Using suprathreshold color-difference ellipsoids to estimate any perceptual color-difference

Samuel Morillas*[a], Mark D. Fairchild[b]

[a]*Instituto Universitario de Matemática Pura y Aplicada, Universitat Politècnica de València, Camino de Vera s/n 46022 Valencia, Spain*
[b]*Munsell Color Science Laboratory, Rochester Institute of Technology, 1 Lomb Memorial Drive, Rochester, NY 14623 (USA)*

Abstract

Relating instrumentally measured to visually perceived colour-differences is one of the challenges of advanced colorimetry. Lately, the use of color difference formulas is becoming more important in the computer vision field as it is a key tool in advancing towards perceptual image processing and understanding. In the last decades, the study of contours of equal color-differences around certain color centers has been of special interest. In particular, the contour of threshold level difference that determines the just noticeable differences (JND) has been deeply studied and, as a result, a set of 19 different ellipsoids of suprathreshold color-difference is available in the literature. In this paper we study whether this set of ellipsoids could be used to compute any color difference in any region of the color space. To do so, we develop a fuzzy multi-ellipsoid model using the ellipsoids information along with two different metrics. We see that the performance of the two metrics vary significantly for very small, small, medium and large color differences. Therefore, we also study how to adapt two metric parameters to optimize performance. The obtained results outperform the currently CIE-recommended color-difference formula CIEDE2000.

*Corresponding author e-mail: smorillas@mat.upv.es
*S. Morillas acknowledges the support of grants PRX16/00050 and PRX17/00384 (Ministerio de Educación, Cultura y Deporte) and MTM2015-64373-P (MINECO/FEDER, UE). The authors thank Dr. Manuel Melgosa, Dr. Luis Gómez-Robledo, Dr. Esther Sanabria-Codesal, Dr. Francisco Montserrat and Mr. Fu Jiang for providing useful materials, information and suggestions.

Keywords: Color-difference formulas, Fuzzy logic, Fuzzy Metrics, STRESS

1. Introduction

Improved correlation between visually perceived (ΔV) and instrumentally measured (ΔE) colour differences under specific illuminating and viewing conditions is an important problem in modern colorimetry. This topic is gaining more and more attention each day in the computer vision field as many recent image processing and computer vision techniques are using colour difference formulas when addressing perceptual processing and understanding of digital images [1]-[3]. For instance, advanced models of perceptual similarity of color images iCAM [1] or S-CIELAB [2] try to represent the Human Visual Systems mechanisms and are based on using appropriate *Constrast Sensitivity Functions* (CSFs) to remove all details in the images that cannot be perceived as well as Color-difference formulas to perceptually characterize the differences which are indeed observed, which has been found to be appropriate in general. In turn, perceptual similarity measures assess how good are filtering methods, compression algorithms or demosicing methods from a perceptual point of view [3].

Research on classical and modern datasets has shown that constant visual differences (ΔV) with respect to a given colour centre do not correspond to constant computed colour-differences (ΔE) in a colour space [4, 5]. Traditionally, points with a constant visual difference with respect to a fixed colour centre are considered to be placed on the surface of an ellipsoid in a given colour space, but the orientation, shape, and size of this ellipsoid change with the fixed colour centre [6, 7, 8, 9]. In short, to date we do not have a uniform colour space that is well related to visual perception. The CIELUV and CIELAB colour spaces, recommended by the CIE in 1976 [11], as well as other recent colour spaces [12, 13] are only approximately uniform.

However, we do have very precise information about threshold color-differences in different regions of the CIELAB color space. In the work [14], 156 tolerances each of equal visual color difference, ΔV, were found around 19 different color centers in the CIELAB color space. Later, in [8], these tolerances where used to derive 19 ellipsoids around the 19 color centers studied. Moreover, these ellipsoids were further tested in [9] and they showed a precise performance. This implies that from each ellipsoid we may derive a local color-difference formula that performs accurately for local threshold differences.

Besides, it is reasonable to assume that colour discrimination in a colour space changes in a smooth and regular way. Thus, for example, experimental colour discrimination ellipses reported in previous experiments [6, 8, 10], in each case follow a quite regular pattern in the CIE x,y chromaticity diagram, although relevant differences (attributable to different parametric factors such as viewing modes, sizes of colour-differences, etc.) may be noted when comparing ellipses from different experiments.

In this paper we study if the information in the set of 19 ellipsoids in [8] can be used to build a general color difference formula just by combining the local difference formulas derived from the ellipsoids. Such a model is based on two assumptions: (i) threshold level color differences not close to any color center may be interpolated from the color difference formulas of nearest color centers; (ii) larger and smaller color differences may be estimated by direct linear scaling of the threshold level difference formula. These two assumptions are described using vague terms which leads us to propose to use a *fuzzy logic*-based approach. Fuzzy logic [15] has been successfully used in many areas of science and engineering [16] including the analysis of color difference datasets [17, 18] as well as for color naming [19].

In so doing, in Section 2 we derive local color difference formulas from the ellipsoid information using two different metrics. The fuzzy model for combining the local difference formulas is described in Section 3. Experimental results using the color difference dataset used at developing the CIEDE2000 color difference formula [20] are provided in Section 4 where we also study how the metric parameters that define size and shape of the ellipsoids can be adapted to improve performance. Finally, conclusions are drawn in Section 5.

2. Local color-difference formulas related to ellipsoids of supra-threshold differences

2.1. Notation

Let us first introduce some notation on color differences as follows:

Experimental observations of color differences are usually given either as datasets or as single observations. Let us denote by $\mathcal{S}$, a dataset consisting of a number of color pairs, denoted as $\mathbf{S}_i$, each of them representing the perceptual colour difference between two colour samples. Each color pair in $\mathcal{S}$ is in turn represented as a set $\mathbf{S}_i = \{\mathbf{A}_i, \mathbf{B}_i, \Delta V_i\}$, where $\mathbf{A}_i$ and $\mathbf{B}_i$ denote the CIELAB coordinates of the two colour samples given by $\mathbf{A}_i = \{a_i^{*1}, b_i^{*1}, L_i^{*1}\}$ and $\mathbf{B}_i = \{a_i^{*2}, b_i^{*2}, L_i^{*2}\}$, and ΔV_i is the perceptual difference between $\mathbf{A}_i$ and $\mathbf{B}_i$. For convenience, let us denote by $\mathbf{C}_i$ the mean point between $\mathbf{A}_i$ and $\mathbf{B}_i$ given by $\mathbf{C}_i = (\mathbf{A}_i + \mathbf{B}_i)/2$, and $\mathbf{D}_i = \mathbf{A}_i - \mathbf{B}_i = \{\Delta a_i^*, \Delta b_i^*, \Delta L_i^*\}$ the difference vector between $\mathbf{A}_i$ and $\mathbf{B}_i$.

Besides we will denote each ellipsoid by $\mathcal{E}_j$ where $j = 1, ..., 19$ are the 19 ellipsoids computed in [8]. The center of ellipsoid j will be denoted by $\mathbf{O}_{\mathcal{E}_\mathbf{j}} = \{a^*_{\mathcal{E}_j}, b^*_{\mathcal{E}_j}, L^*_{\mathcal{E}_j}\}$.

2.2. Classical ellipsoid metric

Thus, according to [8] Eq. (2), we can compute the local color difference between the samples in $\mathbf{S}_i$ according to the ellipsoid $\mathcal{E}_j$, denoted by $\Delta E_i^{\mathcal{E}_j}$, as

$$(\Delta E_i^{\mathcal{E}_j})^2 = \mathcal{E}_j^{11}(\Delta a_i^*)^2 + 2\mathcal{E}_j^{12}(\Delta a_i^*)(\Delta b_i^*) + 2\mathcal{E}_j^{13}(\Delta a_i^*)(\Delta L_i^*) + \mathcal{E}_j^{22}(\Delta b_i^*)^2 + 2\mathcal{E}_j^{23}(\Delta b_i^*)(\Delta L_i^*) + \mathcal{E}_j^{33}(\Delta L_i^*)^2, \quad (1)$$

which, noted in matrix expression, is equivalent to obtaining $(\Delta E_i^{\mathcal{E}_j})^2$ as the matrix product $(\Delta E_i^{\mathcal{E}_j})^2 = \mathbf{D}_i \mathbf{M}_{\mathcal{E}_j} \mathbf{D}_i^T$, where

$$\mathbf{M}_{\mathcal{E}_j} = \begin{bmatrix} \mathcal{E}_j^{11} & \mathcal{E}_j^{12} & \mathcal{E}_j^{13} \\ \mathcal{E}_j^{12} & \mathcal{E}_j^{22} & \mathcal{E}_j^{23} \\ \mathcal{E}_j^{13} & \mathcal{E}_j^{23} & \mathcal{E}_j^{33} \end{bmatrix} \quad (2)$$

is the symmetric scalar product matrix associated to $\mathcal{E}_j$ and ^{T} denotes transpose. The coefficients in $\mathcal{E}_j$ are obtained by fitting to experimental data [6, 7, 8] so that the points at equal distance generate an ellipsoid analogous to the one in Figure 1. It is interesting to note that coefficients $\mathcal{E}_j^{11}, \mathcal{E}_j^{22}, \mathcal{E}_j^{33}$ are related to linear independency of color coordinates whereas $\mathcal{E}_j^{12}, \mathcal{E}_j^{13}, \mathcal{E}_j^{23}$ are related to correlation among the coordinates and geometrically describe the rotation of the ellipsoid with respect to the reference system of the color space.

From another point of view we can also consider that an ellipsoid $\mathcal{E}_j$ is characterized by its center $\mathbf{O}_{\mathcal{E}_\mathbf{j}}$, an ortonormal reference system given by three unitary ortogonal vectors $\mathbf{U}^1_{\mathcal{E}_j}, \mathbf{U}^2_{\mathcal{E}_j}, \mathbf{U}^3_{\mathcal{E}_j}$, and the length of the semi-axis of the ellipsoid in the direction of each unitary vector denoted respectively by $\mathcal{L}^1_{\mathcal{E}_j}, \mathcal{L}^2_{\mathcal{E}_j}, \mathcal{L}^3_{\mathcal{E}_j}$.

Both $\mathbf{U}^k_{\mathcal{E}_j}, k = 1, 2, 3$ and $\mathcal{L}^k_{\mathcal{E}_j}, k = 1, 2, 3$ can be obtained from $\mathbf{M}_{\mathcal{E}_j}$ by diagonalization. $\mathbf{U}^k_{\mathcal{E}_j}, k = 1, 2, 3$ are the eigenvectors of $\mathbf{M}_{\mathcal{E}_j}$ and $\mathcal{L}^k_{\mathcal{E}_j}, k = 1, 2, 3$ can be obtained from the eigenvalues $\lambda^k_{\mathcal{E}_j}, k = 1, 2, 3$ as $\mathcal{L}^k_{\mathcal{E}_j} = \sqrt{\frac{1}{\lambda^k_{\mathcal{E}_j}}}$.

Given $\mathbf{U}^k_{\mathcal{E}_j}$ and $\mathcal{L}^k_{\mathcal{E}_j}$ we can also calculate $\Delta E_i^{\mathcal{E}_j}$ from $\mathbf{D}_i$ by computing the norm-2 of the vector resulting from first rotating $\mathbf{D}_i$ to the reference system of $\mathbf{U}^k_{\mathcal{E}_j}$ of $\mathcal{E}_j$ and then scaling according to the ellipsoid semis axis as

$$\Delta E_i^{\mathcal{E}_j} = ||\lambda_{\mathcal{E}_j} \mathbf{U}^T \mathbf{D}_i||_2, \tag{3}$$

where $\lambda_{\mathcal{E}_j}$ is the diagonal matrix formed by the eigenvalues $\lambda^k_{\mathcal{E}_j}, k = 1, 2, 3$. This expression is equivalent to computing the classical ellipsoid metric, which is related in turn to ellipsoid implicit equation, from the rotated vector of differences as follows:

$$\mathbf{D}_i^{\mathcal{E}_j} = \mathbf{U}^T \mathbf{D}_i, \tag{4}$$

$$\Delta E_i^{\mathcal{E}_j} = \sqrt{\sum_{k=1}^{3} \frac{\mathbf{D}_i^{\mathcal{E}_j}(k)^2}{{\mathcal{L}^k_{\mathcal{E}_j}}^2}}. \tag{5}$$

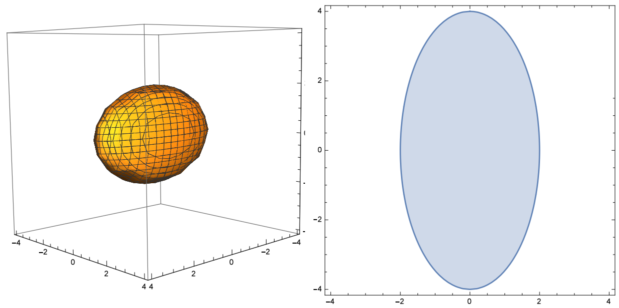

Figure 1: Geometry generated by points of equal distance of 1 in a classical ellipsoid centered at $(0, 0, 0)$ with semi-axis equal to $1, 2, 3$, respectively, and its 2D projection over the Z axis.

In Figure 1 we can see the ellipsoid geometry generated by the points at equal distance of 1 for an ellipsoid centered in $(0, 0, 0)$ with semi-axis equal to $1, 2$, and 3, respectively.

All four equations 1, 2, 3, and 5 are equivalent for computing $\Delta E_i^{\mathcal{E}_j}$ using the classical ellipsoid distance model. Indeed, if we were just interested in using the classical distance we could just use equation 1 and save computations to obtain $\mathbf{U}_{\mathcal{E}_j}^k$ and $\mathcal{L}_{\mathcal{E}_j}^k$. But obtaining them provides us with the needed information to use alternative metrics as the one we use in the following section.

2.3. Standard fuzzy metric in ellipsoid context

Fuzzy metrics are a tool within the fuzzy set and fuzzy logic framework which have been thoroughly studied from the theoretical point of view [21, 22] and that have shown interesting properties with respect to classical metrics.

Basically, a fuzzy metric is a function $M(\mathbf{x}, \mathbf{y}, t)$ in $]0, 1]$ that measures the closeness or similarity of two objects $\mathbf{x}, \mathbf{y}$ with respect to a context parameter t. There are a few performance differences among fuzzy metrics and classical (Minkowski) metrics being the two most relevant ones that: (i) fuzzy metrics use a context parameter t that make the metric be adaptive to context; and (ii) fuzzy metrics use t-norms for conjunction, for instance when used in a vector context, whereas classical metrics use summation. These two points make fuzzy metrics to perform different from classical metrics [16], being more interesting those applications where context information, t, is available.

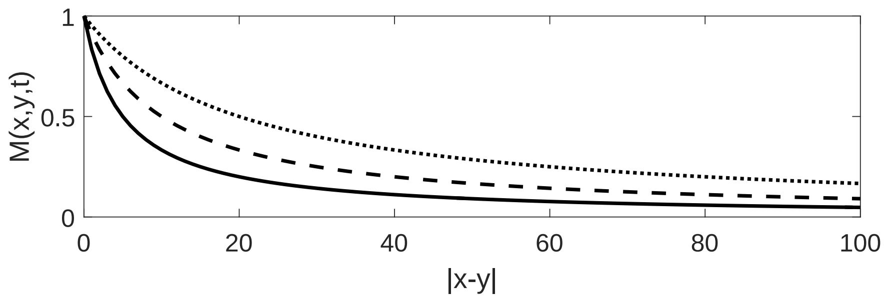

Figure 2: $M(x, y, t)$ as a function of $|x - y|$ for $t = 5$ (solid), $t = 10$ (dashed), $t = 20$ (dotted).

A classical fuzzy metric is the so-called standard fuzzy metric [21] given by

$$M(x, y, t) = \frac{t}{t + |x - y|}, \tag{6}$$

where x and y are two scalar values and $|\cdot|$ denotes the absolute value. It can be seen how the difference between x and y is measured with respect to t. The larger the difference, the larger the dissimilarity. Increasing the value of t makes M to be less sensitive to changes in $|x - y|$ and makes higher the global similarity computed, as can be seen in Figure 2.

In the context of an ellipsoid of suprathreshold color differences $\mathcal{E}_j$ we have available three pieces of information: the center of the ellipsoid $\mathbf{O}_{\mathcal{E}_\mathbf{j}}$, the vectors $\mathbf{U}^k_{\mathcal{E}_j}, k = 1, 2, 3$ that describe rotation of the ellipsoid, and the ellipsoid semi axis $\mathcal{L}^k_{\mathcal{E}_j}, k = 1, 2, 3$ that characterize the sensitivity that should be taken into account relative to the directions of $\mathbf{U}^k_{\mathcal{E}_j}$. The sensitivity in each direction is explicitly taken into account in equation 5 in a linear way, given that differences in each direction are divided by that direction semi axis length. On the other hand, we could make an analogous measurement using the standard fuzzy metric but in this case the sensitivities are taken into account in a non-linear way as the curve in Figure 2 shows. The details of this processing are the following:

First we rotate the vector $\mathbf{D}_i$ to the reference system of $\mathcal{E}_j$ obtaining $\mathbf{D}_i^{\mathcal{E}_j}$ as

$$\mathbf{D}_i^{\mathcal{E}_j} = \mathbf{U}^T \mathbf{D}_i, \tag{7}$$

where $\mathbf{U}$ is the ortonormal rotation matrix having as columns the vectors $\mathbf{U}^k_{\mathcal{E}_j}$. Then, the difference computed using the standard fuzzy metric [21] in direction k is given by

$$\Delta EFM_i^{\mathcal{E}_j}(k) = \frac{\kappa \mathcal{L}^k_{\mathcal{E}_j}}{\kappa \mathcal{L}^k_{\mathcal{E}_j} + \mathbf{D}_i^{\mathcal{E}_j}(k)}, k = 1, 2, 3, \tag{8}$$

where κ is an scaling parameter that we set $\kappa = 9$ so that when $\mathcal{L}^k_{\mathcal{E}_j} = \mathbf{D}_i^{\mathcal{E}_j}(k)$ we have $\Delta EFM_i^{\mathcal{E}_j}(k) = 0.9$.

Finally, according to [21] we need to use a continuous t-norm $*$ to combine the differences in each direction for which we use the classical product t-norm [15] as follows:

$$\Delta EFM_i^{\mathcal{E}_j} = \prod_{k=1}^{3} EFM_i^{\mathcal{E}_j}(k). \tag{9}$$

In this way, this metric takes into account the information of directional sensitivity in an ellipsoid in a different way. The main differences with respect to the classical ellipsoid metric are: First, that the sensitivities are taken into account in a non-linear way, since differences are not scaled with respect to sensitivity but measured *with respect to* the sensitivity; Second, we can see in Figure 3 that the geometry generated by points at equal distance is no longer an ellipsoid. The *diamond-like* shape geometry indicates that this metric is more sensitive to differences appearing in *diagonal* directions in the color space than when the difference is only in one of the axis directions.

In the following, we will consider both $\Delta E_i^{\mathcal{E}_j}$ in Eq. 5 and $\Delta EFM_i^{\mathcal{E}_j}$ as possible local color differences that are combined to derive a general color difference computation using the framework described in the following section.

3. A fuzzy soft-switching model for color differences in a multi-ellipsoid context

Now, let us consider that we have N ellipsoids denoted by $\mathcal{E}_j, j = 1, ..., N$ and so N different $\Delta E_i^{\mathcal{E}_j}, j = 1, .., N$ (and $\Delta EFM_i^{\mathcal{E}_j}$) to predict the color dif-

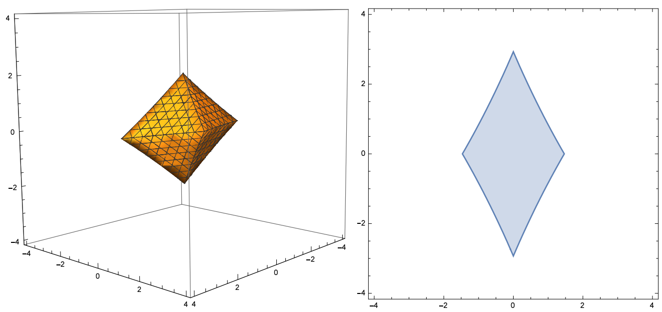

Figure 3: Geometry generated by points of equal fuzzy metric distance of 0.9 centered at $(0, 0, 0)$ with semi-axis equal to $1, 2, 3$, respectively, and its 2D projection over the Z axis.

ference for each color pair $\mathbf{S}_i$. In the following we detail how to combine all $\Delta E_i^{\mathcal{E}_j}$ to obtain the general difference ΔE_i by means of a weighted average that constitutes a soft-switching model. An analogous computation should be made for $\Delta EFM_i^{\mathcal{E}_j}$ to obtain ΔEFM_i.

The general equation for the soft-switching model is given as

$$\Delta E_i = \frac{\sum_{j=1}^{N} W_j R_j \Delta E_i^{\mathcal{E}_j}}{\sum_{j=1}^{N} W_j R_j}, \tag{10}$$

where each local difference $\Delta E_i^{\mathcal{E}_j}$ is weighted according to two criteria: closeness of $\mathbf{S}_i$ to $\mathbf{O}_{\mathcal{E}_\mathbf{j}}$ represented by W_j and reliability of $\mathcal{E}_j$ given by R_j. Since *close* is a vague term, it can be represented as a fuzzy set and so W_j can be set using a S-type fuzzy membership function representing the fuzzy set *close* [15] as

$$W_j = 1 - \mu(||\mathbf{C}_i - \mathbf{O}_{\mathcal{E}_\mathbf{j}}||, \alpha, \gamma), \tag{11}$$

where $||\cdot||$ denotes the Euclidean norm and μ is an S-type membership function given by

$$\mu(x,\alpha,\gamma) = \begin{cases} 0 & \text{if } \ x \leq \alpha \\ 2\left(\frac{x-\alpha}{\gamma-\alpha}\right)^2 & \text{if } \ \alpha < x \leq \frac{\alpha+\gamma}{2} \\ 1-2\left(\frac{x-\gamma}{\gamma-\alpha}\right)^2 & \text{if } \ \frac{\alpha+\gamma}{2} < x \leq \gamma \\ 1 & \text{if } \ x > \gamma \end{cases} \quad (12)$$

We set $\alpha = 1$ and $\gamma = 5$ so that when computing the color difference for a given pair only those ellipsoids at less than 5 CIELAB units of distance are given non-null weights and all ellipsoids at less than 1 unit of distance are given the maximum weight of 1. Figure 4 shows this behaviour.

However, since we do not have enough ellipsoids to densely cover the whole color space, it may happen that no ellipsoids are found at less than 5 units of distance from a given pair. In such a case we set α to the distance of the closest ellipsoid in the space and γ to the distance of the sixth closest ellipsoid so that we use always five ellipsoids with non-null weights, including one of them with weight 1. The setting of these parameters only affects the performance if γ was set so that only one or two ellipsoids were used, which should be avoided. Alternatively, including more ellipsoids does not have a great influence since those additional ones would be further than the rest and so given very little weight, but it seems less reasonable to use many of them. It is important to note that when computing color differences of pairs in a dataset our method has a different behaviour for those pairs with close ellipsoids than for the rest. So, it is interesting to analyze the performance of our method both for the whole dataset as well as separately for pairs with or without close ellipsoids, as we do in the following section.

Furthermore, the weight R_j is included because it has been found [8] that not all ellipsoids fit the data in the same degree and that not all ellipsoids are totally reliable [9] so we use R_j to modulate the importance of each ellipsoid used. Each R_j will be set based on experimental observations in the next section.

Finally it is worth to point out that this model is not a closed color difference

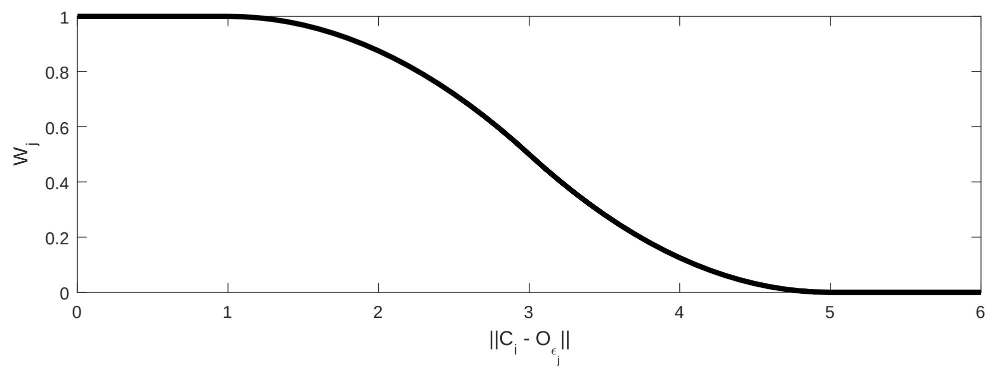

Figure 4: W_j obtained using the fuzzy set *close* over $||\mathbf{C}_i - \mathbf{O}_{\mathcal{E}_\mathbf{j}}||$.

formula but an open model that relies on the color differences predicted by a set of ellipsoids to compute a kind of hybrid color difference measure. This means that if we change the set of ellipsoids, as we do in the next section, the performance may differ significantly.

4. Experimental results and refinements of the method

To assess the performance of the two color difference formulas proposed, ΔE and ΔEFM, we use the so-called *COM* dataset [23]. However, since the data with which the ellipsoids that we used here were fitted [14] is included in the *COM*, we must remove this data from the dataset and use for the assessment the remaining of the dataset that from now on we will call *Reduced COM* (RCOM).

We compare the performance of ΔE and ΔEFM with that of the currently CIE-recommended color difference measure CIEDE2000 (ΔE_{00}) [20]. As figure of merit we use the STRESS measure proposed in [24], which is currently the measure of reference in the literature and for which we can perform statistical significance test to evaluate different performances. Also, we are interested in comparing the performance of the different formulas for the subsets of pairs in RCOM with and without close ellipsoids. For this we also use STRESS but setting the auto-scaling parameter F within STRESS to the value used when computing STRESS for the whole dataset. This implies that we are not using STRESS but a variant of it that illustrates how much a particular subset contributes to the STRESS of the global dataset. If we let F free for the subsets, the results could be significantly different since each subset would use a different scaling and the value of STRESS in the subsets could be inconsistent with that of the whole dataset.

First we study the performance using all 19 ellipsoids found in [8] without assigning any reliability weights R_j in Eq. 10, that is, setting $R_j = 1, \forall j = 1, ..., 19$. STRESS for ΔE and ΔEFM are given in row 1 of Table 1. We can see that performance is better for the subset of pairs with near ellipsoids (WN) than for those without near ellipsoids (WON) as expected. Also, performance of ΔEFM is better than ΔE. We would later discuss about this.

Next, we want to quantify the relative importance of the ellipsoids used. For this, we assess performance by removing one ellipsoid at a time so that we got 19 variants of the original model. When removing en ellipsoid we can assess its importance for the model in terms of STRESS for the RCOM dataset as well as for the WN and WON subsets. These results are also in Table 1 (rows 2 to 20). It is surprising to see that removing some ellipsoids improves performance of the whole model. This does not mean that those ellipsoids are wrong but only that either one or the other assumption behind the model does not hold in the color region of that ellipsoid. That is, that ellipsoid does not predict well either close larger/smaller than threshold color differences or further interpolated color differences. This is reflected in the performance for the WN and WON subsets. We can see for instance that ellipsoid 3 works very well for the WN subset since, if we remove it, STRESS for WN increases significantly. Analogous reasoning can be made for ellipsoid 19 and WON subset. However, color differences near ellipsoid 6 are better predicted from nearby ellipsoids than ellipsoid 6 itself, since removing it makes contribution to STRESS of WN to decrease. We have seen that about half of these color pairs are much larger/smaller than threshold which may indicate that linear scaling of threshold color differences predicted by ellipsoid 6 is inaccurate. Similarly, ellipsoid 9 is not useful to predict further

Table 1: Performance in terms of STRESS of ΔE and ΔEFM for all 19 ellipsoids and for 18 of them removing one at a time. STRESS is computed for the whole dataset and relative contribution to STRESS is given for the subsets of pairs with near ellipsoids (WN) and without near ellipsoids (WON).

Ellipsoid	Datasets and ΔE			Datasets and ΔEFM		
removed	RCOM	WN	WON	RCOM	WN	WON
None	35.59	33.39	35.84	32.73	31.22	32.90
1	35.50	33.29	35.74	32.70	31.07	32.88
2	35.97	33.50	36.25	33.24	31.43	33.45
3	36.93	37.71	36.83	33.76	33.95	33.73
4	35.86	34.07	36.07	33.04	32.29	33.12
5	35.45	33.34	35.69	32.42	31.44	32.53
6	35.54	32.96	35.83	32.65	31.08	32.82
7	35.72	34.31	35.88	32.92	33.22	32.88
8	35.64	33.40	35.89	32.60	31.24	32.75
9	35.29	33.31	35.51	32.49	31.53	32.60
10	35.81	33.40	36.08	32.86	31.24	33.04
11	35.67	33.51	35.91	32.85	30.95	33.06
12	35.39	33.32	35.62	32.60	31.42	32.73
13	35.91	33.53	36.17	32.64	30.99	32.82
14	35.81	33.46	36.07	32.68	31.06	32.87
15	35.56	33.40	35.80	32.68	31.19	32.84
16	35.66	33.57	35.90	32.74	30.47	32.99
17	35.71	33.85	35.92	32.93	31.17	33.12
18	35.79	34.75	35.91	32.82	31.53	32.96
19	36.45	33.69	36.76	33.57	30.70	33.88

color differences since removing it makes contribution to STRESS of WON to decrease. Also, this may be due to the lack of a perfect uniformity in the color space which may make some close ellipsoids to be incompatible in a hybrid model such as the one we propose if the local non uniformity is more acute in their region.

Now, we study whether we could remove several ellipsoids at the same time and improve overall performance. So, what we did was removing recursively more and more ellipsoids until STRESS of the method becomes worse. We did this separately for ΔE and ΔEFM following the STRESS decreasing order according to Table 1. Finally, we select a group of 6 ellipsoids that removing them at the same time improves performance of both ΔE and ΔEFM: the set includes ellipsoids $\{1, 5, 6, 9, 12, 15\}$, so the reduced set of ellipsoids is composed

Table 2: Performance in terms of STRESS of ΔE and ΔEFM for the reduced set of 13 ellipsoids and for 12 of them removing one at a time. Stress is computed for the whole dataset and relative contribution to STRESS is given for the subsets of pairs with near ellipsoids (WN) and without near ellipsoids (WON).

Ellipsoids	Datasets and ΔE			Datasets and ΔEFM		
removed	RCOM	WN	WON	RCOM	WN	WON
$\{1,5,6,9,12,15\}$	34.43	32.96	34.59	31.74	31.89	31.73
$\{1,2,5,6,9,12,15\}$	34.82	33.15	35.01	32.34	32.11	32.37
$\{1,3,5,6,9,12,15\}$	36.56	38.06	36.38	33.45	35.90	33.15
$\{1,4,5,6,9,12,15\}$	34.96	33.74	35.10	32.12	32.89	32.03
$\{1,5,6,7,9,12,15\}$	35.38	36.05	35.30	32.47	34.95	32.17
$\{1,5,6,8,9,12,15\}$	34.73	32.96	34.92	31.68	32.04	31.64
$\{1,5,6,9,10,12,15\}$	34.70	32.99	34.89	31.91	32.04	31.90
$\{1,5,6,9,11,12,15\}$	35.09	32.99	35.32	32.10	31.42	32.18
$\{1,5,6,9,12,13,15\}$	34.60	32.94	34.78	31.68	31.55	31.69
$\{1,5,6,9,12,14,15\}$	34.71	32.89	34.91	31.83	31.71	31.84
$\{1,5,6,9,12,15,16\}$	34.88	32.77	35.11	32.20	30.72	32.37
$\{1,5,6,9,12,15,17\}$	35.54	33.69	35.75	32.98	31.96	33.09
$\{1,5,6,9,12,15,18\}$	34.74	34.07	34.81	31.96	32.24	31.93
$\{1,5,6,9,12,15,19\}$	35.44	33.05	35.71	32.80	31.00	33.00

by 13 ellipsoids $\{2,3,4,7,8,10,11,13,14,16,17,18,19\}$.

Performance of ΔE and ΔEFM using the reduced ellipsoid set is given in first row of Table 2. We can see that performance for both measures has improved in about 1 STRESS unit for the RCOM dataset. Also, we see that performance has increased more for the WON subset than for the WN which means that removing those ellipsoids improve the interpolation capability of the whole system.

Besides, we want to study the importance of each ellipsoid in the reduced set so that we could assign a weight R_j for each ellipsoid in Eq. 10 according to its importance within the model. For this, we remove one additional ellipsoid in the reduced set at a time and see how performance changes (see Table 2 rows 2-14 and column 1 to see the ellipsoids removed in each case). We see that now no significative improvement is obtained in any case. Moreover, there are a few ellipsoids that if we remove them STRESS increases significantly. For instance, ellipsoids 3, 5, 17, and 19. However, there are other ellipsoids that removing them does not influence STRESS that much, as it happens with ellipsoids 8, 13,

Table 3: Two assignments of weights for the reduced set of ellipsoids. In both cases we give more weight to those ellipsoids that have the most influence in STRESS but in the *light* case the relative differences in the weights given are smaller than in the *heavy* case.

Ellipsoid	Light weights	Heavy weights
2	0.90	0.75
3	1.00	1.00
4	0.80	0.50
7	0.90	0.90
8	0.70	0.30
10	0.70	0.30
11	0.90	0.75
13	0.60	0.10
14	0.70	0.30
16	0.90	0.75
17	1.00	1.00
18	0.90	0.75
19	1.00	1.00

or 14. So, we aim to assign a higher R_j weight to the former and a lower one to the latter. We propose two different weight assignments given in Table 3. To come up with this proposal we gave more importance to relative performance in the subset of WON than in the WN since it is in the WON where combination of ellipsoids is more important given that more ellipsoids will be involved in the computation and R_j weights are more relevant. We have assessed performance when including these weights in the model. For ΔE STRESS drops from 34.43 to about 34.24 for both settings of weights whereas for ΔEFM STRESS drops from 31.74 to 31.39 and 30.86 when using the light weights and the heavy weights, respectively, so we decided to use the latter.

Next, we compare the performance of ΔE and ΔEFM with the reference color difference formula CIEDE2000 ΔE_{00}. Results are given in Table 4. We compare the performance in the RCOM dataset but also we compute the relative contribution to this value of STRESS of different subsets of interest: WN, WON and 7 subsets of different ranges of ΔV. We can see that performances of ΔE and ΔEFM are not far from that of ΔE_{00}. In particular, ΔEFM is quite close. By looking at performance in WN and WON we see that ΔE_{00} and ΔE perform better for WN. In particular, ΔE_{00} yields less than 3.5 units of STRESS for WN

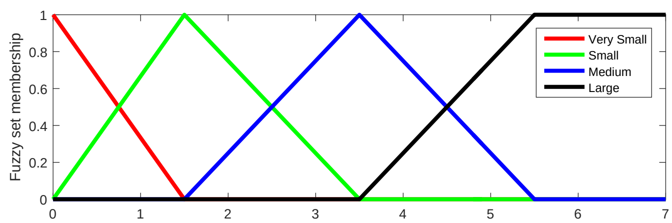

Figure 5: Fuzzy Sets used for classifying color differences.

than for WON. However, performance in ΔEFM is a little better for WON than for WN meaning that the fuzzy metric somehow benefits the combination of different ellipsoids of the color space. If we look at performance for different ranges of ΔV, we see that minimum and maximum relative contribution to STRESS of different subsets differ more in ΔEFM than in ΔE_{00} or ΔE. In fact, ΔEFM performs much worse for smaller color differences, but better for medium differences. In general, a very different performance is observed for ΔE and ΔEFM. This means that they are not adapting the same to different ranges of color differences. This means that if we are able to identify what feature is making each metric behave better in each case, we may be able to improve their performance in general. In the following we pursue this target.

If we look at their definitions we see two differences between Eq. 5 and Eqs. 8-9 that may explain this: one of them is the different shapes of the geometry generated by equally separated samples shown in Figs. 1-3. We saw that Eqs. 8-9 generate a *diamond* shape geometry that models a higher sensitivity to differences appearing in *diagonal* directions in the color space; and the other is that Eq. 5 applies a linear scaling. That is, the square of the differences are divided by the square of the ellipsoid semi axis in each direction. However, Eqs. 8-9 do not apply a linear scaling but a non-linear one according to Figure 2. As a consequence, we think that by appropriately changing geometry and scaling for different ranges of ΔV can improve performance. In the following we detail how can we incorporate this changes into ΔE and ΔEFM and we define two improved metrics that we name $\Delta E*$ and $\Delta EFM*$

Table 4: Performance in terms of STRESS of ΔE_{00}, ΔE and ΔEFM for the reduced set of 13 ellipsoids. STRESS is computed for the whole RCOM dataset and relative contribution to STRESS is given for the subsets of pairs with near ellipsoids (WN), without near ellipsoids (WON), and for different ranges of ΔV color differences. $\Delta E*$ and $\Delta EFM*$ refer to the variants of ΔE and ΔEFM proposed in equations (13)-(14).

Set	#pairs	ΔE_{00}	ΔE	ΔEFM	$\Delta E*$	$\Delta EFM*$
RCOM	3501	29.37	34.24	30.86	32.46	27.61
WN	837	26.13	32.92	31.89	32.95	28.30
WON	2664	29.73	34.39	30.74	32.40	27.13
$\Delta V \in [0, 0.5[$	464	32.51	34.77	57.71	45.44	46.05
$\Delta V \in [0.5, 1.5[$	1528	30.61	35.08	41.33	36.34	34.53
$\Delta V \in [1.5, 2.5[$	899	28.89	34.61	29.22	34.70	29.40
$\Delta V \in [2.5, 3.5[$	329	28.76	33.28	23.83	28.81	24.71
$\Delta V \in [3.5, 4.5[$	162	33.90	38.29	26.02	30.45	21.57
$\Delta V \in [4.5, 5.5[$	71	28.44	35.03	34.23	33.67	27.19
$\Delta V \geq 5.5$	48	23.37	26.56	33.19	30.61	28.08

We can introduce changes in geometry and scaling in Eq. 5 reformulating this equation by including as parameters the power p applied to the differences and the root and adding a scaling factor s as follows:

$$\Delta E_i^{\mathcal{E}_j} = \left(\sum_{k=1}^{3} \frac{\mathbf{D}_i^{\mathcal{E}_j}(k)^p}{\left(s\mathcal{L}_{\mathcal{E}_j}^k \right)^p} \right)^{\frac{1}{p}}. \tag{13}$$

Thus, s controls the scaling factor applied in each case and p allows for different geometries of equally distance samples so that for $p < 2$ geometries tend to a *diamond-like* shape and for $p > 2$ the geometry tends to a *rectangular* shape.

Analogously, we can parametrize Eq. 8 as

$$\Delta EFM_i^{\mathcal{E}_j}(k) = \frac{\kappa \left(s\mathcal{L}_{\mathcal{E}_j}^k \right)^p}{\kappa \left(s\mathcal{L}_{\mathcal{E}_j}^k \right)^p + \mathbf{D}_i^{\mathcal{E}_j}(k)^p}, k = 1, 2, 3, \tag{14}$$

where s and p play similar roles.

We do not have much information to set s and p except that we hypothesized that they should be different for different sizes of color differences so, we propose to use a simple fuzzy rule based system for the setting including these 4 rules:

1. IF ΔV is *very small* THEN set $s = a$ and $p = e$
2. IF ΔV is *small* THEN set $s = b$ and $p = f$
3. IF ΔV is *medium* THEN set $s = c$ and $p = g$
4. IF ΔV is *large* THEN set $s = d$ and $p = h$

Given that ΔV is *very small*, *small*, *medium*, and *large* are vague statements we model them by using 4 different fuzzy sets shown in Figure 5. Given that ΔV is unknown, we need here to use an estimation of it for which we use ΔE_{00}, but any other color difference could be used instead given that we do not need a very accurate estimation at this point. For each color pair, fuzzy logic inference process determines the certainty of the antecedent of the rule using the certainty association depicted in Figure 5. The certainty of the antecedent is assigned to the consequence. Finally s and p are determined by averaging $a - d$ and $e - h$ using as weight the respective certainties in each case. This means that there is a smooth linear transition between the 4 considered values for s and p. We will name the color difference measures derived from this model $\Delta E*$ and $\Delta EFM*$.

We set a, b, c, d, e, f, g, h separately for $\Delta E*$ and $\Delta EFM*$ by finding sub-optimal performance in terms of STRESS through extensive experimentations. Optimal performance is very difficult to determine since Stress is prone to many local minima. In the optimization, we restricted $a-d$ and $e-h$ to be monotonic for consistency and to avoid data over-fitting. We determine that an appropriate setting for $\Delta E*$ is $a = 1.3, b = 1.9, c = 2.1, d = 2.8$ and $e = 1.9, f = 1.5, g = 1.2, h = 1.2$. This means that scaling is set so that global sensitivity decreases as ΔV increases and that the geometry tends from ellipsoidal to *diamond-like* as ΔV increases, as it can be seen in Figure 6. With this setting the performance of $\Delta E*$ is almost 2 STRESS units better than ΔE, as it is shown in Table 4. For $\Delta EFM*$ we found $a = 3.9, b = 3.9, c = 1.1, d = 0.2$ and $e = 0.8, f = 0.5, g = 0.3, h = 0.2$. In this case we see that same pattern with respect to the exponent p that determine the geometry but now the scalings are decreasing. This can be interpreted as the original metric ΔEFM having too much sensitivity for small ΔV which is corrected by using large scales in this cases and too little sensitivity for large ΔV which is increased by using a scaling factor lower than 1. With this setting the performance of $\Delta EFM*$ is more than 3 STRESS units better than ΔEFM and almost 2 units better than ΔE_{00} (see Table 4).

From a statistical point of view we can analyze whether the STRESS differences observed are significant using the F-test of significance at 95% confidence level as explained in [24]. This F-test shows that the differences in terms of STRESS between $\Delta E*$ and $\Delta EFM*$ and ΔE_{00} are statistically significative at 95% confidence level, which means that $\Delta E*$ is not good enough to be considered an alternative to ΔE_{00} and $\Delta EFM*$ is significantly better than ΔE_{00} in terms of STRESS.

A downside of the proposed method is the high number of parameters included which arises some uncertainty on the possibility of data overfitting. ΔE_{00} is already a complex color-difference formula including 20 fitting parameters, although some of them interact and k_l, k_c, k_h are fitted to a special case. In the proposed method we have 13 weights in Table 3 that, despite we did not numerically fit to any data we did set to 6 different levels. Also, we have the 8 scaling-power parameters $a - h$ that were indeed fitted, which increases the number of parameters to 21. We can reduce this number and so increase the confidence on absence of data overfitting by removing the R_j factors. If we do this, STRESS for $\Delta E*$ and $\Delta EFM*$ increase to 33.37 and 27.97, respectively, which, in terms of F-test, is not significative at a 95% confidence level. Thus, the number of parameters drops to just 8.

5. Conclusions

In this paper we have proposed a model to combine local threshold color difference information to build two different color difference formulas. For this we have used the classical ellipsoid metric and a fuzzy metric that models and analogous reasoning but in a different way. We have found that there are significant performance differences between the two metrics for different color difference ranges. This has led us to study how to change the metric parameters to op-

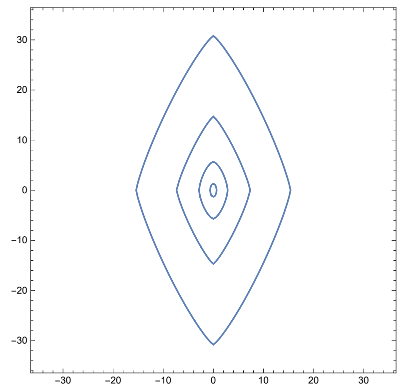

Figure 6: 2D projection of the geometry generated by points of equal $\Delta E*$ of $0.5, 1.5, 3.5$ and 5.5 centered at $(0, 0, 0)$ with semi-axis equal to $1, 2$, respectively.

timize performance. We have proposed a simple fuzzy rule system to set the metric parameters. As a result, we have obtained a significant improvement in global performance that even outperforms the currently CIE-recommended color-difference formula CIEDE2000. This implies a new look at the problem of color-differences since with the parameters adapting to size of the difference being measured it turns into a 4-D problem instead of the common 3-D approach.

References

[1] M.D. Fairchild, G.M. Johnson, iCAM framework for image appearance differences and quality, Journal of Electronic Imaging 13 (1) (2004) 126-138

[2] G.M. Johnson, M.D. Fairchild, A top down descrption of S-CIELAB and CIEDE2000, Color Research and Application, 28 (6) (2003) 425-435.

[3] Svetlana Grečova, Samuel Morillas, Perceptual similarity between color images using fuzzy metrics, Journal of Visual Communication and Image Representation 34 (2016) 230-235.

[4] Melgosa, M.; Hita, E.; Romero, J.; Jiménez del Barco, L. Some classical color differences calculated with new formulas, *J. Opt. Soc. Am. A* **1992**, 1247-1254.

[5] Melgosa, M.; Quesada, J.J.; Hita, E. Uniformity of some recent color metrics tested with an accurate color-difference tolerance dataset, *Appl. Opt.* **1994**, 33, 8069-8077.

[6] M.R. Luo, B. Rigg, Chromaticity discrimination ellipses for surface colors, *Color reserach and application* 11 25-42 (1986).

[7] M. Cheung, B. Rigg, Colour-difference ellipsoids for five CIE color centers, *Color reserach and application* 11 185-195 (1986).

[8] Melgosa, M.; Hita, E.; Poza, A.J.; Alman, D.H.; Berns, R.S. Suprathreshold color-difference ellipsoids for surface colors, *Color Res. Appl.* **1997**, 22, 148-155.

[9] S. Shen, R.S. Berns, Evaluating Color Difference Equation Performance Incorporating Visual Uncertainty, *Color reserach and application* 34 (5) 375-390 (2009).

[10] MacAdam, D.L. Visual sensitivities to color differences in daylight, *J. Opt. Soc. Am.* **1942**, 32, 247-274.

[11] CIE Publication 15:2004. Colorimetry. 3rd Edition. CIE Central Bureau, Vienna, 2004.

[12] Luo, M.R.; Cui, G.; Li, C. Uniform colour spaces based on CIECAM02 colour appearance model, *Color Res. Appl.* **2006**, 31, 320-330.

[13] Oleari, C.; Melgosa, M.; Huertas,R. Euclidean color-difference formula for small-medium color differences in log-compressed OSA-UCS space, *J. Opt. Soc. Am. A* **2009**, 26, 121-134.

[14] Berns, R.S.; Alman, D.H.; Reniff, L.; Snyder, G.D.; Balonon-Rosen, M.R. Visual determination of suprathreshold color-difference tolerances using probit analysis, *Color Res. Appl.* **1991**, 16, 297-315.

[15] Kerre,E.E. *Fuzzy Sets and Approximate Reasoning*; Xian Jiaotong University Press: Jiaotong, 1998, ISBN 7-5605-1091-4.

[16] S. Morillas, “Fuzzy Metrics and fuzzy logic for colour image filtering”, *Ph.D. Thesis*, Technical University of Valencia, June 2007.

[17] S. Morillas, L. Gomez-Robledo, R. Huertas, M. Melgosa, “Fuzzy analysis for detection of inconsistent data in experimental datasets employed at the development of the CIEDE2000 colour-difference formula ”, *Journal of Modern Optics* **56** (13), 1447-1456.

[18] S. Morillas, L. Gomez-Robledo, R. Huertas, M. Melgosa, “Method to determine the degrees of consistency in experimental datasets of perceptual color differences”, *Journal of the Optical Society of America - A* **33** (12), 2289-2296.

[19] Benavente, R.; Vanrell, M.; Baldrich, R. Parametric fuzzy sets for automatic color naming, *J. Opt. Soc. Am. A* **2008**, 25, 2582-2593.

[20] CIE Publication 142-2001. Improvement to industrial colour-difference evaluation. CIE Central Bureau, Vienna, 2001.

[21] A. George, P. Veeramani, “On Some results in fuzzy metric spaces”, *Fuzzy Sets and Systems*, 64 (3), 395-399, 1994.

[22] V. Gregori, S. Romaguera, “Some properties of fuzzy metric spaces”, *Fuzzy Sets and Systems*, 115 (3), 477-483, 2000.

[23] Melgosa, M.; Huertas, R.; Berns, R.S. Performance of recent advanced color-difference formulas using the standardized residual sum of squares index, *J. Opt. Soc. Am. A* **2008**, 25, 1828-1834.

[24] García, P.A.; Huertas, R.; Melgosa, M.; Cui, G. Measurement of the relationship between perceived and computed color differences, *J. Opt. Soc. Am. A* **2007**, 1823-1829.