years quantifying color values from a scene has been the main goal of the imaging technology so that they can be accurately reproduced later. The human visual system is only sensitive to a part of the electromagnetic spectrum, called the visible spectrum. In other words, the human eye responds to the wavelengths from about 380 to 780 nanometers. One way to measure the sensitivities of human photoreceptors, as described in [12], is by projecting a known amount of light and measuring how much light is absorbed by the photopigment molecules. Such experiments have yielded response curves for three different kinds of cones in the retina of the human eye with peaks at long (red, 560 nm – 580 nm), medium (green, 530 nm – 540 nm), and short (blue, 420 nm – 440 nm) wavelengths. The concentration of "green" and "red" cones is higher than "blue" cones in the fovea centralis. The shapes of the curves obtained by these measurements are shown in Figure 1.1. The relative heights for the three types are set equal for lack of detailed data.

$$S = \int_{\lambda=380}^{\lambda=780} r_1(\lambda)P(\lambda)\,d\lambda, M = \int_{\lambda=380}^{\lambda=780} r_2(\lambda)P(\lambda)\,d\lambda, L = \int_{\lambda=380}^{\lambda=780} r_3(\lambda)P(\lambda)\,d\lambda \qquad (1.1)$$

The power spectrum of the light $P(\lambda)$ that reaches our eyes are weighted by the spectral sensitivity of each cone cell $r_i(\lambda)$ according to Equation 1.1 to render three values of the stimulus as a specification of the color of that light. All possible tristimulus values denoted
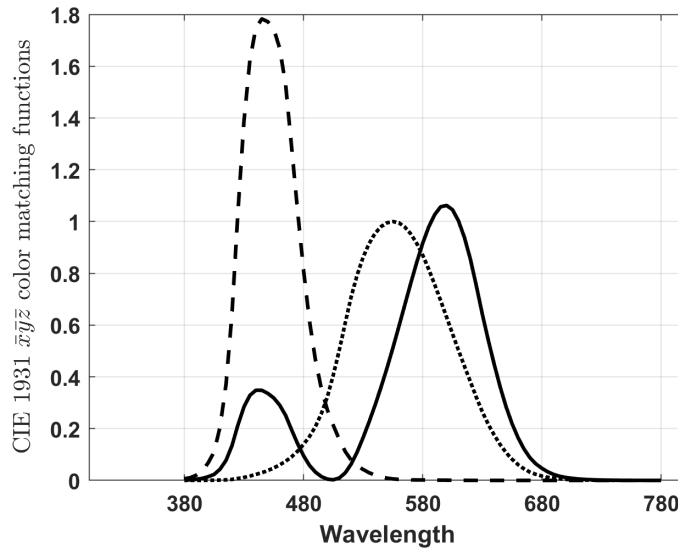
Figure 1.2: The CIE 1931 color matching functions [19] [92]

"L", "M", and "S", indicate a 3-dimensional space called LMS color space that quantifies human color vision, where $r_1(\lambda)$, $r_2(\lambda)$ and $r_3(\lambda)$ are the short (S-), medium (M-) and long (L-) cone sensitivities known as the cone fundamentals.

The tristimulus values, depend on the observer's field of view due to the distribution of the cones in the eye. To overcome this problem, in 1931 the Commission Internationale de l'Éclairage (CIE) [19][92] defined the color-matching functions to represent an average human's chromatic response within a 2° arc inside the fovea as it was believed that most of the color-sensitive cones reside within a 2° arc of the fovea. These functions, known as the CIE 1931 $\bar{x}\bar{y}\bar{z}$ 2° Standard Observer, are plotted in Figure 1.2.

Although the trichromatic process is lossy, meaning that we cannot go from the tristimulus to the original spectral distribution of the light, the CIE color matching functions help to numerically specify a color. It is also worth mentioning that it does not attempt to describe how colors appear to a human observer.

The color of a pixel recorded by a camera is determined as a function of the light source with spectral power distribution $E(\lambda)$, the surface reflectance $x(\lambda)$, and the camera sensitivity functions $r_i(\lambda)$:

$$\varphi_i(x) = \int_{\lambda_{min}}^{\lambda_{max}} E(\lambda)x(\lambda)r_i(\lambda)d\lambda \tag{1.2}$$

$E(\lambda)$ defines the radiant power of the light at each wavelength of the visible spectrum. The CIE also specifies [3] [4] the relative spectral power distributions of typical lights, e.g., different phases of daylight and fluorescent illuminants. These standard illuminants provide a basis for comparing images or colors recorded under different lighting conditions. Some

3
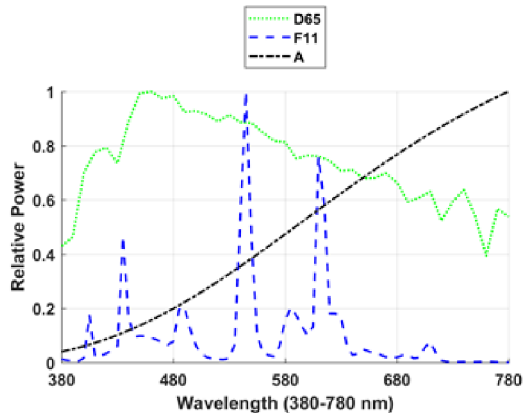
Figure 1.3: The relative spectral power distributions (SPDs) of CIE illuminants D65, F11 and A
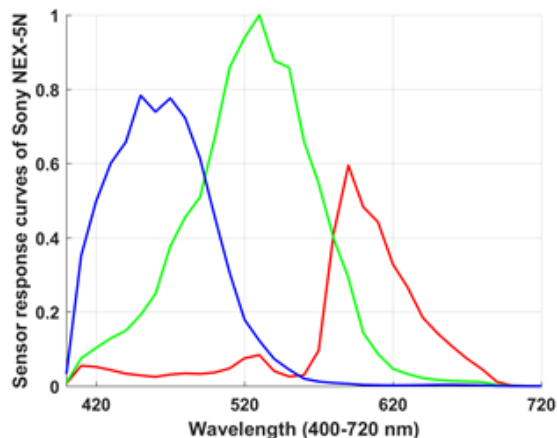
Figure 1.4: The sensor response curves for Sony NEX-5N [41]

of these illuminants that are usually named by a letter-number combination are plotted in Figure 1.3.

$x(\lambda)$, as a function of wavelength, indicates the proportion of the incoming light that is reflected off the surface. In other words, it represents the percentage of reflected radiant energy at each wavelength, where zero and one mean perfect absorption and perfect reflection, respectively. Basically, the reflection depends on both the viewing and illumination direction but here the surface is assumed to obey Lambert's law; meaning it appears equally bright from all viewing directions.

$r_i(\lambda)$, the camera's spectral sensitivity, determines its relative efficiency in detecting the radiation at each wavelength. It relates the light coming from the scene to the captured RGB values. The sensor response curves of a Sony NEX-5N camera measured by Jiang et al. [41] are shown in Figure 1.4 as an instance.

4
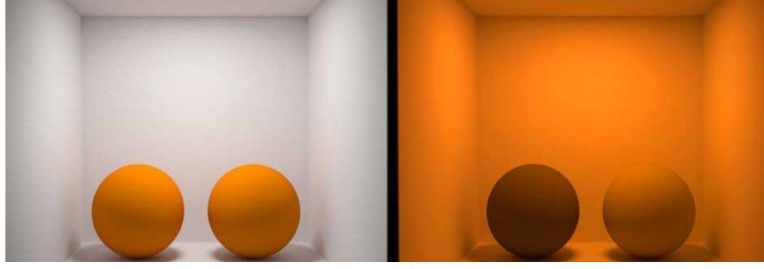
Figure 1.5: Two balls are placed in a lightbooth [2]. Left: the two balls are metamers under one light, meaning they produce the same color signals under that light. Right: the color signals of the balls do not match anymore under a different light. This is called metamer mismatching.

The set of an illuminant $E(\lambda)$ and a sensor's spectral sensitivity $r_i(\lambda)$ is called color mechanism φ_i with spectral sensitivity $s_i(\lambda) = E(\lambda)r_i(\lambda), i = 1, \ldots, n$ [54]. The vector $\Phi(x) = (\varphi_1(x), \varphi_2(x), ..., \varphi_n(x))$ of the sensor responses $\varphi_i(x)$ defined in Equation 1.2 are referred to as the 'color signal' produced by the sensor set in response to the surface reflectance $x(\lambda)$ illuminated by $E(\lambda)$ [54].

As is clear from Equation 1.2, different spectral reflectance functions may result in the same color signal. In that case the reflectance functions are called metamers. Metamerism will be studied in detail in the next section.

1.2 Metamer Mismatching

Two objects with two different spectral reflectance functions $x(\lambda)$ and $x'(\lambda)$ are called metamers if they produce equal tristimulus values with a given color mechanism. However, a change in the color mechanisms may cause the color signals of metameric reflectance functions $x(\lambda)$ and $x'(\lambda)$ to mismatch. This phenomenon is known as metamer mismatching. For instance, the RGB values of two different objects may match when a picture is captured with a given camera under a given light (Figure 1.5). But the color signals of these two objects may mismatch if the image is captured under a different illuminant, or with a different camera. The metamer mismatching resulting from a change either in the illuminant or the sensor sensitivities is referred to as illuminant-induced and observer-induced metamer mismatching [54], respectively.

The set of all color signals that can be produced by a set of n color mechanisms $(\varphi_1, \varphi_2, ..., \varphi_n)$ forms a convex hull in R^n, which is known as the object color solid (OCS) [54].

In general, given the tristimulus values of an object with respect to a given set of color mechanisms, there can be many other surface reflectances that lead to the same tristimulus values with respect to the same color mechanisms. The color signals of these metamers may be dispersed into many different tristimulus values if the color mechanisms change. The set of

all possible colors forms a convex body [107] which is known as Metamer Mismatch Volume (MMV) [56] or Metamer Mismatch Body (MMB). We will use the Metamer Mismatch Body (MMB) terminology proposed by Zhang et al. [97] to avoid the confusion of referring to a volume's volume. Many researchers have tried to address the problem of finding the metamer set or the MMB in different ways. Some of the proposed algorithms are reviewed in the next section.

1.3 Metamer Set

Wyszecki [104] first concluded that a reflectance function x_λ can be decomposed into two parts, the fundamental and metameric black. Let's say that the CIE 1931 tristimulus values XYZ for an object can be approximated by a set of sums as in equations 1.3-1.5:

$$X = \sum_{\lambda=380}^{\lambda=780} x_\lambda E_\lambda r_\lambda^1 \tag{1.3}$$

$$Y = \sum_{\lambda=380}^{\lambda=780} x_\lambda E_\lambda r_\lambda^2 \tag{1.4}$$

$$Z = \sum_{\lambda=380}^{\lambda=780} x_\lambda E_\lambda r_\lambda^3 \tag{1.5}$$

Where x_λ, r_λ^1, r_λ^2 and r_λ^3 are the N-vectors (usually $N = 40$ with the interval of 10nm between 380nm and 770nm) showing the object's spectral reflectance function and CIE XYZ color matching functions, $\overline{x}$, $\overline{y}$ and $\overline{z}$, respectively. E_λ is a $N \times N$ diagonal matrix where the diagonals are the illuminant's spectral energy distribution. Wyszecki suggested that x_λ can be decomposed into two additive components, one being a particular solution $x_\lambda^{(p)}$ of Equations 1.3 to 1.5, and the other being the solution $x_\lambda^{(h)}$ of the homogenous part with $(0, 0, 0)$ as its tristimulus values. Since the tristimulus values of $x_\lambda^{(h)}$ are equal to zero, it can be any N-dimensional vector orthogonal to the three vectors $E_\lambda r_\lambda^1$, $E_\lambda r_\lambda^2$ and $E_\lambda r_\lambda^3$. Therefore $N-3$ linearly independent metameric blacks can be obtained such that any other metameric black can be represented as a linear combination of them. He also suggested a systematic algorithm to compute a set of $N-3$ linearly independent metameric blacks in his article. It should be pointed out that these spectral reflectances have necessarily negative values in some parts of the spectrum and consequently cannot be materialized physically.

Cohen and Kappauf [20] then formalized Wyszecki's decomposition by outlining orthogonal projectors. They introduced a projector matrix R that minimizes the distance between the surface reflectance x_λ and its projection to the columns of color formation matrix $\Lambda = [E_\lambda r_\lambda^1, E_\lambda r_\lambda^2, E_\lambda r_\lambda^3]$. The suggested projector, in what they referred to as matrix-R theory, is defined as follows:

$$R = \Lambda^T (\Lambda \Lambda^T)^{-1} \Lambda \tag{1.6}$$

A projector to the orthogonal complement of the space spanned by the columns of color formation matrix Λ, denoted by $R^\perp$ was also derived via R:

$$R^\perp = I - R \tag{1.7}$$

where I is an $N \times N$ identity matrix. Thus, any reflectance function x_λ can be written as:

$$x_\lambda = R x_\lambda + R^\perp x_\lambda \tag{1.8}$$

$R x_\lambda$ is referred to as fundamental component and $R^\perp x_\lambda$ as the metameric black. Both projectors are invariant to linear transformation so a linear change of the color space will not change its fundamental and metameric black components. The R-matrix operation was later extended by Burns [16] to multiple metamers, i.e., the reflectance functions that remain metamers with respect to several different illuminants.

The above algorithm explains how to generate metameric reflectances, but it does not fully specify the set of all possible metameric reflectances for a given color signal. Schmitt [84] on the other hand, derived a basis for the entire set of metameric reflectances. Given the illuminant E and the CIE 1931 color matching functions r_1, r_2 and r_3, the visible spectrum can be split into M, not necessarily equal, intervals $\Delta \lambda_i, i = 1, .., M$. The metamer reflectances x must satisfy the following equations:

$$X_a = \frac{100 \sum_{i=1}^{M} x_i (E r_a)_i \Delta \lambda_i}{\sum_{i=1}^{M} (E r_2)_i \Delta \lambda_i} \quad (a = 1, 2, 3),$$
$$0 \leq x_i \leq 1, i = 1, ..., M \tag{1.9}$$

where:

$$X_1 = X, X_2 = Y, X_3 = Z \tag{1.10}$$

$$r_1 = \overline{x}, r_2 = \overline{y}, r_3 = \overline{z} \tag{1.11}$$

The denominator and the factor 100 in the numerator for $a = 2$ are added to normalize the color values such that $Y = 100$. A numerical algorithm is proposed to find all independent x elements that satisfy the above equations. These elements, referred to as simple metamers, are finite and less than the binomial coefficient $\binom{M}{3}$. They build a metameric ensemble, F, which is described as a hyperpolyhedron volume in M-dimensional space with

7

simple metamers as its apexes. Schmitt then demonstrated that this volume is convex and any other metamer can be defined as a linear combination of these elements.

Finlayson et al. [28] use the original framework of Wyszecki [104] and R-matrix theory to find all surface reflectances that induce a particular response. They place some constraints on the reflectances to simultaneously make them physically realizable and statistically replicating natural, measured reflectances.

All the aforementioned algorithms are numerical approaches to the problem. The infinite-dimensional set of all the object reflectance spectra is usually sampled to build a finite-dimensional subset and some restrictions will be applied to provide an approximate solution. However, the algorithm proposed by Logvinenko et al. [54] is the first to precisely characterize the metamer mismatching volume. They theoretically find the limits of the MMB for any strictly positive illuminant and sensor sensitivity functions. The boundary points of the MMB, which are computed without placing any restriction on the reflectance functions, are sufficient to specify the whole convex volume.

Schrodinger [85] showed that an optimal reflectance function, i.e., a reflectance function whose color falls on the boundary of the OCS, only takes two values, 0 or 1. He also claimed that for human color vision, an optimal object reflectance function is a step-like function with some transitions between zero and one. Such optimal functions, x_m or $1 - x_m$, are defined below:

$$x_m(\lambda; \lambda_1, ..., \lambda_m) = \sum_{i=1}^{m} (-1)^{i-1} x_1(\lambda; \lambda_i) \tag{1.12}$$

$$1 - x_m(\lambda; \lambda_1, ..., \lambda_m) \tag{1.13}$$

where

$$x_1(\lambda; \lambda_1) = \begin{cases} 0 & \text{if } \lambda < \lambda_1 \\ 1 & \text{if } \lambda \geq \lambda_1 \end{cases} \tag{1.14}$$

Logvinenko [56] showed that an elementary step-like function with transition wavelengths $\lambda_1, \lambda_2, ..., \lambda_m$ is an optimal reflectance function if $\lambda_1, \lambda_2, ..., \lambda_m$ are the only roots of the following equation:

$$k_1 s_1(\lambda) + k_2 s_2(\lambda) + ... + k_n s_n(\lambda) = 0 \tag{1.15}$$

Based on the equation above, Logvinenko et al. [54] showed how, in theory, computing the optimal reflectance function (i.e., reflectance functions whose color signals fall on the OCS boundary) leads to the computation of the MMB boundary points. They showed that in a trichromatic world, if the color mechanisms $(\varphi_1, \varphi_2, \varphi_3)$ change to (ψ_1, ψ_2, ψ_3), the MMBs can be well approximated by optimal reflectance functions with up to five transitions. In other words, elementary step functions with $m = 5$ with the color signal
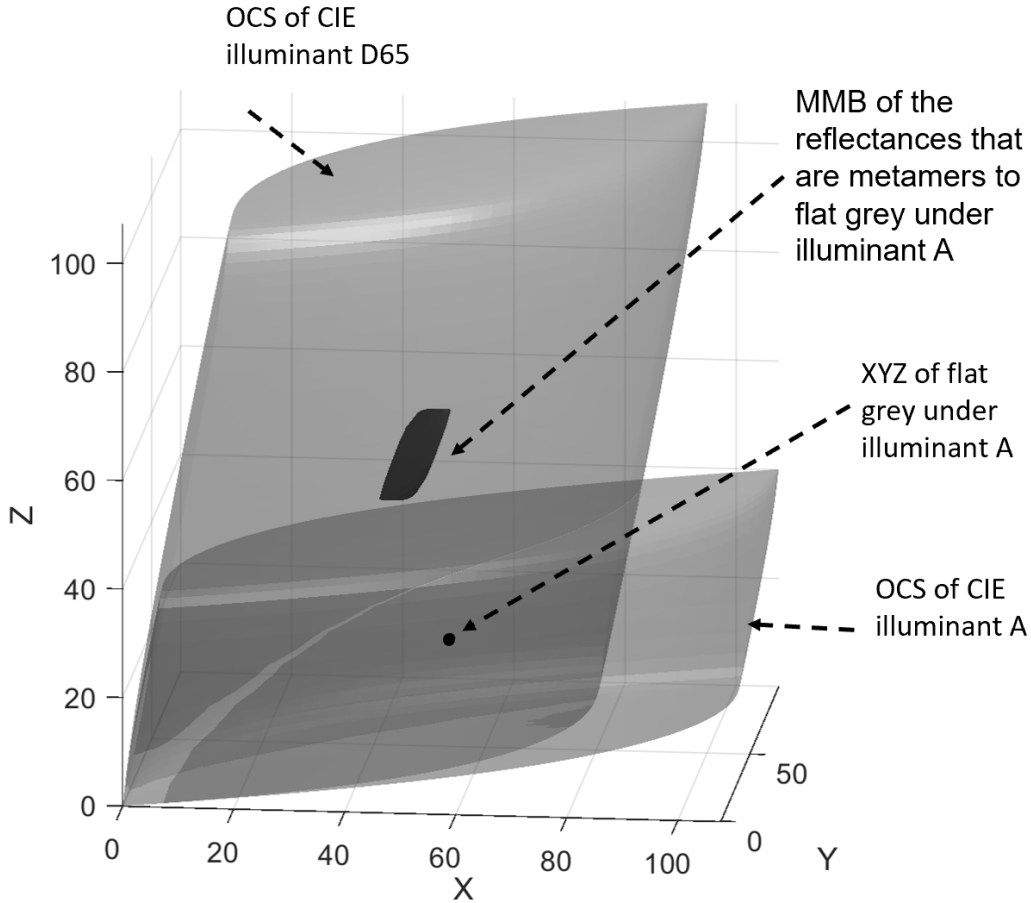
Figure 1.6: Sample MMB for flat grey induced by going from CIE illuminant A to D65.

$(\varphi_1(x_5), \varphi_2(x_5), \varphi_3(x_5))$ estimate the boundary of the metamer mismatch volume with high precision while being computationally efficient. The convex metamer mismatching volume can be specified by computing sufficient boundary points in different directions.

A sample MMB for a flat grey under CIE illuminant A that is induced by going from CIE illuminant A to D65 is plotted in Figure 1.6 along with the object color solids of CIE XYZ 1931 2° Standard Observer under illuminants D65 and A.

Logvinenko's method measures the extent of metamer mismatching in theory. But such steep, step-like reflectance functions likely never happen in nature. Zhang et al. [110] gathered more than 25 million unique reflectance spectra of man-made, natural and industrial objects that were obtained using multi-spectral imaging systems and spectrophotometers. The reflectance functions that have the same CIE XYZ values under the first light are used to define the empirical MMBs under the second light. They showed that the volumes of these empirical MMBs are significantly smaller than the theoretical ones computed by Logvinenko's algorithm. Moreover, they showed that the volumes vary systematically with

Munsell [73] value and chroma just as the theoretical ones do. This is illustrated by the plot in Figure 1.7.
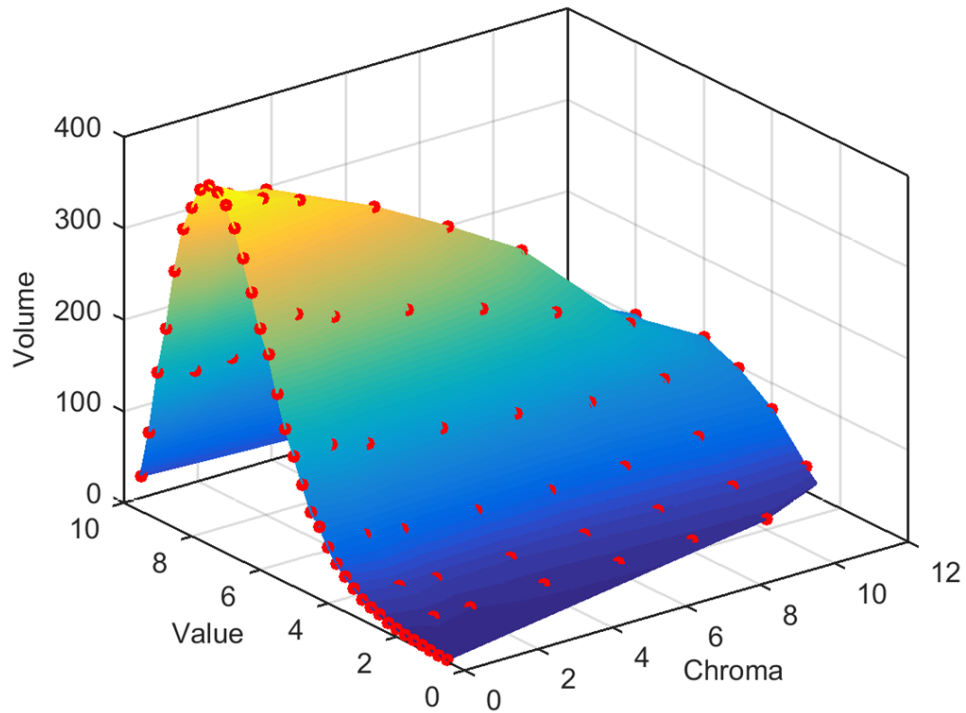
Figure 1.7: The MMB volume averaged over all hues and plotted as a function of Munsell chroma and value, created by going from illuminant A to D65, and showing how the MMB volume decreases with distance in Munsell value and/or chroma from grey (value 6, chroma 0) [110]

.

Chapter 2

Do Computational Color Prediction Methods Explain Asymmetric Color Matching?

As pointed out in the previous chapter, a change in the illuminant of the scene can lead to a change in the object color signal as shown in Figure 2.1. This causes a serious problem for computer vision applications such as color-based object identification and tracking, image reproduction, color feature extraction and scene understanding [26][111][34]. Therefore, it is important to be able to predict the color signal of the same object under different illuminants.

Figure 2.1: The sRGB rendering of the same image from Columbia dataset under different lighting conditions, the illuminants from left to right: CIE A and CIE D65.

A similar issue arises when a color image recorded by a set of sensors is to be converted to the output image produced by another set of sensors. Several algorithms are proposed to answer this question, given a color signal for an unknown surface reflectance under the first light—what its color signal is likely to be under the second light?
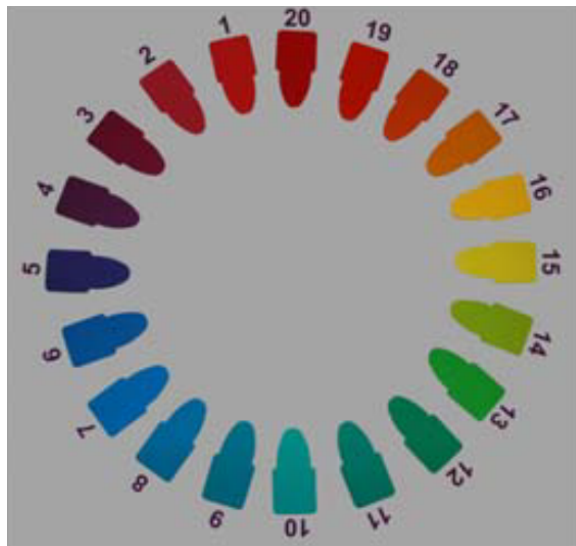
Figure 2.2: The chromatic stimulus papers used in the L&T experiment [58] with Munsell notations of 1) 5RP5/12; 2) 10P4/12; 3) 5P4/12; 4) 10PB4/12; 5) 5PB5/12; 6) 10B5/12; 7) 5B5/10; 8) 10BG5/10; 9) 5BG6/10; 10) 10G5/10; 11) 5G5/10; 12) 10GY6/12; 13) 5GY7/12; 14) 10Y8.5/12; 15) 5Y8/14; 16) 10YR7/14; 17) 5YR7/14; 18) 10R5/16; 19) 5R4/14; 20) 10RP5/14.

Logvinenko et al. [54] proposed a theoretical algorithm to compute the boundary of metamer mismatching volume. They also showed that in theory the extent of metamer mismatching can be very significant. The existence of such a volume means that given a color signal under the first illuminant, it can become any one of the color signals within the metamer mismatch volume under the second illuminant. Therefore, the color signal of an object under the second light is in fact unpredictable when the MMB is not zero. This is an ill-posed problem and there is no unique answer for it because there is no obvious criterion to choose between the possible solutions. Nonetheless, it is natural that one look for an approximate solution that fits the experimental data best on average.

Logvinenko and Tokunaga [58] (L&T henceforth) conducted an asymmetric color matching experiment in which observers view a Munsell paper under one light (the test illuminant) and then choose the least dissimilar paper from a set of 22 papers (Figure 2.2) under a second light (the match illuminant). In this section we analyze the performance of the state-of-the-art color prediction methods in terms of how well they explain L&T's asymmetric color matching experiment results. We address three questions: (i) Are observers generally choosing the physically identical Munsell paper under the match illuminant? (ii) Which computational method most closely corresponds to the observer average? and (iii) How does the performance of individual observers compare to the computational methods in predicting the least-dissimilar matches of the average observer?

This chapter has led to the following publications [80][81]:

- Emitis Roshan and Brian Funt. Evaluation of color prediction methods in terms of least-dissimilar asymmetric matching. Electronic Imaging, 2017(14):140–144, 2017.

- Emitis Roshan and Brian Funt. Computational color prediction versus least-dissimilar matching. JOSA A, 35(4):B292–B298, 2018.

2.1 Color Signal Prediction Methods

Numerous methods for predicting color signal under a change of illumination have been proposed. Given the color signal of a test paper under the first illuminant, what its color signal will be under the second illuminant?

The Relit color signal is simply the color signal of the given object's spectral reflectance under the second illuminant. Computing it requires the full spectral reflectance function of the surface as well as the SPD of the second illuminant. For a matte surface, we assume that the color signal $(\varphi_1(x), \varphi_2(x), \varphi_3(x))$ resulting from the light impinging on sensors $r_i(\lambda), (i = 1, 2, 3)$ from a surface of spectral reflectance $x(\lambda)$ illuminated by light with spectral power distribution $E(\lambda)$ can be obtained using Equation 1.2. The Relit color signal, of course, is not really a prediction at all but rather, under the assumption of matte reflectance, a straightforward calculation of what the actual color signal will be [81].

Derhak and Berns [24] make the distinction between chromatic adaptation transforms (CATs) and material adjustment transforms (MATs). A CAT is intended to predict what color signal under the second illuminant will appear the same as under the first illuminant. Of course, there is the issue of what 'the same' means. Derhak and Berns define the goal of CATs and MATs as "a CAT models appearance constancy via corresponding color experiments, whereas a MAT models material constancy via least dissimilar color matching." They consider a MAT as a material constancy prediction, or how sensor excitations for an object color change with changes in observing conditions [24]. The problem with this definition is, as established by Logvinenko et al. [55], that as a result of metamer mismatching intrinsic object colors that are independent of the illuminant simply do not exist—hence material constancy in the Derhak and Berns sense does not exist either since from such material constancy, intrinsic object color would immediately follow.

They introduced the Wpt coordinate system, where W represents perceptive lightness, and p and t represent perceptive chromaticness, that is, a combination of perceptive chroma and hue at a constant perceptive lightness. They call it "Waypoint" because this color equivalency representation forms a waypoint between source and destination observing conditions. The orientations of the p and t axes are similar to the a and b axes of the CIELAB color space. The Wpt coordinates of a training set consisting of 40 Munsell papers with hues at an intermediate lightness and chroma of Value 5 and Chroma 6, plus a perfect diffuser reflector (PRD), and a vivid yellow are defined such that they achieve orthogonality between lightness and chromaticness, and if possible have linear loci of constant hue and
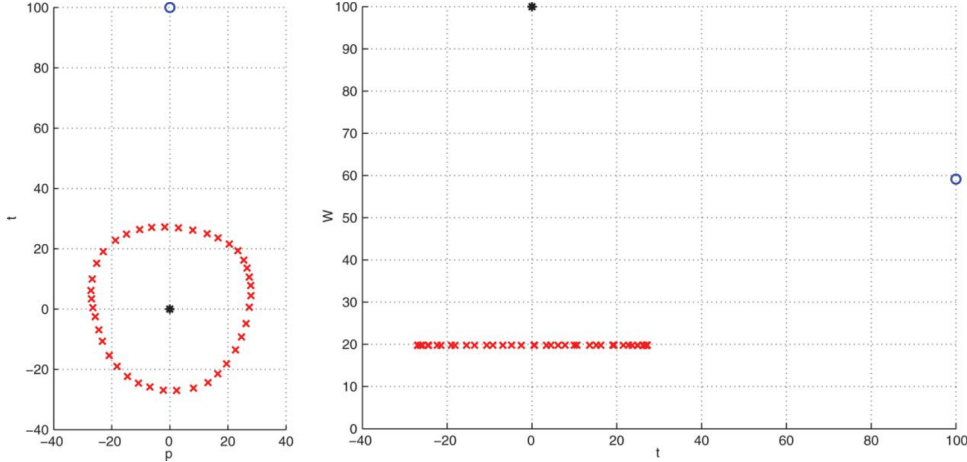
Figure 2.3: The visualization of Wpt coordinates of the 42 Munsell papers [24].

circular contours of constant chroma (Figure 2.3). The PRD is set to have a W value of 100. The color signals of the training set elements under CIE illuminant C and 2° standard observer and also other different observing conditions are mapped to their Wpt coordinates using 3×3 linear matrices. Let's say A_1 and A_2 are the best 3×3 transformation matrices that map the color signals of the training set under the first and second viewing condition to their Wpt coordinates. Given the color signal $\Phi_1(x)$ of a test object under the first viewing condition, the predicted color signal $\Phi_2(x)$ under the second observing condition can be computed as:

$$\Phi_2(x) = A_2^{-1} * A_1 * \Phi_1(x) \tag{2.1}$$

The intermediate space of Wpt causes the transformation matrix $A_2^{-1} * A_1$ in Equation 2.1 to be not the optimal 3×3 matrix that relates the color signals under the first and second viewing conditions. In comparison, the Best Linear Fit method [31] is based on using the optimal 3×3 matrix mapping the color signals from the training set (1600 Munsell papers) under the first illuminant to the second illuminant.

Another approach is the MMB center prediction which is based on computing metamer mismatch volumes. As indicated before, for a given color signal under the first illuminant, the set of color signals it could theoretically become under the second illuminant defines a convex volume in the color signal space called the metamer mismatch body (MMB). Computing the MMB requires full knowledge of the SPDs of both illuminants. Logvinenko et al. [54] propose using the color signal at the geometric center of the MMB as a candidate for what the color signal under the first illuminant is likely to become under the second illuminant. Here we label that prediction method the "MMB center".

The above-mentioned algorithms all need the full SPD of the first (test) and second (reference) illuminants. There is another group of color signal prediction methods that

14

requires only the color signals of the illuminants, not their full SPDs. Von-Kries-based CIECAM02 [69] and KSM^2 [68] belong to this group.

CIECAM02 is a color appearance model. Given a set of tristimulus values in XYZ, a transformation matrix M$_{\text{CAT02}}$ is applied to convert them to the spectrally sharpened CAT02 LMS space:

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = M_{CAT02} * \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, M_{CAT02} = \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{bmatrix} \tag{2.2}$$

After this sharpening transformation [27] [95], at the heart of CIECAM02 is the chromatic adaptation transform CAT02, which applies the standard von Kries (diagonal) transformation. It takes the tristimulus values of the adapting white point, the adapting background, and the surround luminance as the input to compute the degree of adaptation D. In theory, the degree of adaptation can vary from 0, for no adaptation, to 1, for complete adaptation. The values in between can be computed as follows:

$$D = F(1 - \frac{1}{3.6} e^{\frac{-(L_A + 42)}{92}}) \tag{2.3}$$

L_A is the adapting field luminance and F is a factor determining the degree of adaptation, and can be assigned certain values corresponding to different surround conditions.

KSM^2 developed by Mirzaei et al. [68] is another algorithm in the second group that only needs the color signal of a PRD. KSM^2 uses Gaussian-like functions (called wraparound Gaussians) to represent both the illuminations and the reflectance. Given the color signal of a light (its full SPD is not required), a metameric Gaussian SPD can be found that is fully specified by 3 parameters: K the scaling, S the sigma, M the peak wavelength. As illustrated in Figure 2.4, to make a color signal prediction, KSM^2 finds three Gaussian functions, one representing an SPD metameric to the first illuminant, a second metameric to the second illuminant, and a third representing a reflectance metameric to the given test color signal under the Gaussian SPD metameric to the first illuminant. It then computes the color signal of that Gaussian reflectance under the second Gaussian illuminant and uses that color signal as its prediction.

In the next section, the color prediction methods are compared to determine which one best models the observers' performance in the L&T least dissimilar asymmetric color matching experiment.

2.2   L&T Least Dissimilar Asymmetric Matching Experiment

Logvinenko & Tokunaga [58] ran an asymmetric color matching experiment. A set of 22 Munsell papers were shown under two lights (referred to as the test and reference illuminants) simultaneously. A laser pointer was used to indicate one of the papers under the test
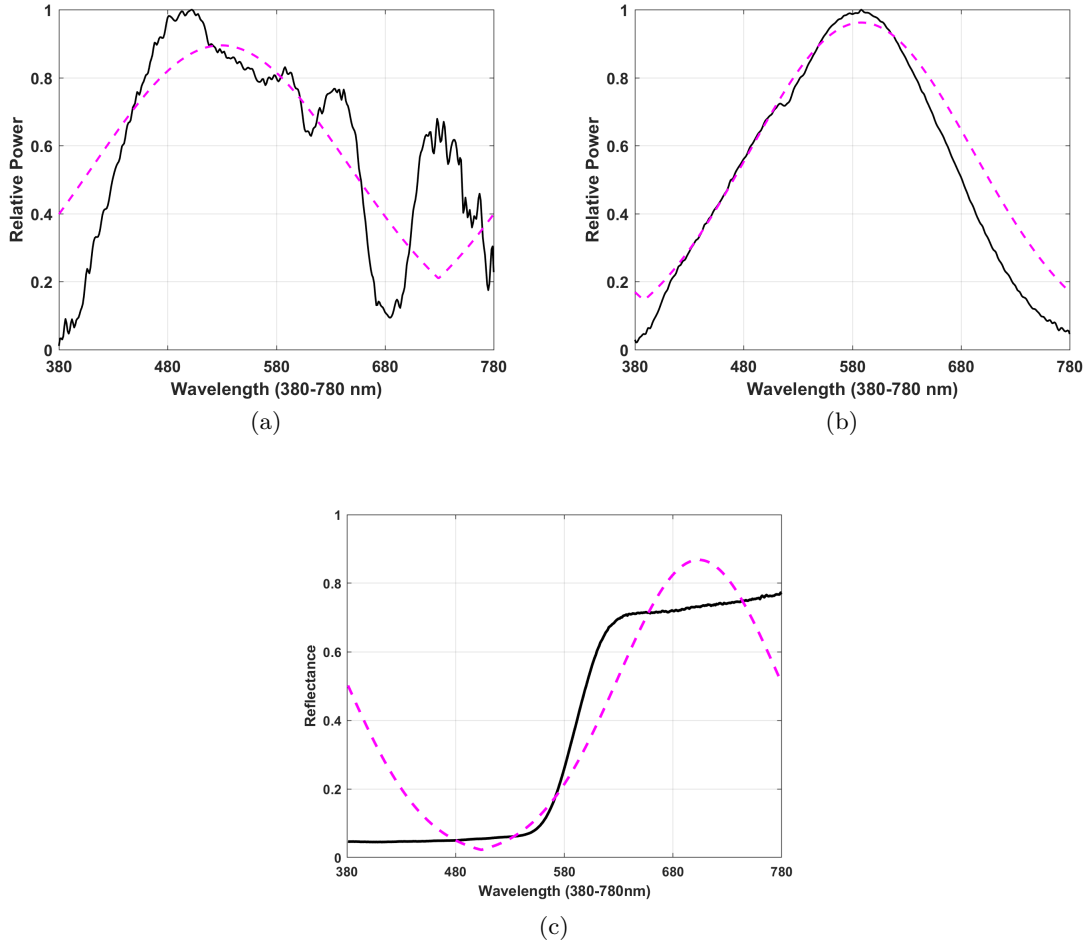
Figure 2.4: Solid black curves are the given spectral power distributions and reflectance. Dashed magenta curves are their Gaussian metamers. (a) Wraparound Gaussian SPD metameric to the first light. (b) Wraparound Gaussian SPD metameric to the second light. (c) Wraparound Gaussian reflectance producing the same color signal when lit by the wraparound Gaussian SPD from (a) as the color signal of the original reflectance (solid black curve in (c)) under the first light [81]
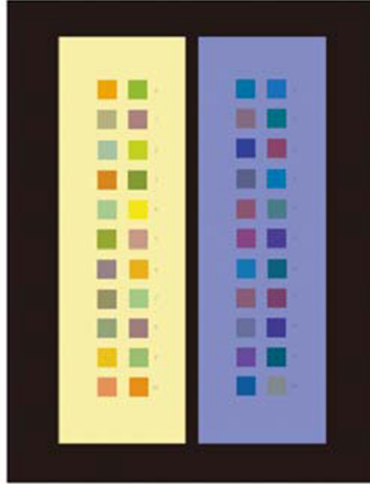
Figure 2.5: The asymmetric matching setup used by L&T [58] showing the example of the left-hand panel in yellowish light and the right-hand panel in bluish light. "Each stimulus array contained 20 papers from every other page in the Munsell book of maximal Chroma completed with grey (N5/) and black (N1/) papers (i.e., 22 papers in all). The stimulus array dimensions were 39 × 55 cm." [58] The papers are rearranged between trials.

illuminant and the observers were asked to choose the least dissimilar paper from the papers under the reference (or the match) light. There were 4 observers and 3 repetitions each. See Figure 2.5 for a photograph of the setup. The papers are rearranged between trials.

Note that these are real papers under real illuminants, not colored patches on a digital display nor colors obtained using hidden illuminants to simulate reflectance changes [13] [23]. The experiment involved 6 illuminants of approximately equal illuminance, green (G), blue (B), neutral (N), yellow (Y), red1 (R1) and red2 (R2), and all 30 possible pairs were used as test/match illuminant conditions. However, since the two red illuminants are very similar, here we exclude one of them (R2). Considering only the non-identical pairs of 5 of the illuminants, there are respectively 5 and 4 possible illuminants as the test and match lights and so 20 illumination conditions.

The L&T experiment differs from many other asymmetric color matching experiments in that subjects are not asked to make exact asymmetric matches, but rather to identify the colored paper that appears least-dissimilar. They argue that the classic asymmetric matching has a major shortcoming in that the observers who set a match report that color matches are not always perceptually identical. They point out that the light-color dimension of object color means that an exact asymmetric color match is impossible in principle. Hence, they ask their observers not to find an exact match but rather a least-dissimilar match [58] [57].

There are other types of color matching, but each has its own shortcomings. In memory matching, the samples under different lights to be compared are shown successively, not at the same time. When there is a delay between successive views this necessarily involves

memory [74]. Allowing time for the eyes to adapt to each illuminant, the observers need to keep the color information in mind, but it is hard to remember it perfectly after a long delay. In Haploscopic matching, a sample under the first light is shown to the right eye. A copy of the same sample under a different light is shown simultaneously (or successively) to the left eye so that each eye becomes adapted to a different light. Haploscopic matching experiments assume that the two eyes are independent with respect to sensitivities and chromatic adaptation mechanisms, which may well be valid for the sensory mechanism but may not hold for cognitive mechanisms. One half of the field of view corresponding to each eye will be projected into the left brain and the other half will be projected into the right region and then the signals will be mixed in a way that is not yet fully understood [18].

As L&T point out, a perfect asymmetric match will usually be impossible due to metamer mismatching (i.e., the fact that two different reflectances may reflect metameric lights under one illuminant, but non-metameric lights under a second illuminant). Further analysis of the effect of metamer mismatching in the context of this experiment is provided by Logvinenko et al. [55].

2.2.1 Do Observers Choose Original Munsell Paper?

Before addressing the issue of how well the various computational methods model the asymmetric matches made by the L&T observers, we consider the question as to whether or not observers are generally choosing the physically identical Munsell paper under the match illuminant as least-dissimilar to the test paper. To answer this question, for each test paper under the test illuminant, we compute the average XYZ under the match illuminant of the Munsell papers chosen as least-dissimilar and calculate how far in terms of Euclidean distance that average is from the actual XYZ of the test paper under the match illuminant, and finally average the results over all 20 test papers.

For each illumination condition, 4 observers with 3 repeats made least-dissimilar matches. All 20 chromatic papers were used as test papers. For each of the 20 test papers, therefore, there are 12 least-dissimilar matches reported, resulting in 240 matches for each illumination condition. Considering the 20 non-identical pairs of lights used in the asymmetric matching experiments, we have 20×240, or 4800 matches in total. The average Euclidean distance between the matched paper and the XYZ of the physically identical Munsell paper under the match illuminant is 6.0. For comparison, the average XYZ difference between a given Munsell paper and the nearest of the other 19 papers under the Neutral illumination is 6.6. In other words, the observers are on average choosing as least dissimilar a paper that is either the physically identical paper or one that is close to it in color.

Our analysis is in agreement with L&T's analysis: "... when the test illuminant was neutral or yellow the average mismatch was roughly one hue step. The mismatch for the other four test illuminants was approximately two hue steps. Therefore, while the exact match rate for these illuminations ... is quite low (less than 30%) the average mismatch

does not exceed two hue steps" ([58] p. 415). An "exact match" is defined as the observer choosing the physically identical paper.

These results suggest, perhaps not surprisingly, that observers generally find the match paper that is physically identical to the test paper to be the least dissimilar one.

2.2.2 Predicting Observer Average Matches

To determine which method most closely predicts observer least dissimilar matching behavior, we consider the 12 (4 observers, 3 repeats) matches made for each test paper under a given illumination condition and compute the average-observer-match as the average of the color signals of the 12 matched papers under the reference illuminant. Each computational method is used to predict the color signal of the test paper under the reference illuminant. A method's prediction error is calculated as the Euclidean distance between the average observer match color signal and the color signal the method predicts.

We compare the performance of the computational color prediction methods to one another using the Wilcoxon signed-rank one sided and two sided tests [89]. The Wilcoxon test is a non-parametric statistical hypothesis test based on the sum of the signed ranks of a set of paired samples. In the present case, the paired samples are the prediction errors for the 20 papers under a given illumination condition of the two methods being compared. All the tests are performed at the 5% significance level.

More specifically, the 20 test papers result in 20 average observer match values for a given pair of test and match illuminants, along with a corresponding set of 20 predictions made by each algorithm. Three tests are performed to compare each pair (Method 1 and Method 2) of methods—one two-sided test and two one-sided tests. The null hypotheses for these tests are as follows.

Two-sided test: the null hypothesis is that the median prediction errors of the two methods are equal.

Right-tailed test: the null hypothesis is that the median prediction error of Method 1 is greater than the median prediction error of Method 2.

Left-tailed test: the null hypothesis is that the median prediction error of Method 2 is greater than the median prediction error of Method 1.

The results of the three Wilcoxon tests will lead to one of the following cases.

Case I: The null hypothesis of the two-sided test cannot be rejected at the 5% significance level. In this case the performance of Method 1 and Method 2 can be considered to be equivalent.

Case II: The null hypothesis of the two-sided test can be rejected and the right-tailed test cannot be rejected, but the null hypothesis of the left-tailed test can be rejected. In this case, Method 2 can be considered to be better (lower median prediction error) than Method 1.

Table 2.1: Comparison of algorithms in predicting the average observer-match in each of the 20 different illumination conditions. The numbers in columns 3-5 indicate how many times across the 20 different illumination conditions that each Case (see text for definition of the Cases) occurs. Informally, Case III indicates Method 1 is 'better' than Method 2, Case II that Method 2 is better than Method 1, and Case I that they perform similarly.

Method 1	Method 2	Case III	Case II	Case I
Relit	KSM^2	10	1	9
Relit	MMB Center	20	0	0
Relit	CIECAM02	10	0	10
Relit	Wpt	8	0	12
Relit	Best Linear	5	0	15
KSM^2	MMB Center	17	0	3
KSM^2	CIECAM02	5	2	13
KSM^2	Wpt	5	5	10
KSM^2	Best Linear	1	6	13
MMB Center	CIECAM02	0	18	2
MMB Center	Wpt	0	19	1
MMB Center	Best Linear	0	20	0
CIECAM02	Wpt	0	6	14
CIECAM02	Best Linear	1	8	11
Best linear	Wpt	1	7	12

Case III: The null hypothesis of the two-sided test can be rejected and the left-tailed test cannot be rejected, but the null hypothesis of the right-tailed test can be rejected. In this case, Method 1 can be considered to be better (lower median prediction error) than Method 2.

Note that the results in Table 2.1 show the relative performance of the methods, not their absolute performance. In other words, the methods might be doing equally poorly rather than equally well. In terms of absolute performance, Table 2.2 lists the accuracy of each method's predictions averaged over the 400 cases. The accuracy is measured in terms of the Euclidean distance between the prediction and the average XYZ of the 12 least-dissimilar matches, and similarly for CIE1976 u'v' coordinates. Although most of the results reported in this study are in terms of XYZ, almost identical ranking results were obtained using Euclidean distances in Hunter-Pointer-Estevez LMS space and the CIEDE2000 metric.

The results in Table 2.1 and Table 2.2 are aggregated over all 20 Munsell papers and all 20 illumination conditions. L&T [58] provide a detailed analysis of how the average 'exact match' rate varies both with the illumination condition and with the test paper.

Table 2.2: Accuracy in predicting average observer matches. Mean and median of the Euclidean distance in XYZ and CIE1976 u'v' between each method's predictions and the average observer match across 400 cases.

Method	Mean XYZ	Median XYZ	Mean u'v'	Median u'v'
Relit	5.21	3.45	0.024	0.015
Best Linear	5.56	4.17	0.040	0.023
Wpt	6.20	4.44	0.096	0.025
KSM^2	8.08	4.50	0.043	0.030
CIECAM02	7.61	5.99	0.040	0.030
MMB Center	39.85	23.44	0.072	0.040

2.2.3 Observers Predicting Other Observers

In the previous section the performance comparison is between computational methods. All those methods might be equally good or bad but how does their performance compare to that of the observers relative to one another? Clearly there will be variability in the least dissimilar matches made by the different observers. To what extent do the observers agree with one another and is a match made by an individual observer any better or worse a predictor of the average observer match than those made by the various computational methods?

To answer this question, we used a leave one observer out comparison in which one observer is excluded and the 9 remaining trials (3 observers, 3 repeats per paper) are combined to create a 3 observer average for each illumination condition. The mean of the excluded observer's 3 trials is then used as a predictor of this 3 observer average. This process is repeated for each of the 4 observers resulting in predictors Obs1,. . .,Obs4 of the 4 different, 3 observer averages.

Table 2.3 compares the individual observers to the computational methods in predicting the 3 observer average. Table 2.3 also includes results based on picking the paper that has the closest 'hue' using M from KSM^2 as the hue measure, which interestingly does slightly better than using all 3 components of KSM^2.

From Table 2.3, it is clear that human observers predict the 3 observer average better than the computational methods do, as indicated by the fact that the numbers in the Case II column are substantially larger than those in the Case III column.

2.2.4 Results Using The Process Of Elimination

In a discussion concerning the results described above, John McCann [65] suggested that perhaps the observers were exploiting the fact that there were only 20 chromatic papers from which to choose and this might in some way be affecting the L&T matching results. In order to address that concern, in this section we provide the computational methods with this

Table 2.3: Observers versus Computational Methods. Similar to the Table 2.1 but in this case comparing via the Wilcoxon test how well each method/observer predicts the 3-observer average of least dissimilar matches. The numbers in columns 3-5 indicate how many times across the 20 different illumination conditions that each Case (see text for definition of the Cases) occurs. Informally, Case III indicates the given method is 'better' than the particular observer, Case II that the observer is better than the method, and Case I that they perform similarly.

Method	Observer	Case III	Case II	Case I
KSM^2	Obs1	0	13	7
	Obs2	0	15	5
	Obs3	0	11	9
	Obs4	0	14	6
Relit	Obs1	0	9	11
	Obs2	0	11	9
	Obs3	2	6	12
	Obs4	0	9	11
Wpt	Obs1	0	13	7
	Obs2	0	14	6
	Obs3	0	9	11
	Obs4	0	12	8
CIECAM02	Obs1	0	14	6
	Obs2	0	16	4
	Obs3	0	13	7
	Obs4	0	18	2
Best Linear	Obs1	0	12	8
	Obs2	0	10	10
	Obs3	0	7	13
	Obs4	0	12	8
MMB Center	Obs1	0	20	0
	Obs2	0	20	0
	Obs3	0	19	1
	Obs4	0	20	0
M of KSM^2	Obs1	0	11	9
	Obs2	0	14	6
	Obs3	0	9	11
	Obs4	0	13	7

additional information to see if they are then able to predict the observers' least-dissimilar matches correctly.

Although the L&T observers were instructed simply to identify the least-dissimilar looking paper, the observers were aware that the same 20 papers were present under both the test and match illuminants so it is conceivable that they used that extra information to do an overall best fit of the least-dissimilar matches for of the 20 papers under the match

Table 2.4: Results corresponding to those in Table 2.3 but allowing the algorithms to include minimizing the total dissimilarity across all 20 papers simultaneously

Method	Observer	Case III	Case II	Case I
KSM^2	Obs1	0	9	11
	Obs2	0	11	9
	Obs3	2	7	11
	Obs4	0	12	8
Relit	Obs1	0	9	11
	Obs2	0	11	9
	Obs3	2	6	12
	Obs4	0	9	11
Wpt	Obs1	0	9	11
	Obs2	0	11	9
	Obs3	2	6	12
	Obs4	0	10	10
CIECAM02	Obs1	0	11	9
	Obs2	0	12	8
	Obs3	0	10	10
	Obs4	0	15	5
Best Linear	Obs1	0	10	10
	Obs2	0	10	10
	Obs3	2	6	12
	Obs4	0	10	10
MMB Center	Obs1	0	9	11
	Obs2	0	11	9
	Obs3	1	10	9
	Obs4	0	14	6
M of KSM^2	Obs1	0	10	10
	Obs2	0	13	7
	Obs3	0	9	11
	Obs4	0	9	11

illuminant to those under the test illuminant. Although we cannot know what observers were doing when they made their least dissimilar matches, we can have the computational methods exploit that extra information.

Table 2.4 shows the results corresponding to those in Table 2.3 but when the algorithms minimize the overall dissimilarity across all 20 papers before deciding on the match for the given test paper.

It is clear from Table 2.4 that the extra information does improve the computational methods' predictions of the 3-observer average (Case I numbers are larger than those in Table 2.3); nonetheless, the individual observers still are statistically better roughly half the time (Case II). In other words, even when the computational methods are modified to

exploit a process of elimination type strategy they are still are not as good as the human observers in predicting the other observers' least-dissimilar matches.

Interestingly, none of the methods is as effective as each individual observer in predicting the 3 observer average of the other observers' matches. This implies that all the methods studied here are not capturing some important aspect of the observers' least-dissimilar matching strategy. L&T [58] argue for the existence of both lighting and material dimensions of object color and propose the concept of an across-illuminant color map. Perhaps once their across-illuminant color map is fully specified it will provide a full model of the L&T asymmetric matching results. All we can say in the meantime, however, is that the computational models we tested do not explain those results adequately.

2.3   Discussion

The use of "least-dissimilar" as opposed to "matching" is an important aspect of the asymmetric color matching experiment by Logvinenko & Tokunaga [58]. Asking the observers to find a "match" tends to abstract color away from what its purpose might be. It also differs from many other experiments as they use real papers under real lights.

Their results raise several questions that we tried to answer by evaluating some of the existing color signal prediction methods. Our analysis shows that observers tend to find the physically identical test paper to be the least-dissimilar match paper. However, this does not mean that observers would always consider that paper to be the least-dissimilar if there were an effectively infinite choice of papers. Note also that because of the possibility of metamer mismatching it is a mistake to interpret the physically identical paper under the match illuminant as the 'correct' answer. An observer is not wrong to find some other paper to be least dissimilar.

A leave-one-observer-out comparison shows that individual observers, somewhat surprisingly, predict the average matches of the remaining observers better than any of the above color prediction methods. This implies that all the computational methods studied are not capturing some important aspect of the observers' least-dissimilar asymmetric matching strategy. Of course, as mentioned before, Best Linear, Wpt, and MMB centers require the full spectra of the test and match illuminants, while KSM^2 and CIECAM02 require only their color signals. In other words, the former ones may or may not predict human performance, but eventually they cannot provide a computational model of any aspect of trichromatic color perception.

Chapter 3

Color Accuracy of Digital Cameras

The desire for accurate color measurement across many fields has led to the development of various color accuracy metrics for digital cameras. For instance, color imaging plays a crucial role in early diagnosis of different diseases such as cancer, dermatological, ophthalmology and dental disorders to name a few [99] [46] [29] [9] [50] [42] [47]. There are further applications within medicine such as anemia, pH, protein and glucose tests [88] [71] [40] [109] [21]. Its application also continues beyond medicine. For example, color is widely used in intelligent agriculture [22] [51], water test quality [43] and marine monitoring [6].

There are many trade-offs involved in digital color camera design in terms of image noise, cost, and physical limitations that means that perfect color accuracy is usually sacrificed. This is especially the case since the usual goal in camera design is to provide good-looking pictures, not to build an imaging colorimeter. Nonetheless, quantifying the colorimetric accuracy of commercial cameras and cellphones, even though they are not initially intended to be used as imaging colorimeters, has become very important due to their image quality, portability and affordability. So there is a definite need to be able to quantify the degree of color accuracy/inaccuracy that a given camera possesses.

In this chapter, a novel method for evaluating the colorimetric accuracy of digital color cameras is proposed based on a new measure of the metamer mismatch body (MMB) created for a 50% flat grey under D65 illuminant induced by the change from the camera as an observer to the human standard observer. Unlike the majority of the existing metrics, the proposed method is based on a theoretical principle rather than a finite set of measurements.

The findings of this chapter has led to the following publications [82][83]:

- Emitis Roshan, Brian Funt, and Hamidreza Mirzaei. Camera color accuracy evaluated via metamer mismatch moments. In Proceedings of the AIC International Colour Association Conference, 2017.

- Emitis Roshan and Brian Funt. Color sensor accuracy index utilizing metamer mismatch radii. Sensors, 20(15):4275, 2020.

3.1 Sensor Quality Metrics

A camera is said to be colorimetric if its sensor sensitivities can be represented as a linear transform of the CIE 1931 2° observer color matching functions. This is known as the Maxwell-Ives criterion (also called the "Luther condition") [61]. If this condition is not satisfied, colorimetric accuracy cannot be guaranteed. In other words there are some spectra that our eyes see as matching, but the camera sees as different, and there are some spectra that the camera sees as matching but our eyes see as different.

The problem with this condition is that if it is not met then it does not determine how accurately a digital camera is able to capture color information from the scene.
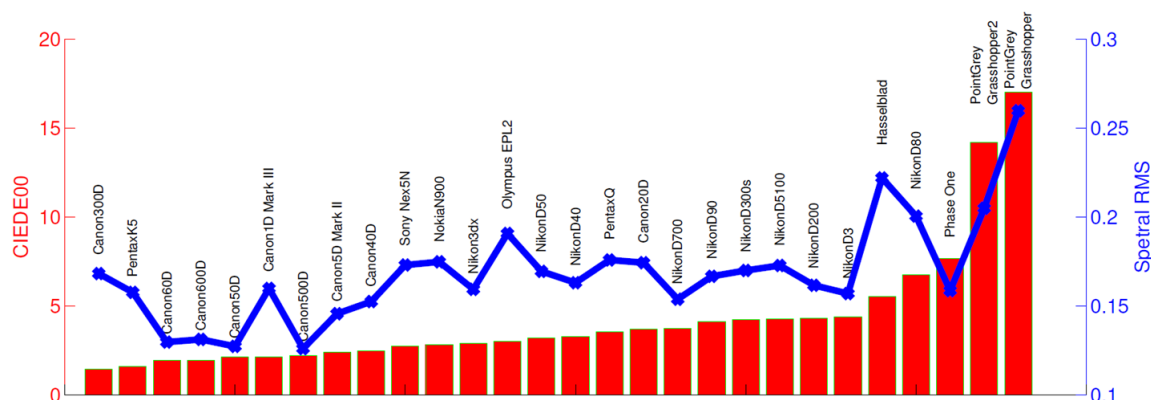
Figure 3.1: The blue crosses show the RMS error between CIE-1931 2° color matching functions (CMF) and the best map of the spectral sensitivities of each camera to the CMFs. The red bars show the average color difference (CIEDE00) between the XYZ of 1269 Munsell chips under D65 and the camera's best estimate of them [41].

Jiang et al. [41] measured the spectral sensitivity functions of 28 digital color cameras in the visible spectrum (i.e., 400nm to 720nm). To see how closely each of the 28 different cameras approximates the Luther condition they measured the RMS error in the best linear fit of the camera sensitivities to the CIE 1931 2° observer color matching functions. This error is zero if the camera perfectly satisfies the Luther condition. They showed that overall most cameras have a deviation from the Luther condition, and that Canon cameras in general have lower RMS than Nikon cameras in this aspect. As an alternative, they calculated the mean color difference (CIEDE00) between the actual XYZs of the 1269 reflectances of the Munsell Book [73] illuminated by D65 and those that are predicted based on using the camera's spectral sensitivity functions. The camera predictions are made by computing the resulting RGB values for the given camera's sensitivity functions and then mapping them to the corresponding XYZ values based on a best linear fit of the camera sensitivity function to the CIE 1931 2° observer $\overline{x}$, $\overline{y}$, $\overline{z}$ color matching functions. The ranking of the cameras based on their average ΔE and RMS errors is shown in Figure 3.1. Obviously, several of

these camera rankings based on RMSE differ significantly from those based on the ΔE measure, which leaves the question as to which ranking to use and why one over the other.

Another approach is the q factor quality measure introduced by Neugebaur [72]. Let $V = [v_1, v_2, v_3]$, where v_i are the N-vectors that represent the illuminant-times-color-matching-function product. The space spanned by these vectors $v_i(i = 1, 2, 3)$ is called the human visual subspace (HVSS). If m is a filter, the q factor can be expressed as:

$$q(m) = \frac{\|P_\nu(m)\|^2}{\|m\|^2} \tag{3.1}$$

where $P_\nu(m)$ is the orthogonal projection of m onto the HVSS and $\|.\|$ denotes the second norm in N-dimensional vector space. Notice that $0 \leq q(m) \leq 1$, and the closer the $q(m)$ value of a filter to one, the better the color scanning filter m. The disadvantage of the q factor is that it is limited to the evaluation of a single filter, and hence is insufficient for the evaluation of color cameras. Three independent filters each with $q(m) = 1$ spanning the HVSS would be perfect; however, the q factor on its own does not indicate whether a set of filters is independent or not. Thus, a set of three identical filters can be deemed perfect even though they clearly are not.

The q factor also does not differentiate among imperfect sets of filters. To be able to handle sets of filters M, Vora et al. [96] extended the q factor to the ν measure which is defined as:

$$\nu(V, M) = \sum_{i=1}^{\alpha} \lambda_i^2(O^T N) \tag{3.2}$$

V represents the illuminant-times-color-matching-function product and O is an orthonormal basis for the space spanned by V. N is also an orthonormal basis for the subspace spanned by the set of filters M under a given illuminant (usually CIE D65). Such an orthonormal basis can be computed by Gram-Schmidt orthogonalization. $\lambda_i(O^T N)$ is the ith singular value of $(O^T N)$. This measure represents the distance between the subspaces. The singular values are related to the principal angles $\theta_1, \theta_2, \ldots, \theta_k \in [0, \pi/2]$ between two subspaces with $\lambda_i^2 = cos^2(\theta_i)$. The ν measure is equal to one if the subspaces are identical, and it is equal to zero if they are orthogonal.

Another extension of Neugebaur's q factor to multiple filters is the CQF, color quality factor, by Trussell et al. [94]. The q factor measures the fraction of the camera filter energy that lies within the HVSS. By reversing the roles of the color matching functions (CMF) and the camera color filters, one can evaluate the fraction of the energy of each CMF that lies in the space spanned by the camera filters. The τ measure is the minimum of the three q factors corresponding to the three CMFs and is defined as:

$$\tau(V, M) = \min_i \{ \frac{\|P_M(\nu_i)\|^2}{\|\nu_i\|^2} \} \tag{3.3}$$

Trussell et al. [94] compare Vora's ν measure with the τ measure. They generated a large number of non-perfect filter sets by adding controlled random deviations to a perfect filter set and then plotted their ν and τ measures versus their average CIE ΔE error. Overall, their plot shows both measures increase when CIE ΔE error decreases but the CIE ΔE error has a higher absolute correlation coefficient with Vora's ν measure than with the τ measure. This confirms the advantage of the ν measure. On the other hand, high ΔE variation around the lower values of both metrics indicates the accuracy decreases when the values of the two metrics decrease.

Sharma et al. [86] criticize the q factor, ν and τ measures, making the point that all the aforementioned algorithms are directly or indirectly computing the mean squared error in CIE XYZ color space, which is known to be perceptually non-uniform. They develop a new figure of merit (FOM) for color scanners/cameras that is based on an error metric in the linearized CIELAB color space to account for both the nonlinearities in color perception process and device noise. The mean squared error for a set of filters is presented as:

$$\epsilon(V, M, B) = E\{\|F(t(x)) - F(\hat{t}(x))\|^2\} \tag{3.4}$$

where $t(x)$, $\hat{t}(x)$, F and $E\{\}$ denote, respectively: the CIE XYZ tristimulus values of object reflectance x; the CIE XYZ tristimulus values estimated as a linear transformation (matrix B) of the scanner/camera measurement plus white noise; a 3×3 transformation of tristimulus values; and the expected value over a set of objects to be scanned. Having the optimal transformation matrix B, the proposed error metric can be computed for a camera/scanner, but a closed form expression cannot be defined for a general non-linear transformation function $F()$. However, it has been shown that if the transformation function $F()$ is differentiable with continuous first partial derivatives, the first-order Taylor series can fairly approximate it locally, and the error metric with optimal transformation matrix B becomes:

$$\zeta(V, M, B_{opt}) = \alpha(V) - \beta(V, M) \tag{3.5}$$

where:

$$\alpha(V) = vec(V^T)^T S_x vec(V^T) \tag{3.6}$$

$$\beta(V, M) = vec(V^T)^T S_x (M \otimes I_3) \times [(M^T \otimes I_3) S_x (M \otimes I_3) + S_\eta]^{-1} \times (M^T \otimes I_3) S_x vec(V^T) \tag{3.7}$$

$$S_x = E\{(xx^T) \otimes (J_F^T(t(x)) J_F(t(x)))\} \tag{3.8}$$

28

$$S_\eta = E\{\eta\eta^T\} \otimes E\{J_F^T(t(x))J_F(t(x))\} \tag{3.9}$$

It can be proved that $0 \leq \beta(V, M) \leq \alpha(V)$. Therefore the ratio

$$q_F(V, M, B_{opt}) = \frac{\beta(V, M)}{\alpha(V)} \tag{3.10}$$

describes a normalized figure of merit that provides a unified framework encompassing the previous measures as a function of the transformation $F()$. $J_F(t(x))$ denotes the Jacobian matrix, and because $F()$ is a linear transformation in Sharma's proposed measure (i.e., they suggested using the linearized CIELAB conversion), the Jacobian matrix is the same as $F()$ and the above approximation is exact. 424 sample reflectances were chosen as a representative dataset of which 240 were from the Kodak Q60 Photographic Scanner Target, 64 from the Munsell chart and 120 from the DuPont paint catalog. A set of 251 Gaussian filters was created to simulate and test the relation between FOM and perceptual error (i.e. CIE ΔE error).

Quan [77] proposed a unified measure of goodness (UMG) that is basically the same as Sharma's FOM. He considers the θ measure for single viewing-recording illuminant pair as:

$$\theta = 1 - \sqrt{1 - q_F(V, M, B_{opt})} \tag{3.11}$$

However, because the viewing illuminant can be different from the capture illuminant in different applications, if the sets of illuminants $\{E_{v_1}, E_{v_2}, \ldots, E_{v_n}\}$ and $\{E_{t_1}, E_{t_2}, \ldots, E_{t_m}\}$ are used as the viewing and capture illuminants, the quality factor matrix M and the unified measure of goodness (UMG) is defined as follows:

$$M = \begin{bmatrix} \theta_{11} & \theta_{12} & \cdots & \theta_{1m} \\ \theta_{21} & \theta_{22} & \cdots & \theta_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{n1} & \theta_{n2} & \cdots & \theta_{nm} \end{bmatrix} \tag{3.12}$$

$$UMG = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \theta_{ij} \tag{3.13}$$

UMG is a weighted average of the matrix M's elements. Quan also argues that dark and shot noise are equally important, and that the white noise modeled in Sharma's FOM is not sufficient for evaluating sensor sensitivities. A representative set of object reflectance spectra that contains 354 samples, including 120 Dupont paint chips, 64 Munsell chips and 170 natural and man-made object reflectance spectra measured in Situ is considered enough for designing and testing the camera spectral sensitivities.

3.2 Proposed Method: Camera Metamer Mismatch Radii Index (CMMRI)

The proposed method here is based on evaluating the degree of metamer mismatching between the camera sensitivity functions and those of the eye. The intuition behind using the degree of metamer mismatching to evaluate color accuracy is that if a human observer sees a pair of lights as matching, then the camera should too, and vice-versa.

For a given camera responding with the value RGB when viewing a given surface reflectance illuminated by a given light, there are many other surface reflectances (from the set of all theoretical surface reflectance functions) for which the camera will record the identical RGB response. The LMS triples for this set of metameric (to the camera) reflectances can become very different. All possible LMS triples construct a 3D volume that is convex [54] and is referred to as the metamer mismatch body (MMB) [110]. The volume of the MMB is a measure of the degree to which matches by the observer and camera differ. In other words, it shows how much the LMS triples of those metameric reflectances can be different. From the fact that they differ it follows that there does not exist a one-to-one mapping between camera RGB and LMS.
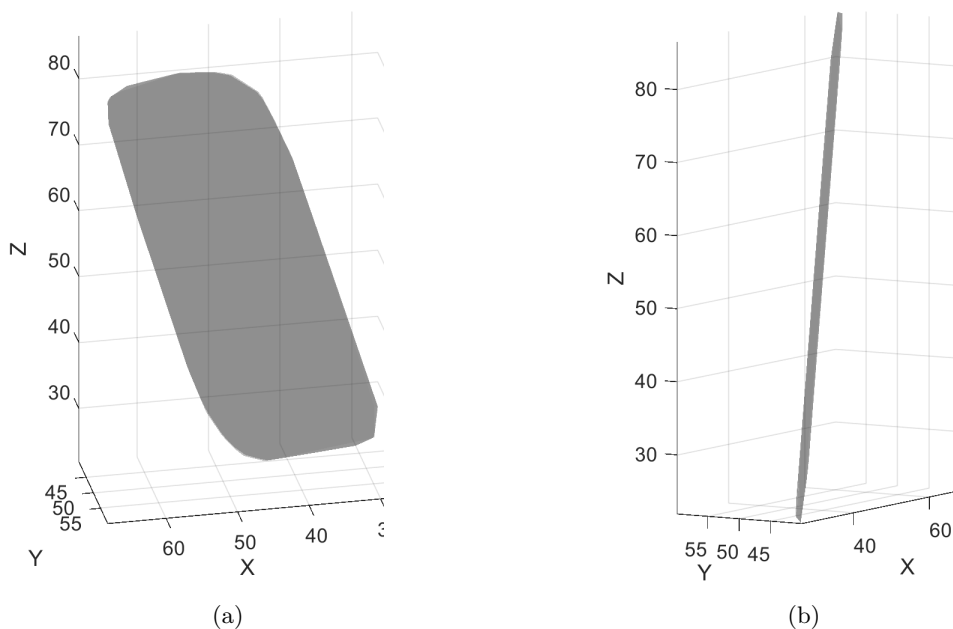
Figure 3.2: (a) A very thin metamer mismatch body (MMB) found for the Point Grey Grasshopper2 camera (and similarly for the Hasselblad H2); (b) The same MMB from a different viewing angle.

The proposed metric must be invariant to any linear transformation of the sensitivity functions. Previous methods for evaluating MMBs in the context of cameras [39] or light

sources [30] normalize the volume of the MMB by the volume of the convex hull of the spectral curve for the second observer (see Equation. 8 of [53] for a formal definition). Since a linear transformation changes both the MMB and the OCS volumes with the same scaling factor (both volumes will be multiplied by the determinant of the transformation matrix), the metric stays unaffected using this normalization step. The normalized volume method is attractive in that is based on a theoretical measure that considers all possible metameric pairs and not a finite sample. However, the MMB can be very thin in one direction and elongated in other directions, which makes the normalized volume an unstable measure. In a case such as that shown in Figure 3.2, the MMB is wide in two directions but narrow in the third. This narrowness means that the volume is small even though the degree of metamer mismatching can be large. To overcome this problem, a measure based on the MMB shape rather than its volume is proposed. Evaluating the MMBs in terms of aspects of its shape rather than its volume helps to keep the benefits of this approach while overcoming the problems created by thin MMBs. Zhang et al. [110] have shown that the MMB of flat grey (i.e., uniform 50% spectral reflectance) typifies the MMBs of other colors and so using only this one case is sufficient for our purposes. Thus, we propose to use the MMB of a 50% flat grey illuminated by D65 that results for a change from camera sensitivities to cone sensitivities. But how can we measure the dimensions of such a wing-like shape? We propose instead to characterize the shape of the MMB in terms of the radii (suitably normalized) of its equivalent ellipsoid. We define an "equivalent ellipsoid" as an ellipsoid with the same principal moments of inertia as the MMB. The advantage of evaluating the MMB in terms of these radii in contrast to using the MMB's volume is that, even in the case of a thin MMB of small volume for which one of the radii will be small, the other two radii may still be large. In other words, the other two radii correctly indicate the possibility of significant metamer mismatching.

Equation 3.14 shows the general formula for computing the moment of inertia tensor, I, of an object Q rotating around a given axis:

$$I = \iiint_Q \rho(x, y, z) ||r||^2 dV \tag{3.14}$$

where $\rho(x, y, z)$ is the mass density function at each point and r is the radius vector from the points to the axis of rotation. To calculate the MMB's equivalent ellipsoid, it is treated as a mass of uniform unit density ($\rho = 1$). For any mass, there exists an equivalent ellipsoid having the same moments of inertia (i.e., characteristics when it is spun) about its principal axes.

An ellipsoid is uniquely characterized by its three principal radii, so they concisely characterize the dominant aspects of the shape of the MMB. A linear transformation of the sensor functions, however, will change the MMB, and consequently changes the principal axes, moments of inertia, and radii of the corresponding equivalent ellipsoid. To obtain radii

that are independent of linear transformations of the sensor space, the MMB is normalized relative to the equivalent ellipsoid of the object color solid (OCS) [52] defined by the 2-transition 'optimal' color reflectances. Specifically, the principal moments of the OCS are used to determine the unique linear transformation, T, that transforms the OCS so that its equivalent ellipsoid becomes the unit sphere. The details of how T is defined are given in the section 3.2.1. The same transformation, T, is then applied to the MMB after which the principal radii of the equivalent ellipsoid of the transformed MMB are computed.

The algorithm of Logvinenko et al. [54] is used here to compute MMBs. The Camera Metamer Mismatch Radii Index (CMMRI), as explained in details in the next section, is defined as the mean of these three principal radii. The orientation of the MMB is not important, so there is no reason to weight one of the radii any more highly than the other two. Clearly, measures based on the median or the maximum of the radii are alternatives, but the mean is used here.

3.2.1 Details of CMMRI Computation

To compute the CMMRI, first consider the OCS, O, (determined as the convex hull of points on its boundary [52]) as a rigid body of a unit density and translate O so that its center of mass lies at the origin. Second, compute the inertia tensor of this centered mass. The diagonal elements of the tensor are the moments of inertia about the x, y and z axes. The off-diagonal elements are the products of inertia. Third, determine the principal moments of inertia from its inertia tensor by rotating O such that all products of inertia become zero.

Eigenvectors of the inertia tensor are ranked in descending order based on the magnitude of their corresponding eigenvalues. They form a 3×3 orthogonal matrix E. Applying E^T to the boundary points of O rotates O to become O^*. The principal axes of O^* align with the coordinate axes, and all its products of inertia are zero. The inertia tensor of O^* is a diagonal 3×3 matrix, where the elements on the diagonal are then the principal moments of inertia I_a, I_b, I_c. Given the principal moments I_a, I_b, and I_c of an ellipsoid of unit density, the ellipsoid's radii a, b and c can be derived from the following equations [101]:

$$I_a = \frac{m}{5}(b^2 + c^2) \tag{3.15}$$

$$I_b = \frac{m}{5}(c^2 + a^2) \tag{3.16}$$

$$I_c = \frac{m}{5}(a^2 + b^2) \tag{3.17}$$

In particular, the mass of an ellipsoid of uniform unit density is:

$$m = \frac{4}{3}\pi abc \tag{3.18}$$

Letting

$$P = I_b + I_c - I_a \tag{3.19}$$

$$Q = I_c + I_a - I_b \tag{3.20}$$

$$R = I_a + I_b - I_c \tag{3.21}$$

and solving for a, b and c yields,

$$a = \sqrt[5]{\frac{15 * P^2}{8\pi\sqrt{QR}}} \tag{3.22}$$

$$b = \sqrt[5]{\frac{15 * Q^2}{8\pi\sqrt{PR}}} \tag{3.23}$$

$$c = \sqrt[5]{\frac{15 * R^2}{8\pi\sqrt{PQ}}} \tag{3.24}$$

$$T = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}^{-1} E^T \tag{3.25}$$

If we apply the transformation T to the boundary points of O, its equivalent ellipsoid will become a unit sphere. Since the columns of E are the eigenvectors ranked in descending order, a,b and c in the matrix $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$ need to be ranked in descending order too. Applying T to the boundary points of the MMB rather than O normalizes the MMB relative to O. Call the normalized MMB, $M_{normalized}$. Now compute α, β and γ as the radii of the equivalent ellipsoid of $M_{normalized}$ (the ellipsoid having the same principal moments of inertia as $M_{normalized}$) in the same way as the equivalent ellipsoid of the OCS was computed. The CMMRI is then defined as the mean of these radii:

$$CMMRI = \frac{\alpha + \beta + \gamma}{3} \tag{3.26}$$

As justification for this normalization, suppose a linear transformation matrix A is applied to the camera's sensitivity functions. This will result in a new object color solid, $A * OCS$, and metamer mismatch body, $A * MMB$. The 3×3 transformation matrix, A, can be decomposed into rotation and scaling matrices using singular value decomposition (SVD):

33

$$A = U * D * V^T \tag{3.27}$$

V^T is a 3×3 orthogonal matrix that rotates the OCS, diagonal matrix D scales the principal axes of the OCS, and U rotates the result. Since rotations U and V^T preserve shape, only D will affect the shapes of the OCS and MBB. Calculating E^T given $A * OCS$ and applying it to the boundary points of $A * OCS$ results in O^* with all the rotations being canceled. The second part of matrix T (i.e., $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}^{-1}$) then cancels the scaling by converting its equivalent ellipsoid to a sphere. The scaling by D may change the order of the principal axes of $A * OCS$. For instance, instead of ijk, the principal axes may become aligned with jik. However, the order of the MMB radii is irrelevant since the order does not affect their mean.

3.2.2 CMMRI Evaluation of 35 Cameras

The sensor sensitivity functions of 28 cameras were measured by Jiang et al. [41]. Prasad et al. [75] provide the sensor spectral sensitivity functions of an additional 6 cameras (Fujifilm XM1, Nikon D5200, Olympus EPL6, Samsung NX 2000, Sony A57 and Panasonic GX1). Their 6 CMMRIs along with those of the previous 28 cameras [41] plus that of an iPhoneX [103] are reported in Figure 3.3. We also compute the RMS errors of the best linear fit of the camera sensitivities to the CIE 1931 2° observer CMF and the mean CIEDE00 color difference between the actual XYZs of the 1,950 NCS papers [36] under D65 and the RGB values of the cameras mapped to the XYZs via a best linear fit.

Figure 3.3 is a combination of three plots of the different metrics. The cameras are sorted by increasing CMMRI (green bars). Figure 3.4 plots the z-scores for each camera. A camera's z-score reflects how many standard deviations it is above or below the mean for the given metric across all the cameras. For example, the SONY NEX-5N is slightly better than average in terms of the root mean squared error (RMSE), average in terms of CMMRI, and worse than average in terms of Mean ΔE.

3.2.3 Effect of the Noise

The CMMRI measures the colorimetric accuracy of a digital color camera. Sharma [86] criticized Vora's measure for considering the camera sensors noiseless and proposed combining the filter properties and noise statistics into the single FOM measure. In particular, he takes white noise into account. The Gaussian white noise in FOM is assumed to be signal-independent and zero-mean. The noise variance is determined based on SNR values of 40, 50 and 60 dB. Quan [77] considered this as a drawback of FOM and proposed using
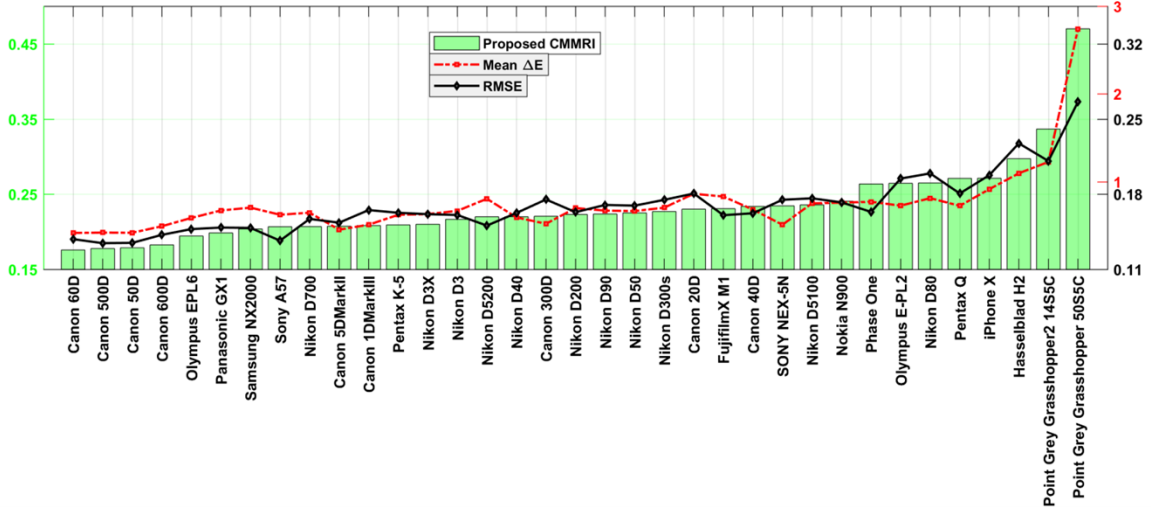
Figure 3.3: A Plot of the camera color accuracy measures. Camera Metamer Mismatch Radii Index (CMMRI), mean ΔE and root mean squared error (RMSE) for each camera sorted by increasing CMMRI. Lower scores are preferred. Note that this figure consists of three plots overlaid on one another. The scales of the ordinate axes of the plots are all different and they are also shifted from zero.
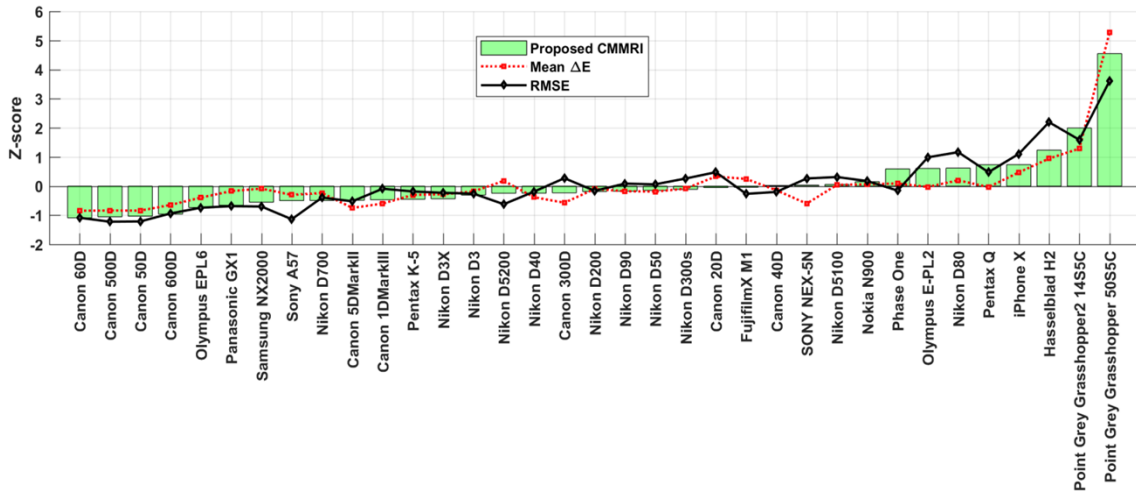
Figure 3.4: Plot of the corresponding z-scores of each of the three accuracy measures for each camera plotted in Figure 3.3. Low (including negative) z-scores are preferred.

the sum of the dark current and shot noise in the signal covariance matrix instead of white noise.

The problem with these two approaches is that in dim light the dark current noise will be the dominant noise factor while in conditions with ample light the shot noise becomes more important. This means that to select the appropriate camera for a certain application these parameters must be specified separately. Hence, a single metric combining colorimetric accuracy with noise is not particularly desirable. Colorimetric accuracy and noise need to be
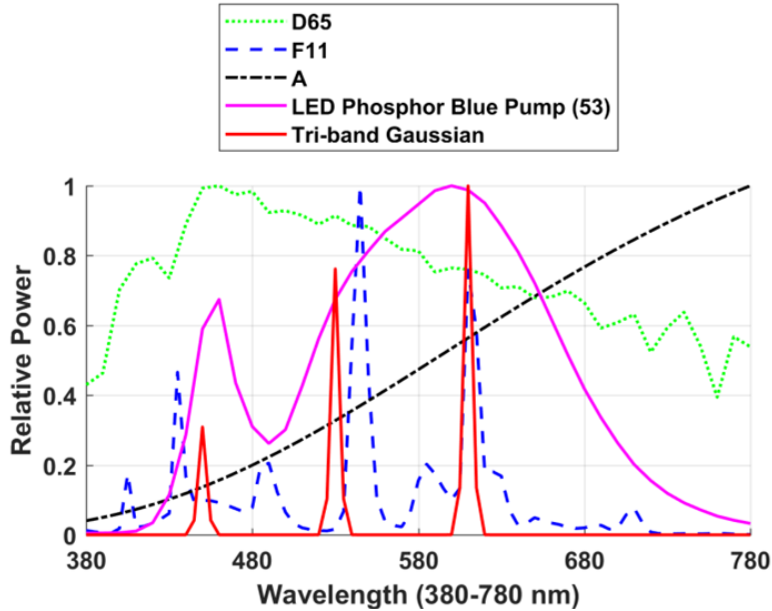
Figure 3.5: Relative spectral power distributions of CIE illuminants A, F11, D65, IES illuminants #221 (LED Phosphor Blue Pump (53)) and #317 (Tri-band Gaussian).

treated as independent variables when evaluating a camera. The focus here is on colorimetric accuracy.

3.2.4 Effect of the Illuminant Choice

As described in the previous sections, the MMB for a change from camera to cone sensitivities is calculated for the case of the flat grey reflectance illuminated by D65. However, will the proposed CMMRI measure change significantly if some other illuminant is used in place of D65? Clearly, a camera is likely to be used under a variety of different illuminants, so it is important that the CMMRI not be limited to one particular illuminant. To evaluate the effect of the choice of light on the results, CMMRIs are calculated using CIE standard illuminants A and F11; and Illuminating Engineering Society (IES) [5] illuminants #221 (light-emitting diode (LED) Phosphor Blue Pump) and #317 (Tri-band Gaussian). Their spectral power distributions are plotted in Figure 3.5. The correlation coefficients between the CMMRIs obtained using the different illuminants are reported in Table 3.1. There is a very strong linear correlation (correlation coefficient 0.97 or greater) between CMMRIs for the 'smooth' illuminants (D65, A, IES 221) (null hypothesis rejected with P-values less than 10^{-10} at the 5% significance level). There is a lesser correlation for the spiky illuminants (F11 and IES 317) but this is of little importance since it is not logical—no matter what the camera is—to evaluate colors under spiky spectra, especially spectra with zero power across a wide range of wavelengths.

Table 3.1: Correlation coefficients between the CMMRIs calculated under different illuminants.

Illuminants	A	D65	F11	IES 221	IES 317
A	1.0	0.97	0.92	0.98	0.86
D65		1.0	0.91	0.97	0.85
F11			1.0	0.96	0.92
IES 221				1.0	0.9
IES 317					1.0

3.3   Camera Accuracy Conclusion

A new color camera accuracy metric based on the degree of metamer mismatching is proposed. The rationale for using the degree of metamer mismatching to assess the color accuracy is that if a human observer sees a pair of lights as matching, then the camera should too, and vice-versa. If a camera sees two objects as the same color but the human observer sees them as different, the greater the difference the poorer the color accuracy. The MMB of 50% flat grey reflectance for a change from color camera sensors to the human cones is computed as it typifies the MMBs of other colors and so using only this one case is sufficient for our purposes. This type of metamer mismatching is referred to as 'observer metamerism' (for a change of observer) and is analogous to 'illuminant metamerism' (for a change of illuminant). The principal radii of the MMB's equivalent ellipsoid, normalized relative to the object color solid, describe the overall shape of the MMB and the mean of the radii is shown to be more stable than its volume. The normalization relative to the OCS makes the method independent of any linear transformation of the sensor space. A key advantage of the proposed method is that it shows how much color error can arise in principle. Moreover, unlike most of the existing metrics (e.g., Jiang's DE measure and Sharma's FOM), it is not dependent on a finite, and necessarily incomplete, set of test reflectances.

Chapter 4

Color Discrimination

Color difference sensitivity as represented by the size of discrimination ellipsoids is known to depend on where the colors reside within color space. Various color spaces and color difference formulas have been developed trying to establish a color coordinate system in which equally discriminable colors are equal distances apart. While fits to the data are valuable in the development of new uniform color spaces and color difference formulas, they do not explain the data. In this chapter, a theory of color discrimination based on the uncertainties reflected in the extent of metamer mismatching is proposed.

The proposed hypothesis explains why color discrimination varies in the way it does. Specifically, the greater the degree of metamer mismatching for a color, the wider the range of spectral reflectances that could have led to it; and, hence, the more finely a color needs to be discriminated in order to reliably identify materials and objects. Available color discrimination datasets for surface colors are gathered and analyzed to test this hypothesis. The strong correlation between color discrimination and metamer mismatching shows metamer mismatching provides an explanation of the variability of color discrimination. The results presented in this chapter have appeared in the following publications [32][33]:

- Brian Funt and Emitis Roshan. Colour discrimination ellipses explained by metamer mismatching. In Proceedings of the AIC International Colour Association Conference, 2018.

- Brian V Funt and Emitis Roshan. Metamer mismatching underlies color difference sensitivity. Journal of Vision, 21(12):11–11, 2021.

4.1   Background

In the study of color perception, one question that usually comes to mind is that given two color signals, how different they are. This question is usually followed by another question of does the distance in 3D color space represent the perceived color difference or not?
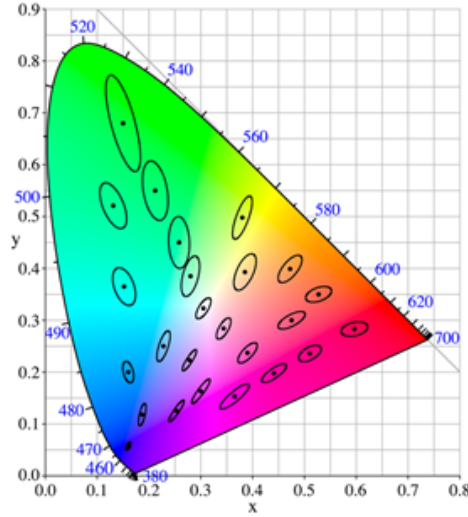
Figure 4.1: MacAdam ellipses plotted in xy-chromaticity plane (enlarged 10 times) showing the non-uniformity of CIE XYZ color space [1]. Figure licensed under CC BY-SA 3.0.

MacAdam [62] set up an experiment in which he asked the observer to change the test color in one side of a colorimeter by adjusting the intensity of three primary lights until it matches the reference color on the other side of the instrument. Then he plotted the matches on the xy-chromaticity plane and fitted ellipses to them. The ellipses are shown in Figure 4.1.

These chromaticity discrimination ellipses describe the set of colors surrounding a given color in CIE xy-chromaticity space that were indistinguishable to the observer. The variation in the ellipses' sizes reveals the non-uniformity of the CIE 1931 color space. In other words, Euclidean distance in CIE XYZ color space does not reflect the perceived color difference between two given colors for a human observer.

Silberstein and MacAdam [91] showed that the distribution of the matches in their color matching experiments were statistically normal with different standard deviations and covariances in different directions about different colors. From this they posit that the surfaces of standard deviation in 3D color spaces can be represented by ellipsoids. Silberstein [90] devised formulas determining the coefficients and the axes of color matching ellipsoids from the experimental data. Subsequently other researchers performed similar experiments aimed at measuring the standard deviation of color matches made in different regions of color space. The experimental data is later used in testing and modification of different color spaces.

CIECAM02 is a color appearance model (CAM) introduced by the CIE that provides a description of how a color stimulus appears to a human observer. It models different aspects of human color perception based on the context in which a color sample is observed, including viewing conditions such as the surrounding colors and the color of the ambient light.

Luo et al.[59] employed a combination of different datasets based on surface color samples and CRT colors to test the performance of the CIECAM02 color appearance model. They introduced three modified versions known as CAM02-SCD, CAM02-LCD and CAM02-UCS to respectively fit what are known as the SCD (small color difference) dataset, LCD (large color difference) dataset, and the combination of both. While all these modifications have the same structure as the original CIECAM02 version, Li et al. [49] [48] derived a new color appearance model named CAM16 by performing both chromatic and luminance adaptation in the same space rather than two different spaces. The proposed model avoids the unexpected problems that may occur in CIECAM02 lightness computation and at the same time outperforms it in predicting the corresponding color datasets and color appearance datasets.

Different color discrimination experiments have been carried out to gather new datasets. In the next section, some of the light-based and surface-color-based experiments will be reviewed.

4.1.1 Experiments with A Mixture of Primary Lights or Display Colors

Brown et al. [15] investigated the color discrimination for the general case of combined chromaticity and luminance differences. They followed Silberstein's method to derive the coefficients of 39 ellipsoids in one color space and presented the methods to transform the results to other coordinate systems such as CIE X, Y, Z and x, y, log Y. Their investigation involved a 2-degree monocular field of view with dark background. In another study Brown [14] used a 10-degree binocular colorimeter with white broad surrounding field. Different apparatuses utilizing a mixture of primary lights as the test or reference colors are used to gather the data which are later used in modification of the color spaces or color difference formulas [105] [106] [79] [108].

Sharma et al. [87] provided a data set for additional tests of the CIEDE2000 formula. CIEDE2000, introduced by International Commission on Illumination (CIE), is a color difference formula based on the CIELAB color space that models the color difference between any two CIELAB color values. Subsequent testing by Sharma et al. [87] revealed 3 sources of discontinuity in the CIEDE2000 equations. Wen [100] proposed calculating the color difference by counting the number of just noticeable differences between two colors and showed that it outperforms CIEDE2000 in predicting threshold color differences. The proposed algorithm is described in more detail in Chapter 5. Pridmore and Melgosa [76] analyzed four different data sets and observed that the difference between colors appears smaller at lower luminance levels. Wyszecki and Fielder [106] state that their results "... show remarkable discrepancies between ellipses obtained by the same observer at different occasions (separated by weeks or months) under otherwise identical observing conditions."

4.1.2 Experiments with Colored Surface Samples

While the majority of the experiments have been conducted using mixtures of lights, there are a few that have used colored surfaces. Cheung et al. [17] prepared one standard pair along with 59 to 82 sample pairs made of dyed wool fabric for each of the five CIE reference color centers Grey, Red, Yellow, Green and Blue [78] and asked the observers to express the color difference for each of the sample pairs as a ratio of the perceived color difference to that of a fixed standard pair. The fitted ellipsoid parameters are reported in xyY color space. The results obtained from ratio assessments with two illuminants D65 and A were shown to be in agreement with CMC (1:1) color difference formula.

Witt [102] used painted samples around four of the five CIE reference color centers: Yellow, Red, Blue and Grey. Observers were asked if the color difference was perceptible in the sample pairs or not. The correlation between the perceptibility of color differences and the colorimetric measures, and also the inter-observer and inter-group variabilities were investigated. The coefficients of the fitted ellipsoids are reported in xyY color space.

Huang et al. [38] prepared 446 pairs of printed color patches surrounding 17 color centers for a greyscale psychophysical experiment to scale the color differences of the sample pairs. Although the parameters of the fitted ellipsoids in CIELAB color space are reported in Table VII of their article, Huang et al. considered the ellipsoid's parameters less reliable than the ellipses' parameters because their research was focused on chromatic differences and the sample pairs were selected such that, compared to the variations in chromatic directions (axes a and b), they had small variations in the lightness direction (axis L). The results were later used to evaluate 10 color difference formulas.

Berns et al. [11] prepared a grey anchor pair with a color difference of 1.02 ΔE_{ab} unit in CIELAB color space using color coated aluminum panels and asked the observers to compare the magnitude of the sample color difference pairs to that of the anchor pair. Probit analysis was then used to compute 156 median tolerances around 19 color centers in different directions. Melgosa et al. [66] then used the 156 median tolerances reported in the RIT-DuPont dataset [11] to compute the ellipsoid parameters in x, y, Y/100 color space. In another study, Luo and Rigg [60] combined the data from different sources to produce a consistent set of ellipses.

Many researchers have repeated similar experiments, trying to quantify and model the color discrimination pattern. The varying sizes and orientations of the ellipses and ellipsoids show that the threshold for discriminating one color from a very similar one varies as a function of the color involved. Data obtained from these experiments are usually used in the development and testing of various new uniform color spaces and color difference formulas but no explanation is provided regarding the underlying cause. The hypothesis investigated here is that it is due to metamer mismatching.

All of the models and uniform color spaces derived from the above experiments are based on fits to the experimental data. Many provide good fits to the data, but they remain data models. Smet et al. (2016) argue that the existing algorithms for specifying the basic structure of color appearance "...are designed only by describing empirical measurements of color discrimination or similarity ratings, and not by asking what causes color appearances to be as they are." In line with our goal here, they derive a color appearance model by making general assumptions about the physiological and neural mechanisms of color encoding. While their model predicts an organization of color experience that is qualitatively similar to that of the Munsell system, no quantitative analysis is provided nor is color discrimination explicitly addressed; and, in particular, it does not explain the fundamental underlying reason as to why color discrimination thresholds vary as they do. The Smet et al. (2016) model and the earlier models by Eskew (1994) and Stockman and Brainard (2010) build upon the three-stage Müller zone model (1930) that describes cone-opponent and color-opponent pathways and model additional factors such as noise and signal compression. These studies aim to explain how neural mechanisms implement the computation of color discrimination. In contrast, the metamer mismatching hypothesis aims to help explain why the visual system computes the differences between colors the way it does.

4.2   Hypothesis

Metamer mismatching refers to the extent to which two physically distinct reflectances that match (i.e., lead to identical color signals) under one light fail to match under a second light. In general, given the color signal of an object with respect to a given illuminant, there can be many other surface reflectances that lead to the same color signal with respect to the same illuminant. Metamer mismatching arises from the fact that normal trichromatic color vision is based on only 3 weighted sum measurements of the reflected light's spectrum impinging at any given point on the retina, whereas that spectrum—the product of the illuminating light's spectrum and the surface's underlying spectral reflectance function—is much more complex. This lack of information allows two reflectances that induce the same color signal (i.e., cone response triple) under one light, and hence are a metameric match, to differ in their color signals (i.e., mismatch) under a second light.

The proposed hypothesis is that in order to be able to reliably discriminate physically distinct surfaces from one another observers must be more sensitive to the differences between colors for which metamer mismatching creates significant uncertainty (i.e., when metamer mismatching is extensive), and least sensitive for colors for which metamer mismatching creates little uncertainty.

MMB for a given color signal is a measure of the possible variability in the nature of the underlying physical response under a given light. The larger the MMB, the larger and more varied is the set of reflectances that are all metameric (i.e., create the same LMS cone

response) under a given light. Hence, for colors with large MMBs there is more uncertainty as to the exact nature of the underlying surface reflectance function. Intuitively, it is clear that there are likely more reflectance functions that lead to a mid-grey where the entire range of the visible spectrum is likely to be involved than there are to a saturated red, for example, where mainly the long-wave portion of the spectrum is likely to be involved. For an observer wishing to identify a given physical surface by its color, it is therefore more important to distinguish the shade of a gray surface as precisely as possible and less important to distinguish the exact tint of a red one. Similarly, there are very few reflectances leading to pure white, with the limit being the ideal white created by a uniform 100% reflectance. In fact, for any color signal on the boundary of the object color solid, in contrast to the grey, which is at the center of the OCS, there is only one possible reflectance creating it, so the volume of the MMB drops to zero for such color signals.

Metamer mismatching is usually discussed from the point of view of two physically distinct surfaces that appear identical under some illuminant. From the reverse perspective, however, the degree of metamer mismatching can be considered as the likelihood that two different color signals corresponding to two different reflectances under one illuminant could become indistinguishable under some other illuminant. Since the degree of metamer mismatching varies throughout color space, to reliably identify similarly colored, but physically different surfaces, from one another the visual system needs to be more discriminating in some regions of color space than others.

As an illustrative case, let us consider the uncertainty with which a color identifies a particular surface S of unknown reflectance. For example, suppose that under D65, S has color C_{D65}. The MMB of C_{D65} for a change of illumination to CIE A, for example, represents the set of all colors that could result when S is lit by CIE A instead of D65. From the reverse perspective, under CIE A, any color C_A in that MMB is a candidate for matching C_{D65} under D65. Now suppose that we observe C_A, does it correspond to S? The answer is "Almost certainly not." because C_A could have arisen from any one of an infinite set of metameric reflectances, only one of which is reflectance S. Hence, the MMB represents the 'uncertainty' in being able to identify a specific surface such as S by its color under some other illuminant (i.e., CIE A in this example).

The MMB itself represents the minimum degree uncertainty. When inaccuracy in the matching a specific color is added then the uncertainty increases. For example, suppose that under D65, color C'_{D65} is similar enough to C_{D65} that it 'matches' C_{D65}. The uncertainty then becomes the union of the MMB of C_{D65} and the MMB of C'_{D65} along with the MMBs of all colors in between. In other words, it is the set of all colors that either C_{D65} or C'_{D65} (or those in between) could become under CIE A. Given a threshold for an acceptable level of uncertainty (keeping in mind that metamer mismatching means that some uncertainty is unavoidable), how does the tolerance for error in color discrimination vary as a function of color? To provide some further intuition, Figure 4.2 shows the trend from grey to both
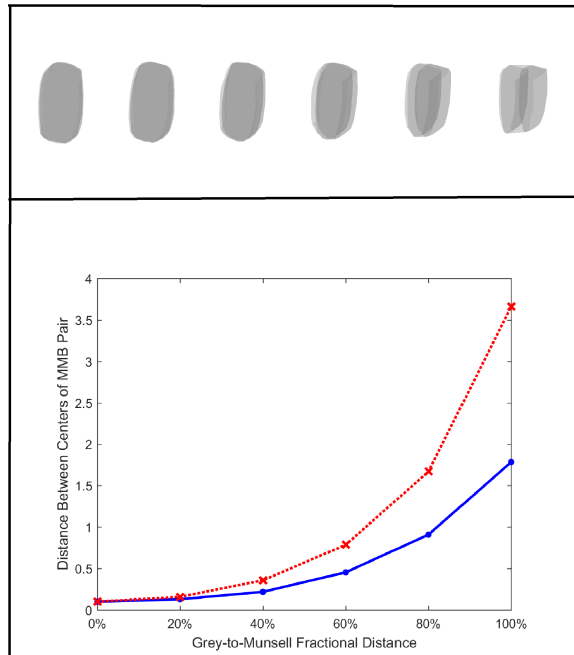
Figure 4.2: Example illustrating how the distance between two points in XYZ needs to increase in order to keep the volume of the convex hull of their MMBs equal to a constant. (Top) The two-dimensional projection of the MMB pairs showing their fixed-volume convex hulls containing MMBs that get progressively smaller as the color is moved from grey on the left to blue Munsell 5B 5/6 on the right. The distance between the two samples making up each pair is adjusted to make all the convex hull volumes the same. (Bottom) Plot of the distance between colors in a pair for flat grey (i.e., uniform 50% spectral reflectance) to blue (solid blue curve) corresponding to the MMBs (top), and in addition grey to Munsell red 5R 5/8 (dashed red curve).

blue and red (Munsells 5B 5/6 and 5R 5/8). The figure is based on keeping the volume of the convex hull of the two MMBs fixed. That choice of volume is quite arbitrary other than needing to be a somewhat larger than the volume of the MMB for flat-spectrum grey under a change in illuminant from D65 to CIE A. The convex hull of the two MMBs provides a good approximation to the union of the infinite set of MMBs for all points between the two colors. The qualitative upward trend in the Figure 4.2 (bottom) is unaffected by the precise number. Note that this example is only an illustration, not a complete model (e.g., it models distance, not volume; and it will fail for colors approaching the boundary of the object color solid where in the limiting case the volume of the MMB tends to zero). The figure is intended to provide some intuition as to how the uncertainty reflected in metamer mismatching could affect the size of discrimination ellipsoids, but intuition only. A formal statistical analysis of the evidence of the relationship between metamer mismatching and ellipsoid volume is presented in the next section.

4.3 Statistical Analysis

Since the proposed hypothesis is based on the MMB volumes, only the datasets that are based on colored surface samples and report the parameters of color discrimination ellipsoids, not just ellipses, are useful for evaluation. Since the existing datasets use a variety of different color spaces, they are first converted to a common color space, and their ellipsoid coefficients are updated correspondingly. Details of the conversion are provided in Appendix A. The 4 datasets explained in section 4.1.2 that are useful for our study are described in more detail below.

Cheung et al. [17] mounted dyed wool samples and standard pairs on a stiff card and the differences were assessed against a grey background with x, y, and Y coordinates of 0.299, 0.322 and 13.1 under D65. The samples around five color centers were chosen such that the color differences within each set range from 1 to 9 ΔE_{ab} in CIELAB units. The coordinates of the five color centers are reported in Table 4.1. The sample pairs were selected so as to give a good coverage in each of the three planes x-y, x-Y and y-Y. The fitted ellipsoid parameters for the 5 color centers were reported in x, y, Y color space.

Table 4.1: Color coordinates of the five color centers in Cheung's dataset [17]

Color Center	x	y	Y
Grey	0.314	0.331	30.0
Red	0.484	0.342	14.1
Yellow	0.388	0.428	69.3
Green	0.248	0.362	24.0
Blue	0.219	0.216	8.8

Witt [102] prepared 50 to 64 sample pairs around four of the five CIE color centers Yellow, Red, Blue and White (light grey), with color differences ranging from near zero to just clearly perceptible. The sample pairs were put on a circular shape card with 10° viewing field under the illuminant D65. The background was grey with a lightness of about Y=20. The question to be answered was "Is a color difference perceptible in the sample pair?" The coefficients of the fitted ellipsoids for the 4 color centers reported in Table 4.2 were measured in x, y, Y color space .

Table 4.2: Color coordinates of the color centers in Witt's dataset [102]

Color Center	x	y	Y
Yellow	0.3865	0.4274	69.53
Red	0.4845	0.3427	14.35
Blue	0.2185	0.2146	8.68
White	0.3143	0.3310	31.15

Huang et al. [38] conducted a greyscale psychophysical experiment in which they pre-pared 20 to 30 color samples around each of the 17 widely spread color centers in CIELAB color space. This study was done with a D65 simulator on a grey background with CIE L, a, b coordinates of 50.3, 0.2 and 1.3. Each sample pair had small contribution of lightness difference as the focus of this research was on chromatic differences. The observers were instructed to assess the chromatic differences in the sample pairs using the five greyscale pairs as the references. The XYZ coordinates of the color centers are reported in Table 4.3. In their study, rather than the ellipsoids' g_{ij} coefficients, the semimajor axes of the ellipsoids along with their orientations in CIELAB color space are reported. Nevertheless, we compute the boundary points of the ellipsoids using the given parameters in CIELAB color space and then transform them to CIE XYZ to be consistent with other datasets.

Table 4.3: XYZ coordinates of the color centers, Huang's dataset [38]

Color Center	X	Y	Z
Grey	30.1	31.8	33.9
Red	22.1	15.7	8.2
High Chroma Red	24.9	14.4	4.2
Orange	35.4	33.3	21.1
High Chroma Orange	38	29.9	4.7
Yellow	62.7	69	28.6
H-Yellow	60.7	68.4	14.9
Yellow-Green	31.1	36	28.3
High Chroma Yellow-Green	25.5	35.2	13.3
Green	17.7	25.9	27.1
High Chroma Green	17.4	26.7	28.3
Blue-Green	15.6	19.6	27.2
High Chroma Blue-Green	14.7	20.9	33.8
Blue	10.2	10	26
High Chroma Blue	9.6	9.4	33.2
Purple	17	16	24
High Chroma Purple	19.8	15.8	32.6

Berns et al. [11] on the other hand prepared only one greyscale anchor pair with CIELAB coordinates of ($L = 49.53, a = -0.08, b = -5.65$) and ($L = 48.89, a = 0.17, b = -4.90$). Showing the test pairs next to the anchor pair, the observers were instructed to judge whether the color difference magnitude of the test pair was larger or smaller than that of the anchor pair. 19 color centers were selected such that they included the 5 colors recommended by the CIE, 5 colors above (higher L values) and 5 colors under them (lower L values) and some near white and black colors. The Probit analysis was then used to compute 156 median tolerances around the color centers. The 19 color centers are reported in Table 4.4.

Table 4.4: XYZ coordinates of the color centers, RIT-DuPont dataset [11]

Color Center	X	Y	Z
Moderate blue	8.1	8.7	21.2
Moderate greenish blue	14.8	18.6	26.5
Medium grey	25.8	27.4	28.7
Moderate bluish green	16.8	23.5	23.8
Light brown	33.3	31.5	20.7
Greyish purple	16.8	15.6	23.5
Dark reddish orange	18	12.8	6.7
Moderate yellow	51.2	53.2	27.1
Greyish yellow green	29.5	33.9	27
Black	1.6	1.7	1.8
Light bluish green	28.3	38.9	46.2
Moderate reddish brown	7.5	5.8	2.8
Dark bluish green	3.8	6.9	8.9
Brilliant greenish blue	23.4	28	51.7
Very dark red	3.7	2.4	2.1
Moderate purplish pink	32.8	26	28.1
Dark blue	6.6	6.3	18.2
Light grey	59.9	63	67.4
Strong orange yellow	54.2	49.9	6.6

No ellipsoid coefficients are reported in this study. Melgosa et al. [66] used the 156 median tolerances reported in the RIT-DuPont dataset [11] to compute the ellipsoid parameters in x, y, Y/100 color space. By putting $a_{11} = a_{22} = 1, a_{33} = 0.01$ and $a_{ij} = 0$ for $i \neq j$ in the conversion equations 3.2 to 3.4 the ellipsoid parameters in x, y, Y can be obtained. Then the Equations 3.10 to 3.15 can be applied to compute the ellipsoid parameters in XYZ color space. The 4 datasets explained above are summarized in Table 4.5.

To test the metamer mismatching hypothesis the volumes of both the discrimination ellipsoids and the MMBs are needed. The discrimination ellipsoids' coefficients from all four datasets converted to XYZ color space are used to compute the volumes of the color discrimination ellipsoids, E_{vol}, in XYZ as described in Appendix A.

For each color center, the volume, M, of the corresponding MMB for a change in illuminant from CIE D65 to CIE A is computed directly in XYZ space using the algorithm of Logvinenko et al. [54]. M is then normalized by C^3, the cube of the Euclidean distance, C, from the origin to the given color center. This normalization eliminates the effect of the intensity/luminance on the volumes. If the hypothesis that metamer mismatching underlies the variability in color discrimination as a function of color center is correct, then there should be a high correlation between E_{vol} and C^3/M.

Table 4.5: The summary of the four datasets used in our statistical analysis

	Sample Type	Number of Color Centers	Number of Observers	Number of Samples
Witt	Painted samples	4	22 to 24	50 to 64 per color center
Cheung	Dyed plain wool serge	5	20	59 to 82 per color center
Huang	Samples produced using EPSON Stylus PRO 7800 ink-jet printer	17	9	20 to 30 per color center
Melgosa	Acrylic-lacquer automotive coating sprayed on primed aluminum panels	19	50	642 pairs in total

(a)   (b)
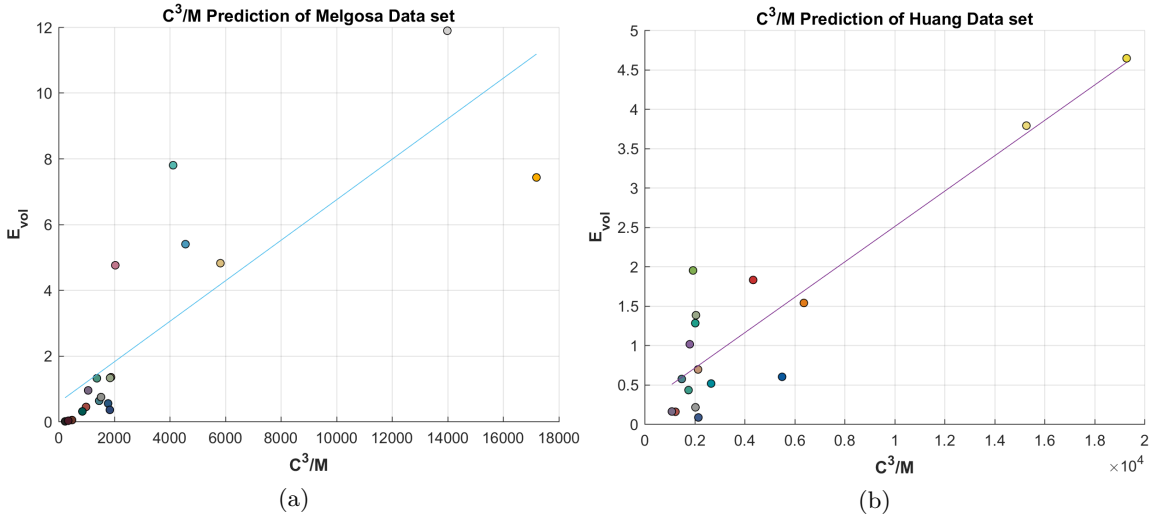
Figure 4.3: Plots of the volumes, E_{vol}, of the color discrimination ellipsoids in XYZ space as a function of the inverse of the normalized volume M of the corresponding metamer mismatch bodies, (i.e., C^3/M), for the two color discrimination datasets having samples with a minimum of 17 color centers. Left: C^3/M fit to the Melgosa dataset, $r = 0.83$, mean jackknife estimate of $r = 0.83$, bias=0.03, STD=0.13. Right: C^3/M fit to the Huang dataset, $r = 0.9$, mean jackknife estimate of $r = 0.9$, bias=-0.05, STD=0.11.

Note that C^3/M is dimensionless. The Jackknife method is used to examine the accuracy of the correlation coefficient estimates. Jackknife uses a leave-one-out strategy to derive the bias in an estimator, resulting in a bias-corrected estimate of the original statistic. The correlations between E_{vol} and C^3/M for the Melgosa (i.e., 19 color centres based on the Berns et al. data) and Huang (17 color centres) datasets are shown in Figure 4.3. The figure caption includes the Pearson correlation coefficient (r), mean jackknife estimate of

r, bias and standard error (STD) in jackknife replicates. A y-intercept is included in the linear regression model. The null hypothesis is rejected at the 5% significance level, with P-values in all cases being less than 10^{-5}. The results for the Melgosa dataset are: $r = 0.83$ with mean jackknife estimate of $r = 0.83$, bias = 0.03, STD = 0.13; and for the Huang dataset $r = 0.9$ with mean jackknife estimate of $r = 0.9$, bias = -0.05, STD = 0.11. The correlation between E_{vol} and C^3/M includes the non-linearity of the inverse $1/M$. The corresponding, simpler (negative) linear correlation results between $1/E_{vol}$ and M/C^3 are significantly weaker: -0.52 and -0.7 respectively.

4.4 Combining the datasets

In the previous section, the statistical analysis is conducted separately for the two color discrimination datasets that contain at least 17 color centers each. Two of the other datasets include discrimination ellipsoids for only 4 color centers, in one case, and 5 in the other. The goal of this section is to combine the data from all four different datasets into one larger dataset. The difficulty in doing so is that the datasets are all based on different experimental protocols resulting in different scales. For instance, in a paired-comparison experiment the color difference of a test pair is compared to an anchor pair. Grey patches with different units of ΔE color difference are prepared as the standard pairs. The observers are asked to express the difference of the test pairs as a ratio of the standard pair. Clearly different anchor pairs will result in ellipses or ellipsoids with different sizes.

We combine the ellipsoid measurements from the various datasets following the basic strategy that Luo and Rigg [60] used when combining ellipse data. Luo and Rigg [60] measured color discrimination ellipses (not ellipsoids, unfortunately) for 70 color centers and plotted the discrimination ellipses from their experiment along with those from 13 other datasets (all measurements are based on physical samples, not lights) and observed that the main discrepancy was in the relative sizes of the ellipses. They introduced the Individual Set Factor (referred to as $\overline{R}$) as a scaling factor for each ellipse, and the mean of $\overline{R}$ values for each group (referred to as S) to adjust the ellipses onto a common scale. They showed that adjusting the individual ellipses with $\overline{R}$ values results in a more consistent plot than using the group mean for each dataset.

Following the basic approach of Luo and Rigg [60], but modified for ellipsoids rather than ellipses, we use scaling factors to bring the different datasets onto a common scale. To combine different color discrimination datasets, we use the ratio of the ellipsoid volumes for the color centers that the datasets have in common. For each pair of datasets, we find the color centers that exist in both, average the ratios of their ellipsoid volumes, and then use that average ratio to normalize the datasets with respect to one another. As one example, the xyY coordinates of the color centers that are in common in the Cheung et al. and Melgosa et al. datasets are listed in Table 4.6.

Table 4.6: xyY coordinates of the common color centers in the Cheung and Melgosa datasets.

	Cheung et al.			Melgosa et al.		
	x	y	Y	x	y	Y
Grey	0.314	0.331	30.0	0.315	0.335	27.4
Red	0.484	0.342	14.1	0.481	0.341	12.7
Green	0.248	0.362	24.0	0.263	0.367	23.53

Figure 4.4 shows a linear fit for the data in the merged dataset derived from the Cheung, Melgosa, Witt, and Huang datasets. The null hypothesis is again rejected at the 5% significance level with $r = 0.88$ and P-value 7e-16.
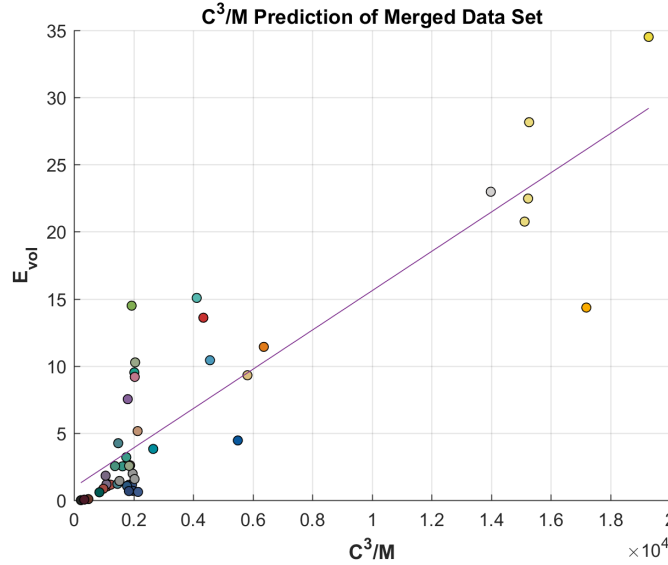
Figure 4.4: Linear fit of the ellipsoid volume (E_{vol}) versus the inverse of the normalized metamer mismatch volume (i.e., C^3/M) for the merged dataset of 45 color centres, with the null hypothesis again rejected at the 5% significance level with P-value 7e-16, $r = 0.88$, mean jackknife estimate of $r = 0.88$, bias=-0.01, STD=0.06.

4.5 Ellipsoid Volume prediction Using Normalized MMB volume

The statistics reported in the two previous sections show that there is a strong correlation between the volumes of the color discrimination ellipsoids reported in the literature and the inverse of the normalized MMB volumes. Overall, the statistics do not reject the hypothesis that the uncertainty introduced by the presence of metamer mismatching explains the variation in color discrimination thresholds. In other words, for a given color, the inverse of the normalized MMB volume predicts the volume of that color's discrimination ellipsoid.
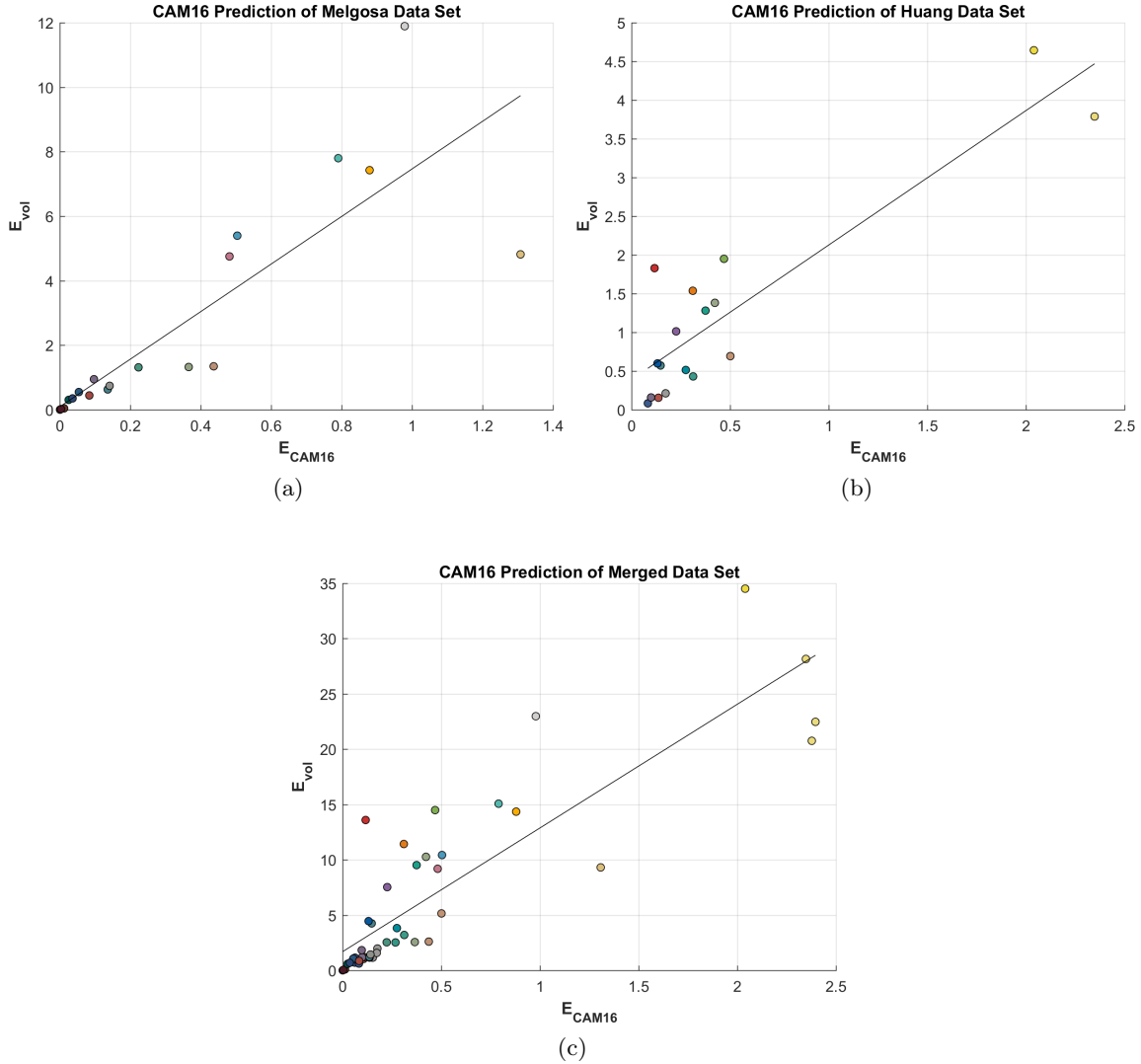
Figure 4.5: Plots of the ellipsoid volumes, E_{vol}, of the experimental color discrimination ellipsoids in XYZ space versus the ellipsoid volumes, E_{CAM16}, of the unit CAM16-UCS ΔE spheres, for both the individual and combined color discrimination datasets. The statistics for the linear fits (summarized in Table 4.7) are very similar to those in Figures 4.3 and 4.4. (4.5a): CAM16 prediction of Melgosa dataset, $r = 0.84$, mean jackknife estimate of $r = 0.85$, bias=0.04, STD=0.11. (4.5b): CAM16 prediction of Huang dataset, $r = 0.89$, mean jackknife estimate of $r = 0.89$, bias=-0.014, STD=0.07. (4.5c): CAM16 prediction of merged dataset, $r = 0.87$, mean jackknife estimate of $r = 0.87$, bias=0.0001, STD=0.04.

Since CAM16-UCS is one of the most uniform color spaces developed thus far, it is natural to consider whether or not it predicts the color discrimination ellipsoid volumes of data sets that were not used in its development any better than the proposed MMB hypothesis. CAM16-UCS describes color appearance in terms of six attributes: lightness, brightness, chroma, colorfulness, saturation, and hue. Its coefficients are based on a direct fit to experimental data. In an ideal uniform color space, the color discrimination thresholds

would be equal in all directions and about all colors; and color discrimination ellipsoids would become spheres. Therefore, we consider spheres of unit ΔE in CAM16-UCS around each of the color centers included in the 4 datasets, convert them to XYZ coordinates, and fit the volumes of the resulting ellipsoids in XYZ to those of the color discrimination ellipsoids. Figure 4.5 shows the Pearson correlation coefficient (r) between the unit ΔE spheres in CAM16-UCS color space (ellipsoids in XYZ space) and the discrimination ellipsoid volumes reported in the datasets, the mean jackknife estimate of r, the bias and standard error (STD) in jackknife replicates.

The CAM16-UCS fits can be compared to those shown in Figures 4.3 and 4.4 obtained using the inverse of normalized MMB volume. The results are also summarized in Table 4.7. The reported statistics are very close to the ones calculated with normalized MMB volumes. Given that the CIE CAM16-UCS model is optimized to fit experimental data that includes the RIT-DuPont dataset [11] (i.e., the same experimental data as the Melgosa dataset), it is significant that the normalized MMB volume performs equally well, based as it is, on an underlying theoretical principle rather than a fit to psychophysical data.

Table 4.7: Correlation coefficients between the experimental ellipsoids and (1) the ellipsoid volumes predicted by the inverse of the normalized MMB volume; and (2) the ellipsoid volumes predicted by CAM16-UCS.

	Inverse of Normalized MMB Volume	CAM16-UCS
Melgosa dataset	0.83	0.84
Huang dataset	0.9	0.89
Merged 4 datasets	0.88	0.87

4.6 Discussion

Zhang et al. [110] showed that metamer mismatching is most significant for ideal grey, high for colors of low saturation, decreases with increasing saturation, and tends to zero for colors on the boundary of the object color solid. The volume of the color discrimination ellipsoids obtained through different experiments are shown to have an inverse trend. In particular, as the extent of metamer mismatching increases, color discrimination thresholds decrease. A hypothesis is proposed that the uncertainty created by metamer mismatching underlies color discrimination thresholds. In other words, the need for the visual system to overcome the uncertainty due to metamer mismatching is the reason why there is more precise discrimination between colors in some regions of color space than others.

Four sets of experimental data measuring color discrimination ellipsoids are available for testing. The results shown in Figures 4.3 and 4.4 indicate a strong correlation between color discrimination and metamer mismatching. The fits shown are not perfect, but they

do indicate the hypothesized relationship. Of course, the experimental data contains noise, but other unaccounted factors need to be considered and investigated. Furthermore, as the results in Table 4.7 and Figure 4.5 show, CAM16-UCS, even though based on direct fits to similar experimental data, is no better a predictor of color discrimination than metamer mismatching. The strong correlation found between the experimental data and the inverse of the metamer mismatch volumes, while not proof, is evidence supporting the idea that metamer mismatching provides an explanation for as to why color discrimination varies in the way it does. It correctly predicts that color discrimination is finest near grey and becomes coarser and coarser for more and more saturated colors.

Chapter 5

Color Discrimination Ellipsoid Prediction

Color discrimination thresholds and suprathresholds are usually represented by ellipses in 2D chromaticity planes or ellipsoids in 3D color spaces. A hypothesis was investigated in the previous chapter that the Metamer Mismatching can be a primary cause for the color discrimination threshold variations throughout the color space. The statistical analysis showed that the normalized MMB volume is inversely correlated with the color discrimination ellipsoid volume. In other words, the normalized MMB volume can be used to predict the discrimination ellipsoid volume. In this chapter we investigate if the normalized MMB can be utilized to predict not only the discrimination ellipsoid volume, but also its shape and orientation as well.

5.1   Background

Over the years, numerous experiments have been conducted to measure color discrimination thresholds for different colors. MacAdam [62] showed that the Just Noticeable Differences (JNDs) in color are proportional to the standard deviation of color matchings, meaning the ellipsoids fitted to the discrimination data represent the JNDs in different directions. A Just Noticeable Difference, also referred to as the difference threshold, is the minimum amount of change required in a stimulus to be detected by 50% of the observers. Macadam [62] measured the noticeable chromaticity differences at constant luminance value. Desjardin et al. [25] measured the JNDs in luminance for the red, green, and blue primaries of a particular display to find the number of effective levels for each channel. Bedford et al. [10] measured the JNDs as a function of wavelengths. In other words, they measured the wavelength change needed in a monochromatic stimulus, while keeping the brightness level fixed, to elicit the least noticeable difference in the hue space. Surface materials such as painted ceramic, fabric, printed samples, paint, and also CRT or LCD display colors were used in the various experiments. Huang et al. [37] measured two sets of discrimination ellipses, one

with printed reflective samples and one with self-luminous color samples displayed on a CRT monitor. They observed that the ellipses obtained using self-luminous color pairs were generally bigger than the ones measured with the printed samples.

Many researchers have tried to explain the neural mechanisms in color discrimination. Kuehni [45] argues that the threshold color difference perception functions at the cone level but the supra-threshold color difference perception operates at the opponent-color level of the human visual system. Alleysson et al. [8] [7] propose a three-layer model of the retina that includes photoreceptors' nonlinear adaptive sensitivity, opponent encoding that provides normalization and contrast gain control, and the adaptive nonlinearity of ganglion cells. The adaptation factor in Naka-Rushton function used in their model accounts for the chromatic adaptation. The parameters of the proposed model are then optimized for each observer and each experimental condition to fit the experimental ellipsoids.

The discrimination data obtained from different experiments has broad application in displays quality tests and visual artifact detection created by different image processing methods. The Spatial Standard Observer (SSO) is an algorithm developed by NASA [98] to describe display mura on a JND scale. It starts with a test and a reference image (usually the low-pass filtered version of the test image). The difference between the test and the reference images is filtered by a contrast sensitivity function (CSF) and multiplied by a border aperture function to correct the visibility estimate. A JND image is then generated that relates to the percentage of human observers who will notice a difference and reflects the visibility of artifacts around a pixel. Kostal et al. [44] proposed a system for automated repeatable measurements to grade the visual quality of the displays in a way that directly correlated with human perception using SSO algorithm. Wandell [97] conducted an experiment and concluded that if the observer's state of adaptation is held fixed and the test lights a and b do not disturb the state of adaptation, in the absence of luminance component, their discriminability can be predicted from their vector difference.

Another widespread use of the discrimination data is the test and development of new color spaces and color difference formulas. CAM16 [49] [48] is one of the most recent color appearance models, obtained by optimizing a 3x3 matrix to replace the CAT02 and HPE matrices of the CIECAM02 model to better fit the experimental data. The Lightness, Colorfulness and Hue attributes have been modified in a nonlinear fashion in the definition of an improved CIELAB-like Uniform Color Space called CAM16-UCS. CIELAB is a device-independent color space based on the opponent color model of human vision. Unlike CAM16-UCS, it does not take any of the appearance phenomena into account other than the white point. Nevertheless, it is extensively used in the industry as it is less complicated and computationally simpler than CAM16-UCS. Stone et al. [93] conducted a study and observed that the smaller color samples look less colorful. They also observed that the discriminability threshold changes with the stimuli size and proposed to modify the distance in CIELAB color space as a function of size. Wen [100] proposed a color difference metric by counting

the number of just noticeable differences on a vector V that connects two color points. Starting from the first color center, it finds the intersection of the chromaticity discrimination ellipse of the first color center and vector V. The distance between the first color center and the intersection point is considered as 1 JND. Then the discrimination ellipse for the new point on V is estimated, the method explained is repeated and the number of JNDs will be counted until it reaches the second color point. The discrimination ellipse parameters for each color are estimated using polynomial regression over MacAdam's chromaticity discrimination ellipses.

Color discrimination thresholds (or JNDs) are usually measured in very time-consuming psychophysical experiments. Limited available experimental data lead the $1\Delta E$ distance in different uniform color spaces to be used as a method to predict the experimental ellipsoids for different colors. Despite the advances in color appearance models, and the uniformity in their corresponding color spaces, the Euclidean distance in CIELAB color space is widely used as the JND in several studies. Mahy et al.'s [63] evaluation of uniform color differences offered an average value of $DE_{ab} = 2.3$ for the JND in CIELAB space. Martin et al. [64] utilize color information in scan matching, a particular aspect of mapping in Simultaneous Localization and Mapping (SLAM) problem. A change in the orientation between scans causes failure in the matching process for mobile robots. A DE_{ab} of 2.3 is used in their study to detect a noticeable color transition. Hao et al. [35] introduced a method to hide a binary image in the gradient domain of a host image. They considered a DE_{ab} of 2.3 as the JND to make sure the embedded data is imperceptible. Zhang et al. [112] proposed a real-time rendering scheme for laser-beam-scanning-optical see-through head mounted displays, enhancing the color contrast between rendered images and the physical environment, while saving the power consumption of the display. The colors of the virtual content on the display are optimized based on the background color while keeping the difference between the original and optimized color under 2.3 DE_{ab}.

It will be shown, in the next section, that the inverse relationship between the normalized MMB volume and the experimental ellipsoid volume shown in the previous chapter can be used to predict the shape and orientaion of discrimination ellipsoids for different colors.

5.2 Color Discrimination Ellipsoid Prediction

It is shown in the previous chapter that for the color centers with larger normalized MMBs, the color discrimination ellipsoids are smaller and vice versa. The MMB for each color signal is created by going from illuminant D65 to illuminant A, and is normalized by C^3, the cube of the Euclidean distance, C, from the origin to the given color center. The inverse relationship between the normalized MMB volume ($\frac{M}{C^3}$) and the discrimination ellipsoid volume (E_{vol}) is demonstrated by the strong correlation coefficients between $\frac{C^3}{M}$ and E_{vol}. To model the general shape of an MMB, we use its equivalent ellipsoid as described in Section 3.2. An

MMB's equivalent ellipsoid is an ellipsoid that has the same moments of inertia (i.e., inertia tensor) as the MMB. The moments of inertia are calculated based on the MMB being a 3-dimensional solid of uniform density and they describe the radii and principal axes of the equivalent ellipsoid representing its dimensions and orientation, respectively. A typical MMB and its equivalent ellipsoid are shown in Figure 5.1.

(a)   (b)
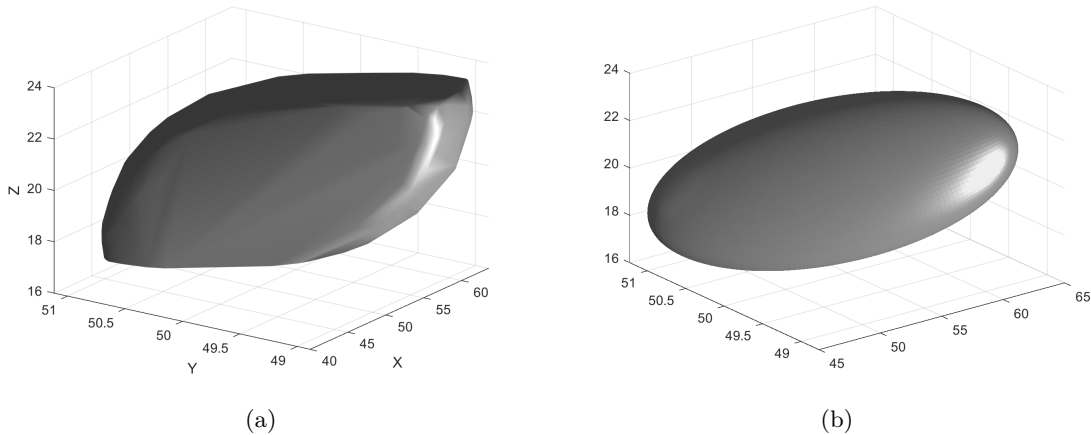
Figure 5.1: (a) A sample MMB for the Grey color center induced by going from CIE D65 to CIE A and (b) its equivalent ellipsoid.

Given an MMB, let the radii of its equivalent ellipsoid be r_1, r_2 and r_3. Because of the strong correlation coefficients between C^3/M and E_{vol} it is reasonable to expect an ellipsoid with the radii C/r_1 , C/r_2 and C/r_3, to be similar to the discrimination ellipsoid in terms of their relative magnitudes. However, the orientation of the MMB, and consequently its EE, depends on the choice of the second illuminant. So how can this orientation dependence be modelled?

The MMBs are computed for the CIE 10° standard observer for a shift from CIE illuminant D65 to CIE A. D65 is used as the first illuminant because the color centers and the color discrimination thresholds in the available experimental datasets useful for our study are all measured and reported under D65. However, how should the second illuminant be chosen? In terms of predicting ellipsoid volumes, the second illuminant has little effect on the results as shown in Table 5.1.

However, the choice of the second illuminant does affect the MMB's orientation, and hence the prediction of the corresponding ellipsoid's orientation. Figure 5.2 shows the equivalent ellipsoids of 3 MMBs computed by going from (i) D65 to Horizon; (ii) D65 to U30; and (iii) D65 to CWF. The ellipsoids are translated to the origin and their sizes are normalized. The long, medium, and short axes of each ellipsoid are shown with red, green, and blue lines, respectively. It is clear from Figure 5.2 that the orientation of the MMB, and consequently its equivalent ellipsoid, changes considerably with the second light, so it is necessary to redefine the orientation.

Table 5.1: The correlation coefficients between C^3/M and E_{vol}, and the mean Jackknife estimates as a function of the second illuminant used in predicting ellipsoid volumes.

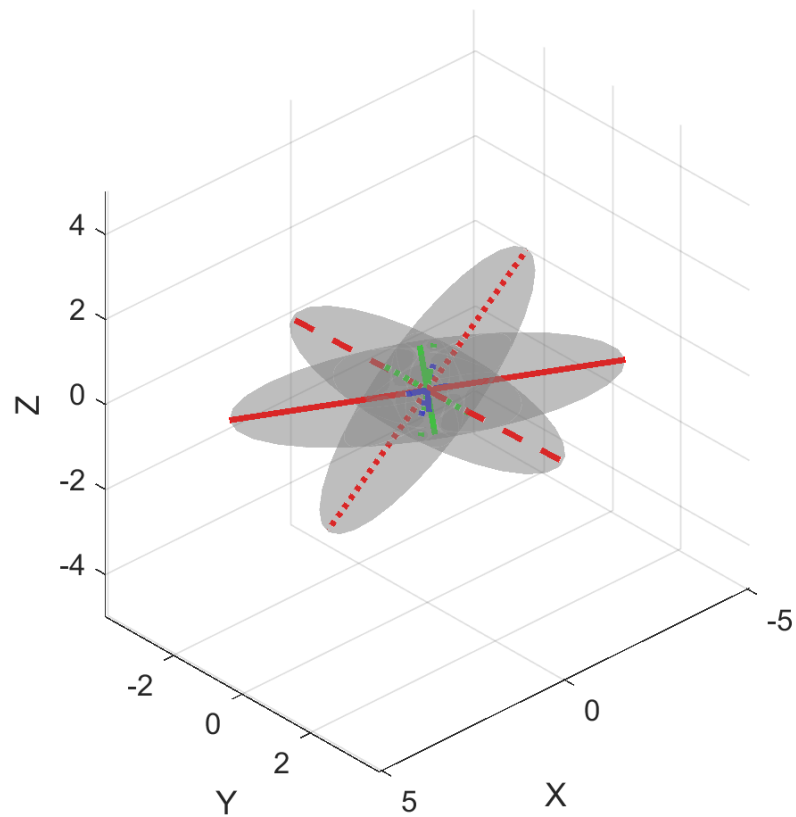| Second Illuminant	Pearson Correlation Coefficient	Mean Jackknife Estimate
Horizon 2300K	0.84	0.84
A 2800K	0.83	0.83
U30 3000K	0.90	0.90
TL84 3800K	0.91	0.91
CWF 4100K	0.89	0.89
F2 4230K	0.89	0.89
F8 5000K	0.90	0.90
F11 4000K	0.90	0.90
C 6774K	0.88	0.88

Figure 5.2: Three ellipsoids representing the MMB orientations computed by going from D65 to Horizon; D65 to U30; and D65 to CWF.

The hypothesis proposed in the previous chapter is that metamer mismatching volume reflects the ambiguity in the nature of the underlying surface reflectance function. But how

does it relate to discrimination threshold variations in different directions about a color center? The proposed hypothesis implies that the MMB volume reveals how varied the reflectances are that result in the same color signal under a given illuminant. If it holds true, then given two color signals, the larger the intersection of their corresponding MMBs, the more varied is the set of reflectances that are metameric to both color signals. Therefore, the intersection of the MMBs indicates the degree of uncertainty, i.e., the possibility of confusing their underlying surface reflectances. As a result, given a color center P, a human observer must be more sensitive when the colors in one direction create significant uncertainties with P, and less sensitive to the colors in other directions with little chance to be confused with the color center P.

Although the intersection of the MMBs of 2 color signals signifies how likely it is to confuse their surface reflectances, the MMB orientation, and hence its overlap with other MMBs, varies with the choice of the second illuminant and needs to be normalized with respect to it. Also, the wing-like shape of the MMB can make the overlap computation very sensitive to the boundary point computation, especially near the tips. So, to compute the MMBs intersection volumes, we suggest: 1. Use the equivalent ellipsoid of the MMB, and 2. Rotate the equivalent ellipsoids of the MMBs to be aligned with D65 OCS to cancel the rotation created by the second illuminant.

To integrate the above idea in our ellipsoid orientation prediction, we define 14 vector directions about each color center. The vector directions are selected in the same way that Berns et al. [11] selected the vectors to measure the color difference tolerance in their experiment. The only difference is that the vectors are defined in CIE XYZ space in our case, as opposed to CIELAB that was used in Berns et al. experiment. The vectors are plotted in Figure 5.3.
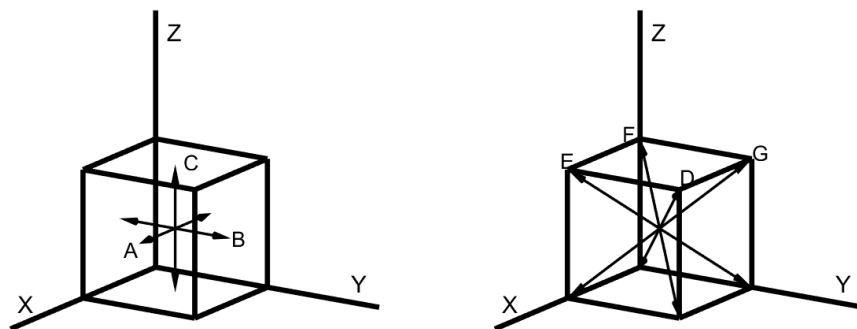

Figure 5.3: The 14 vector directions (vectors A, B, C, D, E, F and G, and their reverse directions) are defined in CIE XYZ color space. Specifically, A=[1,0,0], B=[0,1,0], C=[0,0,1], D=[1,1,1], E=[1,-1,1], F=[-1,-1,1], G=[-1,1,1].

Assuming the norm of the vector directions to be equal to 0.1 with their tails on a color center P, the tips/heads of the vectors define 14 neighboring colors around P. The MMBs for these 14 color neighbors plus one for the color center P are computed by going from D65 to illuminant A. The equivalent ellipsoids of the 15 MMBs are computed and rotated to be aligned with D65 OCS. The length of each of these vectors will then be set equal to the volume of the intersection of the two equivalent ellipsoids corresponding to the two colors at the tail and head of each vector. The tips of the new vectors (the same vector directions but with lengths equal to the overlap volumes) are used to fit an ellipsoid around P using the algorithm explained by Brown and MacAdam [15]. The 3 principal axes of the computed ellipsoid are used as the principal axes of the predicted discrimination ellipsoid, but in the reverse order. The rationale for using the reverse order is that a larger MMB overlap volume means there are likely more reflectance functions that may lead to P or the neighbor color signals in that direction, so a human observer must be more sensitive in that direction to be able to distinguish the colors, and consequently the discrimination threshold becomes smaller.

The proposed method, illustrated in Figure 5.4, is described in detail as follows:

1. Given a color center, P, in CIE XYZ space compute the MMB induced by an illuminant change from D65 to A.

2. Compute the equivlent ellipsoid of the MMB as described in Chapter 3. Call it EE_P. The radii (r_1, r_2, r_3) from the longest to the shortest, and their corresponding axes $V = [V_1, V_2, V_3]$) of the equivlent ellipsoid represent the lengths and directions of the principal axes of the MMB, respectively.

3. Compute the CIE D65 OCS for the standard observer (2° or 10°, depending on the test dataset) and CIE illuminant D65 using the 2-transition optimal color reflectances as defined by Logvinenko [54]. Modelling the CIE D65 OCS as a rigid body of unit density, compute its equivlent ellipsoid and principal axes.

4. Compute the 14 color signals around P using the 14 vector directions plotted in Figure 5.3. Call these 14 neighbor color signals N_1 to N_{14}. Compute the MMB for each color signal for a change from D65 to A.

5. Compute the equivlent ellipsoids of the 14 MMBs. Call them $EE_{N1}, \ldots, EE_{N14}$.

6. Rotate each of the 15 ellipsoids, $EE_{N1}, \ldots, EE_{N14}$ and EE_P to align with the principal axes of the D65 OCS.

7. Set the length of each vector equal to the volume of the intersection of the two ellipsoids for the two colors P and N_i at the tail and head of the vector.

8. Considering the 14 color signals at the tips of the length-adjusted vectors from the previous step, use the algorithm described in Brown and MacAdam [15] to fit an ellipsoid to them. Call the axes of the new ellipsoid (V_1', V_2', V_3') from the longest to the shortest.

9. Compute C as the Euclidean distance between the color center P and the origin. Adjust the size of EE_p such that its radii (r_1, r_2, r_3) become $(C/r_1, C/r_2, C/r_3)$, respectively. Now the order is reversed, with C/r_1 being the shortest and C/r_3 being the longest.

10. Rotate the size-adjusted ellipsoid such that $(C/r_1, C/r_2, C/r_3)$ radii will become aligned with (V_1', V_2', V_3') respectively. The new size-adjusted and rotated ellipsoid is our predicted discrimination ellipsoid for color center P.

The algorithm explained above is illustrated in Figure 5.4.

5.3   Ellipsoid Similarity Measures

To evaluate the proposed method, an ellipsoid similarity measure is needed to compare the predicted ellipsoids to those derived from the psychophysical data. Moshtaghi et al. [70] proposed an ellipsoid similarity metric, referred to as "compound similarity", for ellipsoid clustering and anomaly detection in wireless sensor networks. The proposed metric consists of three exponential factors, representing the difference in the location, orientation, and dimensions of the ellipsoid pair:

$$CS = e^{-\|\mu_1 - \mu_2\|} e^{-\|sin\theta\|} e^{-\|\alpha^* - \beta^*\|} \tag{5.1}$$

The first component represents the positional similarity, where the exponent $\|\mu_1 - \mu_2\|$ measures the Euclidean distance between the centers of the ellipsoids. The $sin\theta = (sin\theta_1, sin\theta_2, sin\theta_3)$ in the second term measures the sin of the angles between the associated eigen vector pairs. The $\alpha^* = (\alpha_1^*, \alpha_2^*, \alpha_3^*)$ and $\beta^* = (\beta_1^*, \beta_2^*, \beta_3^*)$ represent the radii of the ellipsoids, from the longest to the shortest. Since both predicted and experimental ellipsoids are centered at the same point, the first exponential component is always equal to 1; hence, it is only the angles between their major axes and the difference between their lengths that matter. The value of the compound similarity (CS) metric can vary between 0 and 1, with 1 indicating two identical ellipsoids.

   Another metric, proposed by Merritt [67], is based on computing the overlap integral between two Gaussian distributions. Since both the experimental and predicted ellipsoids are centered at the same color center, only the shape and rotational similarities are relevant. Each ellipsoid can be described by a symmetric 3×3 matrix, U, such that its eigenvectors represent the axes, and the inverse of the square root of the eigenvalues represent their
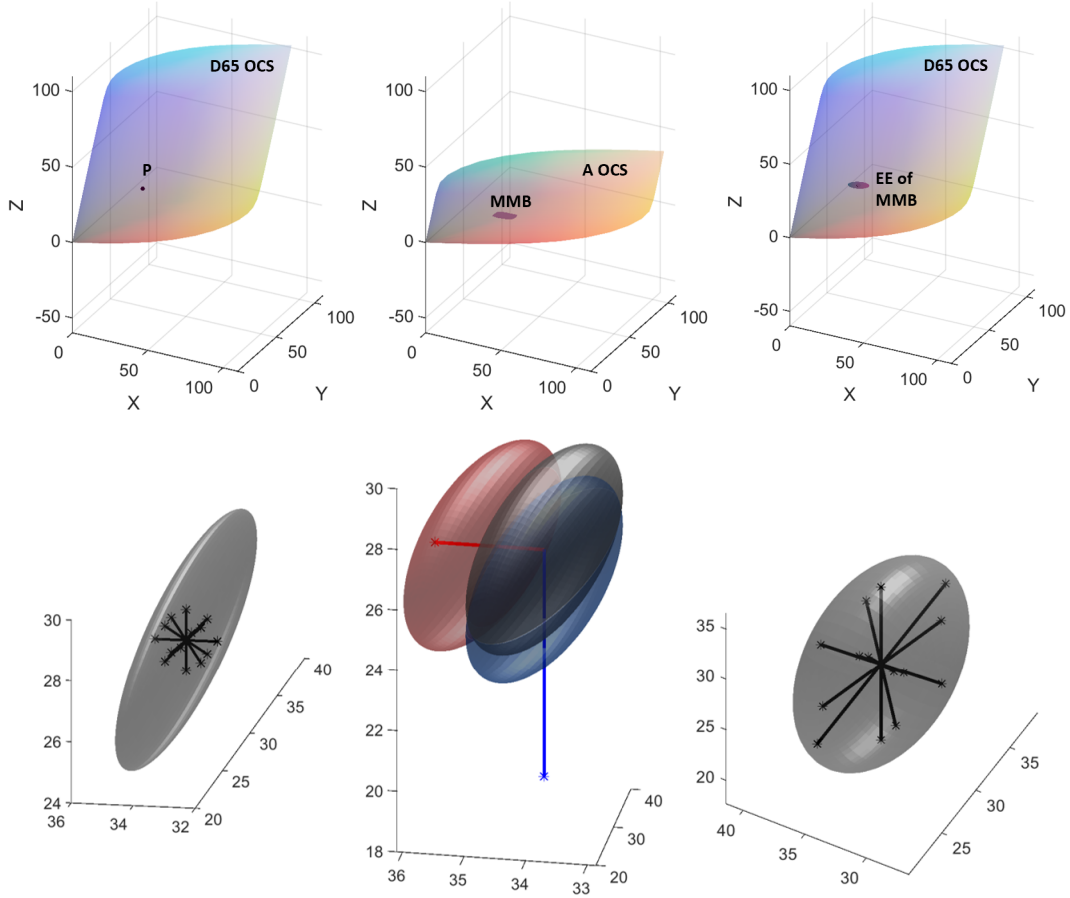
Figure 5.4: Illustration of the proposed algorithm: Top left: D65 OCS and an arbitrary color center P within the OCS. Top middle: The MMB for color center P for a change from illuminant D65 to A. The MMB is located inside the OCS for CIE illuminant A. Top right: The equivalent ellipsoid (EE) of the MMB is calculated and then translated such that its centroid coincides with color center P within D65 OCS. Bottom left: The 14 vectors defined in Figure 5.3 are centered at P within the EE of the MMB. The tips of the vectors (shown as black asterisks) show the N_1 to N_{14} color signals around P. Bottom middle: The EE of the MMB for P is shown in grey, with the EEs of the MMBs for two example color centers N_2 and N_{10} shown as the red and blue ellipsoids. The length of N_2 vector is adjusted to be equal to the overlap volume of the N_2 EE and the P EE. The length adjusted vector is shown in red. The length of N_{10} vector (the blue line) is also adjusted to be equal to the overlap volume of the N_{10} and P EEs. Bottom right: The length adjusted vectors (black lines centered at P) are then used to fit an ellipsoid. The principal axes of the fitted ellipsoid are used as the principal axes of the predicted discrimination ellipsoid.

lengths. Matrix U can be regarded as the covariance matrix of a Gaussian distribution. Let U and V be the covariance matrices corresponding to two ellipsoids. The Merritt correlation coefficient for these ellipsoids can then be computed since the Merritt correlation coefficient between two Gaussian distributions U and V is defined as:

$$Merritt(U,V) = \frac{[det(U^{-1})det(V^{-1})]^{1/4}}{[(1/8)det(U^{-1}+V^{-1})]^{1/2}} \qquad (5.2)$$

The closer the value to 1, the more similar the ellipsoids. Sample pairs of ellipses with their similarity measures are plotted in Figure 5.5.
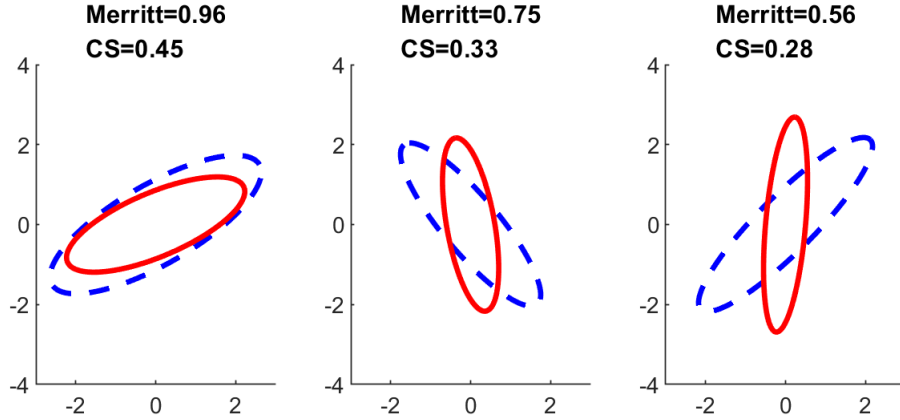
Figure 5.5: Sample pairs of ellipses with their corresponding CS and Merritt similarity measures.

5.4 Summary

Since the predictions follow from MMBs based on reflectances, only the discrimination ellipsoids that were measured using physical color patches (i.e., not colors on a display or created from mixture of primary lights) are useful for testing. The datasets summarized in Table 4.5 (Melgosa et al. [66], Huang et al. [38], Witt [102] and Cheung et al. [17]) are useful for the test.

Table 5.2: Ellipsoid similarity measures between the experimental ellipsoid versus MMB-based ellipsoid prediction, and experimental ellipsoid versus CIELAB-based ellipsoid prediction.

	MMB-Based Prediction		CIELAB-Based Prediction	
	Merritt	CS	Merritt	CS
Melgosa et al.	0.83	0.41	0.85	0.44
Huang et al.	0.82	0.36	0.77	0.33
Cheung et al.	0.81	0.37	0.82	0.38
Witt	0.82	0.55	0.83	0.62
Mean Over All Datasets	0.82	0.40	0.81	0.41

The algorithm of Logvinenko et al. [54] is used to compute the MMBs for the 45 color centers for the change from illuminant D65 to A. The equivalent ellipsoids of the MMBs are then calculated, and the algorithm explained in Section 5.2 is used to predict the experimental ellipsoids. For comparison, a unit ΔE_{ab} sphere is computed in CIELAB color space around each color center and converted to XYZ, where it becomes an ellipsoid. The mean similarity measures between the MMB-based predicted ellipsoids and the experimental data, along with the mean similarity between the ellipsoids resulting from unit ΔE_{ab} spheres in CIELAB and the experimental data for each dataset are reported in Table 5.2.

From Table 5.2, it is interesting that MMB-based prediction that is only based on a principle —the hypothesis that the uncertainty created by metamer mismatching underlies the variation in difference thresholds— is as accurate as CIELAB-based prediction which is based on a model constructed from a simple fit to experimental data. It is worth mentioning that we do not claim that the proposed hypothesis is the sole cause for the color discrimination variations. However, considering noise and possibly some other unaccounted factors, the strong correlation coefficients between the shape and orientation of the predicted ellipsoids and the experimental data demonstrate further evidence for the proposed hypothesis and reinforce the idea that the uncertainty created by metamer mismatching underlies the color difference thresholds.

Chapter 6

Conclusion

Color is not an intrinsic and independent feature of an object, but rather is an attribute of the combination of light and surface reflectance pair [55]. In other words, object color can change as a result of a change in the illumination. One of the most important phenomena that can pose big challenges in that regard is "metamer mismatching". Metamer mismatching shows that the color of an object is not a stable and can disperse into many different colors under different illuminants. An algorithm was proposed by Logvinenko et al. [54] to compute the extent of this effect by finding the boundary points of the metamer mismatching body (MMB). The convex metamer mismatching volume that can be fully described by its boundary points has interesting features that can be used to overcome some of the existing problems in color vision.

This thesis makes several contributions to the implications of metamer mistmatching. The thesis represents joint work involving extensive collaboration with my thesis supervisor. The author order given on the jointly published papers reflects the relative contribution of each author.

The first contribution is an analysis of the results of an asymmetric color matching experiment done by Logvinenko & Tokunaga. [58]. The performance of the existing color prediction methods is compared in terms of how well they explain the results of that experiment. The discrepancy between the matches made by the human observers and the computational methods reveals that human's asymmetric matching strategy is not yet well-understood. This work can serve as a foundation for further research in the field.

A color accuracy metric for digital cameras based on the MMB dimensions is introduced in the third chapter. The moments of the MMBs are used to quantify the dimensions of their wing-like shapes, and to define a new metric that is invariant to the linear transformations of the sensor space. The advantage of using MMB's shape rather its volume is that it is not subject to instability in cases in which the MMB is very narrow. In contrast to the existing methods, the proposed metric is founded upon a theoretical approach and hence does not need to measure the average error over a finite representative and necessarily incomplete set of reflectances.

A theory about the underlying reason for variations in color discrimination is explored in the fourth chapter. We showed that the greater the extent of metamer mismatching, the more finely a color needs to be discriminated. In other words, a larger MMB for a given color means more varied is the set of reflectances that are all metameric. Therefore, people become more sensitive in order to reliably discriminate physically distinct surfaces from one another, and the corresponding color discrimination ellipsoid becomes smaller. Since Zhang et al. [110] showed that metamer mismatching is greatest for grey, high for colors of low saturation, and decreases with increasing saturation, the hypothesis correctly predicts that color discrimination is finest near grey and coarsest for the saturated colors near the object color solid boundary. It is worth mentioning that the proposed theory is not intended to serve as a comprehensive model of color discrimination. Instead, it is shown to be a dominant underlying reason.

In the last chapter, the inverse relationship between the normalized MMB volume and discrimination ellipsoid volume presented in the fourth chapter is extended to predict the shape of color discrimination ellipsoids. The volume of the MMBs overlap is considered as the likelihood that two distinct colors under one illuminant will become indistinguishable under another illuminant. The greater the overlap volume in one direction, the greater the likelihood, and so a human observer become more sensitive in that direction. Thus metamer mismatching is shown to predict not only the volume, but also the shape and orientation of these ellipsoids as well. Thereby further demonstrating that the inherent uncertainty resulting from metamer mismatching underlies color discrimination thresholds. The proposed method can be utilized to predict the just noticeable color differences in different directions about a color signal that are not measured in the existing data sets.

Bibliography

[1] Macadam ellipses plotted in xy-chromaticity plane. `https://en.wikipedia.org/wiki/MacAdam_ellipse`.

[2] Testronix testing instruments. https://www.testronixinstruments.com/blog/phenomenon-of-metamerism-and-solution-to-eradicate-metamerism/.

[3] CIE technical report: Colorimetry, 3rd ed. 15:2004, CIE central bureau, Vienna. 2004.

[4] Joint CIE/ISO standard: Colorimetry – part 2: CIE standard illuminants for colorimetry. ISO 11664-2. 2007.

[5] IES method for evaluating light source color rendition. Technical report, Technical memorandum series, Illuminating Engineering Society, New York, 2018.

[6] Derya Akkaynak, Eric Chan, Justine J. Allen, and Roger T. Hanlon. Using spectrometry and photography to study color underwater. In *Oceans'11 MTS/IEEE KONA*, pages 1–8. IEEE, 2011.

[7] David Alleysson and Jeanny Hérault. Variability in color discrimination data explained by a generic model with nonlinear and adaptive processing. *Color Research & Application*, 26(S1):S225–S229, 2001.

[8] David Alleysson and Sabine Süsstrunk. On adaptive non-linearity for color discrimination and chromatic adaptation. In *Conference on Colour in Graphics, Imaging, and Vision*, volume 2002, pages 190–195. Society for Imaging Science and Technology, 2002.

[9] M. Shamsul Arifin, M. Golam Kibria, Adnan Firoze, M. Ashraful Amini, and Hong Yan. Dermatological disease diagnosis using color-skin images. In *2012 international conference on machine learning and cybernetics*, volume 5, pages 1675–1680. IEEE, 2012.

[10] R.E. Bedford and Günter W. Wyszecki. Wavelength discrimination for point sources. *JOSA*, 48(2):129–135, 1958.

[11] Roy S. Berns, David H. Alman, Lisa Reniff, Gregory D. Snyder, and Mitchell R. Balonon-Rosen. Visual determination of suprathreshold color-difference tolerances using probit analysis. *Color Research & Application*, 16(5):297–316, 1991.

[12] James Bowmaker and H.J.A. Dartnall. Visual pigments of rods and cones in a human retina. *The Journal of physiology*, 298(1):501–511, 1980.

[13] David H. Brainard, Wendy A. Brunt, and Jon M. Speigle. Color constancy in the nearly natural image. 1. asymmetric matches. *JOSA A*, 14(9):2091–2110, 1997.

[14] W.R.J. Brown. Color discrimination of twelve observers. *JOSA*, 47(2):137–143, 1957.

[15] W.R.J. Brown and David L. MacAdam. Visual sensitivities to combined chromaticity and luminance differences. *JOSA*, 39(10):808–834, 1949.

[16] Scott A. Burns, Jozef B. Cohen, and Edward N. Kuznetsov. Multiple metamers: preserving color matches under diverse illuminants. *Color Research & Application*, 14(1):16–22, 1989.

[17] M. Cheung and B. Rigg. Colour-difference ellipsoids for five CIE colour centres. *Color Research & Application*, 11(3):185–195, 1986.

[18] A.K.R. Choudhury. *Principles of colour and appearance measurement: Object appearance, colour perception and instrumental measurement.* Elsevier, 2014.

[19] CUPC CIE. Commission internationale de l'eclairage proceedings, 1931. *Cambridge University, Cambridge*, 1932.

[20] Jozef B. Cohen and William E. Kappauf. Color mixture and fundamental metamers: Theory, algebra, geometry, application. *The American journal of psychology*, pages 171–259, 1985.

[21] Shaun Collings, Oliver Thompson, Evan Hirst, Louise Goossens, Anup George, and Robert Weinkove. Non-invasive detection of anaemia using digital photographs of the conjunctiva. *PloS one*, 11(4):e0153286, 2016.

[22] Meenu Dadwal and V.K. Banga. Color image segmentation for fruit ripeness detection: a review. In *2nd International Conference on Electrical, Electronics and Civil Engineering (ICEECE'2012)*, 2012.

[23] Vasco M.N. De Almeida, Paulo T. Fiadeiro, and Sérgio M.C. Nascimento. Color constancy by asymmetric color matching with real objects in three-dimensional scenes. *Visual Neuroscience*, 21(3):341–345, 2004.

[24] Maxim W. Derhak and Roy S. Berns. Introducing wpt (waypoint): A color equivalency representation for defining a material adjustment transform. *Color Research & Application*, 40(6):535–549, 2015.

[25] Daniel D. Desjardins and Patrick Gardner. Just noticeable color difference: implications for display systems. In *Display Technologies and Applications for Defense, Security, and Avionics IX; and Head-and Helmet-Mounted Displays XX*, volume 9470, pages 116–127. SPIE, 2015.

[26] Mark D. Fairchild. *Color appearance models.* John Wiley & Sons, 2013.

[27] Graham D. Finlayson, Mark S. Drew, and Brian Funt. Spectral sharpening: sensor transformations for improved color constancy. *JOSA A*, 11(5):1553–1563, 1994.

[28] Graham D. Finlayson and Peter Morovic. Metamer sets. *JOSA A*, 22(5):810–819, 2005.

[29] Toshihisa Fujiyoshi, Ryoji Miyahara, Kohei Funasaka, Kazuhiro Furukawa, Tsunaki Sawada, Keiko Maeda, Takeshi Yamamura, Takuya Ishikawa, Eizaburo Ohno, Masanao Nakamura, et al. Utility of linked color imaging for endoscopic diagnosis of early gastric cancer. *World journal of gastroenterology*, 25(10):1248, 2019.

[30] Brian Funt, Ben Hull, and Xiandou Zhang. Evaluation of the IES method for evaluating light source color rendition in terms of metamer mismatching. In *Color and Imaging Conference*, volume 2016, pages 192–197. Society for Imaging Science and Technology, 2016.

[31] Brian Funt and Hao Jiang. Nondiagonal color correction. In *Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429)*, volume 1, pages I–481. IEEE, 2003.

[32] Brian Funt and Emitis Roshan. Colour discrimination ellipses explained by metamer mismatching. In *Proceedings of the AIC International Colour Association Conference*, 2018.

[33] Brian V. Funt and Emitis Roshan. Metamer mismatching underlies color difference sensitivity. *Journal of Vision*, 21(12):1–11, 2021.

[34] Theo Gevers and Arnold W.M. Smeulders. Pictoseek: Combining color and shape invariant features for image retrieval. *IEEE transactions on Image Processing*, 9(1):102–119, 2000.

[35] Lu Hao, Jie Feng, and Bingfeng Zhou. Gradient domain binary image hiding using color difference metric. In *SIGGRAPH Asia 2015 Technical Briefs*, pages 1–4. 2015.

[36] Anders Hård and Lars Sivik. NCS—natural color system: a swedish standard for color notation. *Color Research & Application*, 6(3):129–138, 1981.

[37] Min Huang, Hao-Xue Liu, and Ning-Fang Liao. Comparative study of just-noticeable color differences using reflective and self-luminous color stimuli. In *2008 International Conference on Computer Science and Software Engineering*, volume 6, pages 298–301. IEEE, 2008.

[38] Min Huang, Haoxue Liu, Guihua Cui, and M. Ronnier Luo. Testing uniform colour spaces and colour-difference formulae using printed samples. *Color Research & Application*, 37(5):326–335, 2012.

[39] Ben Hull and Brian Funt. Comparing colour camera sensors using metamer mismatch indices. In *Proceedings of the AIC International Colour Association Conference*, 2015.

[40] Ming-Yan Jia, Qiong-Shui Wu, Hui Li, Yu Zhang, Ya-Feng Guan, and Liang Feng. The calibration of cellphone camera-based colorimetric sensor array and its application in the determination of glucose in urine. *Biosensors and Bioelectronics*, 74:1029–1037, 2015.

[41] Jun Jiang, Dengyu Liu, Jinwei Gu, and Sabine Süsstrunk. What is the space of spectral sensitivity functions for digital color cameras? In *2013 IEEE Workshop on Applications of Computer Vision (WACV)*, pages 168–179. IEEE, 2013.

[42] Lijun Jiang, Wang Xu, and Jianfeng Chen. Digital imaging system for physiological analysis by tongue colour inspection. In *2008 3rd IEEE Conference on Industrial Electronics and Applications*, pages 1833–1836. IEEE, 2008.

[43] Volkan Kilic, Gazihan Alankus, Nesrin Horzum, Ali Y. Mutlu, Abdullah Bayram, and Mehmet E. Solmaz. Single-image-referenced colorimetric water quality detection using a smartphone. *ACS omega*, 3(5):5531–5536, 2018.

[44] H. Kostal, G. Pedeville, and R. Rykowski. Methods for measuring display defects as correlated to human perception. In *Image Quality and System Performance VI*, volume 7242, pages 151–159. SPIE, 2009.

[45] Rolf G. Kuehni. Threshold color differences compared to supra-threshold color differences. *Color Research & Application*, 25(3):226–229, 2000.

[46] Lauren Kunde, Erin McMeniman, and Malcolm Parker. Clinical photography in dermatology: Ethical and medico-legal considerations in the age of digital and smartphone technology. *Australasian Journal of Dermatology*, 54(3):192–197, 2013.

[47] Terence S. Leung, Felix Outlaw, Lindsay W. MacDonald, and Judith Meek. Jaundice eye color index (JECI): quantifying the yellowness of the sclera in jaundiced neonates with digital photography. *Biomedical optics express*, 10(3):1250–1256, 2019.

[48] Changjun Li, Zhiqiang Li, Zhifeng Wang, Yang Xu, Ming Ronnier Luo, Guihua Cui, Manuel Melgosa, Michael H. Brill, and Michael Pointer. Comprehensive color solutions: CAM16, CAT16, and CAM16-UCS. *Color Research & Application*, 42(6):703–718, 2017.

[49] Changjun Li, Zhiqiang Li, Zhifeng Wang, Yang Xu, Ming Ronnier Luo, Guihua Cui, Manuel Melgosa, and Michael Pointer. A revision of CIECAM02 and its CAT and UCS. In *Color and Imaging Conference*, volume 2016, pages 208–212. Society for Imaging Science and Technology, 2016.

[50] Qingli Li and Zhi Liu. Tongue color analysis and discrimination based on hyperspectral images. *Computerized Medical Imaging and Graphics*, 33(3):217–221, 2009.

[51] Guoxu Liu, Shuyi Mao, and Jae Ho Kim. A mature-tomato detection algorithm using machine learning and color analysis. *Sensors*, 19(9):2023, 2019.

[52] Alexander D. Logvinenko. An object-color space. *Journal of Vision*, 9(11):5–5, 2009.

[53] Alexander D. Logvinenko. Colour variations arising from observer-induced metamer mismatching. *Unpublished manuscript. Available at https://www. researchgate. net/publication/262122838*, 2014.

[54] Alexander D. Logvinenko, Brian Funt, and Christoph Godau. Metamer mismatching. *IEEE Transactions on Image Processing*, 23(1):34–43, 2013.

[55] Alexander D. Logvinenko, Brian Funt, Hamidreza Mirzaei, and Rumi Tokunaga. Rethinking colour constancy. *PLoS One*, 10(9):e0135029, 2015.

[56] Alexander D. Logvinenko and Vladimir L. Levin. *Foundations of Colour Science: From Colorimetry to Perception.* John Wiley & Sons, 2022.

[57] Alexander D. Logvinenko and Laurence T. Maloney. The proximity structure of achromatic surface colors and the impossibility of asymmetric lightness matching. *Perception & Psychophysics*, 68(1):76–83, 2006.

[58] Alexander D. Logvinenko and Rumi Tokunaga. Colour constancy as measured by least dissimilar matching. *Seeing and perceiving*, 24(5):407–452, 2011.

[59] M. Ronnier Luo, Guihua Cui, and Changjun Li. Uniform colour spaces based on CIECAM02 colour appearance model. *Color Research & Application*, 31(4):320–330, 2006.

[60] M. Ronnier Luo and B. Rigg. Chromaticity-discrimination ellipses for surface colours. *Color Research & Application*, 11(1):25–42, 1986.

[61] Robert Luther. Aus dem gebiet der farbreizmetrik (on color stimulus metrics). *Z. Tech. Phys*, 12:540–558, 1927.

[62] David L. MacAdam. Visual sensitivities to color differences in daylight. *JOSA*, 32(5):247–274, 1942.

[63] Marc Mahy, Luc Van Eycken, and André Oosterlinck. Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV. *Color Research & Application*, 19(2):105–121, 1994.

[64] Fernando Martín, Jaime Valls Miró, and Luis Moreno. Towards exploiting the advantages of colour in scan matching. In *ROBOT2013: First Iberian Robotics Conference*, pages 217–231. Springer, 2014.

[65] John McCann. personal communication, 2017.

[66] M. Melgosa, E. Hita, A.J. Poza, David H. Alman, and Roy S. Berns. Suprathreshold color-difference ellipsoids for surface colors. *Color Research & Application*, 22(3):148–155, 1997.

[67] Ethan A. Merritt. Comparing anisotropic displacement parameters in protein structures. *Acta Crystallographica Section D: Biological Crystallography*, 55(12):1997–2004, 1999.

[68] Hamidreza Mirzaei and Brian Funt. Gaussian illuminants and reflectances for colour signal prediction. In *Color and Imaging Conference*, volume 2014, pages 212–216. Society for Imaging Science and Technology, 2014.

[69] Nathan Moroney, Mark D. Fairchild, Robert W.G. Hunt, Changjun Li, M. Ronnier Luo, and Todd Newman. The CIECAM02 color appearance model. In *Color and Imaging Conference*, volume 2002, pages 23–27. Society for Imaging Science and Technology, 2002.

[70] Masud Moshtaghi, Timothy C. Havens, James C. Bezdek, Laurence Park, Christopher Leckie, Sutharshan Rajasegarar, James M. Keller, and Marimuthu Palaniswami. Clustering ellipses for anomaly detection. *Pattern Recognition*, 44(1):55–69, 2011.

[71] Ali Y. Mutlu, Volkan Kılıç, Gizem Kocakuşak Özdemir, Abdullah Bayram, Nesrin Horzum, and Mehmet E. Solmaz. Smartphone-based colorimetric detection via machine learning. *Analyst*, 142(13):2434–2441, 2017.

[72] H.E.J. Neugebauer. Quality factor for filters whose spectral transmittances are different from color mixture curves, and its application to color photography. *JOSA*, 46(10):821–824, 1956.

[73] Jussi P.S. Parkkinen, Jarmo Hallikainen, and Timo Jaaskelainen. Characteristic spectra of munsell colors. *JOSA A*, 6(2):318–322, 1989.

[74] J. Pérez-Carpinell, V.J. Camps, and J.A. Diaz. Color memory in protanomals and deuteranomals: matching time effect. *Color Research & Application*, 28(2):151–154, 2003.

[75] Dilip Prasad. personal communication, 2018.

[76] Ralph W. Pridmore and Manuel Melgosa. Effect of luminance of samples on color discrimination ellipses: Analysis and prediction of data. *Color Research & Application*, 30(3):186–197, 2005.

[77] Shuxue Quan. *Evaluation and Optimal Design of Spectral Sensitivities for Digital Color Imaging*. PhD thesis, PhD dissertation, Rochester, NY: Rochester Institute of Technology, 2002.

[78] A. R. Robertson. CIE guidelines for coordinated research on color-difference evaluation. *Color Research & Application*, 2:149–152, 1978.

[79] Javier Romero, José A. García, Luis Jiménez del Barco, and E. Hita. Evaluation of color-discrimination ellipsoids in two-color spaces. *JOSA A*, 10(5):827–837, 1993.

[80] Emitis Roshan and Brian Funt. Evaluation of color prediction methods in terms of least-dissimilar asymmetric matching. *Electronic Imaging*, 2017(14):140–144, 2017.

[81] Emitis Roshan and Brian Funt. Computational color prediction versus least-dissimilar matching. *JOSA A*, 35(4):B292–B298, 2018.

[82] Emitis Roshan and Brian Funt. Color sensor accuracy index utilizing metamer mismatch radii. *Sensors*, 20(15):4275, 2020.

[83] Emitis Roshan, Brian Funt, and Hamidreza Mirzaei. Camera color accuracy evaluated via metamer mismatch moments. In *Proceedings of the AIC International Colour Association Conference*, 2017.

[84] Francis J.M. Schmitt. A method for the treatment of metamerism in colorimetry. *JOSA*, 66(6):601–608, 1976.

[85] Erwin Schrödinger. Theorie der pigmente von größter leuchtkraft [theory of pigments of greatest lightness]. *Annalen der Physik*, 62:603–622, 1920.

[86] Gaurav Sharma and H. Joel Trussell. Figures of merit for color scanners. *IEEE Transactions on Image Processing*, 6(7):990–1001, 1997.

[87] Gaurav Sharma, Wencheng Wu, and Edul N. Dalal. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application*, 30(1):21–30, 2005.

[88] Li Shen, Joshua A. Hagen, and Ian Papautsky. Point-of-care colorimetric detection with a smartphone. *Lab on a Chip*, 12(21):4240–4243, 2012.

[89] S. Siegal and N.J. Castellan. Nonparametric statistics for the behavioral sciences, new york; pcgraw-hill, 1988.

[90] Ludwik Silberstein. On two accessories of three-dimensional colorimetryi. The probable error of colorimetric tensor components as derived from a number of color matchingsii. the determination of the principal colorimetric axes at any point of the color threefold. *JOSA*, 36(8):464–468, 1946.

[91] Ludwik Silberstein and David L. MacAdam. The distribution of color matchings around a color center. *JOSA*, 35(1):32–39, 1945.

[92] Thomas Smith and John Guild. The CIE colorimetric standards and their use. *Transactions of the optical society*, 33(3):73, 1931.

[93] Maureen Stone, Danielle Albers Szafir, and Vidya Setlur. An engineering model for color difference as a function of size. In *Color and Imaging Conference*, volume 2014, pages 253–258. Society for Imaging Science and Technology, 2014.

[94] H.J. Trussell, G. Sharma, P. Chen, and S.A. Rajala. Comparison of measures of goodness of sets of color scanning filters. In *Proc. IEEE Ninth Multidimensional Signal Processing Workshop*, pages 98–99, 1996.

[95] Johannes Von Kries. Chromatic adaptation. *Sources of color vision*, pages 145–148, 1970.

[96] Poorvi L. Vora and H. Joel Trussell. Measure of goodness of a set of color-scanning filters. *JOSA A*, 10(7):1499–1508, 1993.

[97] Brian A. Wandell. Color measurement and discrimination. *JOSA A*, 2(1):62–71, 1985.

[98] Andrw B. Watson. Spatial standard observer. Technical report, 2010.

[99] Alvin G. Wee, Delwin T. Lindsey, Shanglun Kuo, and William M. Johnston. Color accuracy of commercial digital cameras for use in dentistry. *Dental Materials*, 22(6):553–559, 2006.

[100] Senfar Wen. A color difference metric based on the chromaticity discrimination ellipses. *Optics express*, 20(24):26441–26447, 2012.

[101] Benjamin Williamson. *An Elementary Treatise on the Integral Calculus: Containing Applications to Plane Curves and Surfaces, and Also Chapters on the Calculus of Variations, with Numerous Examples.* Longmans, Green, and Company, 1896.

[102] Klaus Witt. Three-dimensional threshold of color-difference perceptibility in painted samples: variability of observers in four CIE color regions. *Color Research & Application*, 12(3):128–134, 1987.

[103] Dietmar Wuelller. personal communication, 2018.

[104] Günter Wyszecki. Evaluation of metameric colors. *JOSA*, 48(7):451–454, 1958.

[105] Günter Wyszecki. Matching color differences. *JOSA*, 55(10):1319–1324, 1965.

[106] Günter Wyszecki and G.H. Fielder. New color-matching ellipses. *JOSA*, 61(9):1135–1152, 1971.

[107] G. Wyzecki and W.S. Stiles. Concepts and methods, quantitative data and formulae. *Color Science, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA*, 1982.

[108] A. Yebra, J.A. Garcia, J.L. Nieves, and J. Romero. Chromatic discrimination in relation to luminance level. *Color Research & Application*, 26(2):123–131, 2001.

[109] Ali K. Yetisen, J.L. Martinez-Hurtado, Angel Garcia-Melendrez, Fernando da Cruz Vasconcellos, and Christopher R. Lowe. A smartphone algorithm with inter-phone repeatability for the analysis of colorimetric tests. *Sensors and actuators B: chemical*, 196:156–160, 2014.

[110] Xiandou Zhang, Brian Funt, and Hamidreza Mirzaei. Metamer mismatching in practice versus theory. *JOSA A*, 33(3):A238–A247, 2016.

[111] Yinda Zhang, Shuran Song, Ersin Yumer, Manolis Savva, Joon-Young Lee, Hailin Jin, and Thomas Funkhouser. Physically-based rendering for indoor scene understanding using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5287–5295, 2017.

[112] Yunjin Zhang, Rui Wang, Evan Yifan Peng, Wei Hua, and Hujun Bao. Color contrast enhanced rendering for optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 28(12):4490–4502, 2021.

Appendix A

Ellipsoid Parameters Conversion Equations

In general, given g_{ik}; the ellipsoid parameters in one coordinate system such as the instrument R, G, B; the following relation recommended by Brown et al. [15] can be used in order to obtain the parameters in another color space such as X, Y, Z:

$$G_{jl} = (\frac{\partial x_i}{\partial x'_j})(\frac{\partial x_k}{\partial x'_l}).g_{ik}, i, j, k, l = 1, 2, 3 \tag{A.1}$$

where $x_i, x_k = R, G, B$ and $x'_j, x'_l = X, Y, Z$, (i.e. $x_1 = R, x_2 = G, etc$)

G_{jl} values are the ellipsoid parameters in XYZ color space, and the partial derivatives are obtained using the transformation equations:

$$R = a_{11}X + a_{12}Y + a_{13}Z \tag{A.2}$$

$$G = a_{21}X + a_{22}Y + a_{23}Z \tag{A.3}$$

$$B = a_{31}X + a_{32}Y + a_{33}Z \tag{A.4}$$

For instance

$$\frac{\partial x_1}{\partial x'_1} = a_{11}, \frac{\partial x_1}{\partial x'_2} = a_{12} \tag{A.5}$$

and writing one equation in full gives

$$\begin{aligned} G_{11} =& (\frac{\partial x_1}{\partial x'_1})(\frac{\partial x_1}{\partial x'_1})g_{11} + (\frac{\partial x_1}{\partial x'_1})(\frac{\partial x_2}{\partial x'_1})g_{12} + (\frac{\partial x_1}{\partial x'_1})(\frac{\partial x_3}{\partial x'_1})g_{13}+ \\ & (\frac{\partial x_2}{\partial x'_1})(\frac{\partial x_1}{\partial x'_1})g_{21} + (\frac{\partial x_2}{\partial x'_1})(\frac{\partial x_2}{\partial x'_1})g_{22} + (\frac{\partial x_2}{\partial x'_1})(\frac{\partial x_3}{\partial x'_1})g_{23}+ \\ & (\frac{\partial x_3}{\partial x'_1})(\frac{\partial x_1}{\partial x'_1})g_{31} + (\frac{\partial x_3}{\partial x'_1})(\frac{\partial x_2}{\partial x'_1})g_{32} + (\frac{\partial x_3}{\partial x'_1})(\frac{\partial x_3}{\partial x'_1})g_{33} \end{aligned} \tag{A.6}$$

The transformation between x, y, Y and XYZ is given by:

$$x = \frac{X}{X+Y+Z} \tag{A.7}$$

$$y = \frac{Y}{X+Y+Z} \tag{A.8}$$

$$Y = Y \tag{A.9}$$

Proceeding similar to the example above, we find that

$$G_{11} = A^2 g_{11} + 2AC g_{12} + C^2 g_{22} \tag{A.10}$$

$$G_{12} = AB g_{11} + (AD+BC) g_{12} + A g_{13} + CD g_{22} + C g_{23} \tag{A.11}$$

$$G_{13} = AB g_{11} + (AC+BC) g_{12} + C^2 g_{22} \tag{A.12}$$

$$G_{22} = B^2 g_{11} + 2BD g_{12} + 2B g_{13} + D^2 g_{22} + 2D g_{23} + g_{33} \tag{A.13}$$

$$G_{23} = B^2 g_{11} + (BC+BD) g_{12} + DC g_{22} + B g_{13} + C g_{23} \tag{A.14}$$

$$G_{33} = B^2 g_{11} + 2BC g_{12} + C^2 g_{22} \tag{A.15}$$

where

$$\frac{\partial x}{\partial X} = \frac{(X+Y+Z)-X}{(X+Y+Z)^2} = A \tag{A.16}$$

$$\frac{\partial x}{\partial Y} = \frac{\partial x}{\partial Z} = \frac{-X}{(X+Y+Z)^2} = B \tag{A.17}$$

$$\frac{\partial y}{\partial X} = \frac{\partial y}{\partial Z} = \frac{-Y}{(X+Y+Z)^2} = C \tag{A.18}$$

$$\frac{\partial y}{\partial Y} = \frac{(X+Y+Z)-Y}{(X+Y+Z)^2} = D \tag{A.19}$$

$$\frac{\partial Y}{\partial X} = \frac{\partial Y}{\partial Z} = 0, \frac{\partial Y}{\partial Y} = 1 \tag{A.20}$$

The ellipsoids' coefficients converted to XYZ color space are then used to compute the discrimination ellipsoids' volumes in the same color space. Silberstein [90] and Brown et al. [15] suggested that the radii of the discrimination ellipsoid defined by $G_{11}, G_{12}, G_{13}, G_{21}, G_{22}, G_{23}, G_{31}, G_{32}, G_{33}$ coefficients are equal to $\frac{1}{\sqrt{\sigma_i}}$ when $\sigma_i (i = 1, 2, 3)$ are the roots of the following equation:

$$\sigma^3 - G_2 \sigma^2 + G_1 \sigma - G_0 = 0 \tag{A.21}$$

where

$$G_2 = G_{11} + G_{22} + G_{33} \tag{A.22}$$

$$G_1 = G_{22} G_{33} - G_{23}^2 + G_{33} G_{11} - G_{31}^2 + G_{11} G_{22} - G_{12}^2 \tag{A.23}$$

$$G_0 = det(\Gamma), \Gamma = \begin{bmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{bmatrix} \tag{A.24}$$

To compute the ellipsoid's volume, we need to compute the product of its radii ($\frac{1}{\sqrt{\sigma_1}} * \frac{1}{\sqrt{\sigma_2}} * \frac{1}{\sqrt{\sigma_3}} = \frac{1}{\sqrt{\sigma_1 * \sigma_2 * \sigma_3}}$). This is equivalent to finding the product of the roots of the Equation A.21 ($\sigma_1 * \sigma_2 * \sigma_3$). The roots of Equation A.21, σ_i, are in fact the eigenvalues of the matrix Γ. The product of the eigenvalues of a matrix is equal to the determinant of that matrix. Therefore, rather than solving the Equation A.21, the determinant of matrix Γ can be used in the ellipsoid volume calculations:

$$det(\Gamma) = \sigma_1 \sigma_2 \sigma_3, radii = \begin{bmatrix} \frac{1}{\sqrt{\sigma_1}} & \frac{1}{\sqrt{\sigma_2}} & \frac{1}{\sqrt{\sigma_3}} \end{bmatrix} \rightarrow Ellipsoid_{vol} = \frac{4}{3}\Pi\frac{1}{\sqrt{det(G)}} \tag{A.25}$$