

Fuzzy C-Mean: A Statistical Feature Classification of Text and Image Segmentation Method

Somporn Chuai-aree, Chidchanok Lursinsap, Peraphon Sophatsathit, and Suchada Siripant
Advanced Virtual and Intelligent Computing Center (AVIC)
Department of Mathematics
Chulalongkorn University
Bangkok, 10330, Thailand
E-mail: {csomporn | lur | sperapho}@math.sc.chula.ac.th, ssuchada@chula.ac.th

Abstract: Classification of text and image using statistical features (mean and standard deviation of pixel color values) is found to be a simple yet powerful method for text and image segmentation. The features constitute a systematic structure that segregates one from another. We identified this segregation in the form of class clustering by means of Fuzzy C-Mean method, which determined each cluster location using maximum membership defuzzification and neighborhood smoothing techniques. The method can then be applied to classify text, image, and background areas in optical character recognition (OCR) application for elaborated open document systems.

Key words: segmentation, clustering, statistical classification.

1. Introduction

Document often composes of different artifacts ranging from the essential text in the forms of paragraphs, tables, and inserts/sidebars, to various illustrations such as charts, diagrams, and pictures/images. The problem of documentation occurs when editing is called for of the document with the absence of the electronic sources. This is where optical character recognition (OCR) technique comes into play. Dealing with mixture of the aforementioned artifacts poses a big challenge to researchers. As depicted in Figure 1, one must resort to a means for separating image and textual information, known as *image segmentation*.

processing stages to identify text, image, and background objects by means of various well-known techniques such as wavelet transform, segmentation, or feature extraction. This research proposes a method called Fuzzy C-Mean (FCM) which employs two simple and straightforward statistical features, namely, mean and standard deviation of block pixel gray scale level¹ to distinguish those objects from one another. The principles rest on the observation that image pixel colors are lighter than those of background in gray scale level. In addition, every pixel's feature values belonging to the same object block are relatively close to those of its neighbors.

We further observed that the standard deviation of text block pixel was approximately

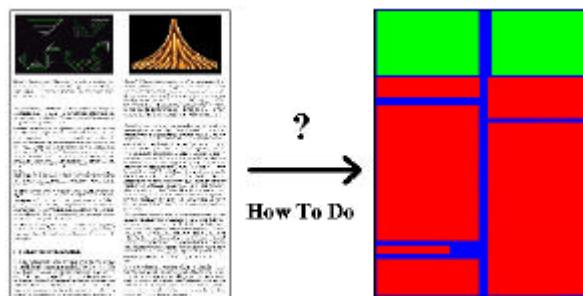


Figure 1: Problems of image segmentation

One of the fundamental principles of conventional image segmentation is the use of attribute characteristics of text, image, and background objects. These distinctions, henceforth referred to as features, are extracted during image

¹ The gray scale level of each pixel under investigation was computed from the corresponding RGB value.

twice as high as the corresponding mean in gray scale. The background pixel, on the contrary, had high mean (bright) and low standard deviation values. Such findings have led to clustering classification of the three objects by virtue of their respective dispersion.

The remaining of this paper will discuss the proposed FCM method in greater detail as follows: Section 2 describes the underlying principles of FCM clustering method. Section 3 establishes the formality of the proposed FCM method. Section 4 elucidates the detail of FCM as apply to image processing. Section 5 shows the experimental results obtained from selected input images. Final thoughts and future work are summarized in the concluding section.

2. Fuzzy C-Mean clustering

A proposed Fuzzy C-Mean (FCM) method is a simple statistical feature comparison of pixel attributes that distinctively characterize the object these pixels constitute. The features employed by the proposed method encompass mean and standard deviation of gray scale measurements of pixel blocks. The values obtained from feature measurement are subject to two basic observations:

1. Image pixel colors are lighter than those of background in gray scale level, and
2. Pixels that differ slightly in mean value or standard deviation are considered belonging to the same object.

Using conventional statistical mean described by the following relation [7]:

$$\bar{p} = \frac{1}{q} \sum_{i=1}^q p_i$$

and the standard deviation

$$s = \sqrt{\frac{1}{q-1} \sum_{i=1}^q (p_i - \bar{p})^2}$$

where q denotes the number of pixels in each block. These statistics were utilized as the feature values of object pixel colors. However, it was found that gray scale feature values offered better discernable results than the RGB counterpart. As such, a color-to-gray scale conversion scheme was devised according to the following straightforward mapping shown in Figure 2.

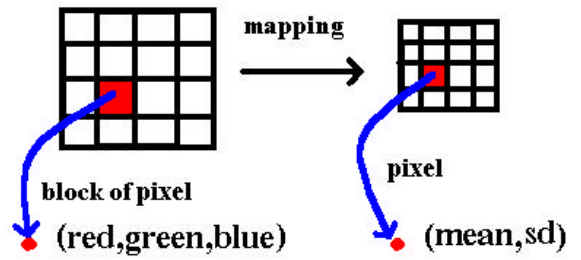


Figure 2: Color-to-gray mapping

In this Figure, a window of proper size is superimposed onto the document in a non-overlapping manner to outline the mapping transformation. Hence, there will be a number of such windows covering the entire document page. An RGB pixel will then be converted to a corresponding gray scale level pair (mean, sd). Details on how mapping is formalized are given in the section that follows.

Based on these observations, we could establish the following classification of pixel clusters for text, image, and background objects as follows:

1. Pixels representing a textual object have distinctively higher feature values than their background (which is usually lighter to enhance legibility), hence high standard deviation.
2. Pixels representing an image object have relatively low mean and standard deviation since the gray scale level usually appear darker than those of textual pixels.
3. Pixels representing a background object have relatively close to zero standard deviation and lighter background color.

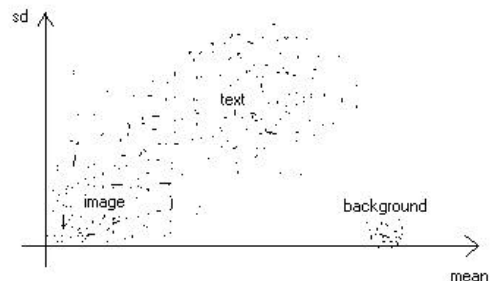


Figure 3: Cluster of image, text, and background

Figure 3 illustrates the above relationship classification guideline based on light color

background and dark color image to enhance the visual contrast. For dark color background and light color image, cluster positions of image and background will be reversed, i.e., image on right and background on left, whilst text cluster remains unchanged as shown in Figure 4.

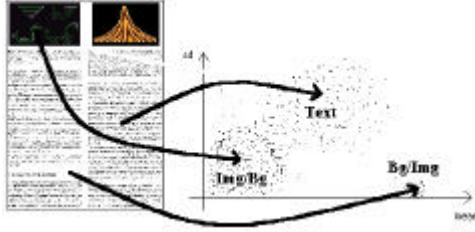


Figure 4: Image and background cluster interchange

3. Statistical classification procedures

As described earlier, the first task of FCM image segmentation method is transformation of color input image (RGB) to gray scale image, which is a three-element tuple of the form (Y, Cb, Cr) [4], according to the following transformation:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299R & 0.587G & 0.114B \\ -0.169R & -0.331G & 0.500B \\ 0.500R & -0.419G & -0.081B \end{bmatrix}$$

where Y denotes the pixel value of the gray scale image, Cb and Cr denote the B and R component, respectively. The G component is represented by the relationship

$$Cg = -0.169R + 0.500G - 0.331B$$

Figure 5 illustrates the result of transforming an RGB image of size $L \times M$ to YCbCr and Cg images of size $p \times p$ pixels, as well as a composite gray scale image.



Original image (RGB)

Gray scale image



Cb image



Cr image



Cg image

Composite gray scale image

Figure 5: Results of color to gray image transformation

Minor adjustments in gray scale image construction (obtained from the Y value) were required to compensate for any image having no textual pixel in black or no color. If the value of any Cb, Cr, or Cg component exceeded 100 (approaching 255), the value of Y would be set to zero since the pixel is likely an image pixel (see Figure 3). The composite gray scale image was the union of Cb, Cr, and Cg images for the sole purpose of boundary enhancement, which would facilitate subsequent region division processing (see Figure 1).

The above transformation inadvertently introduces inexactness of pixel mapping from RGB to gray scale image. Thus, we employed fuzzy set technique to help determine to which image region the pixel belonged [5]. Define a universal set $R = \{\text{image, text, background}\}$, representing the three regions of any given document. Denote the three proper subsets of R, namely, $R_1 = \{\text{image}\}$, $R_2 = \{\text{text}\}$, and $R_3 = \{\text{background}\}$ as fuzzy sets having associated fuzzy membership function

$$U_{ik}: X \rightarrow [0, 1]$$

where X denote the values of pixel block belonging to R. The value of U_{ik} describes membership grade of pixel block k in region R_i , $i = 1, 2, 3$.

3.1 Fuzzy classification

Let A be a set of vectors x_i , $i = 1, 2, \dots, n$, where n is the number of image blocks to be determined their membership, in this case, $n = (L/p) \times (M/p)$. Each vector x_i consists of d feature attributes, in this case $d = 2$, i. e., mean and standard deviation [1, 3, 7]. The set A becomes

$$A = \{x_1, x_2, \dots, x_n\}$$

Grouping pixel blocks according to their fuzzy membership by region yields

$$U = [U_{ik}], 1 \leq i \leq c, 1 \leq k \leq n$$

where U denotes the fuzzy membership matrix of pixel block k in region i , c denotes the number of regions. The matrix thus becomes

$$U = \begin{matrix} & \begin{matrix} \text{blocks} \\ u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ u_{21} & u_{22} & u_{23} & \dots & u_{2n} \\ u_{31} & u_{32} & u_{33} & \dots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{c1} & u_{c2} & u_{c3} & \dots & u_{cn} \end{matrix} \\ \begin{matrix} \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{matrix} & \end{matrix}$$

In this study, c is equal to 3. Each row of matrix U thus represents the membership grade of all blocks belonging to region R_i , i. e., pixels constituting the same object.

The elements of U must satisfy the following constraints:

$$\sum_{i=1}^c u_{ik} = 1, \forall k; \quad \sum_{k=1}^n u_{ik} > 0, \forall i; \quad u_{ik} \in [0,1], \forall i, k$$

The value of U_{ik} denotes the probability of block k being mapped to region i . The FCM method minimizes the dispersion of feature attributes between the center of region i and block k to determine the pixel membership as follows:

$$J_m(U, V) = \sum_{k=1}^n \sum_{i=1}^c u_{ik}^m d_{ik}^2 = \sum_{k=1}^n \sum_{i=1}^c u_{ik}^m \|x_k - v_i\|^2$$

where m is a real number greater than 1, being employed to control the fuzziness of the cluster, i.e., if $m \gg 1$, all regions have high fuzziness; matrix $V = [v_1, v_2, \dots, v_c]$ holds the vectors (having $d=2$ as in x_i) representing the center of each region, in this case $V = [v_1, v_2, v_3]$; and d_{ik} denotes the dispersion of x_i from v_i , i. e., $d_{ik} = \|x_k - v_i\|$ and $d_{jk} = \|x_k - v_j\|$. Local minimization of the function J_m is accomplished by repeatedly adjusting the values of U_{ik} and V_i according to the following relation:

$$u_{ik} = \left(\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}} \right)^{-1}, \forall i, k; \quad v_i = \frac{\sum_{k=1}^n u_{ik}^m x_k}{\sum_{k=1}^n u_{ik}^m}, \forall i$$

Computations began by initializing a fuzzy pseudo-partition $U = [U_{ik}]$, $1 \leq i \leq 3$, $1 \leq k \leq n$, and x_k , $1 \leq k \leq n$, having $m = 1.25$. The value of v_i and d_{ik} could be obtained from the above relations. Iteration tolerance was set at 0.0000000001 since initial error was still small. The membership of X_k to region i was iteratively adjusted through a new value of U , denoted U_{new} as follows:

$$\text{error} = \|U_{\text{new}} - U_{\text{old}}\| = \max_{1 \leq i \leq 3, 1 \leq k \leq n} \|U_{ik, \text{new}} - U_{ik, \text{old}}\|$$

As J_m was iteratively minimized, v_i became more stable, i. e., v_1 distinctively differed from v_2 and v_3 , and v_2 from v_3 , successively. Iteration terminated when the error was smaller than the predefined

tolerance. According to the sample document, the appropriate tolerance was set at 1.0 to yield fast convergence.

The FCM statistical classification process can be summarized as follows:

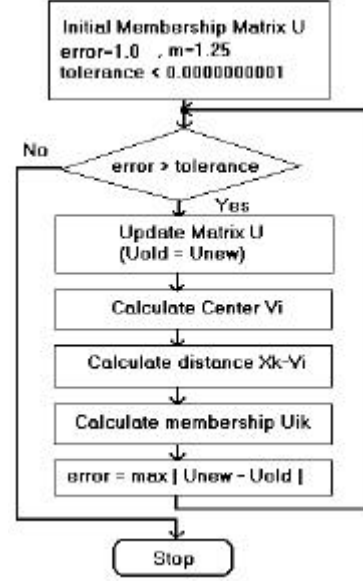


Figure 6: The FCM Classification Process

4. The FCM segmentation method

The underlying principle of FCM segmentation method is to distinctively separate all fuzzy classifications discussed in previous section, forming segmented image identifiable by subsequent processing such as OCR, morphing, enhancement, etc. In effect, the FCM segmentation method offers an alternative to image segmentation and enhancement.

Segmenting fuzzy regions encompasses two processing stages, namely, defuzzification and neighborhood smoothing. Details are described in the sections that follow.

4.1 Defuzzification

When the transient iterative mapping became stable, all the pixel blocks were considered belonging to one of the predetermined regions, in this case text, image, and background regions [2]. The process began by transforming the final U matrix back to image by deciding to which region x_k belonged. The decision on how to assign x_k to region i was based on the *winning* U_{ik} having the highest value in the group. In so doing, x_k would be painted the same color code for region i , whereby the resulting images were classified by color-coded regions. This reversed mapping process is called defuzzification.

The next step of defuzzification is to map these color-coded regions to the original image clusters. Color-to-image mapping is an arbitrary process arranged at one's discretion. In this study, R, G, B were mapped to region 1, 2, 3,

encompassing the winning U_{ik} , respectively.

The final step of defuzzification is to tie each region to the appropriate image categories, i.e., text, image, and background [6]. Adhering to the aforementioned pixel cluster classification guidelines (Figure 3), for region i having the smallest v_i , $i = 1, 2, 3$, region i became the image. By the same token, the largest v_i of region i became the background, and the remaining became text, respectively. The result yielded a conversion viewport for use in original-sized image reconstruction. By applying this conversion viewport to every gray scale level pair (mean, sd), the original-sized image could be straightforwardly obtained.

4.2 Neighboring smoothing

Smoothing is called for to render pixel level correction caused by the above mapping inexactness (fuzzy classification). Conventional neighborhood influence and dominance made up for the primary smoothing criteria. By arbitrarily assigned color code to a given region as illustrated in Figure 7, coloring the resulting image is carried out at pixel level by means of neighborhood smoothing as follows:

Denote the colors R, G, and B for text, background, and image pixels, respectively, pixel color assignment is accomplished according to the following criteria:

- if there are text pixels juxtaposing to the left or right of image pixels as shown in Figure 7(a) or 7(b), the color of such text pixels is replaced by the color of the background pixels.
- if there are either text or background pixels being surrounded by image pixels as shown in Figure 7(c), the color of such text or background pixels is replaced by the color of the image pixels.
- if there are background pixels being surrounded by text pixels as shown in Figure 7(d), the color of such background pixels is replaced by the color of the text pixels.

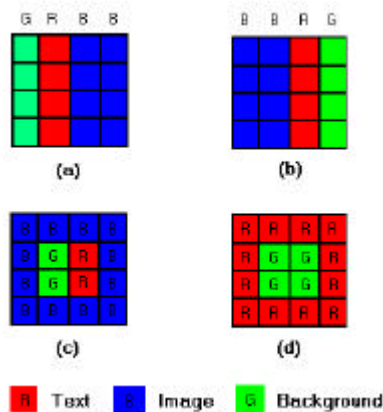


Figure 7: Neighborhood Smoothing

Once neighborhood smoothing is complete, the result yields a segmented image of the original image as shown in Figure 8.

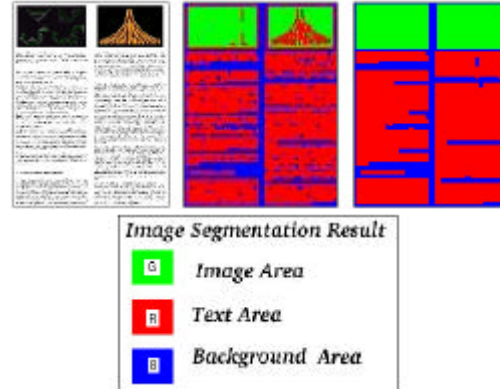


Figure 8: Segmented image of a document

The flowchart of segmentation method can be summarized as shown below:

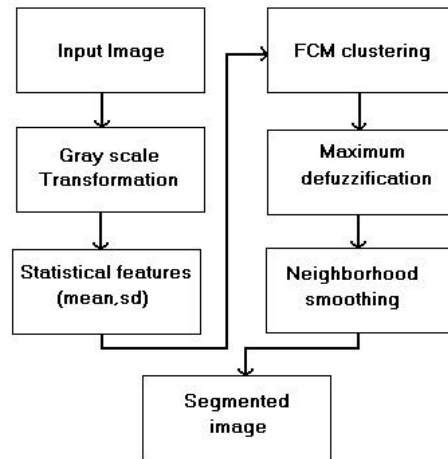


Figure 9: Segmentation method

5. Experimental results

The FCM method was put to use with two selected samples, namely, one picture-embedded document and one image. The input parameters of the sample picture-embedded document are given in Table 1 and the result is depicted in Figure 8. Table 2 shows the input parameters for the sample breast image. The corresponding result is depicted in Figure 10.

Table 1: Sample document input parameters

Window size	8 x 8 pixels
Weighted exponent (m)	1.25
Iteration tolerance	0.000000000000001
Number of iterations	54
Number of clusters	3 classes
Feature value of cluster 1	[210.528834,196.523061]
Feature value of cluster 2	[13.960429,27.061843]
Feature value of cluster 3	[252.112987,19.311612]
Coloring assignment	R=Text, G=Image, B=Background

Table 2: Sample image input parameters

Window size	5 x 5 pixels
Weighted exponent (m)	1.25
Iteration tolerance	0.000000000000001
Number of iterations	33
Number of clusters	2 classes
Feature value of cluster 1	[4.810647,1.750243]
Feature value of cluster 2	[173.999710,8.273517]
Coloring assignment	R=Text, G=Image, B=Background

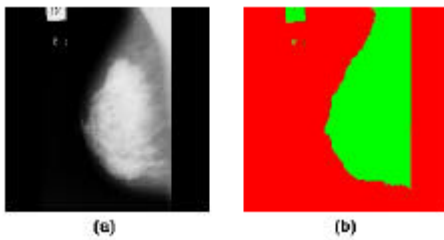


Figure 10: Segmented image of a breast image

6. Conclusion

We applied the FCM method to perform text and image segmentation based on simple statistical features. The results were satisfactory. As a consequence, the method will be extended to accommodate on-going OCR projects as part of future open document processing systems.

References

- [1] Klir, G. J., and Yuan, B. Fuzzy sets and fuzzy logic theory and applications, *Prentice Hall*, 1995.
- [2] Tobias, O.J., Seara, R., and Soares, F.A.P. Automatic Image Segmentation Using Sets, *Proceedings of the 38th Midwest Symposium*, 2: 921-924, 1996.
- [3] Ruben, M., Mireille, G., Diego, J., Carlos, C., and Javier, T. Segmentation of Ventricular Angiographic Images Using Fuzzy Clustering, *IEEE 17th Annual conference*, 1: 405-406, 1997.
- [4] Chai, D., and Ngan, K. N. Face Segmentation Using Skin-Color Map in Videophon Applications, *Circuits and Systems for Video Technology, IEEE Transactions on Circuits and Systems for Video Technology*, 9: 551-564, June 1999.
- [5] Hong, T., and Lee, C. Introduction of Fuzzy rules and membership functions from training examples, *Fuzzy Sets and Systems*, 84: 33-47, 1996.
- [6] Li, J., and Gray, R.M. Text and Picture Segmentation by the Distribution Analysis of Wavelet Coefficients. *Proceedings of International Conference on Image Processing*, 790-794, October 1998.
- [7] Runkler, T.A., and Bezdek, J.C. Image segmentation using fuzzy clustering with fractal features, *Proceedings of the Sixth IEEE International Conference on Fuzzy Systems*, 3: 1393-1398, 1997.