Clustering Hierarchical Fuzzy Colors and its Application on Image segmentation

Been-Chian Chien and Ming-Cheng Cheng
Department of Information Engineering
I-Shou University
cbc@isu.edu.tw

Abstract

The purpose of this research is to analyze the low-level features of images as a representation of semantic concept. This paper proposes a flexible approach for clustering features of images and mapping the low-level image features to the high-level concept recognition. Owing to the uncertainty of image features for human’s recognition, we use the approach of fuzzy colors clustering to analyze image features based on fuzzy entropy. The proposed approach first analyzes color features using HLS color space to determine the best number of fuzzy colors and cluster the colors into a hierarchical concept for human’s recognition. We also apply the clustered fuzzy colors to segment out meaningful regions from an image automatically. In this paper, the application of fuzzy colors on segmentation is demonstrated and compared with other methods. The experimental results show that the proposed method can extract meaningful regions from images as effectively as human visual perception.

Keywords: image analysis, features clustering, fuzzy entropy, similarity measure.

1. Introduction

Distinguishing and identifying an object from a picture or an image is easy for human. However, even a fast computer cannot recognize a human face correctly and effectively. To realize how human can recognize object easily, we try to cluster pixels with similar features as a region in an image. Clustering image features is a process of partitioning images with homogeneous or meaningful representation for users. The image low-level features usually include colors, shapes, textures and spatial relationships. However, it is difficult to understand for human’s recognition in such low-level features. For clustering image features as human’s concept recognition, we propose a clustering approach based on fuzzy entropy. This approach can be used to develop hierarchical clustering algorithms for various low-level features. We will show the approach by clustering of color feature and apply the clustering result to segment out meaningful regions for human’s recognition.

A color image is usually represented by the RGB color coordinate system, and most of the traditional segmentation algorithms for color images use Minkowski’s $r$-metric [11], e.g. Euclidean distance, to compute the similarity of colors in the process of segmentation. However, the RGB color coordinate system and Euclidean metric cannot model human visual perception on colors. Human eyes recognize colors by hue, lightness and saturation and are not as sensitive as machine on recognizing colors. That is the reason why the segmented results of traditional segmentation algorithms are usually far inferior to human visual perception and unable to capture high-level concepts of real objects appropriately.

This paper is organized as follows. We briefly introduce the HLS color space in Section 2. Section 3 depicts the proposed clustering algorithm of a low-level feature. In Section 4, we define hierarchical fuzzy color and design a fuzzy similarity measure. Section 5 gives the detailed algorithm of image segmentation based on the feature space analysis and fuzzy colors similarity measure. The experimental results are shown in Section 6. Finally, conclusions are summarized in Section 7.

2. The HLS Color Coordinate System

The color coordinate system we used for measuring similarity of colors is the HLS [8]. Hue is the attribute of a visual sensation according to which an area appears to be similar to one of the perceived colors. Saturation is the relative purity of hue. Lightness is the intensity of color components red, green and blue. Given a color represented by the RGB color coordinate system, $(R, G, B)$, the values of $R, G$, and $B$ are all integers in the range of $[0, 255]$ and represent the values of red, green and blue, respectively. We define the relative intensity value $(r, g, b)$ for $(R, G, B)$ as:

$$r = \frac{R}{255}, g = \frac{G}{255}, b = \frac{B}{255}.$$ 

Let $I_{\text{max}} = \max\{r, g, b\}$ and $I_{\text{min}} = \min\{r, g, b\}$. The values of hue(H), lightness(L) and saturation(S) can be transformed from the values of $R, G$ and $B$ by the following formulas:

---

* This research was supported in part by the National Science Council of Taiwan, R.O.C under Grant NSC-91-2213-E-214-018
Let universal set \( U \) be defined as
\[
\mathcal{U} \equiv \{ r_1, r_2, \ldots, r_n \}
\]
(1)

Then, we assign a corresponding membership function to each interval. Applying fuzzy entropy to each data, we can determine a best membership function to each interval. Applying several intervals. Then, we assign a corresponding membership function to each interval. Applying fuzzy entropy to each data, we can determine a best number of intervals. The fuzzy entropy is defined as follows.

**Definition 1:** Fuzzy entropy of each feature dimension:

1. Let universal set \( \mathcal{U} \) be the support of a feature space and let \( N_m \) be the number of support elements \( r_i \) where \( i = 1, 2, \ldots, n \).
2. Let \( n \) elements of support set \( R \) be divided into \( j \) intervals named \( C_1, C_2, \ldots, C_j \).
3. Let \( C_j(r) \) denote a set of elements of interval \( j \) on the support set \( R \).
4. Let feature space, where support set \( R \) distributed, be defined \( k \) fuzzy set named \( A_l \) where \( l = 1, 2, \ldots, k \). For an element \( r_i \) of fuzzy set \( A_l \) the membership degree is denoted by \( \mu_{A_l}(r_i) \).
5. The fuzzy entropy \( FE_{A_l} \) of a fuzzy set \( A_l \) is defined as
\[
FE_{A_l} = - \sum_{j=1}^{k} \sum_{m \in C_j(r)} N_m \times \mu_{A_l}(m) \times \log_2 \mu_{A_l}(m)
\]
(5)

where \( N_m \) is the number of occurrence of element \( m \).
6. The fuzzy entropy \( FE \) on the feature space with support set \( R \) is defined as
\[
FE = \sum_{i=1}^{k} FE_{A_i}
\]
(6)

In equation (5), the fuzzy entropy \( FE_{A_i} \) is dependent on membership grade but not probabilistic entropy.

### 3.2 Membership Functions of Intervals

We use a triangular form membership function to represent each fuzzy set. There are two kinds of different environments: acyclic and cyclic. In the acyclic environment, the left-most and right-most elements we assign the value 0.5. The membership grade for an element \( x \) belonging to a fuzzy set is defined as follows:

**Case 1:** The left-most case:
\[
\mu(x) = \begin{cases} 
1 - \frac{|c_1 - x|}{2c_1}, & \text{for } x \leq c_1 \\
\max(0,1 - \frac{|x - c_1|}{c_2 - c_1}), & \text{for } x > c_1
\end{cases}
\]
(7)

**Case 2:** The internal case:
\[
\mu(x) = \begin{cases} 
\max(0,1 - \frac{|x - c_1|}{c_1 - c_{i-1}}), & \text{for } x \leq c_i \\
\max(0,1 - \frac{|x - c_1|}{c_1 - c_{i+1}}), & \text{for } x > c_i
\end{cases}
\]
(8)

**Case 3:** The right-most case:
\[
\mu(x) = \begin{cases} 
\max(0,1 - \frac{|x - c_1|}{r_n - c_j}), & \text{for } x \leq c_j \\
1 - \frac{|x - c_1|}{2x(r_n - c_j)}, & \text{for } x > c_j
\end{cases}
\]
(9)

We assign the highest membership value 1.0 to the center of an interval and the lowest value 0.0 to the centers of this interval’s neighbors in the cyclic environment. The membership grade for an element \( x \) belonging to a fuzzy set is defined as follows:

**Case 1:** The left-most case:
\[
\mu(x) = \begin{cases} 
\max(0,1 - \frac{|x - c_1|}{r_n - c_j + c_1}), & \text{for } x \leq c_1 \\
\max(0,1 - \frac{|x - c_1|}{c_1 - c_{i-1}}), & \text{for } x > c_1
\end{cases}
\]
(10)

**Case 2:** The internal case:
\[
\mu(x) = \begin{cases} 
\max(0,1 - \frac{|x - c_1|}{c_1 - c_{i-1}}), & \text{for } x \leq c_i \\
\max(0,1 - \frac{|x - c_1|}{c_1 - c_{i+1}}), & \text{for } x > c_i
\end{cases}
\]
(11)

**Case 3:** The right-most case:
\[
\mu(x) = \begin{cases} 
\max(0,1 - \frac{|x - c_1|}{c_1 - c_{i-1}}), & \text{for } x \leq c_j \\
\max(0,1 - \frac{|x - c_1|}{r_n - c_j + c_1}), & \text{for } x > c_j
\end{cases}
\]
(12)
3.3 Feature Clustering Algorithm

In the proposed system, each pixel of an image is transformed from RGB color space into HLS color space. Then we produce the histogram of an image on each 1-D HLS color space and obtain hue histogram, lightness histogram, and saturation histogram. We then use clustering technique and fuzzy entropy on each obtained 1-D histogram to analyze the information of pattern distribution. For each 1-D histogram, the clustering algorithm is described as follows:

**Algorithm:** 1-D histogram clustering algorithm  
**Input:** A 1-D histogram.  
**Output:** The fuzzy sets of hue, lightness, saturation.

1. Set the initial number of intervals \( I = 2 \).
2. \( \textbf{do} \)
3. Set the initial center of each interval.
4. \( \textbf{do} \)
5. Assign the membership function for each interval.
6. Assign interval label to each element.
7. Re-computing the interval centers.
8. \( \textbf{while} \) ( Each interval center does not change )
9. Compute fuzzy entropy \( FE(I) \) for \( I \) intervals.
10. \( \textbf{if} \) \( ( FE(I) < FE(I-1) ) \)
11. \( I \leftarrow I + 1 \).
12. \( \textbf{while} \) \( ( FE(I) < FE(I-1) ) \)

In line 1, we set initial number of intervals \( I = 2 \). In line 3 we set initial centers of each interval. The initial interval centers \( c_i \) can be randomly selected where \( i = 1, 2, ..., I \). In the proposed system, the interval center \( c_i \) is assigned as follows:

\[
c_i = \frac{i-1}{I}, \quad i = 1, 2, ..., I.
\]  

In line 5, we assign a corresponding membership function for each interval. The assignment of membership functions is detailed in subsection 3.2. After determining the interval centers and assigning membership functions, line 6 assigns element \( x \) to an interval label which has the highest membership degree. Since the initial centers are selected randomly, we have to reevaluate each center in line 7. In line 4 to 8, we refine the interval centers repeatedly until that every center of interval is not changed. After all interval centers are determined, line 9 computes the total fuzzy entropy of all \( I \) intervals. In line 10 and 11, we compare total fuzzy entropy of all \( I \) intervals \( FE(I) \) with all \( I-1 \) intervals \( FE(I-1) \). When \( FE(I) < FE(I-1) \), it means that the \( I \) interval’s division is better than the \( I-1 \) interval’s. In this case, we have to repeat entire algorithm to know whether \( I+1 \) intervals will be better than \( I \) intervals. When \( FE(I) > FE(I-1) \), the \( I \) interval division is worse than the \( I-1 \) interval division, we stop the algorithm. After the feature clustering, we obtain \( I-1 \) intervals, the center of each interval, and their membership functions.

4. Analysis of Hierarchical Fuzzy Colors

Kawamura et al. [13] utilized color attributes hue and tone to design a fuzzy based approach for color specification. Sugano [14] made use of color attributes including hue and tone to design expression model of subjective color impressions. The usage of color attributes hue and tone to reflect human’s subjectivity is easier than the RGB color space system. Hence, we define fuzzy colors based on the combination of hue and tone to construct fuzzy colors for achieving human visual perception. Based on the feature clustering of Section 3, we got the fuzzy set of hue, fuzzy set of lightness, and fuzzy set of saturation. Then we use the fuzzy set of hue, fuzzy set of lightness, and fuzzy set of saturation to construct the set of hierarchical fuzzy colors. Fuzzy colors are composed by fuzzy set of hue and fuzzy set of tone. We combine fuzzy set of lightness and fuzzy set of saturation into the fuzzy set of tone. After that, the fuzzy set of hue is combined with the fuzzy set of tone to construct fuzzy colors.

4.1 The Fuzzy Set of Hue

Let \( H = \{ H_i \mid i = 1, 2, ..., h \} \) be the set of hue obtained by feature clustering in Section 3. For a hue \( H \) the fuzzy set of hue \( \tilde{H} \) is represented as

\[
\tilde{H} = (\mu_{H_i}(x) / H_i + \mu_{H_2}(x) / H_2 + \ldots + \mu_{H_h}(x) / H_h)
\]

\[= \sum_{i=1}^{h} \mu_{H_i}(x) / H_i, \tag{14}\]

where \( x \) is the value of the hue \( H \) defined in equations (1) and (2).

4.2 The Fuzzy Set of Tone

Since tone is constructed by lightness and saturation, we first define the fuzzy sets of lightness and saturation and then combine the two fuzzy sets to be the fuzzy set of tone.

Let \( L = \{ L_i \mid i = 1, 2, ..., l \} \) be the set of lightness obtained by feature clustering in Section 3. For a lightness \( L \), the fuzzy set of \( \tilde{L} \) is represented as

\[
\tilde{L} = (\mu_{L_1}(y) / L_1 + \mu_{L_2}(y) / L_2 + \ldots + \mu_{L_l}(y) / L_l)
\]

\[= \sum_{i=1}^{l} \mu_{L_i}(y) / L_i, \tag{15}\]

where \( y \) is the value of the lightness \( L \) defined in equation (3).

Let \( S = \{ S_i \mid i = 1, 2, ..., s \} \) be the set of saturation obtained by feature clustering. For a saturation \( S \), the fuzzy set \( \tilde{S} \) is represented as
\[ S = (\mu_S(z)/S_1 + \mu_{S_2}(z)/S_2 + \ldots + \mu_{S_k}(z)/S_k) \]
\[ \tilde{T} = \sum_{i=1}^{\text{t}} \sum_{j=1}^{\text{t}} \mu_{T_j}(y, z)/T_{ij} \cdot \mu_{S_j}(z) \cdot \mu_{S_j}(z) \]  
(16)

where \( z \) is the value of the saturation \( S \) defined in equation (4).

After the fuzzy sets of lightness and saturation are defined, we use the two fuzzy sets to compose the fuzzy set of tone. Let \( T_{ij} \) be the region composed by \( L_i \) and \( S_j \) inside the tone plane. The fuzzy set of tone \( \tilde{T} \) is represented as

\[ \tilde{T} = \sum_{i=1}^{\text{t}} \sum_{j=1}^{\text{t}} \mu_{T_j}(y, z)/T_{ij} \cdot \mu_{S_j}(z) \cdot \mu_{S_j}(z) \]  
(17)

We further define the membership functions for the fuzzy set of tone as the product of the membership grade of lightness and the membership grade of saturation, as follow:

\[ \mu_{T_j}(y, z) = \mu_{T_j}(y) \cdot \mu_{S_j}(z) \]  
(18)

where \( y \) and \( z \) are the values of lightness and saturation obtained in equations (3) and (4), respectively. \( \mu_{L_i} \) and \( \mu_{S_j} \) are the membership functions of lightness \( L_i \) and saturation \( S_j \), respectively.

### 4.3 Fuzzy Colors

A fuzzy color is constructed by an element of fuzzy set hue and an element of fuzzy set tone. Assume that there are \( h \) elements of fuzzy set hue and \( t \) elements of fuzzy set tone. Let \( C = \{ C_{i} \mid 1 \leq i \leq h \) and \( 1 \leq j \leq t \) \} be the set of fuzzy colors constructed by the \( h \) hues and \( t \) tones. The fuzzy set of fuzzy color \( \tilde{C} \) is represented as

\[ \tilde{C} = \sum_{i=1}^{h} \sum_{j=1}^{t} \mu_{C_{i}}(x, y, z)/C_{ij} \]  
(19)

where \( x \) is defined in equation (1) and (2), \( y \) and \( z \) are defined in equation (3) and (4), respectively.

The membership degree for the fuzzy color are defined as follow

\[ \mu_{C_{i}}(x, y, z) = \mu_{H_i}(x) \cdot \mu_{T_j}(y, z) \]  
(20)

where \( \mu_{H_i} \) is the membership function of hue \( H_i \) and \( \mu_{T_j} \) is defined in equation (18).

### 4.4 Fuzzy Color Similarity Measure

We design a fuzzy similarity measure to evaluate the similarity between two fuzzy colors. The fuzzy similarity measure is based on the membership grades of fuzzy colors. For two given specified colors \( Color_1 = (R_1, G_1, B_1) \) and \( Color_2 = (R_2, G_2, B_2) \), the corresponding values in HLS space are \( C_1 = (h_1, l_1, s_1) \) and \( C_2 = (h_2, l_2, s_2) \), respectively. From equations (19) the fuzzy color \( \tilde{C}_1 \) of color \( C_1 \) and \( \tilde{C}_2 \) of color \( C_2 \) are

\[ \tilde{C}_1 = \sum_{i=1}^{h} \sum_{j=1}^{t} \mu_{C_{i}}(h_1, l_1, s_1)/C_{ij} \]  

\[ \tilde{C}_2 = \sum_{i=1}^{h} \sum_{j=1}^{t} \mu_{C_{i}}(h_2, l_2, s_2)/C_{ij} \]

The fuzzy similarity measure is defined as follow:

\[ \text{Sim}(\tilde{C}_1, \tilde{C}_2) = \frac{\sum_{i=1}^{h} \sum_{j=1}^{t} \min(\mu_{C_{i}}(h_1, l_1, s_1), \mu_{C_{i}}(h_2, l_2, s_2))}{\sum_{i=1}^{h} \sum_{j=1}^{t} \max(\mu_{C_{i}}(h_1, l_1, s_1), \mu_{C_{i}}(h_2, l_2, s_2))} \]  
(21)

The value of \( \text{Sim}(\tilde{C}_1, \tilde{C}_2) \) is in the range of \([0, 1]\).

### 5. The Application on Image Segmentation

We apply the feature clustering algorithm, fuzzy colors and fuzzy color similarity measure to develop an unsupervised image segmentation algorithm. In line 1, we transform the color of each pixel in input image \( I_{in} \) from RGB color space to HLS color space. In line 2, we produce 1-D histograms of hue, lightness, and saturation. In line 3, we execute the feature clustering algorithm on each obtained 1-D histogram. In line 4, we use the results of feature clustering algorithm to construct a set of fuzzy color. When lines 5 to 11 are completed, each pixel in input image will be represented by a fuzzy color with color attribute hue, lightness, and saturation. Initially, each pixel in input image \( I_{in} \) is regarded as a region. In line 16 we apply the fuzzy similarity measure, \( \text{Sim} \in [0, 1] \), to find two adjacent regions with similar color. In line 18 the adjacent regions with similar color are assigned the same region identifier. At last, lines 12 to 20 are repeated to automatically segment out regions in input image.

**Algorithm:** Unsupervised image segmentation algorithm

**Input:** An image \( I_{in} \)

**Output:** A segmented image

1. \( \text{RGB}(I_{in}) \rightarrow \text{HLS}(I_{in}) \)

2. Produce 1-D histogram of hue, lightness, and saturation.

3. Feature space analysis algorithm.  
   // Detailed in subsection 4.3

4. Define a set of fuzzy color.  
   // Detailed in subsection 4.3

5. for \( h \leftarrow 1 \) to \text{height}(I_{in})

6. \{ 

7. for \( w \leftarrow 1 \) to \text{width}(I_{in})

8. \( \tilde{C}(I_{in}(w, h)) \);  // equation (19) & (20)

9. \} 

10. for \( h \leftarrow 1 \) to \text{height}(I_{in})

11. \{ 

12. for \( w \leftarrow 1 \) to \text{width}(I_{in})

13. \{ 

16. Compute \( \text{Sim} \), equation (21), by of pixel \( I_{in}(w, h) \) correspondence to its eight adjacent pixels.

17. If \( \text{Sim} > 0 \) then
Assign the same region id to \( I_w (w, h) \) and its corresponding adjacent pixels.

6. Experimental Results

For demonstrating the effectiveness of clustering fuzzy colors segmentation mechanism, the proposed method of image segmentation in Section 5 was implemented to show and compare the performance. The system is written in C++ language on a personal computer with Intel Celeron 1.2GHz CPU and 256MB RAM. The tested images are in size of 256×256.

The experimental results are shown in Figure 1. The original images are shown in the first row and our experimental results are shown in the second row. We also print out the segmentation results of Lin’s method [7] for giving a comparison in the third row. In the listed image, we use block size 4×4 to down sample the image. A size of 256×256 image is down sampled into size 64×64 image. The segmented results in our experiments show that segmentation of Lin’s method has block effect and is far from the human perception. The segmentation results of the proposed method based on the clustered fuzzy colors are more natural than block segmentation method and more close to human’s recognition.

7. Conclusions

In this paper, we present a clustering algorithm based on fuzzy entropy to analyze the fuzzy colors in an image. The proposed approach tries to map low-level image features into high-level concept with human semantics. For example, people recognize colors by green, red, blue, black, white, brown and so on. The proposed clustering algorithm can automatically analyze the color features in HLS color space and granule them as fuzzy colors. Based on the fuzzy colors, we also develop an image segmentation method to show the effectiveness of the clustering results. Each pixel in an image is assigned a representative fuzzy color. Then, the segmentation algorithm uses a fuzzy similarity measure to evaluate the degree of similarity between two fuzzy colors of pixels and merge the adjacent pixels recursively. The meaningful regions will then be segmented out from an image. The work can be extended to the application other image features such as shapes, texture and spatial relationships.

References

Committee on Data Engineering, Volume 19, Number 4, pp. 19-27, 1996.


<table>
<thead>
<tr>
<th>Source Images</th>
<th>Segmented Results</th>
<th>Boundaries of regions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Our Method</strong></td>
<td>The numbers of regions</td>
<td>5  7  8</td>
</tr>
<tr>
<td></td>
<td>Boundaries of regions</td>
<td></td>
</tr>
<tr>
<td><strong>Lin’s Method</strong></td>
<td>Segmented Results</td>
<td>The number of regions</td>
</tr>
<tr>
<td></td>
<td>Boundaries of regions</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: The results of region segmentation.