Metamer mismatching underlies color difference sensitivity

Brian V. Funt
School of Computing Science, Simon Fraser University, Vancouver, BC, Canada

Emitis Roshan
School of Computing Science, Simon Fraser University, Vancouver, BC, Canada

Color difference sensitivity as represented by the size of discrimination ellipsoids is known to depend on where the colors reside within color space. In the past, various color spaces and color difference formulas have been developed as parametric fits to the experimental data with the goal of establishing a color coordinate system in which equally discriminable colors are equal distances apart. These empirical models, however, provide no explanation as to why color discrimination varies in the way it does. This article considers the hypothesis that the variation in color discrimination tolerances reflects the uncertainty created by the degree of metamer mismatching for a given color. 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. To test this hypothesis, the available color discrimination data sets for surface colors are gathered and analyzed. A strong correlation between color discrimination and the degree of metamer mismatching is found. This correlation provides evidence that metamer mismatching provides an explanation as to why color discrimination varies throughout color space as it does.

Introduction

Two objects with different spectral reflectance functions are metamers under a given light if they induce the identical cone response triples. When the light is changed, their cone responses may no longer be identical to one another. If so, they are no longer metamers under the second light. In other words, they match one another under the first light but fail to match under the second light. This phenomenon is sometimes referred to as *metamerism*, but we prefer the less ambiguous term *metamer mismatching* (Logvinenko et al., 2013). This article explores the relationship between metamer mismatching and how accurately observers can distinguish pairs of very similar colors.

Chromaticity discrimination ellipses, initially measured and plotted by MacAdam (1942), describe the set of colors surrounding a given color in CIE xy-chromaticity space that are indistinguishable to the average observer. The variation in the ellipses' sizes reveals the nonuniformity 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. Subsequently, many researchers have conducted related experiments to quantify and model the color discrimination pattern in order to define more uniform color spaces.

It is common to represent color discrimination thresholds in terms of ellipses in two-dimensional chromaticity planes or ellipsoids in three-dimensional color spaces. 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. 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. The hypothesis that metamer mismatching is the underlying cause of the variation in color difference sensitivity was posited and briefly outlined before (Funt & Roshan, 2018). The present article explores the hypothesis in detail and thoroughly tests it by bringing to bear all the available experimental data on color discrimination ellipsoids. This includes a comparison to the CAM16-UCS uniform color space showing that—despite being based on a direct fit to experimental data—CAM16-UCS models color discrimination no better than predictions based solely on the metamer mismatching hypothesis.

Background

MacAdam (1942) was the first to fit ellipses for 25 color stimuli throughout the CIE 1931 chromaticity space. Silberstein and MacAdam (1945) found that the distribution of the matches in their color-matching

Citation: Funt, B. V., & Roshan, E. (2021). Metamer mismatching underlies color difference sensitivity. *Journal of Vision*, *21*(12):11, 1–11, https://doi.org/10.1167/jov.21.12.11.

https://doi.org/10.1167/jov.21.12.11 Received February 5, 2021; published November 23, 2021 ISSN 1534-7362

experiments was 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 three-dimensional color spaces can be represented by ellipsoids. Subsequently, other researchers performed similar experiments aimed at measuring the standard deviation of color matches made in different regions of color space. Silberstein (1946) devised formulas determining the coefficients and the axes of color-matching ellipsoids from the experimental data.

Brown and MacAdam (1949) followed Silberstein's method and derived the coefficients of 39 ellipsoids in one color space and presented methods for converting the ellipsoid coefficients to other color spaces. Their investigation involved a 2-degree monocular field of view and a dark background. In a subsequent study, Brown (1957) used a 10-degree field of view, binocular matching, and a white surround.

There has been a great deal of study into threshold color differences. For example, many different apparatuses using a mixture of primary lights as the test or reference colors have been used to gather the data that were later used in modifying color spaces and color difference formulas (Romero et al., 1993; Wyszecki, 1965; Wyszecki & Fielder, 1971; Yebra et al., 2001). Sharma et al. (2005) provided a data set for additional tests of the CIEDE2000 formula. CIEDE2000, introduced by the 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. (2005) revealed three sources of discontinuity in the CIEDE2000 equations. Wen (2012) 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. Wyszecki and Fielder (Wyszecki & Fielder, 1971) 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" (p. 1140).

Another group of studies has used colored surface samples. Huang et al. (2012) conducted an experiment using 466 pairs of printed samples surrounding 17 color centers to evaluate 10 color difference formulas. In another study, Luo and Rigg (1986) combined the ellipses for surface colors from different sources to produce a consistent set of ellipses. The RIT-DuPont data set (Berns et al., 1991) is based on color-coated aluminum panels and provides color difference data for 19 color centers. Melgosa et al. (1997) subsequently derived a set of color difference ellipsoids from the RIT-DuPont data. Witt (1987) and Cheung and Rigg (1986) reported color discrimination ellipsoids for four and five CIE reference color centers, respectively, using printed samples and dyed wools.

Pridmore and Melgosa (2005) analyzed the ellipse area and dimensions from four different data sets and observed that the principal semiaxes of the discrimination ellipses increased at lower luminance levels. Luo et al. (2006) employed a combination of different data sets based on surface color samples and CRT colors to test the performance of CIECAM02. 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. (2006) introduced three modified versions of CIECAM02 known as CAM02-SCD, CAM02-LCD, and CAM02-UCS by fits, respectively, to three data sets: the SCD (small color difference) data set (Luo et al., 2001) of 3,657 color pairs from different 4 data sets, the LCD (large color difference) data set (Zhu et al., 2002) of 2,954 color pairs from 6 different data sets, and the combined SCD plus LCD data set. While all these modifications have the same structure as the original CIECAM02 version, Li et al. (2016) and Li et al. (2017) derived a new color appearance model named CAM16 by performing both chromatic and luminance adaptation in the same space rather than in two different color spaces. CAM16 avoids the unexpected problems that sometime occur with the CIECAM02 lightness computation and at the same time outperforms it in predicting the corresponding color data sets and color appearance data sets.

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 and Kortick (1994) and Stockman and Brainard (2010) build on 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. These are parallel, but fundamentally different, questions. However, progress in answering one can be expected to be useful to answering the other.

Hypothesis

Metamer mismatching arises from the fact that normal trichromatic color vision is based on only three 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.

In general, given the color signal of an object with respect to a given illuminant, many other surface reflectances can lead to the same color signal with respect to the same illuminant. The color signals of these metamers will disperse into many different color signals under a second, different illuminant. The set of all such color signals forms a convex body referred to as the metamer mismatch volume (Logvinenko et al., 2013) or metamer mismatch body (MMB) (Zhang et al., 2016). The study by Zhang et al. (2016) on metamer mismatching showed that it is most severe for gray and least severe for highly saturated colors (see Figure 1). Our hypothesis is that 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.

The volume of the MMB (i.e., the volume of the set of all possible color signals that can arise under the second light given the color signal under the first light) for a given color signal is a measure of the possible variability in the nature of the underlying physical reflectance. 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-gray, where the entire range of the visible spectrum

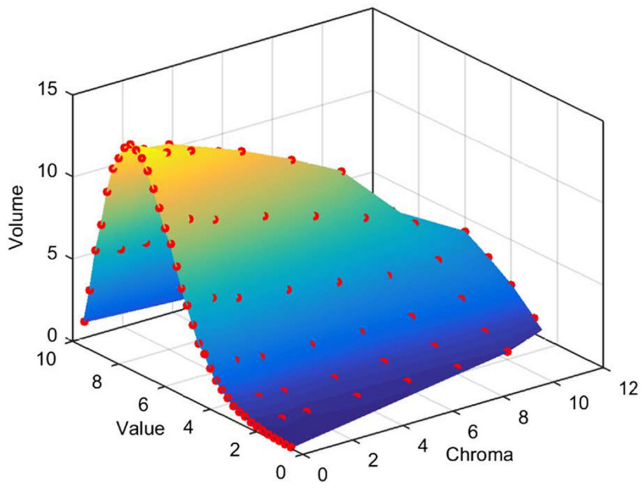

Figure 1. Plot of MMB volume averaged over all hues showing how the MMB volume decreases with distance in Munsell value and/or chroma from gray (value 6, chroma 0).

is likely to be involved, than there are to a saturated red, for example, where mainly the long-wave portion of the spectrum is involved. Color discrimination experiments typically use a fixed illumination condition, a rather artificial circumstance. However, for an observer wishing to identify a given physical surface by its color under typical but varied lighting conditions, it is therefore more important to distinguish the shades of a grey 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. The set of all color signals that a set of sensors can produce under a given light forms a convex hull in R^n, which is known as the object color solid (OCS). In contrast to the gray, which is at the center of the OCS, for any color signal on the boundary of the OCS, there is only one possible reflectance creating it, so the volume of the MMB drops to zero for such color signals. This is illustrated by the plot in Figure 1.

Metamer mismatching is usually discussed from the point of view of two physically distinct surfaces that appear identical (i.e., match in the sense of creating identical cone response triples) under some first illuminant become distinguishable (i.e., mismatch) under some second illuminant. From the reverse perspective, however, the degree of metamer mismatching can be considered the likelihood that two different color signals (i.e., cone response triples) corresponding to two different reflectances under one illuminant could become indistinguishable under some other illuminant. Why might this matter? Consider, for example, the green leaves of two different plants, one poisonous, the other not, that appear different at noon but become almost indistinguishable at sunset. Clearly, it is important to be able to distinguish them at all

times of day. 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 reflectance, S. For example, suppose that under D65, S has color $\boldsymbol{C}_{D65}$. The MMB of $\boldsymbol{C}_{D65}$ 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 $\boldsymbol{C}_A$ in that MMB is a candidate for matching $\boldsymbol{C}_{D65}$ under D65. Now suppose that we observe $\boldsymbol{C}_A$; does it correspond to S? The answer is "almost certainly not" because $\boldsymbol{C}_A$ 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 of uncertainty. When inaccuracy in matching a specific color is added, then the uncertainty increases. For example, suppose that under D65, color $\boldsymbol{C}'_{D65}$ is similar enough to $\boldsymbol{C}_{D65}$ that it matches $\boldsymbol{C}_{D65}$. The uncertainty then becomes the union of the MMB of $\boldsymbol{C}_{D65}$ and the MMB of $\boldsymbol{C}'_{D65}$ along with the MMBs of all colors in between. In other words, it is the set of all colors that either $\boldsymbol{C}_{D65}$ or $\boldsymbol{C}'_{D65}$ (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 2 shows the trend from gray to both 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 somewhat larger than the volume of the MMB for flat-spectrum gray 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 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.

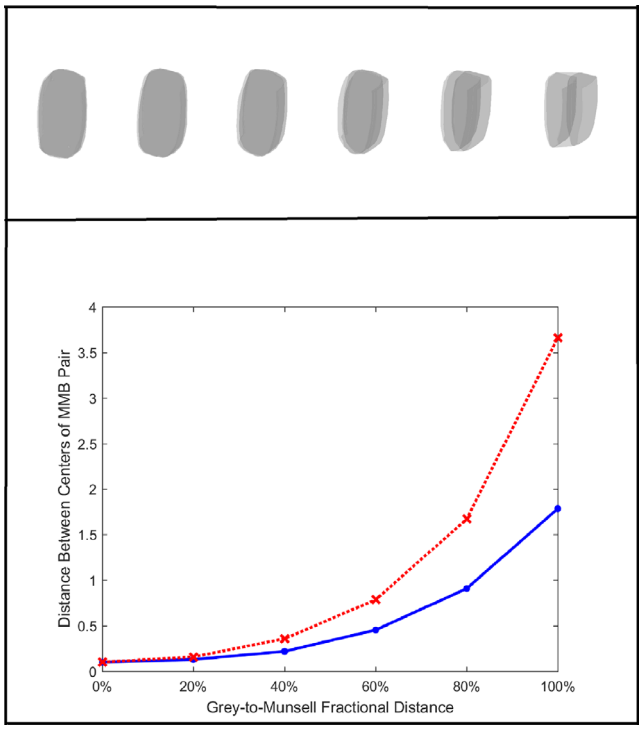

Figure 2. Example illustrating how the distance between two points in XYZ needs to increase in order to keep the volume of the convex hull of their MMBs equal to a constant. (Top) The two-dimensional projection of the MMB pairs showing their fixed-volume convex hulls containing MMBs that get progressively smaller as the color is moved from gray 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 gray (i.e., uniform 50% spectral reflectance) to blue (solid blue curve) corresponding to the MMBs (top) and, in addition, gray to Munsell red 5R 5/8 (dashed red curve).

Color discrimination data sets and statistical analysis

Four color discrimination data sets of Cheung and Rigg (1986), Witt (1987), Huang et al. (2012), and Melgosa et al. (1997) report the parameters of the color discrimination ellipsoids, not just ellipses, and are therefore useful for evaluating the metamer mismatching hypothesis. Since the data sets use a variety of different color spaces, they are first converted to a common color space and their ellipsoid coefficients updated correspondingly. Details of the conversion are provided in the Appendix.

Data set	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

Table 1. Summary of the four data sets used in our study.

The details of the data sets are as follows. Cheung and Rigg (1986) 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 of Gray, Red, Yellow, Green, and Blue (Robertson, 1978) 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, which we convert to CIE XYZ color space.

Witt (1987) used painted samples around four of the five CIE reference color centers: Yellow, Red, Blue, and Gray. Observers were asked if the color difference was perceptible in the sample pairs or not. The coefficients of the fitted ellipsoids are reported in xyY color space.

Huang et al. (2012) prepared 446 pairs of printed color patches surrounding 17 color centers for a grayscale 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). 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 data sets.

Berns et al. (1991) prepared a gray anchor pair with a color difference of 1.02 $\Delta \boldsymbol{E}^*_{ab}$ units (CIELAB color space). They asked the observers to compare the magnitude of the color difference of the sample pair 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. (1997) then used the 156 median tolerances reported in the RIT DuPont data set (Berns et al., 1991) to compute the ellipsoid parameters in x, y, Y/100 color space. For our purposes, we convert the Melgosa ellipsoids' coefficients to XYZ color space. The four data sets explained above are summarized in Table 1.

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 data sets converted to XYZ color space are used to compute the volumes of the color discrimination ellipsoids, $\boldsymbol{E}_{vol}$, in XYZ, as

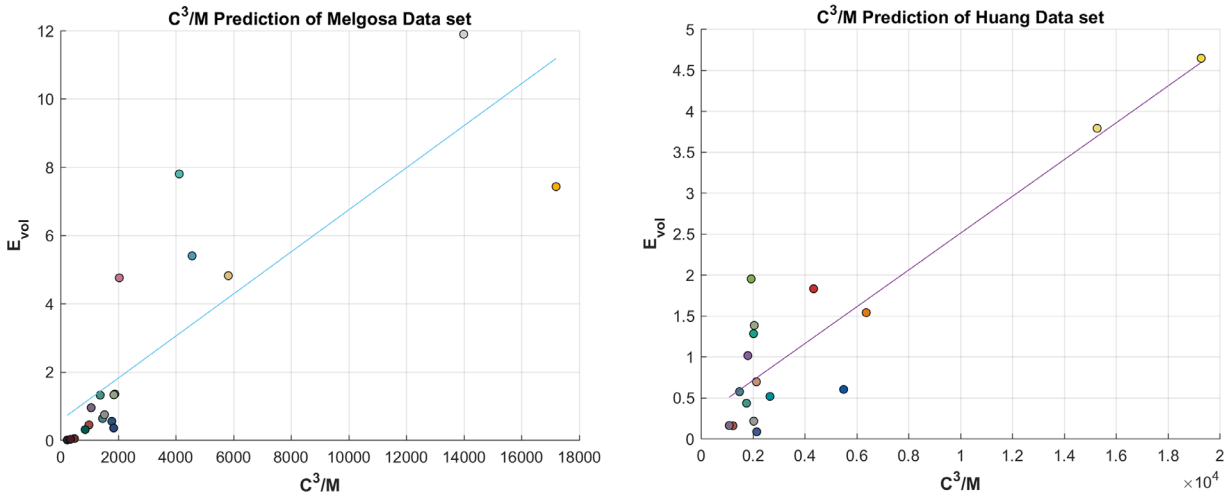

Figure 3. Plots of the volumes, $\boldsymbol{E}_{vol}$, of the color discrimination ellipsoids in XYZ space as a function of the inverse of the normalized volume $\boldsymbol{M}$ of the corresponding metamer mismatch bodies (i.e., $\boldsymbol{C}^3/\boldsymbol{M}$) for the two color discrimination data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

described in the Appendix. 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. (2013). $\boldsymbol{M}$ is then normalized by $\boldsymbol{C}^3$, the cube of the Euclidean distance, $\boldsymbol{C}$, 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 $\boldsymbol{E_{vol}}$ and $\boldsymbol{C}^3/\boldsymbol{M}$. Note that $\boldsymbol{C}^3/\boldsymbol{M}$ 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 $\boldsymbol{C}^3/\boldsymbol{M}$ for the Melgosa (i.e., 19 color centers based on the Berns et al. [1991] data) and Huang (17 color centers) data sets are shown in Figure 3. The figure caption includes the Pearson correlation coefficient (r), mean jackknife estimate of r, bias, and standard error (SE) 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 are as follows: $r = 0.83$ with a mean jackknife estimate of $r = 0.83$, bias $= 0.03$, and $SE = 0.13$ for the Melgosa data set; and $r = 0.9$ with a mean jackknife estimate of $r = 0.9$, bias $= -0.05$, and $SE = 0.11$ for the Huang data set. The correlation between E_{vol} and $\boldsymbol{C}^3/\boldsymbol{M}$ includes the nonlinearity of the inverse $1/\boldsymbol{M}$. The corresponding, simpler (negative) linear correlation results between E_{vol} and $\boldsymbol{M}/\boldsymbol{C}^3$ are significantly weaker: -0.52 and -0.7, respectively.

Merging the data sets

In the previous section, the statistical analysis is conducted separately for the two color discrimination data sets that contain at least 17 color centers each. Two of the other data sets include discrimination ellipsoids for only four color centers, in one case, and five in the other. The goal of this section is to combine the data from all four different data sets into one larger data set. The difficulty in doing so is that the data sets are all based on different experimental protocols resulting in different scales. We combine the ellipsoid measurements from the various data sets following the basic strategy that Luo and Rigg (1986) used when combining ellipse data.

Luo and Rigg (1986) 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 data sets (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 $\bar{\mathbf{R}}$) as a scaling factor for each ellipse and the mean of $\bar{\mathbf{R}}$ values for each group (referred to as $\mathbf{S}$) to adjust the ellipses onto a common scale. They showed that adjusting the individual ellipses with $\bar{\mathbf{R}}$ values results in a more consistent plot than using the group mean for each data set.

	Cheung et al.			Melgosa et al.		
Color	x	y	Y	x	y	Y
Gray	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

Table 2. xyY coordinates of the common color centers in the Cheung and Melgosa data sets.

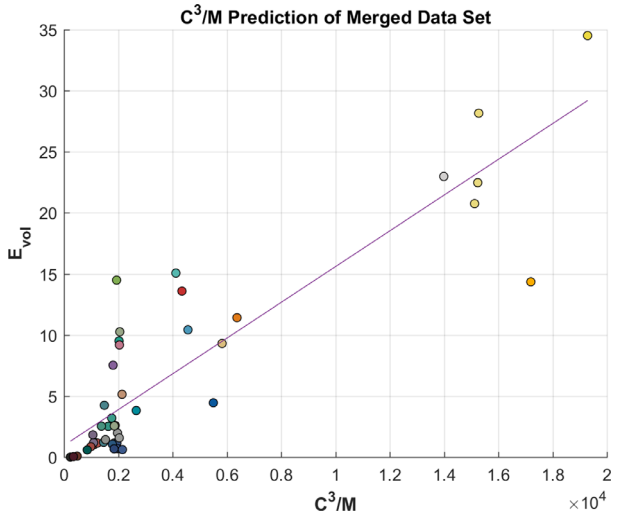

Figure 4. Linear fit of the ellipsoid volume ($\boldsymbol{E_{vol}}$) versus the inverse of the normalized MMB volume (i.e, $\boldsymbol{C}^3/\boldsymbol{M}$) for the merged data set of 45 color centers, with the null hypothesis again rejected at the 5% significance level with p value 7e-16, $r = 0.88$, mean jackknife estimate of $r = 0.88$, bias $= -0.01$, $SE = 0.06$.

Dataset	Inverse of normalized MMB volume	CAM16-UCS
Melgosa data set	0.83	0.84
Huang data set	0.9	0.89
Merged four data sets	0.88	0.87

Table 3. Correlation coefficients between the experimental ellipsoids and (a) the ellipsoid volumes predicted by the inverse of the normalized MMB volume and (b) the ellipsoid volumes predicted by CAM16-UCS.

Following the basic approach of Luo and Rigg (1986), but modified for ellipsoids rather than ellipses, we use scaling factors to bring the different data sets onto a common scale. To combine different color discrimination data sets, we use the ratio of the ellipsoid volumes for the color centers that the data sets have in common. For each pair of data sets, 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 data sets 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. data sets are listed in Table 2.

Figure 4 shows a linear fit for the data in the merged data set derived from the Cheung, Melgosa, Witt, and Huang data sets. The null hypothesis is again rejected at the 5% significance level with $r = 0.88$ and a p value of 7e-16.

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 support 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

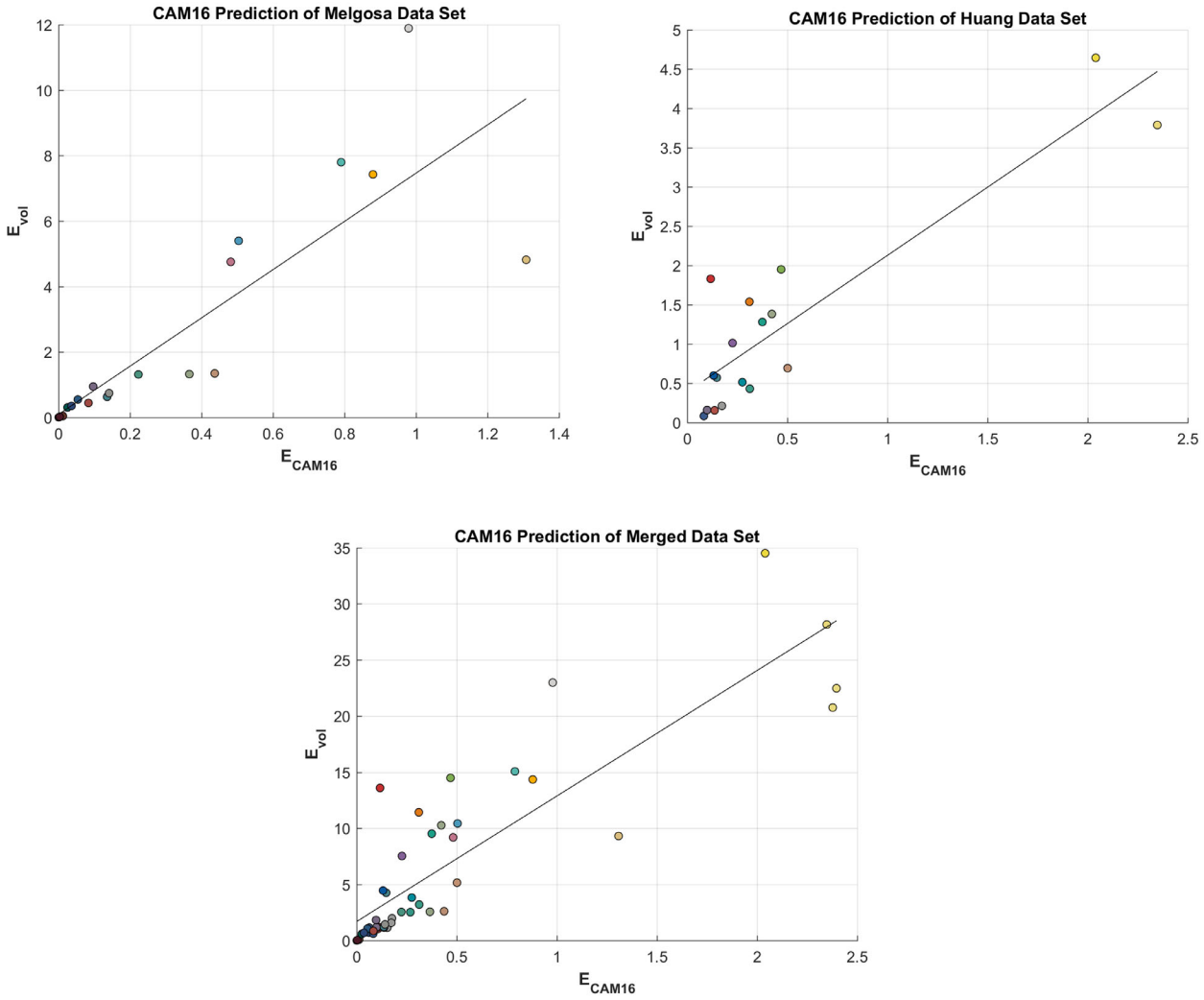

Figure 5. Plots of the ellipsoid volumes, E_{vol}, of the experimental color discrimination ellipsoids in XYZ space versus the ellipsoid volumes, E_{CAM16}, of the unit CAM16-UCS ΔE spheres, for both the individual and combined color discrimination data sets. The statistics for the linear fits (summarized in Table 3) are very similar to those in Figures 3 and 4. (Upper left) CAM16 prediction of Melgosa data set, $r = 0.84$, mean jackknife estimate of $r = 0.85$, bias $= 0.04$, $SE = 0.11$. (Upper right) CAM16 prediction of Huang data set, $r = 0.89$, mean jackknife estimate of $r = 0.89$, bias $= -0.014$, $SE = 0.07$. (Bottom) CAM16 prediction of merged data set, $r = 0.87$, mean jackknife estimate of $r = 0.87$, bias $= 0.0001$, $SE = 0.04$.

color, the inverse of the normalized MMB volume predicts the volume of that color's discrimination ellipsoid.

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 the 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 $\Delta \boldsymbol{E}$ in CAM16-UCS around each of the color centers included in the four data sets, convert them to XYZ coordinates, and fit the volumes of the resulting ellipsoids in XYZ to those of the color discrimination ellipsoids. Figure 5 shows the Pearson correlation coefficient (r) between the unit $\Delta \boldsymbol{E}$ spheres in CAM16-UCS color space (ellipsoids in XYZ space) and the discrimination ellipsoid volumes reported in the data sets, the mean jackknife estimate of r, and the bias and standard error (SE) in jackknife replicates.

The CAM16-UCS fits can be compared to those shown in Figures 3 and 4 obtained using the inverse of normalized MMB volume. The results are also summarized in Table 3. 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 (Berns, Alman, Reniff, Snyder & Balonon-Rosen, 1991) data set (i.e., the same experimental data as the Melgosa data set), 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 the psychophysical data.

Discussion and conclusion

The results shown in Figures 3 and 4 indicate a strong correlation between color discrimination and metamer mismatching. In particular, as the extent of metamer mismatching increases, color discrimination thresholds decrease. Zhang et al. (2016) showed that metamer mismatching is most significant for ideal gray, is 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 strong correlation supports the hypothesis that the uncertainty created by metamer mismatching underlies color discrimination thresholds since it correctly predicts that color discrimination is finest near gray and becomes coarser and coarser for more and more saturated colors.

Four sets of experimental data measuring color discrimination ellipsoids are available for testing. The fits shown in Figures 3 and 4 are not perfect, but they do indicate the hypothesized relationship. Furthermore, as the results in Table 3 and Figure 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.

This article has explored the hypothesis that the need for the visual system to overcome the uncertainty (as illustrated by Figure 2) due to metamer mismatching is the reason why there is more precise discrimination between colors in some regions of color space than others. The strong correlation found between the experimental data and the inverse of the metamer mismatch body volumes, while not proof, is an evidence supporting the idea that metamer mismatching provides an explanation for as to why color discrimination varies in the way it does.

Keywords: color discrimination, metamer mismatching, discrimination ellipsoid

Acknowledgments

The authors thank to the reviewers who provided especially insightful criticisms and valuable suggestions.

Supported by the Natural Sciences and Engineering Research Council of Canada.

Commercial relationships: none.
Corresponding author: Brian V. Funt.
Email: funt@sfu.ca.
Address: School of Computing Science, Simon Fraser University, Burnaby, B.C. V5A 1S6, Canada.

References

Berns, R. S., Alman, D. H., Reniff, L., Snyder, G. D., & Balonon-Rosen, M. R. (1991). Visual determination of suprathreshold color-difference tolerances using probit analysis. *Color Research & Application, 16*(5), 297–316.

Brown, W. R. (1957). Color discrimination of twelve observers. *Journal of the Optical Society of America, 47*(2), 137–143.

Brown, W. R. J., & MacAdam, D. L. (1949). Visual sensitivities to combined chromaticity and luminance differences. *Journal of the Optical Society of America, 39*(10), 808–834.

Changjun, L., Li, Z., Wang, Z., Xu, Y., Luo, M., Cui, G., Melgosa, M., Brill, M., . . . Pointer, M. (2017). Comprehensive color solutions: CAM16, CAT16,

and CAM16-UCS. *Color Research & Application,* 42(6), 703–718, doi:10.1002/col.22131.

Changjun, L., Li, Z., Wang, Z., Xu, Y., Luo, M., Cui, G., Melgosa, M., . . . Pointer, M. (2016). A revision of CIECAM02 and its CAT and UCS. *Color and Imaging Conference, 2016*(1), 208–212, Society for Imaging Science and Technology, doi:10.2352/ISSN.2169-2629.2017.32.208.

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

Eskew, R. T., & Kortick, P. M. (1994). Hue equilibria compared with chromatic detection in 3D cone contrast space. *Investigative Ophthalmology & Visual Science, 35*(4), 1555.

Funt, B., & Roshan, R. (2018). Color discrimination ellipses explained by metamer mismatching. In *Proceedings of AIC 2018 International Colour Association Conference,* Lisbon.

Huang, M., Liu, H., Cui, G., & Luo, M. R. (2012). Testing uniform colour spaces and colour-difference formulae using printed samples. *Color Research & Application, 37*(5), 326–335.

Logvinenko, A. D., Funt, B., & Godau, C. (2013). Metamer mismatching. *IEEE Transactions on Image Processing, 23*(1), 34–43.

Luo, M. R., Cui, G., & Li, C. (2006). Uniform colour spaces based on CIECAM02 colour appearance model. *Color Research & Application, 31*(4), 320–330.

Luo, M. R, Cui, G., & Rigg, B, (2001). The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research and Application, 26*(5), 340–350.

Luo, M. R., & Rigg, B. (1986). Chromaticity-discrimination ellipses for surface colours. *Color Research & Application, 11*(1), 25–42.

MacAdam, D. L. (1942). Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America, 32*(5), 247–274.

Melgosa, M., Hita, E., Poza, A., Alman, D. H., & Berns, R. S. (1997). Suprathreshold color-difference ellipsoids for surface colors. *Color Research & Application, 22*(3), 148–155.

Müller, G. E. (1930). Über die Farbenempfindungen. Zeitschrift für Psychologie und Physiologie der Sinnesorgane [On colour sensations]. *Ergänzungsband, 17*, 1–430.

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

Robertson, A. R. (1978). Guidelines for coordinated research on colour difference equations. *Color Research & Application, 3*, 149–151.

Romero, J., García, J. A., Barco, L. J., & Hita, E. (1993). Evaluation of color-discrimination ellipsoids in two-color spaces. *Journal of the Optical Society of America, 10*(5), 827–837.

Sharma, G., Wu, W., & Dalal, E. N. (2005). The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application, 30*(1), 21–30.

Silberstein, L. (1946). On two accessories of three-dimensional colorimetry I. the probable error of colorimetric tensor components as derived from a number of color matchings II. the determination of the principal colorimetric axes at any point of the color threefold. *Journal of the Optical Society of America, 36*(8), 464–468.

Silberstein, L., & MacAdam, D. L. (1945). The distribution of color matchings around a color center. *Journal of the Optical Society of America, 35*(1), 32–39.

Smet, K. A., Webster, M. A., & Whitehead, L. A. (2016). A simple principled approach for modeling and understanding uniform color metrics. *Journal of the Optical Society of America, 33*(3), A319–A331.

Stockman, A., & Brainard, D. H. (2010). Color vision mechanisms. In M. Bass (Ed.), *OSA Handbook of Optics: Vol. 3. Vision and Vision Optics* (Third ed.). (pp. 11.1–11.104). New York: McGraw-Hill, Inc.

Wen, S. (2012). A color difference metric based on the chromaticity discrimination ellipses. *Optics Express, 20*(24), 26441–26447.

Witt, K. (1987). 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.

Wyszecki, G. (1965). Matching color differences. *Journal of the Optical Society of America, 55*(10), 1319–1324.

Wyszecki, G., & Fielder, G. (1971). New color-matching ellipses. *Journal of the Optical Society of America, 61*(9), 1135–1152.

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

Zhang, X., Funt, B., & Mirzaei, H. (2016). Metamer mismatching in practice versus theory. *Journal of the Optical Society of America, 33*(3), A238–A247.

Zhu, S. Y., Luo, M. R., Cui, G. H., & Rigg, B. (2002) Comparing different colour discrimination data sets. In *Color and Imaging Conference, 2002*(1), 51–54, Society for Imaging Science and Technology.

Appendix

The conversion of ellipsoid parameters from one color space to another is as follows. In general, given $\mathbf{g_{ik}}$, the ellipsoid parameters in one coordinate system, such as the instrument's RGB, the following relation recommended by Brown and MacAdam (1949) can be used to convert them to another color space such as XYZ:

$$G_{jl} = \left(\partial x_i/\partial x_j\right)\left(\partial x_k/\partial x_l\right)\ .\ g_{ik};\quad i,j,k,l = 1,2,3 \quad (1)$$

where

$$x_i,\ x_k = R,\ G, B \text{ and } x_j, x_l = X, Y,\ Z \quad \text{(i.e., } x_1 = R;\ x_2 = G, \text{ etc.)}$$

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

$$R = a_{11}X + a_{12}Y + a_{13}Z \quad (2)$$

$$G = a_{21}X + a_{22}Y + a_{23}Z \quad (3)$$

$$B = a_{31}X + a_{32}Y + a_{33}Z \quad (4)$$

For instance,

$$\partial R/\partial X = a_{11};\quad \partial R/\partial Y = a_{12} \quad (5)$$

As an example, writing one equation in full gives

$$\begin{aligned} G_{11} = {} & (\partial R/\partial X)(\partial R/\partial X)g_{11} + (\partial R/\partial X)(\partial G/\partial X)g_{12} \\ & + (\partial R/\partial X)(\partial B/\partial X)g_{13} + (\partial G/\partial X)(\partial R/\partial X)g_{21} \\ & + (\partial G/\partial X)(\partial G/\partial X)g_{22} + (\partial G/\partial X)(\partial B/\partial X)g_{23} \\ & + (\partial B/\partial X)(\partial R/\partial X)g_{31} + (\partial B/\partial X)(\partial G/\partial X)g_{32} \\ & + (\partial B/\partial X)(\partial B/\partial X)g_{33} \end{aligned} \quad (6)$$

Cheung and Rigg (1986) and Witt (1987) report the ellipsoid parameters in xyY space so in our case, R, G, B in Equation 2 to Equation 4 should be replaced with x, y, Y. The transformation between xyY and XYZ is given by

$$x = X/(X + Y + Z) \quad (7)$$

$$y = Y/(X + Y + Z) \quad (8)$$

$$Y = Y \quad (9)$$

Proceeding similarly to the example above, we find that

$$G_{11} = A^2 g_{11} + 2ACg_{12} + C^2 g_{22} \quad (10)$$

$$G_{12} = ABg_{11} + (AD + BC)\,g_{12} + Ag_{13} + CDg_{22} + Cg_{23} \quad (11)$$

$$G_{13} = ABg_{11} + (AC + BC)\,g_{12} + C^2 g_{22} \quad (12)$$

$$G_{22} = B^2 g_{11} + 2BDg_{12} + 2Bg_{13} + D^2 g_{22} + 2Dg_{23} + g_{33} \quad (13)$$

$$G_{23} = B^2 g_{11} + (BC + BD)\,g_{12} + DCg_{22} + Bg_{13} + Cg_{23} \quad (14)$$

$$G_{33} = B^2 g_{11} + 2BCg_{12} + C^2 g_{22} \quad (15)$$

where

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

$$\partial x/\partial Y = \partial x/\partial Z = \frac{-X}{(X + Y + Z)^2} = B \quad (17)$$

$$\partial y/\partial X = \partial y/\partial Z = \frac{-Y}{(X + Y + Z)^2} = C \quad (18)$$

$$\partial y/\partial Y = \frac{(X + Y + Z) - Y}{(X + Y + Z)^2} = D \quad (19)$$

$$\partial Y/\partial X = \partial Y/\partial Z = 0,\quad \partial Y/\partial Y = 1 \quad (20)$$

Silberstein (1946) and Brown and MacAdam (1949) suggested that the radii of the discrimination ellipsoid defined by $G_{11}, G_{12}, G_{13}, G_{22}, G_{23}, G_{33}$ coefficients are equal to $\frac{1}{\sqrt{\sigma_i}}$ when $\boldsymbol{\sigma_i}$ $(\mathbf{i} = 1, 2, 3)$ are the roots of the following equation:

$$\sigma^3 - G_2\sigma^2 + G_1\sigma - G_0 = 0 \quad (21)$$

where

$$G_2 = G_{11} + G_{22} + G_{33} \quad (22)$$

$$G_1 = G_{22}G_{33} - G_{23}^2 + G_{33}G_{11} - G_{31}^2 + G_{11}G_{22} - G_{12}^2 \quad (23)$$

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

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 Equation 21 ($\sigma_1 * \sigma_2 * \sigma_3$). The roots of Equation 21 ($\boldsymbol{\sigma}_\mathbf{i}$) 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 Equation 21, the determinant of matrix Γ can be used in the ellipsoid volume calculations:

$$\det(\Gamma) = \sigma_1\sigma_2\sigma_3, \ radii = \left[\frac{1}{\sqrt{\sigma_1}}, \frac{1}{\sqrt{\sigma_2}}, \frac{1}{\sqrt{\sigma_3}}\right] \overset{yields}{\rightarrow} Ellipsoid_{vol} = \frac{4}{3}\pi \frac{1}{\sqrt{det(\Gamma)}}.$$