Investigating Euclidean Mappings for CIEDE2000 Color Difference Formula

Lorenzo Ridolfi & Marcelo Gattass, Department of Informatics, Hélio Lopes, Department of Mathematics, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brazil

Abstract

In recent years, various color difference formulas were developed for the CIELAB space, such as CMC, CIE94 and CIEDE2000. Although these formulas have achieved greater accuracy in perceptual measurement between colors, many applications cannot take advantage of this greater precision, because the Euclidean distances in CIELAB are not isometric in accordance with these new formulas. Thus, applications such as gamut mapping and color interpolation need a color space that is isometric in relation to the latest color difference formulas. This paper studies the mapping of the CIELAB space, particularly the ab plane of this space according to the metrics of the CIEDE2000 formula, through multidimensional scaling (MDS) techniques, more specifically ISOMAP and an optimization based on Sammon Mapping.

Introduction

In recent years, various color difference formulas were developed for the CIELAB space, such as CMC [1], CIE94 [2] and CIEDE2000 [3]. Although these formulas have achieved greater accuracy in perceptual measurement between colors, many applications cannot take advantage of this greater precision, because the Euclidean distances in CIELAB are not isometric in accordance with these new formulas. Thus, applications such as gamut mapping and color interpolation need a color space that is isometric in relation to the latest color difference formulas.

The creation of colometric spaces that are isometric to the most recent color difference formulas has been the subject of several works. Regarding the CIE94 formula, we note the space DIN99 [4], which became a standard in Germany, and an improvement of this space was proposed by Cui et al. [5]. Isometric mapping from the CIEDE2000 formula was partially studied by Völz [6], who investigated only the first quadrant, and through a more comprehensive approach by Urban et al. [7, 8], who also produced mappings for the CIE94 and CMC formulas.

This paper studies the isometric mapping of the CIELAB space, particularly the *ab* plane of this space according to the metrics of the CIEDE2000 formula, through multidimensional scaling (MDS) [9, 10] techniques, more specifically, ISOMAP [11] and an optimization based on Sammon Mapping [12]. MDS-based techniques have been employed in color and image science in many different contexts, such as a three-dimensional representation of the Munsell color chips spectrum [13], analysis of the geometry of the Munsell color space [14] and perceptual analysis of image quality [15]. However, we could not find any studies on the use of MDS to generate Euclidean metric spaces based on color difference formulas.

This paper is structured as follows: first we describe more precisely the problem of mapping a space with a non-Euclidean

metric in a Euclidean metric space. Then, we present our approach to construct this mapping, describing the mapping of the lightness coordinate and ab plane. Following, we show the results obtained, an analysis of the dimensionality of the isometric mapping of the ab plane and the accuracy of the two- and three-dimensional mappings achieved in this plane. Finally, we present our conclusions.

Description of the Problem

In this section we describe mathematically the problem of creating a mapping that approximates the CIELAB space, under the non-Euclidean distance metrics of the CIEDE2000 formula, onto a new Euclidean metric space.

Mathematical Formulation

Let $c_1 = (L_1, a_1, b_1)$ and $c_2 = (L_2, a_2, b_2)$ be coordinates in the CIELAB space, where $||c_2 - c_1|| \le D_0$, for a small D_0 . The Euclidean mapping M_{2k} is defined as:

$$M_{2k} : \mathbf{R}^3 \to \mathbf{R}^{d+1}$$

$$\Delta E_{2k}(c_1, c_2) = \|M_{2k}(c_2) - M_{2k}(c_1)\|$$
(1)

where *d* is the dimension of an Euclidean perceptual space correspondent to the *ab* plane, $\Delta E_{2k}(c_1, c_2)$ is the CIEDE2000 color difference equation and D_0 is the largest color difference where equation 1 can be applied.

In the CIEDE2000 formula (2), component L from the CIELAB space is independent from variables a and b.

$$\Delta E_{2k}(c_1, c_2) = \sqrt{\Delta E_{2k}^L(L_1, L_2)^2 + \Delta E_{2k}^{ab}((a_1, b_1), (a_2, b_2))^2}$$
(2)

In the same way as in [8, 7], mapping M_{2k} can be expressed in terms of mapping M_{2k}^L for the *L* coordinate and in terms of mapping M_{2k}^{ab} for the *ab* plane.

$$M_{2k}^{L} : \mathbf{R} \to \mathbf{R}$$

$$\Delta E_{2k}^{L}(L_{1}, L_{2}) = \|M_{2k}^{L}(L_{2}) - M_{2k}^{L}(L_{1})\|$$
(3)

$$M_{2k}^{ab} : \mathbf{R}^{2} \to \mathbf{R}^{d}$$

$$\Delta E_{2k}^{ab}(c_{1}, c_{2}) = \|M_{2k}^{ab}(a_{2}, b_{2}) - M_{2k}^{ab}(a_{1}, b_{1})\|$$
(4)

Thus, we can say that $c_m = M_{2k}(c)$, where c = (L, a, b), $c_m = (L_m, m_1, \dots, m_d)$, $L_m = M_{2k}^{L}(L)$ and $(m_1, \dots, m_d) = M_{2k}^{ab}(a, b)$.

Approach

Euclidean mapping M_{2k} was generated from the partial mappings M_{2k}^L and M_{2k}^{ab} , and each of these partial mappings was obtained separately, as described below.

Lightness Mapping

The mapping of the lightness coordinate is determined by calculating the following integral derived from the CIEDE2000 formula, as described in [8, 7].

$$M_{2k}^{L}(L) = \int_{0}^{L} \frac{dt}{k_{L}S_{L}(t)}$$
(5)

where $S_L(t)$ is defined by the CIEDE2000 formula as:

$$S_L(t) = 1 + \frac{0.015(t-50)^2}{\sqrt{20 + (t-50)^2}}$$
(6)

The integral (5) is used for the construction of a look-up table in the form $L_i \rightarrow M_{2k}^L(L_i)$, where $L_i = 0, L_p, \dots, 100$, and L_p is the desired precision, typically ranging from 0.1 to 1.

Ab Plane Mapping

The CIEDE2000 function presents a complex behavior regarding the *ab* plane. This behavior was described in [7], by analyzing the Gaussian curvature of the *ab* plane according to the metric induced from the CIEDE2000 function, which revealed significant variations in the plane's curvature. This result motivated us to study the intrinsic dimensionality of the mapping in question. As we know in advance that the mapping is not perfectly isometric, the mapping techniques based on multidimensional scaling [9] were adequate to analyze the dimensionality and to obtain an almost isometric mapping of the *ab* plane based on the metrics induced by the CIEDE2000 function. Since this function is only defined for small distances, we focused on techniques designed for multidimensional scaling in manifolds, which are based only on local distances, such as ISOMAP [11] and LLE [16].

Aiming to increase the accuracy of the Euclidean mapping obtained by ISOMAP, an additional optimization stage based on Sammon Maps was conceived, using a minimization algorithm based on simplex search [17]. The minimization was applied to an error function, which is also described ahead.

ISOMAP

ISOMAP is a method that takes as input a graph whose nodes are a set of points that discretize the space to be mapped and whose edges are the measurements of the distances between a point and its neighbors. The output of ISOMAP is a new set of points, representing the isometric mapping in *d* dimensions that most closely approximates the set of input points subject to the distance measurements provided. In this work, the set of points was generated from a mesh of triangles that discretizes the *ab* plane and the distance measures were naturally obtained from the CIEDE2000 formula itself.

More specifically, let $X_i = (a_i, b_i)$ and $X_j = (a_j, b_j)$, $i, j \in \{1, \dots, n\}$, be points in the mesh that discretizes the *ab* plane. X_i is a neighbor of X_j if and only if $||X_j - X_i|| \le D_0$. We say that *j* belongs to the N_i neighborhood of X_i , if and only if X_i is a neighbor of X_j .

Graph G = (V, E) is defined as $V = \{X_1, \dots, X_n\}, (i, j) \in E$, if and only if X_i is a neighbor of X_j . Function (7) maps the graph edges onto CIEDE2000 distances in the *ab* plane:

$$F_{2k}^{ab}: E \to [0,\infty),$$

$$(i,j) \to \Delta E_{2k}^{ab}(X_i, X_j)$$
(7)

ISOMAP has three processing steps. The first step performs the generation of graph G, the calculation of function F_{2k}^{ab} for all edges in the graph in question and the creation of the $n \times n$ distance matrix D_{2k}^{ab} with the following initial values:

$$D_{2k}^{ab}(i,j) = \begin{cases} \Delta F_{2k}^{ab}(i,j) & \text{if } ||X_j - X_i|| \le D_0, \\ \infty & \text{if } ||X_j - X_i|| > D_0 \end{cases}$$
(8)

In the following step, the elements of the D_{2k}^{ab} matrix that were initialized with ∞ , are updated by an algorithm that calculates the shortest distances among all nodes in the graph, such as the Johnson's [18] or Floyd-Warshall's [19, 20] algorithm.

The final step of ISOMAP is the application of classical multidimensional scaling algorithm in the matrix D_{2k}^{ab} . Classical MDS performs a linear mapping that preserves the distance relationships present in this matrix, using a projection known as classing scaling [21, 10, 9]. The result of classical MDS is a matrix $Y \in \mathbf{R}^{n \times d}$, d < n, where the Y columns are the resulting Euclidean mapping of the X_i points.

To determine Y, classical MDS explores the relationships between distances and inner product. The Gram matrix G corresponding to the Euclidean space is defined by:

$$G = YY^T \tag{9}$$

According to classical MDS, matrix G can also be expressed in terms of distance matrix D_{2k}^{ab} , as follows:

$$G = -\frac{1}{2}J(D_{2k}^{ab})^{\star 2}J$$
(10)

$$J = I_{n \times n} - \frac{1}{n} \mathbf{1}_{n \times n} \tag{11}$$

where $\star 2$ is the element-wise square operation, $I_{n \times n}$ is the $n \times n$ identity matrix and $1_{n \times n}$ is the $n \times n$ matrix of value 1 elements.

Matrix G is created from equation (10) and the Y values are obtained by the eigendecomposition of G, as follows:

$$G = Q\Lambda Q^T \tag{12}$$

where Λ is a diagonal matrix formed with the *G* eigenvalues and *Q* is an orthogonal matrix with the eigenvectors. To obtain a *d*-dimensional Euclidean space, the corresponding *Y* values are:

$$Y = Q_d \Lambda_d^{0.5}, d \le n \tag{13}$$

where Q_d is a matrix with the *d* leading eigenvectors, corresponding to the *k* largest eigenvalues, and $\Lambda_d^{0.5} \in \mathbf{R}^{n \times d}$ contains the square roots of the corresponding eigenvalues. Thus, the Euclidean mapping of point X_i is point Y_i , which is the *i* column of matrix *Y*.

Optimization

Aiming to improve the accuracy of the mapping obtained by ISOMAP, an additional stage of optimization for the results was developed. The objective of this optimization is similar to that of the non-linear projection technique known as Sammon Mapping, but employing an error function empirically developed in this work. The error function (14) is based on the traditional relative error formula, calculated for every mesh point in relation to its neighborhood.

$$E_{r}(i,j) = \frac{|||Y_{j} - Y_{i}|| - \Delta E_{2k}^{ab}(X_{i}, X_{j})|}{\Delta E_{2k}^{ab}(X_{i}, X_{j})}, j \in N_{i}$$
(14)

In our understanding, the optimization step must seek to improve both the average error and the maximum error of mapping. Consequently, the function to be minimized was constructed in terms of the average and maximum relative error of each grid point in relation to its neighborhood.

The average error is weighted by Gaussian weight W_{ij} (15), which decreases as the neighbors move away from point *i*. The weighting is designed to increase the importance of the points near point *i*, which are usually less numerous than the points near the edge of the N_i neighborhood in a regular mesh.

$$W_{ii} = e^{-\beta \|X_i - X_i\|^2} \tag{15}$$

Thus, the weighted average error for point *i* is defined by:

$$E_{mean}(i) = \frac{\sum_{j \in N_i} W_{ij} E_r(i,j)}{\sum_{j \in N_i} W_{ij}}$$
(16)

Likewise, the maximum error for point *i* is defined by:

$$E_{max}(i) = \arg\max_{j \in N_i} E_r(i, j) \tag{17}$$

Finally, the final error function is a linear combination of the weighted average error and maximum error:

$$E_{opt}(i) = \alpha E_{mean}(i) + (1 - \alpha) E_{max}(i), \alpha \in [0, 1]$$
(18)

The projection techniques based on Sammon Mapping support various minimization algorithms [9]. The minimization algorithm chosen was the simplex search developed by Nelder and Mead [17], because it is available in mathematical programming packages, such as MATLAB [22], and also because it does not require the calculation of partial derivatives.

In this study, the algorithm searches small displacements $\Delta m_1, \dots, \Delta m_d$ for each point $Y_i = (m_{i1}, \dots, m_{id})$ in order to minimize the error $E_{opt}(i)$, according to the pseudo-code below:

$$errglobal = \sum_{i=1}^{n} E_{opt}(i);$$

repeat

$$errglobalold = errglobal;;$$

for $i = 1$ to n do
Find $\Delta m_{i1}, \dots, \Delta m_{id}, where \arg\min_{Y_i + \Delta m_{i1}, \dots, \Delta m_{id}} E_{opt}(i)$
 $Y_i = Y_i + \Delta m_{i1}, \dots, \Delta m_{id};$
end for
 $errglobal = \sum_{i=1}^{n} E_{opt}(i);$
until $errglobalold < errglobal$

Results

The results illustrate the two key points of this work. The first is the dimensionality analysis of the mapping of the *ab* plane in multidimensional spaces whose Euclidean distances approximate the CIEDE2000 formula in relation to this plane. The second focus is the mapping itself, in which Euclidean spaces with two and three dimensions are obtained from the same formula in the *ab* plane with high precision. These results are available online at www.tecgraf.puc-rio.br/~color.

Dimensionality Analysis

The dimensionality analysis of the Euclidean mapping of *ab* plane based on the CIEDE2000 formula was performed based on the eigenvalues resulting from the multidimensional scaling step of ISOMAP in a triangle mesh with an hexagon shape. This mesh was constructed adaptively by successively refining an initial mesh with six equilateral triangles whose edge sizes was 150, centered at the origin. The adaptive algorithm [23, 24] successively refined the mesh by dividing each triangle when the CIEDE2000 difference between the two vertices from an edge was greater than an established threshold. The final refined mesh has 16,220 vertices, 31,991 triangles with all the edges corresponding to a CIEDE2000 difference that is less or equal to 1.1.



Figure 1: Cumulative variance by dimension

Figure 1 shows the cumulative percentage variance of the dimensions. This figure clearly shows that there are two dominant dimensions, confirming the predominant bi-dimensional nature of the mapping. We also note that the relative contributions of other dimensions are more uniform. The presence of this greater unifor-



Figure 2: Two-dimensional mapping of a simplified version of the mesh



Figure 3: Three-dimensional Euclidean mapping of the *ab* plane



(a) Unity ellipses plotted in the *ab* plane

(b) Mapping into the new Euclidean space

Figure 4: Two-dimensional Euclidean mapping of unity ellipses

mity of the eigenvalues after the second dimension has motivated us to examine the accuracy of the mapping in two and three dimensions.

Mapping in Two and Three Dimensions

Accurate Euclidean mappings in two dimensions of the *ab* plane according to the CIEDE2000 formula are an important practical result, since these spaces can be used in various applications such as gamut mapping and color interpolation. Our discrete approach, however, produces results only at the nodes of the supporting triangular mesh. Interpolation techniques can be used for forward and reverse mappings to obtain values throughout the mesh. In our results, we used the CGAL [25] library to enforce C^1 continuity.

Although the Euclidean spaces in three dimensions may also be used in the same applications as spaces of two dimensions, in this work, they were generated with the main purpose of understanding the contribution of additional dimensions to the accuracy of the mapping as well as to obtain insights on the complexity of the interpolation to/from the *ab* plane and spaces with more than two dimensions.

The same mesh used in the dimensionality analysis was used to obtain Euclidean mappings in two and three dimensions. For clarity, Figure 2 only presents a simplified version of the original mesh. To illustrate the three-dimensional mapping, two different views of the three-dimensional mesh are shown in Figure 3. The performance of the two-dimensional mapping is also illustrated in Figure 4, where unity color difference ellipses are mapped from the *ab* plane to the new Euclidean space.

The relative error using Formula (14) was calculated for two million pairs of points randomly generated in the *ab* plane, where the Euclidean distance between each pair of points was less than 5. The relative error of this set of points was calculated for two

and three dimensions with interpolations based on the CGAL library for the 16,220-point mesh. For comparison, we applied the same interpolation and set of testing points to the mapping data by Urban et al. [7] obtained online at [26]. The mean and maximum errors for these mappings are shown in Table 1.

Mesh	Dim.	Mean Error	Max. Error
Mesh 16,220	2	3.98%	26.5%
Mesh 16,220	3	0.88%	49.5%
Urban et al.	2	4.96%	154.0%

Table 1: Mean and maximum errors for different mappings

Conclusions

We studied the mapping of the CIELAB space, particularly the *ab* plane, according to the metrics defined by the CIEDE2000 formula, focusing on multidimensional scaling techniques. We performed a dimensionality analysis of the Euclidean mapping of the *ab* plane and investigated the mappings of this plane in two and three dimensions.

The dimensional analysis revealed that, although the mapping of the *ab* plane according to the metrics of the CIEDE2000 function is essentially a two-dimensional space, the first two dimensions capture only about 89% of the variance. To capture 95%, for instance, about eight dimensions would be necessary. This result, in our opinion, may be another evidence of the inherent complexity of the human perception of colors and could become an input for the creation of more accurate formulas, in line with the significant academic interest on this subject.

Regarding the two-dimensional mapping, our approach had achieved better accuracy than the work by Urban et al. [7] in terms of average and maximum errors. In the three-dimensional mapping, the accuracy improvement in the mean error was significant, but we believe that there is a limited number of applications requiring such precision.

We found that Euclidean mappings according to metrics of the CIEDE2000 formula between the ab plane and threedimensional Euclidean spaces result in some isolated areas where the third dimension is not a function of the first two dimensions. The presence of these regions makes the inverse mapping from three-dimensional space to the ab plane a nontrivial task, which is only worthwhile if the gains in precision are relevant in face of the precision requirements of the application. We plan to better understand the precision requirements in a future research, by studying the Euclidean mapping obtained directly from the psychophysics data used in the construction of the CIEDE2000 formula, as suggested by one of our reviewers.

As the regions that hinder inverse mapping are small, another future work is the creation of an optimization function with a restriction ensuring that the third dimension is always a function of the first two dimensions. This optimization has the potential to achieve an accuracy level similar to the one obtained with the three-dimensional mapping presented herein, but without imposing difficulties to the inverse mapping of the *ab* plane.

Finally, we believe that the generation of a 3D surface that is isometric to the *ab* plane according to the metrics of the CIEDE2000 formula was another significant contribution of this work. The simple three-dimensional visualization of the original surface shows sharp bends and a rugged relief that will be useful for future analyses of this color difference formula or for the generation of new, improved formulas.

Acknowledgments

This work was partially funded by the CNPq and had the support of Tecgraf/PUC-Rio. Tecgraf is mainly funded by PETRO-BRAS.

References

- Committee of the Society of Dyers and Colorists. BS 6923: Method for calculation of small color differences. Technical Report BS 6923, British Standards Institution, 1988.
- [2] CIE. Industrial color-difference evaluation. Technical Report CIE Publication No. 116, Bureau Central de la CIE, 1995.
- [3] M R Luo, G Cui, and B Rigg. The development of the CIE 2000 colour-difference formula: CIEDE2000. Color Research & Application, 24(5):340--350, 2001.
- [4] DIN. Farbmetrische bestimmung von farbabständen bei körperfarben nach der DIN99-formel. Technical Report 6176, Deutsches Institut für Normung e.V, 2000.
- [5] G. Cui, M. R. Luo, B. Rigg, G. Roesler, and K. Witt. Uniform colour spaces based on the DIN99 colour-difference formula. *Color Research & Application*, 27(4):282--290, 2002.
- [6] Hans G. Völz. Euclidization of the first quadrant of the CIEDE2000 color difference system for the calculation of large color differences. *Color Research & Application*, 31(1):5--12, 2006.
- [7] Philipp Urban, Mitchell R. Rosen, Roy S. Berns, and Dierk Schleicher. Embedding non-Euclidean color spaces into Euclidean color spaces with minimal isometric disagree-

ment. Journal of the Optical Society of America, 24(6):1516--1528, 2007.

- [8] Philipp Urban, Mitchell R. Rosen, and Roy S. Berns. Constructing Euclidean color spaces based on color difference formulas. In *IS&T/SID 15th Color Imaging Conference* (*CIC*), Albuquerque, New Mexico, 2007. Society of Imaging Science and Technology (IS&T) and the Society for Information Display (SID).
- [9] Elzbieta Pekalska and Robert P. W. Duin. The Dissimilarity Representation for Pattern Recognition: Foundations And Applications (Machine Perception and Artificial Intelligence). World Scientific Publishing Co., Inc., River Edge, NJ, USA, 2005.
- [10] Trevor F. Cox, Michael A. A. Cox, and Trevor F. Cox. *Multidimensional Scaling, Second Edition*. Chapman & Hall/CRC, September 2000.
- [11] J. B. Tenenbaum, V. Silva, and J. C. Langford. A Global Geometric Framework for Nonlinear Dimensionality Reduction. SCIENCE, 290(5500):2319--2323, 2000.
- [12] John W. Sammon. A nonlinear mapping for data structure analysis. *IEEE Transactions on computers*, 18(5):401--409, 1969.
- [13] A. Kimball Romney. Relating reflectance spectra space to Munsell color appearance space. *Journal of the Optical Society of America*, 25(3):658--666, 2008.
- [14] Eric W. Farmer, Robert M. Taylor, and Andrew J. Belyavin. Large color differences and the geometry of Munsell color space. *Journal of the Optical Society of America*, 70(2):243--245, 1980.
- [15] Ethan D. Montag and Hirokazu Kasahara. Multidimensional analysis reveals importance of color for image quality. In *IS&T/SID 9th Color Imaging Conference (CIC)*, pages 17--21, Scottsdale, Arizona, United States, 2001. Society of Imaging Science and Technology (IS&T) and the Society for Information Display (SID).
- [16] Sam T. Roweis and Lawrence K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *SCIENCE*, 290:2323--2326, 2000.
- [17] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308--313, January 1965.
- [18] Donald B. Johnson. Efficient algorithms for shortest paths in sparse networks. *Journal of ACM*, 24:1--13, January 1977.
- [19] Robert W. Floyd. Algorithm 97: Shortest path. *Communications of ACM*, 5:345--, June 1962.
- [20] Stephen Warshall. A theorem on boolean matrices. *Journal* of ACM, 9:11--12, January 1962.
- [21] Gale Young and A. Householder. Discussion of a set of points in terms of their mutual distances. *Psychometrika*, 3(1):19--22, March 1938.
- [22] Thomas Coleman, Mary Ann Branch, and Andrew Grace. Optimization Toolbox For Use with MATLAB User's Guide Version 5. The Mathworks, Inc, 2010.
- [23] Tomas Akenine-Möller, Eric Haines, and Natty Hoffman. *Real-Time Rendering 3rd Edition*. A. K. Peters, Ltd., Natick, MA, USA, 2008.
- [24] A. J. Chung and A. J. Field. A simple recursive tessellator for adaptive surface triangulation. *Journal of Graphics*,

GPU, and Game Tools, 5(3):1--9, 2000.

- [25] Lutz Kettner and Stefan N\u00e4her. Two computational geometry libraries: LEDA and CGAL. In Jacob E. Goodman and Joseph O'Rourke, editors, *Handbook of Discrete and Computational Geometry*, chapter 64, pages 1435--1463. CRC Press LLC, Boca Raton, FL, second edition, 2004.
- [26] Philipp Urban. Supplemental material: Embedding non-Euclidean color spaces into Euclidean color spaces with minimal isometric disagreement. http://munsell.cis. rit.edu/~pmupci.

Author Biography

Lorenzo Ridolfi received his B.S. degree in Computer Engineering and his M.S. degree in Computer Science from Pontificia Universidade Católica do Rio de Janeiro in 1991 and 1995 respectively. He had worked with digital images since 1991 and since 2001, also with color control applications. Since 2007, he is a D.Sc. student in Pontificia Universidade Católita do Rio de Janeiro, with emphasis in digital image processing and color science.