

Image Segmentation by Gaussian Mixture Models and Modified FCM Algorithm

Karim Kalti and Mohamed Mahjoub

Department of Computer Science, University of Sousse, Tunisia

Abstract: *The Expectation Maximization (EM) algorithm and the clustering method Fuzzy-C-Means (FCM) are widely used in image segmentation. However, the major drawback of these methods is their sensitivity to the noise. In this paper, we propose a variant of these methods which aim at resolving this problem. Our approaches proceed by the characterization of pixels by two features: the first one describes the intrinsic properties of the pixel and the second characterizes the neighborhood of pixel. Then, the classification is made on the base on adaptive distance which privileges the one or the other features according to the spatial position of the pixel in the image. The obtained results have shown a significant improvement of our approaches performance compared to the standard version of the EM and FCM, respectively, especially regarding about the robustness face to noise and the accuracy of the edges between regions.*

Keywords: *EM algorithm, FCM algorithm, image segmentation, adaptive distance.*

Received May 4, 2011; accepted December 30, 2012; published online January 29, 2013

1. Introduction

Image segmentation constitutes an important step and an essential process of image analysis. Fuzzy-C-Means (FCM) [2], is one of the most famous unsupervised fuzzy clustering techniques that are applied with success in image segmentation [1, 5, 9, 10]. Although the original FCM algorithm yields good results for segmenting noise free images, it fails to segment images corrupted by noise or containing inaccuracy edges. This sensitivity is essentially due to the absence of utilization of the information on the spatial position of pixels to be classified. Several authors tried to overcome this drawback by the integration of spatial information [5]. Chuang *et al.* [4], proposed a novel fuzzy clustering algorithm that uses a spatial membership degree representing the summation of the membership degree in the neighbourhood of each pixel. Tolia and Panas [12], developed a Sugeno type rule based system that imposes spatial constraint by modifying the membership degree of clustering results obtained after FCM algorithm. In our paper we propose a novel version of FCM that integrates the spatial information. The novelty concerns essentially the manner of calculating the distance of similarity between the pixels of the image and the centers of the classes.

Recently, finite mixture models have attracted considerable interest for image segmentation [6, 7, 16]. However, the application of finite mixtures model to image segmentation presents some difficulties. For the classical mixture statistical model each pixel must be associated with exactly one class. This assumption may

be not realistic. Thus, several methods have been proposed to circumvent this problem.

The segmentation algorithm developed in this paper is based on a parametric model in which the probability density function of the gray levels in the image is a mixture of Gaussian density functions. This model has received considerable attention in the development of segmentation algorithms and it has been noted that the performance is influenced by the shape of the image histogram and the accuracy of the estimates of the model parameters. However, the model-based segmentation not allows a good results if the histogram of an image is a poor approximation of a mixture of two Gaussian density functions. The application of this model in image segmentation is, therefore, limited to the images which are a good approximations of Gaussian mixtures with well-defined modes.

Among the techniques of classification, we used in this work models of Gaussian mixtures. An Expectation Maximization (EM) algorithm is developed to estimate parameters of the Gaussian mixtures. The finite mixture is a flexible and powerful probabilistic modelling tool. It can be used to provide a model-based clustering in the field of pattern recognition. However, the application of finite mixtures to image segmentation presents some difficulties; especially it's sensible to noise. In this paper we propose a variant of this method which aims to resolve this problem. The main objective of our work is the integration of a fuzzy distance in EM and FCM algorithms for unsupervised segmentation of a noisy image.

The rest of this paper is organized as follows: in sections 2 and 3, we will present respectively the principle of EM and FCM algorithms. The limits of these methods are presented in Section 4. In the following section, we will present our fuzzy clustering approach algorithm that incorporates a spatial constraint for image segmentation. Our segmentation methods are tested on several images, especially MRI images. The results are illustrated in section 6. Finally, in section 7 conclusions are given.

2. Segmentation Method Based on EM

The EM algorithm proposed by Dempster *et al.* [6], is a class of algorithms that find the maximum likelihood parameters of probabilistic models where the model depends on unobserved variables [11]. The finite mixture of distributions has provided a mathematical approach to the statistical modeling of a wide variety of random phenomena. In the past decades, the extent and the potential of the applications of finite mixture models have widened considerably. In the field of pattern recognition, the finite mixture can be used to provide a model-based clustering. A finite mixture density has the form:

$$f(x/\theta) = \sum_{i=1}^K p_i f_i(x/\alpha_i) \tag{1}$$

where:

- p_i is the proportion of the class i ($p_i \geq 0$ and $\sum_i p_i = 1$).
- $\alpha_i = (\mu_i, \Sigma_i)$, μ_i et Σ_i are respectively the center and the variance matrix of k^{th} normal component $f_i(\cdot/\alpha_i)$.

The evaluation of all these parameters can be calculated by maximizing the log-likelihood of the global parameter θ :

$$\ln f(X/\theta) = \sum_{j=1}^n \ln f(x_j/\theta) \tag{3}$$

where $X=(x_1, \dots, x_n)$, this maximization can be made by the EM algorithm of Dempster. Besides, the k number of components can also be estimated while keeping the value between $k=1$ and k_{sup} (k_{sup} to choose a priori) that minimizes the Bayesian BIC criteria:

$$BIC(K) = -2 \ln f(X/\hat{\theta}_K) + v_K \ln n \tag{4}$$

where $\hat{\theta}_K$ and v_K are respectively the maximum likelihood estimator and the number of degrees of freedom model.

2.1. Classification

Any classification model is defined on the space N of maps from the image domain to the set N of classes (each class n corresponds to an entity of interest in the scene). Thus each classification $v \in N$ assigns a class $n=v(p) \in N$ to each pixel p giving the class of that

pixel. By defining a posterior probability distribution on N , and using a suitable loss function, an optimal classification can be chosen. The loss function is more often than not taken to be the negative of a delta function, the resulting estimate then being a Maximum A Posterior (MAP) estimate. The posterior distribution is expressed as the normalized product of a likelihood, such as the Gaussian Mixture Models (GMM) models that we will discuss in this paper, which gives the distribution of images corresponding to a given class, and a prior probability distribution on the classifications [14, 16].

2.2. Parameters Estimation

The EM algorithm is an iterative approach to compute maximum-likelihood estimates when the observations are incomplete [6]. In the mixture density estimation, the information that indicates the component from which the observable sample originates is unobservable. In EM, alternating steps of Expectation (E) and Maximization (M) are performed iteratively till the results converge. The E step computes an expectation of the likelihood by including the latent variables as if they were observed, and a Maximization (M) step, which computes the maximum likelihood estimates of the parameters by maximizing the expected likelihood found on the last E step. The parameters found on the M step are then used to begin another E step, and the process is repeated until convergence.

Let's suppose that we have K components in the model of mixtures. The shape of the density of probability of this mixture is given by:

$$f(x/\theta) = \sum_{i=1}^k \alpha_i f_i(x/\theta_i) \tag{5}$$

where x is the characteristic vector, α_i is the weight of the mixture as $\sum_{i=1}^k \alpha_i = 1$, θ represents the parameters $(\alpha_1, \alpha_2, \dots, \alpha_k, \theta_1, \theta_2, \dots, \theta_k)$ and f_i the density of the gaussian parameterized by θ_i that is to say (μ_i, σ_i) :

$$P(x/\theta_i) = \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left\{-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right\} \quad i = 1, 2, \dots, N \tag{6}$$

where $\theta_i=(\mu_i, \sigma_i)$ is the Gaussian mixture distribution parameter. Assume that the density is derived from a mixture of Gaussians. That is to say that:

$$f(x) = \sum_{i=1}^k p_i \phi(x, \mu_i, \Sigma_i) \tag{7}$$

We will then estimate the parameters by maximizing the likelihood. For this, we must start with a number of K Gaussian fixed a priori and then seek a local maximum in the first order conditions. Thus, the EM algorithm is summarized as follows:

Algorithm 1. EM Algorithm:

Input: H =histogram, k =gaussian number, ε =Error

Output: model parameters $(p_1, p_2, \dots, p_k, \alpha_1, \alpha_2, \dots, \alpha_k)$

Steps:

1. Evaluation expectancy (E):

$$\varphi(i/x_j, \theta) = \frac{p_i f_i(x_j/\alpha_i)}{\sum_{k=1}^K p_k f_k(x_j/\alpha_k)} \quad (8)$$

2. Maximization step (M): GMM parameters update

$$p_i^{new} = \frac{1}{N} \sum_{j=1}^N \varphi(i/x_j, \theta^{old}) \quad (9)$$

$$\mu_i^{new} = \frac{\sum_{j=1}^N x_j \varphi(i/x_j, \theta^{old})}{\sum_{j=1}^N \varphi(i/x_j, \theta^{old})} \quad (10)$$

$$\sum_i^{new} = \frac{\sum_{j=1}^N \varphi(i/x_j, \theta^{old}) (x_j - \mu_i^{new})(x_j - \mu_i^{new})^T}{\sum_{j=1}^N \varphi(i/x_j, \theta^{old})} \quad (11)$$

3. The process stops when $\|\theta^{new} - \theta^{old}\| \leq \varepsilon$

However, during the classification phase, we can modify this algorithm to classify a given pixel by calculating the distance mahalonobis distance to the class instead of a probability calculation:

$$D_M(X_i) = \text{sqrt} \left((X_i - \mu_k)^T \sum_k^{-1} (x_i - \mu_k) \right) \quad (12)$$

The modified algorithm is called Distance EM (DEM).

3. Fuzzy-C-Means Clustering

FCM clustering was developed by Bezdek [2, 11] and it's very large used in image segmentation [3, 8, 10]. It can be described as follows: Let $X = \{x_1, x_2, \dots, x_n\}$ denoted a set of n objects to be partitioned into C clusters, where each x_j has d features. The FCM algorithm minimizes the objective function defined as follows:

$$J = \sum_{i=1}^C \sum_{j=1}^n (u_{ij})^m D(x_j, v_i) \quad (13)$$

where:

- u_{ij} represents the membership degree of j^{th} object in the i^{th} cluster.
- v_i represents the i^{th} cluster center.
- D represents a distance metric generally the square of Euclidian distance that measures the similarity between an object and cluster center.
- $m \geq 1$ the degree of fuzzyfication.

The membership degree of x_j to i^{th} cluster is determined by calculating the gradient of J with respect to u_{ij} . Thus, these membership degrees are given by equation 13:

$$u_{ij} = \left(\frac{\sum_{k=1}^C (D(x_j, v_i) / D(x_k, v_i))^{1/m-1}}{\sum_{k=1}^C (D(x_j, v_i) / D(x_k, v_i))^{1/m-1}} \right)^{-1} \quad (14)$$

The cluster centers $v_i, i:1..C$ are determined by calculating the gradient of J with respect to v_i . These centers are given by equation 14:

$$v_i = \frac{\sum_{j=1}^n (u_{ij})^m x_j}{\sum_{j=1}^n (u_{ij})^m} \quad (15)$$

The FCM algorithm can be summarized in the following steps [3, 4]:

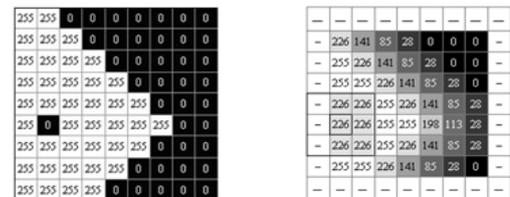
- *Step 1:* Fix the cluster number. Initialize the centers by random points from data set.
- *Step 2:* Update the membership degrees by using equation 13.
- *Step 3:* Update centers using equation 14.
- *Step 4:* Repeat steps 2 and 3 until convergence.

The convergence of this algorithm will be reached when the change in membership values is less than the threshold.

4. Limits of FCM and EM Algorithms

FCM clustering algorithm is using on image segmentation successfully for many years [13]. However, traditional FCM did not consider the spatial information of pixels or different effects of different sample vectors. Therefore, it is very sensitive to noise and the results will be affected when different vectors contribute not the same to cluster. Considering problems above, some researchers proposed spatially Weighted FCM (SWFCM), which based on image histogram and spatial information of pixels. The weights of the algorithm are given by the ratio of every gray value in histogram [15].

Figure 1-a shows a grey level synthetic image formed by two regions: black region (0) and white region (255) that includes a black noisy pixel (0). The application of standard FCM using the grey level as a single feature of pixels on this image yields good segmentation of pixels inside regions and pixels of edges. However, it provokes a bad clustering of noisy pixel of white region. This clustering drawback is essentially due to the only utilization of the intrinsic feature of pixel to be classified (grey level) without taking into account the information relative to spatial position in the image. This information was proved very important in the context of segmentation.



a) Grey levels of image.

b) The means of grey levels of image (a).

Figure 1. Simulation of an image with two gray levels.

To overcome this limitation, one of solutions consists to integrate the neighborhood effect of pixel to be classified. There are several statistic estimators to accomplish this effect. In this work, we have chosen the spatial feature: arithmetic means estimator denoted

μ . Figure 1-b represents the image of the means obtained by replacing the Grey Level (GL) of the pixels of the image of Figure 1-a with the means of the GL of their neighborhood calculated on a window of size 3x3. The application of the FCM on this image engenders a good clustering of pixels inside regions as well as noisy pixel, but it produces a degradation of edges between regions. This result is essentially due to the smoothing effect of the spatial feature used in the clustering processes. Table 1 shows the advantages and the inconveniences of using the GL and the spatial feature for the clustering of noise and edges.

Table 1. Advantages and inconveniences of using grey level and spatial feature in the clustering process.

	Spatial Feature	Gray level
Noise	Bad clustering	Good clustering
Edges	Good clustering	Bad clustering

5. Proposed Methods

The complementarity of the grey level feature and the spatial feature as regards the FCM clustering can let envisage a joint use of these two features in image segmentation. In this section we will present a new version of FCM baptized Adaptive Distance based FCM (ADFCM) which takes the advantages of both features while avoiding their inconveniences and by using the one or the other in an adaptive way according to the spatial configuration of each pixel.

5.1. Specification of the Considered Spatial Configurations

5.1.1. Presentation of Spatial Configurations

In our work, we distinguished four possible spatial configurations for pixels asking each a specific choice of the clustering criterion show in Figure 2. These are the configurations: Pixel belonging to a Region (PR), Pixel belonging to an Edge (PE), Noisy Pixel (NP) and Neighbour of a Noisy pixel (NN).

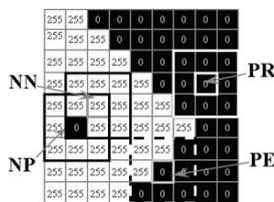


Figure 2. Spatial configurations of pixels.

5.1.2. Characterization of the Spatial Configurations

Formally, the spatial configurations are characterized by two statistical descriptors of decision that are presented as follows:

- The standard deviation σ which characterizes the dynamics of the distribution around the pixel to be classified. This feature is defined as follows:

$$\sigma(x_j) = \sqrt{\frac{1}{N} \sum_{k=1}^N (x_k - \mu(x_j))^2} \tag{16}$$

- The knn which represents the number of the closest neighbours in term of grey levels with regard to the considered pixel. The knn is defined as follows:

$$knn(x_j) = Card\{x_p \in Neighborhood(x_j) / |x_p - x_j| < S\} \tag{17}$$

where S designates a threshold which is generally chosen in an empirical way. From these two features we can characterize the various possible spatial configurations of pixels. In case of a PR, the standard deviation σ is generally low, it is null for the constant regions. However, the σ becomes high for the PE, NP and NN. The distinction between these three configurations is made by using the knn feature. This number is generally low for a NP, moderate for a PE and high for a NN.

5.1.3. Choice of the Criterion of Clustering According to the Spatial Configuration

The clustering of a PR or a NP has to privilege the spatial feature spatial because the decision must be taken on the basis of the information of its neighborhood. On the other hand the clustering of a PB or a NN has to privilege the grey level of the pixel respectively to protect better contours and to avoid the Influence of the noise. The choice of the criterion of clustering and the characterizations of the spatial configurations are summarized in Table 2.

Table 2. Spatial configurations, their characterizations according to both descriptors: σ and knn, their privileged feature.

Cases of Spatial Configurations	Characteristic		Privileged Feature
	σ	knn	
Pixel of Region (PR)	Low	---	Spatial feature
Pixel of Edges (PE)	High	Moderate	Grey level
Noisy Pixel (NP)	High	Low	Spatial feature
Nearby Pixel of a Noise (NN)	High	High	Grey level

5.2. Proposition of a New Distance of Similarity

The standard FCM uses generally to measure the similarity between an object x_j and a class given by its center v_i a distance which grants the same importance for the features taken into account in the clustering process. To introduce the adaptive effect for the selection of the features, we suggest using a dynamic and weighted distance derived from the Euclidian distance. This new distance is given by equation 17.

$$D(x_j, v_i) = (1 - p_j)(x_j^{GL} - v_i^{GL})^2 + p_j(x_j^{Spatial} - v_i^{Spatial})^2 \tag{18}$$

D is a bi-dimensional distance based on both features GL and spatial. In this equation, the weight p_j allows to control the importance of each feature for the clustering of the pixel x_j . So, if the p_j is high then we privilege the spatial feature otherwise we privilege the grey level. The term p_j must be calculated for each

pixel to be classified according to its spatial configuration in the image. From the configurations presented in Table 2, we can deduce that the weight p_j must be maximized (tender verse 1) when the pixel to be classified is a PR or a NP, because the decision of its membership in the various classes must be taken only on the basis of the spatial feature. However, this weight p_j must be minimized (tender verse 0) in case of a PE or a NN, because the grey level in these cases will constitute a good criterion of clustering.

5.3. Estimation of the Spatial Weight

The choice of the spatial weight p_j is very important for calculating the new distance (6). We propose in this paragraph a fuzzy method of the estimation of this weight. Therefore, we use a fuzzy system [12, 17] which possesses as entries two linguistic variables of decision σ and knn to give results onto the linguistic variable of exit p . By considering the choices of p according to the spatial configurations, it is possible to define for each of these configurations, a fuzzy rule in the form of If-Then (denoted R_i). Then we can deduce four rules characterizing the relations between the classes of entries (σ and knn) with the class of corresponding exit (p) to determine all the necessary consequences so that we can calculate the value of p [17]. The linguistic rules are defined as follows:

R_1 : If σ is Low Then p is High.

R_2 : If σ is High A knn is Low Than p is Low.

R_3 : If σ is High and knn is High Then p is High.

R_4 : If σ is High and knn is Moderate Then p is Low.

The membership functions of variables σ , knn and p used in our system are represented in the Figure 3.

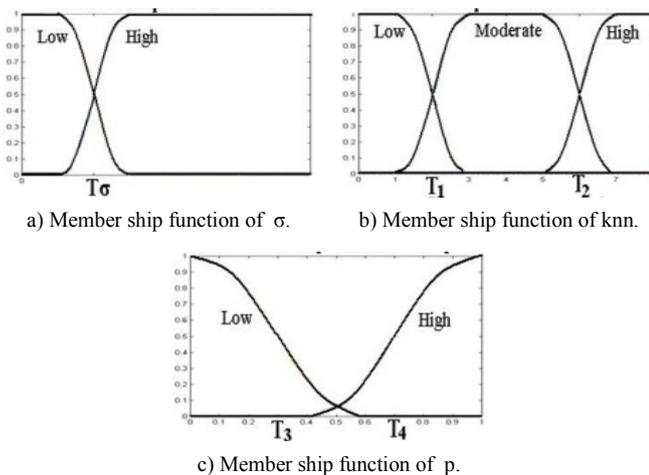


Figure 3. Membership functions used for the estimation.

The thresholds used in these curves (T_σ , T_1 , T_2 , T_3 and T_4) are fixed in an empirical way.

6. Analysis of Exam Results

In this section, we present the results of the application

of improved FCM and EM algorithms. The performance of these algorithms are compared to the standard version of the FCM and EM algorithms. The EM algorithm requires the initialization of model parameters of Gaussian mixture. The covariance matrix is initialized by the identity matrix, and K mean vectors are initialized by the various centers of Gaussian mixture estimated by the algorithm K-means. Both techniques are tested on "Lena" image IN Figure 4 corrupted by 7% of Gaussian noise, and an MRI cerebral image corrupted by 5% of Gaussian noise. These techniques are experimented in the same conditions (a factor of fuzzyfication $m=2$ and a convergence error=0.001). The ADFCM and ADEM use as spatial feature the means μ calculated on an analysis window of size 3×3 . The segmentation results of ADFCM and ADEM are illustrated respectively in Figure 5 and Figure 6.



Figure 4. Lena image.

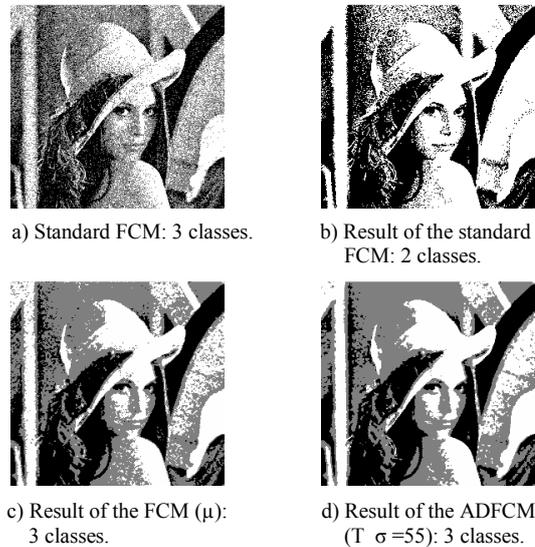


Figure 5. Segmentation results of FCM algorithm and its variants.

Figure 5-a shows the segmented Lena image containing three classes with FCM standard algorithm. Figure 5-b shows the result of applying the standard FCM to the original image using as a criterion of clustering the grey levels. This result clearly illustrates the limitations of this method for classification of noisy pixels. However, the application of the FCM based on the means feature can resolve the problem of noise as shown in Figure 5-c. Contrariwise, it engenders inaccurate edges segmentation. The application of ADFCM (with $T_\sigma=55$) yields the segmentation shown in Figure 4-d. The three classes are better detected.

This result confirms the good performance of the ADFCM compared to the standard FCM. Indeed, by using the adaptive distance, the ADFCM has achieved a compromise that allowed one can make the same argument with the EM algorithm as shown in Figure 6.

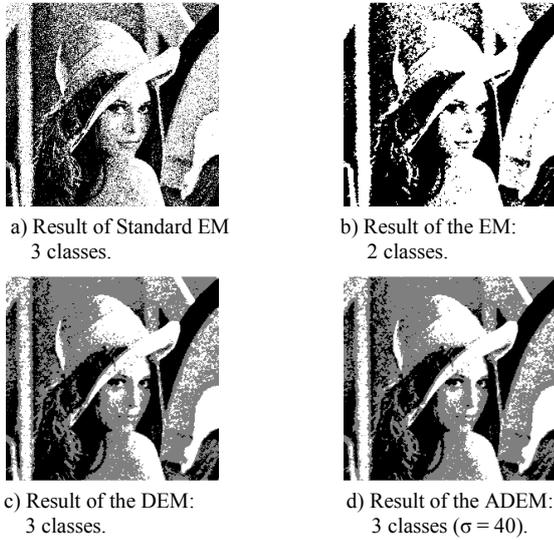


Figure 6. Segmentation results of EM algorithm and its variants.

Cerebral image segmentation consists in bounding three cerebral structures: Grey Matter (GM), White Matter (WM) and the Cerebro Spinal Fluid (CSF). Our tests are realized on the cerebral image of Figure 7. The application of the FCM or EM based on the GL feature on this image gives noisy and overlapped classes particularly between both classes GM and WM shown in Figures 8-a and 8-b. The use of the FCM based on the spatial feature provokes a degradation of the obtained edges show in Figure 8-d. Whereas the use of the ADEM and ADFCM for a threshold equal to 15 help enormously to reduce the noisy pixels while obtaining good identified regions and having continuous edges that are closer to reality shown in Figures 8-e and 8-f.

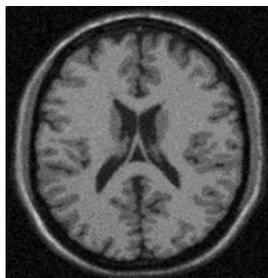


Figure 7. Original cerebral image.

On the other hand, for a better comparison of the two algorithms and ADFCM ADEM, we used the image of Figure 9-a and its segmented ground truth. The Table 3 gives the performance of both methods. We note that the ADEM method is slightly better than ADFCM method as illustrated in Figures 9-c and 9-d.

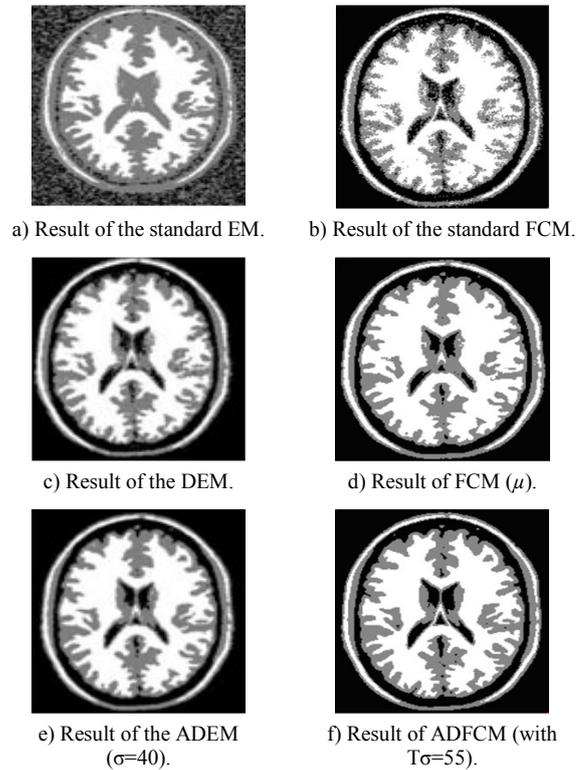


Figure 8. Segmentation results for different approaches.

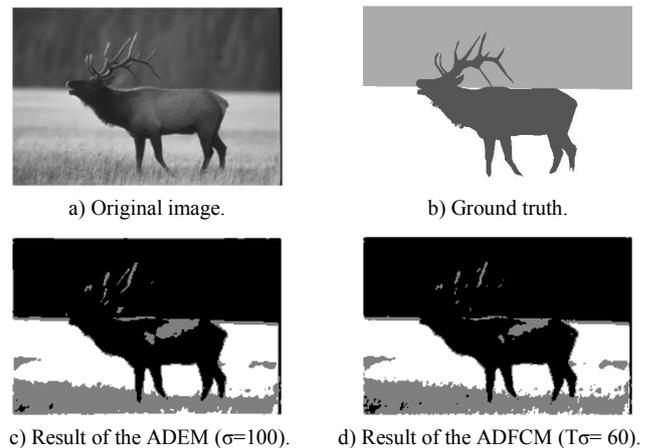


Figure 9. Comparison with ground truth.

Table 3. Algorithms performance.

Algorithms	Treatment Time (s)	% of well Classified Pixels
ADFCM	66.81	61,94%
ADEM	72.23	63 %

7. Conclusions

In this paper we have proposed a novel version of FCM and EM algorithms based on dynamic and weighted distance of similarity. They are new approaches that we proposed of pixels classification based on a dynamic and weighted similarity distance. The main idea is to use a manner conjoined and adaptive two attributes of classification: level of gray of the pixels and a spatial attribute the local average. The adaptation must privilege one or the other of these attributes according to the spatial configuration of the

pixel to classify. The implementation of this adaptive effect is gotten through the level-headedness of these attributes. The originality of our approaches resides in the manner with which is calculated these weights and that takes as a basis on a mechanism of fuzzy inference. The new distance that we proposed thus permitted to get a new variant of the FCM and EM methods that is adapted more to the segmentation of images

Our approach is tested on different images, especially MRI cerebral images. The obtained results have shown a significant improvement of our approach performance compared to the standard FCM and EM, especially concerning the robustness face to noise and the accuracy of the edges between regions. However, the choice of the threshold $T\sigma$ is strongly dependent on the used image. This problem is not solved in this paper, this may construct further steps of research. Indeed, these last showed a clean improvement of the performances of our approaches notably in relation to the classic versions FCM and EM with regard to the hardness in relation to the noise and the precision of the contours gotten.

References

- [1] Ahmed N., Yamany M., Mohamed N., and Farag N., "A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data," *IEEE Transaction on Medical Imaging*, vol. 21, no. 3, pp. 193-199, 2002.
- [2] Bezdek C., *Pattern Recognition with Fuzzy Objective Functions Algorithms*, Plenum Press, New York, 1981.
- [3] Cai W., Chen S., and Zhang D., "Fast and Robust Fuzzy Cmeans Clustering Algorithms Incorporating Local Information for Image Segmentation," *Pattern Recognition*, vol. 40, no. 3, pp. 825-838, 2007.
- [4] Chuang S., Tzeng L., Chen S., Wu j., and Chen T., "Fuzzy C-Means Clustering with Spatial Information for Image Segmentation," *Elsevier Science*, vol. 30, no. 1, pp. 9-15, 2006.
- [5] Chen C. and Zhang Q., "Robust Image Segmentation using FCM with Spatial Constraints Based in New Kernel-Induced Distance Measure," *IEEE Transaction. Systems Man Cybernetics*, vol. 34, no. 4, pp.1907-1916, 2004.
- [6] Dempster A., Larid N., and Rubin D., "Maximum Likelihood From Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society*, vol. 39, no. 1, pp. 1-38, 1977.
- [7] Figueiredo A. and Jain K., "Unsupervised Learning of Finite Mixture Models," *IEEE Transaction on Pattern Analysis and Machin Intelligence*, vol. 24, no. 3, pp. 381-396, 2002.
- [8] Hemanth J., Selvathi D., and Anitha J., "Effective Fuzzy Clustering Algorithm for Abnormal MR Brain Image Segmentation," in *Proceedings of IEEE International Advance Computing Conference*, Patiala, pp. 609-614, 2009.
- [9] Jain K., Murty N., and Flynn J., "Data Clustering: A Review," *ACM Computing Surveys*, vol. 31, no. 3, pp. 264-323, 1999.
- [10] Kawa J. and Pietka E., "Image Clustering with Median and Myriad Spatial Constraint Enhanced FCM," in *Proceedings of the 4th International Conference on Computer Recognition Systems*, Berlin, vol. 30, pp. 211-218, 2005.
- [11] Oussema A., "A Comparison Between Data Clustering Algorithms," *International Arab Journal of Information Technology*, vol. 5, no. 3, pp. 320-325, 2008.
- [12] Tolia A. and Panas M., "On Applying Spatial Constraints in Fuzzy Image Clustering using a Fuzzy Rule Based System," *IEEE Signal Processing Letters*, vol. 5, no. 10, pp. 245-247, 1998.
- [13] Wang J., Dou L., Che N., Liu D., Zhang B., and Kong J., "Local Based Fuzzy Clustering for Segmentation of MR Brain Images," in *Proceedings of the 8th IEEE International Conference on Bioinformatics and Bioengineering*, Athens, pp. 1-5, 2008.
- [14] Xu L. and Jordan I., "On Convergence Properties of the EM Algorithm for Gaussian Mixture," *Neural Computation*, vol. 8, no. 1, pp. 129-151, 1996.
- [15] Xian G., Chen K., and Yuan M., "Medical Brain MRI Images Segmentation by Improved Fuzzy C-Means Clustering Analysis," *Journal of Jilin University Engineering and Technology Edition* vol. 39, no. 2, pp.381-385, 2009.
- [16] Yamazaki T., "Introduction of EM Algorithm into Color Image Segmentation," in *Proceedings of International Consultation on Incontinence Research Society*, pp. 368-371, 1998.
- [17] Zadeh L., "Fuzzy Sets," *Information and Control*, vol. 8, no. 3, pp. 338-353, 1965.



Karim Kalti is an assistant professor in the Department of Computer Science of the Faculty of Science of Monastir, and member of research unit Sage (Advanced System in Electrical Engineering) team signals, image and document

National Engineering School of Sousse, University of Sousse. His research interests include Computer vision, pattern recognition, medical imaging and data retrieval. He is a member of IEEE and his main results have been published in international journals and conferences.



Mohamed Mahjoub is an assistant professor in the Department of Computer Science and mathematics at the Preparatory Institute of Engineering of Monastir and member of research unit Sage (Advanced System in Electrical Engineering), team signals, image and document National Engineering School of Sousse, University of Sousse. His research interests include dynamic bayesian network, computer vision, pattern recognition, HMM, and data retrieval. He is a member of IEEE and his main results have been published in international journals and conferences.