New Fuzzy Color Clustering Algorithm Based on *hsl* Similarity

Vasile Pătraşcu

Department of Informatics Technology, Tarom Company
Bucharest, Romania
Email: patrascu.v@gmail.com

Abstract*—In this paper, one presents a fuzzy color clustering algorithm that is based on a new measure of similarity. This new measure of color similarity is defined on a perceptual color system called hsl.*

Keywords— Color similarity, fuzzy color clustering, *hsl* perceptual color system.

1 Introduction

The color clustering algorithms play an important role in image analysis. The obtained results are very dependent on the coordinate system and similarity/dissimilarity functions used for color separation [1], [4], [5] and [6]. In this paper, one proposes the use of a perceptual system hsl. In the framework of this system, a new measure is proposed for the color similarity. In the following sections, the paper is thus organized: section 2 presents the system hsl for color representation based on new measures of luminosity L and saturation S; section 3 presents distances and similarities for each component of the hsl system; section 4 presents the new color similarity; section 5 presents the color clustering algorithm based on the proposed similarity; section 6 presents the experimental results and conclusions are in section 7.

2 The perceptual color system *hsl*

The most part of the color images are represented by the RGB color system. Image analysis is done in the perceptual coordinate systems. One of them is the hSL system where h is the hue and an angular value; S is the saturation and lastly, L is the luminosity. In this paper, we will use the following formulae for the definition of the hSL components:

$$\begin{cases} L = \dfrac{M}{1+M-m} \\ S = \dfrac{2\cdot(M-m)}{1+|M-0.5|+|m-0.5|} \\ h = \operatorname{atan2}\left(\dfrac{B-G}{\sqrt{2}}, \dfrac{2R-B-G}{\sqrt{6}}\right) \end{cases} \tag{1}$$

where:

$$M = \max(R,G,B)$$
$$m = \min(R,G,B)$$

We suppose that $R,G,B \in [0,1]$. According to the considered formulae for the definition of the components hSL, it results that $L,S \in [0,1]$ and $h \in (-\pi,\pi]$. In order to have a unitary consideration for the three color components, hSL, we will scale the luminosity L and the saturation S from $[0,1]$ to $\left[0,\frac{\pi}{2}\right]$ by formulae:

$$\begin{cases} l = \dfrac{\pi}{2}\cdot L \\ s = \dfrac{\pi}{2}\cdot S \end{cases} \tag{2}$$

Thus, any color q from RGB space will have the following representation in the hsl space:

$$\begin{cases} l = \dfrac{\pi}{2}\cdot\dfrac{M}{1+M-m} \\ s = \pi\cdot\dfrac{M-m}{1+|M-0.5|+|m-0.5|} \\ h = \operatorname{atan2}\left(\dfrac{B-G}{\sqrt{2}}, \dfrac{2R-B-G}{\sqrt{6}}\right) \end{cases} \tag{3}$$

3 Distances and similarities for the values *hsl*

We know that for two angular values α,β a good measure for distance is based on the *sin* function, namely:

$$d(\alpha,\beta) = \left|\sin\left(\frac{\alpha-\beta}{2}\right)\right| \tag{4}$$

Thus, for the hue, the distance is given by the function:

$$d_h(h_1,h_2) = \left|\sin\left(\frac{h_1-h_2}{2}\right)\right| \tag{5}$$

Using the formula (4) for the saturation s and luminosity l we obtain their values in the interval $\left[0,\frac{1}{\sqrt{2}}\right]$, because $l,s \in \left[0,\frac{\pi}{2}\right]$. Consequently, we will multiply the formula (4) with the factor $\sqrt{2}$ and thus we will obtain the values in

ISBN: 978-989-95079-6-8

the whole interval $[0,1]$. In this way, the distances for luminosity and saturation will be defined by:

$$d_l(l_1,l_2)=\sqrt{2}\cdot\left|\sin\left(\frac{l_1-l_2}{2}\right)\right| \quad (6)$$

$$d_s(s_1,s_2)=\sqrt{2}\cdot\left|\sin\left(\frac{s_1-s_2}{2}\right)\right| \quad (7)$$

In this paper, we will take into account that the square of distance is a good measure of dissimilarity and the negation of dissimilarity is a good measure of similarity. In other words, if d is the distance and σ is the similarity, then the following relation exists between them:

$$\sigma=1-d^2 \quad (8)$$

Having the distance definitions, we will then define the similarity function for luminosity and saturation,

$$\begin{cases}\sigma_l(l_1,l_2)=1-d_l^2(l_1,l_2)=\cos(l_1-l_2)\\ \sigma_s(s_1,s_2)=1-d_s^2(s_1,s_2)=\cos(s_1-s_2)\end{cases} \quad (9)$$

and next for the hue:

$$\sigma_h(h_1,h_2)=1-d_h^2(h_1,h_2)=\cos^2\left(\frac{h_1-h_2}{2}\right) \quad (10)$$

4 The *hsl* color similarity

We will initially define the chromaticity c and the achromaticity a for a color having as saturation the value s:

$$\begin{cases}c=\sin(s)\\ a=\cos(s)\end{cases} \quad (11)$$

The following relation between these two defined parameters is obvious:

$$a^2+c^2=1 \quad (12)$$

We will consider two colors q_1,q_2 defined by parameters $(h_1,s_1,l_1),(h_2,s_2,l_2)$. We will add to these parameters the chromaticity coefficients c_1,c_2 and achromaticity coefficients a_1,a_2, computed by formulae (11). Now, we define the similarity between two colors by formula:

$$\sigma(q_1,q_2)=c_1c_2\cdot\sigma_h(h_1,h_2)+a_1a_2\cdot\sigma_l(l_1,l_2) \quad (13)$$

Seeing the formula (13), one remarks that this new color similarity measure is a linear and adaptive combination, between the hues similarity and luminosities similarity. Thus for two chromatic colors, the main component is given by the hues similarity and for two achromatic colors the main component is given by the luminosities similarity. We can see that, when a color is chromatic and the other is achromatic, the similarity has small values. This is nothing else than the two considered colors are quite non-similar. Also, we can see that in formula (13), the saturations similarity does not appear directly, but it exists somewhere in background. The formula (13) has the following equivalent form:

$$\sigma(q_1,q_2)=\sin(s_1)\cdot\sin(s_2)\cdot\cos^2\left(\frac{h_1-h_2}{2}\right)+ \\ +\cos(s_1)\cdot\cos(s_2)\cdot\cos(l_1-l_2) \quad (14)$$

Taking into account that:

$$\begin{cases}\cos^2\left(\frac{h_1-h_2}{2}\right)\le 1\\ \cos(l_1-l_2)\le 1\end{cases} \quad (15)$$

it results from (14) this inequality:

$$\sigma(q_1,q_2)\le\sin(s_1)\cdot\sin(s_2)+\cos(s_1)\cdot\cos(s_2)$$

namely

$$\sigma(q_1,q_2)\le\cos(s_1-s_2)$$

and finally we obtain that:

$$\sigma(q_1,q_2)\le\sigma_s(s_1,s_2) \quad (16)$$

The formula (16) shows that if the saturations of the two colors have a low similarity then the two colors have a low similarity too and if the two colors are strongly similar then their saturations are strongly similar as well. Also, from (15) and (16) it results the following implications:

$$\sigma(q_1,q_2)=1 \;\Rightarrow\; \begin{cases}\sigma_h(h_1,h_2)=1\\ \sigma_l(l_1,l_2)=1\\ \sigma_s(s_1,s_2)=1\end{cases} \;\Rightarrow\; \begin{cases}h_1=h_2\\ l_1=l_2\\ s_1=s_2\end{cases}$$

Using the formula (8) we can define the distance between two colors in the *hsl* space as:

$$d(q_1,q_2)=\sqrt{1-\sigma(q_1,q_2)} \quad (17)$$

We can state that formula (17) defines a metric because it verifies the following three metric properties:

- $d(q_1,q_2)=0 \;\Leftrightarrow\; q_1=q_2$
- $d(q_1,q_2)=d(q_2,q_1)$
- $d(q_1,q_2)+d(q_2,q_3)\ge d(q_1,q_3)$

A detailed proof of these three properties is not a subject of this paper and thus not considered here.

ISBN: 978-989-95079-6-8

5 The fuzzy color clustering algorithm

We consider n colors $q_1, q_2, ..., q_n$ that must be separated into k clusters. Each cluster i is characterized by the membership coefficients $w_{i1}, w_{i2}, ..., w_{in}$ for the considered n colors and the cluster center defined by the color $\overline{q}_i = \left(\overline{h}_i, \overline{s}_i, \overline{l}_i\right)$. For color clustering we will construct an algorithm that is similar to the fuzzy c-means algorithm [1]. We will consider the following objective function:

$$J = \sum_{i=1}^{k} \sum_{j=1}^{n} w_{ij}^{\alpha} \cdot \sigma\left(\overline{q}_i, q_j\right) \quad (18)$$

where α is a fuzzification-defuzzification parameter and also, $\alpha \in (0,1)$. If α approaches 1, then the fuzzy algorithm approaches a crisp one. The objective function J (18) must be maximized. In order to have fuzzy partitions, we must add the following conditions, for $j = 1,2,...,n$:

$$w_{1j} + w_{2j} + ... + w_{kj} = 1 \quad (19)$$

In this case, considering for the condition (19) the Lagrange multipliers $\lambda_1, \lambda_2, ..., \lambda_n$, the objective function (18) becomes:

$$J = \sum_{i=1}^{k} \sum_{j=1}^{n} w_{ij}^{\alpha} \cdot \sigma\left(\overline{q}_i, q_j\right) + \sum_{j=1}^{n} \lambda_j \cdot \left(\sum_{i=1}^{k} w_{ij} - 1 \right) \quad (20)$$

The maximum value of the objective function J (20) results from the following conditions:

- $\forall i \in [1,k], \forall j \in [1,n]$,

$$\frac{\partial J}{\partial w_{ij}} = 0 \quad (21)$$

It results:

$$w_{ij} = \frac{\left(\sigma(\overline{q}_i, q_j)\right)^{\frac{1}{1-\alpha}}}{\sum_{m=1}^{k} \left(\sigma(\overline{q}_m, q_j)\right)^{\frac{1}{1-\alpha}}} \quad (22)$$

- $\forall i \in [1,k]$,

$$\frac{\partial J}{\partial \overline{h}_i} = 0 \quad (23)$$

It results:

$$\overline{h}_i = \operatorname{atan2}\left(\frac{\sum_{j=1}^{n} u_{ij} \sin\left(h_j\right)}{\sum_{j=1}^{n} u_{ij}}, \frac{\sum_{j=1}^{n} u_{ij} \cos\left(h_j\right)}{\sum_{j=1}^{n} u_{ij}} \right) \quad (24)$$

where

$$u_{ij} = w_{ij}^{\alpha} \cdot c_j \quad (25)$$

- $\forall i \in [1,k]$,

$$\frac{\partial J}{\partial \overline{l}_i} = 0 \quad (26)$$

It results:

$$\overline{l}_i = \operatorname{atan2}\left(\frac{\sum_{j=1}^{n} v_{ij} \sin\left(l_j\right)}{\sum_{j=1}^{n} v_{ij}}, \frac{\sum_{j=1}^{n} v_{ij} \cos\left(l_j\right)}{\sum_{j=1}^{n} v_{ij}} \right) \quad (27)$$

where

$$v_{ij} = w_{ij}^{\alpha} \cdot a_j \quad (28)$$

- $\forall i \in [1,k]$,

$$\frac{\partial J}{\partial \overline{s}_i} = 0 \quad (29)$$

It results:

$$\overline{s}_i = \operatorname{atan2}\left(\frac{\sum_{j=1}^{n} x_{ij} \sin\left(s_j\right)}{\sum_{j=1}^{n} w_{ij}^{\alpha}}, \frac{\sum_{j=1}^{n} y_{ij} \cos\left(s_j\right)}{\sum_{j=1}^{n} w_{ij}^{\alpha}} \right) \quad (30)$$

where

$$\begin{cases} x_{ij} = w_{ij}^{\alpha} \cdot \sigma_h\left(\overline{h}_i, h_j\right) \\ y_{ij} = w_{ij}^{\alpha} \cdot \sigma_l\left(\overline{l}_i, l_j\right) \end{cases} \quad (31)$$

6 Experimental results

The proposed algorithm was applied to the image „flower” (Fig. 1) and the clustered image can be seen in Fig. 2. For comparison, the fuzzy c-means algorithm [1], [3], [8] was applied to image „flower”, in the coordinate systems RGB [3], [9], Lab [7], Luv [3], [9], hSL [9], I1I2I3 [2], [3] and the results can be seen in Figs. 3, 4, 5, 6 and 7.

Figure 1: The image „flower”.

ISBN: 978-989-95079-6-8

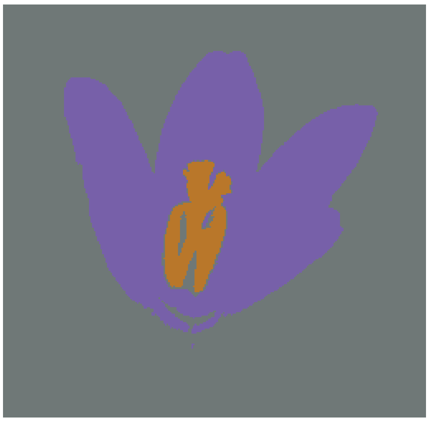

Figure 2: The image „flower" clustered with the proposed algorithm and *hsl*.

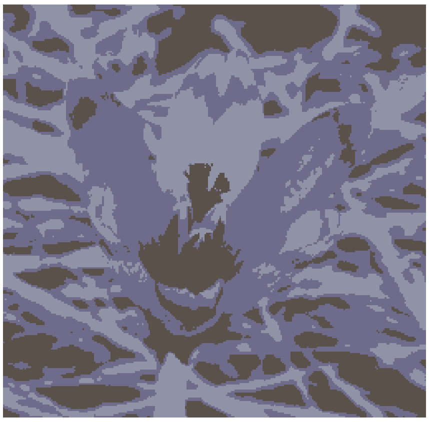

Figure 3: The image „flower" clustered with FCM algorithm and *RGB*.

Figure 4: The image „flower" clustered with FCM algorithm and *Lab*.

Figure 5: The image „flower" clustered with FCM algorithm and *Luv*.

Figure 6: The image „flower" clustered with FCM algorithm and *hSL*.

Figure 7: The image „flower" clustered with FCM algorithm and *I1I2I3*.

7 Conclusions

In this paper, a new measure for color similarity was defined. This new measure was defined in a new perceptual system *hsl*. Using this measure, an algorithm for color clustering was constructed.

In the section dedicated to the experimental results, one can see the advantage of using this new color similarity measure.

References

[1] J.C. Bezdek. Pattern Recognition with Fuzzy Objective Functions. New York: Plenum Press, 1981.

[2] Y.I. Ohta, T. Kanade and T. Sakai. Color information for region segmentation, *Computer and Image Processing*, vol. 13, pp.222-241, 1980.

[3] Y.W. Lim and S.U. Lee. On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques, *Pattern Recognition*, vol. 23, no.9 , pp.935-952, 1990.

ISBN: 978-989-95079-6-8

[4] V. Pătraşcu. A generalization of fuzzy c-means algorithm using a new dissimilarity function, The 11th International Fuzzy Systems Association World Congres, IFSA 2005, *Fuzzy Logic, Soft Computing and Computational Intelligence*, Vol II, pp. 591-596, Beijing China July, 28-31,2005.

[5] V. Pătraşcu. A Generalization of Gustafson-Kessel Algorithm Using a New Constraint Parameter*, In Proceedings of the 4th Conference of the European Society for Fuzzy Logic and Technology and 11th Rencontres Francophones sur la Logique Floue et ses Applications*, EUSFLAT-LFA 2005, pp.1250-1255, Barcelona, Spain 7-9 September, 2005.

[6] V. Pătraşcu. Fuzzy Image Segmentation Based on Triangular Function and Its n-dimensional Extension, in the volume *Soft Computing in Image Processing. Recent Advances, Series: Studies in Fuzziness and Soft Computing*, pp 187-208, Vol. 210 , (Eds.) M. Nachtegael, D. Van der Weken, E. E. Kerre, W. Philips, ISBN: 3-540-38232-1, Springer-Verlag, 2007.

[7] S. Tominga. A colour classification method for color images using a uniform color space, *Proc. 10th Int. Conf. On Pattern Recognition*, vol. I, Atlantic City, new Jersey, pp. 803-807, 16-21 June 1990.

[8] M. Trivedi and J.C. Bezdek. Low-level segmentation of aerial images with fuzzy clustering, *IEEE Trans. On Systems, Man, and Cybernetics*, vol. 16, no. 4, pp.589-598, 1986.

[9] S.E. Umbaugh, R.H. Moss, W.V. Stoecker and G.A. Hance. Automatic colour segmentation algorithms with application to skin tumor feature identification, IEEE Engineering in Medcine and Biology, vol. 12, no. 3, pp. 75-82, 1993.

ISBN: 978-989-95079-6-8