Improved Possibilistic C-Means

Mohamed Fadhel SAAD[1], Adel M. ALIMI[2]
[1]Department of Computer Science. ISET Gafsa
University Compus Sidi Ahmed Zarroug
Gafsa, Tunisia
[2]Research Group on Intelligent Machines
University of Sfax, ENIS, BP W – 3038
Sfax, Tunisia
{fadhel.saad, adel.alimi}@ieee.org

ABSTRACT: *Clustering has been widely used in pattern recognition, image processing, and data analysis. It aims to organize a collection of data items into clusters, such that items within a cluster are more similar to each other than they are in other clusters. An Improved possibilistic clustering algorithm was developed based on the conventional possibilistic c-means (PCM) to obtain better quality clustering result. The new approach uses a parameter â based on the factor of separation between a point and a cluster. This coefficient is fed into the objective function; it defines the relative importance of typicality values in the algorithm.*

Keywords: Fuzzy Clustering, Fuzzy C-means, Possibilistic C-means

Received: 20 February 2013, Revised 18 March 2013, Accepted 28 March 2013

1. Introduction

Cluster analysis is a technique used for classifying data, i.e., for dividing a given dataset into a set of classes or clusters. The goal is to divide the dataset in such a way that two cases from the same cluster are as similar as possible and two cases from different clusters are as dissimilar as possible [1,16]. In this way one tries to model the human ability to group similar objects or cases into classes and categories. Most clustering algorithms do not rely on assumptions common to conventional statistical methods, such as the underlying statistical distribution of data, and therefore they are useful in situations where little prior knowledge data can be exploited in a wide variety of applications, including classification, image processing, pattern recognition, modeling and identification. The conventional clustering methods consider each point of the data set to match exactly one cluster. Since 1965, Zadeh has proposed fuzzy sets in order to come closer to the physical world [2, 22, 23]. Zadeh introduced the idea of partial memberships described by membership functions. Fuzzy sets could allow membership functions to all clusters in a data set so it was very suitable for cluster analysis. Ruspini (1969) first proposed fuzzy c-partitions as a fuzzy approach to clustering [3]. Later, the fuzzy c-means (FCM) algorithms with a weighting exponent $m = 2$ proposed by Dunn (1974) [4], and then generalized by Bezdek (1981) with $m > 1$ became popular [24]. The FCM used the probabilistic constraint that the memberships of a data point across classes summed to one. While this was useful in creating partitions, the memberships resulting from FCM and its derivatives, however, do not always generally correspond to the intuitive concept of degree of belongingness or compatibility. Fan et al. [5] have proposed the algorithm S-FCM (suppressed fuzzy c-means) to overcome the slow convergence of FCM to the optimum by introducing a new parameter á. After Wen-Liang et al. [6] have proposed a new algorithm called MS-FCM (modified suppressed fuzzy c-means) by modifying the parameter α.

Moreover, the FCM was sensitive to noise. To mitigate such an effect, Krishnapuram and Keller (1993) throw away the constraint of memberships in FCM and proposed the possibilistic c-means (PCM) algorithm [19]. The advantages of PCM were that it overcomes the need to specify the number of clusters and were highly robust in a noisy environment. However, a few weaknesses remained in the PCM, i.e., it highly depended on a good initialization and had the undesirable tendency to produce coincident clusters [20].

The clustering techniques are used to maximize or minimize an objective function. So the clustering problem is an optimization problem, which groups similar objects into clusters and satisfy some additional conditions [29]. The choice of an appropriate objective function is the key to the success of the cluster analysis and to obtain better quality clustering results [10, 27, 28]. To meet a suitable objective function, we started from the following set of requirements: The distance between clusters and the data objects assigned to them should be minimized and the distance between clusters should be maximized [12, 14]. The attraction between data and clusters is modeled by an objective function.

The proposed algorithm called IPCM (improved possibilistic c-means) combines the principle of MS-FCM approach with possibilistic approach by adding a new parameter β, which improves the quality of classification, and to accelerate the convergence of the algorithm.

The remainder of this paper is organized as follows. In section II, preliminary theory algorithms are presented. In section III, the improved possibilistic c-means is proposed.The proposed IPCM can solve these drawbacks mentioned in section II, and obtain better quality clustering results. In section IV, we present several examples to assess the performance of IPCM. The comparisons are made between FCM, MS-FCM, PCM and IPCM. Finally, conclusions are made in Section V.

2. Preliminary Theroy

2.1 Fuzzy c-means algorithm

In the literature of fuzzy clustering, the algorithm fuzzy c-means (FCM) and its variants are the best known and most used methods [7, 8]. However, it is necessary to assume the number c of these clusters for fuzzy algorithms. With the work of Bezdek, fuzzy clustering methods reach a certain maturity. In 1973, Bezdek submitted a doctoral thesis on fuzzy math for classification [17] which is necessary conditions for the minimization of the general criterion that defines the family of algorithms known as the fuzzy c-means (FCM) [9]. Bezdek also studied the convergence of FCM [21]. Other variants of this algorithm that were then developed to improve performance [10]. These upgrades are often dedicated to a particular application, and even today the FCM is generally useful in many situations. Versions change the metric used, others alter the objective function, incorporate additional constraints or support the work of the environment. The fuzzy c-means (FCM) can be seen as a method of clustering which allows one piece of data to belong to two or more clusters. The FCM algorithms clustering was first presented by Dunn in 1973 [4] and improved by Bezdek in 1981 [24]. The algorithm is an iterative clustering method that produces an optimal c partition by minimizing the weighted within group sum of squared error objective function:

$$J_{FSM}(V, U, X) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} d^{2}(x_j, v_i), \quad 1 \prec m \prec +\infty \tag{1}$$

Where $X = \{x_1, x_2, ..., x_n\} \subseteq R^p$ is the data set in the p-dimensional vector space, p is the number of data items, c is the number of clusters with $2 \le c \le n-1$. $V = \{v_1, v_2, \ldots, v_c\}$ can be regarded as c prototypes for the clusters represented by the membership grades. For the purpose of minimizing the objective function, the cluster centers and membership grades are chosen so that a high degree of membership that occurs for samples closer to the corresponding cluster centers, v_i is the p-dimension center of the cluster i, and $d^2(x_j, v_i)$ is a distance measure between object x_j and cluster center v_i. $U = \{\mu_{ij}\}$ is a fuzzy partition matrix which must satisfy:

$$0 \prec \sum_{j=1}^{n} \mu_{ij} = 1, \forall i \in \{1,..., c\} \quad (2)$$

$$\sum_{i=1}^{c} \mu_{ij} = 1, \forall j \in \{1,..., n\} \quad (3)$$

The parameter *m* is a weighting exponent on each fuzzy membership and determines the amount of fuzziness of the resulting classification; it is a fixed number greater than one. The objective function J_{FCM} can be minimized under the constraint of *U*. Specifically, taking of J_{FSM} with respect to μ_{ij} and v_i and zeroing then respectively, two necessary but not sufficient conditions for J_{FCM} to be at its local extrema will be as the following:

$$u_{ij} = \left[\sum_{k=1}^{c} \left(\frac{d^2(x_j, v_i)}{d^2(x_j, v_k)} \right)^{2/(m-1)} \right]^{-1}, 1 \le i \le c, \quad 1 \le j \le n \quad (4)$$

$$v_i = \frac{\sum_{k=1}^{n} \mu_{ij}^{m} x_k}{\sum_{k=1}^{n} \mu_{ik}^{m}}, 1 \le i \le c \quad (5)$$

Thus, we have the *FCM* algorithm as follows:

FCM algorithm

$S1$ Fix $m > 1, 2 \le c \le n - 1$

Initialize the *c* cluster centers v_i randomly.

REPEAT

$S2$ Update $U^t = [\mu_{ij}]$ by (4).

$S3$ Update $V^t = [v_i]$ by (5).

UNTIL (cluster centers stabilized)

- Suppressed fuzzy c-means algorithm

The goal of S-FCM algorithm is to speed up the convergence of FCM to the optimum. That's why it adds a new step that on the one hand magnifies the largest membership degree and on the other hand suppresses the second largest membership degree [5]. The new step is as follows:

$$\mu_{pj} = 1 - \alpha \sum_{i \ne p} \mu_{ij} = 1 - \alpha + \alpha\mu_{pj} \quad (6)$$

Where $0 \le \alpha \le 1$.

The S-FCM algorithm is as follows:

S-FCM algorithm

$S1$ Fix $m > 1, 2 \le c \le n - 1$

Initialize the c cluster centers v_i randomly.

$S2$ Give a value of the parameter α in [0, 1]

REPEAT

$S3$ Update $U^t = [\mu_{ij}]$ by (4).

$S4$ modify $U^t = [\mu_{ij}]$ by (6).

$S3$ Update $V^t = [v_i]$ by (5).

UNTIL (cluster centers stabilized)

- Modified suppressed fuzzy c-means algorithm

S-FCM speeds up FCM but its performance depends on which is selected randomly. Moreover, when $\alpha = 0$, the algorithm is equivalent to HCM and when $\alpha = 1$, the algorithm is equivalent to the FCM. That's why, Hung et al. modified it by adding a method which determined the value of and gave by consequence the new algorithm MS-FCM [6]. It performs clustering and also selects the parameter in S-FCM with a prototype-driven learning approach. The idea of selecting α is based on the separation strength given by $\min_{i \neq k} = \|v_i - v_k\|^2$. This term indicates the separation strength between clusters. If the value of $\min_{i \neq k} = \|v_i - v_k\|^2$ is large, the result is a lower degree of overlapping and a greater separation between the clusters is produced. That's a small value of α will be a better selection for S-FCM and vice-versa.

According to the above, we know that α is better assigned as a monotone decreasing function of $\min_{i = k} = \|v_i - v_k\|^2$. The researchers propose an exponential function in selecting α with

$$\alpha = \exp\left(-\min_{i \neq k} \frac{\|v_i - v_k\|^2}{\beta}\right) \tag{7}$$

$$\text{and } \beta = \frac{\sum_{j=1}^{n} \|x_i - \bar{x}\|^2}{n}$$

Thus, the MS-FCM objective function $J_{MS\text{-}FCM}$ is defined as:

$$J_{MS\text{-}FSM}(V, U, X) = \sum_{i=1}^{c} \sum_{j=1}^{n} \alpha u_{ij}^{m} d^{2}(x_j, v_i) \tag{8}$$

The S-FCM and MS-FCM have the same structure of the basic algorithm FCM, but a new step is inserted at algorithm; and this to give better classification results by using parameter α, that is constant in S-FCM and variable in MS-FCM.

The MS-FCM algorithm is summarized as follows:

MS-FCM algorithm

*S*1 Fix $m > 1, 2 \leq c \leq n - 1$

Initialize the c cluster centers v_i randomly.

REPEAT

*S*2 Compute α by (7).

*S*3 Update $U^t = [\mu_{ij}]$ by (4).

*S*4 modify $U^t = [\mu_{ij}]$ by (6).

*S*3 Update $V^t = [v_i]$ by (5).

UNTIL (cluster centers stabilized)

2.2 Possibilistic c-means algorithm

Krishnapuram and Keller proposed a possibilistic approach of c-means called possibilistic c-means, or PCM [19]. Their approach is expected to lead to better performance in the presence of noise [15]. This method permits to obtain clusters that correspond more closely to the intuitive concept of typicality or compatibility, and this by the discovery of the fuzzy partitions that do not satisfy the fuzzy constraint. The component generated by the PCM corresponds to a dense region in the data set; each cluster is independent of the other clusters in the PCM strategy. The objective function of the PCM can be formulated as follows:

$$J_{PCM}(V, U, X) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} d^{2}(x_j, v_i) + \sum_{i=1}^{c} \eta_i \sum_{j=1}^{n} (1 - u_{ij})^{m} \tag{9}$$

where

$$\eta_i = \frac{\sum_{j=1}^{n} \mu_{ij}^{m} \|x_j - v_i\|^2}{\sum_{j=1}^{n} \mu_{ij}^{m}} \tag{10}$$

is the scale parameter at the i^{th} cluster.

μ_{ij} is the typicality value of training sample x_j belonging to the cluster i. It is defined as:

$$u_{ij} = \frac{1}{1 + \left(\frac{d^2(x_j, v_i)}{\eta_i} \right)^{\frac{1}{m-1}}} \quad (11)$$

PCM reformulates FCM membership as a function of the distance of an object from its cluster. Krishnapuram and Keller [19] have followed the ideas of Zadeh [13] by relaxing the membership constraint of (3). The degrees of membership should only belong to the interval [0, 1]. Thus, a new set of constraints is defined:

$$\forall i \in [1..c], \forall j \in [1..n], \; \mu_{ij} \in [0..1] \; and \sum_{j=1}^{n} \mu_{ij} \prec n \quad (12)$$

$$\forall j \in [1..n] \max_{i} \mu_{ij} \succ 0 \quad (13)$$

The PCM depends on initialization [26]; the clusters do not have a lot of mobility, since each data point is classified as only one cluster at a time rather than all the clusters simultaneously. Therefore, a suitable initialization is required for the algorithms to converge to nearly global minimum.

3. Proposed Improved Possibilistic C-Means

MS-FCM previously mentioned, has proved effective in the classification of medical images and synthetic images. This was based on a parameter α founded by his predecessor S-FCM to accelerate FCM convergence to the optimal.

Studies [18] showed that the PCM method is preferred in the field of medical imaging and in a noisy environment. So our approach will be based on the classification possibilistic PCM with the integration of a new parameter β. The weight β will play the same role as á in MS-FCM.

S-FCM used α as a parameter that defines the relative importance of membership in computing of clusters. This parameter is characterized by a random choice of value. If you make a bad choice, the result of algorithm will thus be incorrect. MS-FCM suppresses the random selection of α and uses a formula for calculating a value of this parameter; but α is common to all points and all clusters of the data set. Or this reasoning is not logical because the context differs from one object to another, in relation to clusters.

The new approach called Improved Possibilistic C-Means will use the PCM algorithm by adding a coefficient β. The weight β will play the same role as α in MS-FCM. Or every object of the data has a weight to belong to a class. For this reason, this coefficient is calculated for each object in relation to the current class. It is called a degree of sharpness, with the following hypothesis: If the degree of membership of a point to a class is high, so it presents a point clear to the class in question. If the degree of membership of a point to a class is small, then it presents a vague point to the class in question. In MS-FCM, the calculation of α is based on the separation between the object and the class with the overall average compactness. Now each point in PCM x_j has a membership degree with respect to the class v_i, all alone. Moreover â will have the following formula:

$$\beta_{ij} = \exp\left(\frac{-\|x_j - v_i\|^2}{\sqrt{c}} \right) \quad (14)$$

The weight β_{ij} measures at what degree the object represents a clear point to whatever class. This coefficient modifies the typical partition.

This new algorithm IPCM is iterative, based on an objective function, which is the same as PCM function by adding to it the weight β_{ij}. A new relation, enabling a more rapid decrease in the function and increase in the membership when they tend toward 1 and decrease this degree when they tend toward 0. This relation is to add Weighting exponent as exhibitor of parameter β_{ij} in the objective function. The objective function of IPCM can be formulated as follows:

$$J_{IPCM}(V,U,X) = \sum_{i=1}^{c} \sum_{j=1}^{n} \beta_{ij}^{m} u_{ij}^{m} d^2(X_j, v_i) + \sum_{i=1}^{c} \eta_i \sum_{j=1}^{n} (1 - \beta_{ij}^{m} u_{ij}) \quad (15)$$

Therefore, we propose the improved possibilistic c-means (IPCM) algorithm that performs clustering and parameter selection simultaneously as follows:

IPCM algorithm

$S1$ Fix m > 1, $2 \le c \le n-1$

Initialize the c cluster centers v_i.

REPEAT

$S2$ Compute β_{ij} by (14).

$S3$ Update $U^t = [\mu_{ij}]$ by (11).

$S4$ modify $U^t = [\mu_{ij}]$ by (6).

$S3$ Update $V^t = [v_i]$ by (5).

UNTIL (cluster centers stabilized)

4. Experimental results

To show the feasibility of the methodology, we perform some experiments to compare the performances of these algorithms with some numerical datasets. All algorithms are implemented under the same initial values and stopping conditions.

We work with matlab version R2012b, a computer with operating system Windows 7 Professionnel.

The parameters we used for the different algorithms are: $m = 2$, maximal number of iterations = 800, $\varepsilon = 1e-5$.

For the data set we work with synthetic data sets and some MR images.

4.1 Results on synthetic images

We take two images composed of two symmetric classes. The second image represents the first one with an outlier. We show the results given by: PCM and IPCM.

Example 1: In the first experiment, we use a two-cluster data set ($X11$) as presented in [25]. We denote $\{x_1, x_2, \ldots, x_{11}\}$ by $X11$, to demonstrate the quality of classification of our approach (IPCM) in relation to the PCM algorithm in a case data set without outlier. The clustering results of these algorithms are shown in Figure 1, where two clusters from the clustering algorithms are with symbols + and o; also the figure shows that our approach is better than PCM.

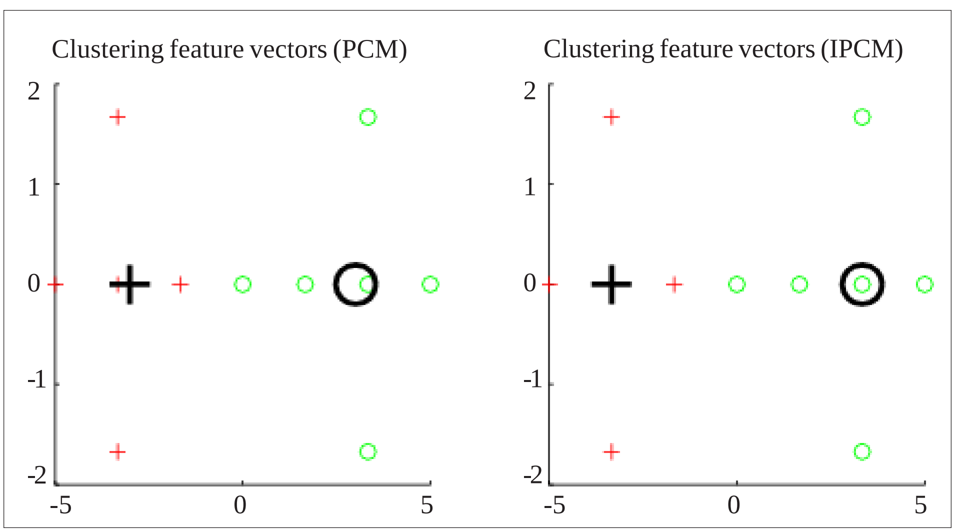

Figure 1. PCM, IPCM clustering results for the two-cluster data set without an outlier, X11

Example 2: In the second experiment, we use a two-cluster data set with outlier ($X12$) as presented in [25]. . $X12$ is $X11 \cup \{x_{12}\}$, x_{12} is called an outlier point. The clustering results of these algorithms are shown in Figure 2, shows that our approach is better than PCM.

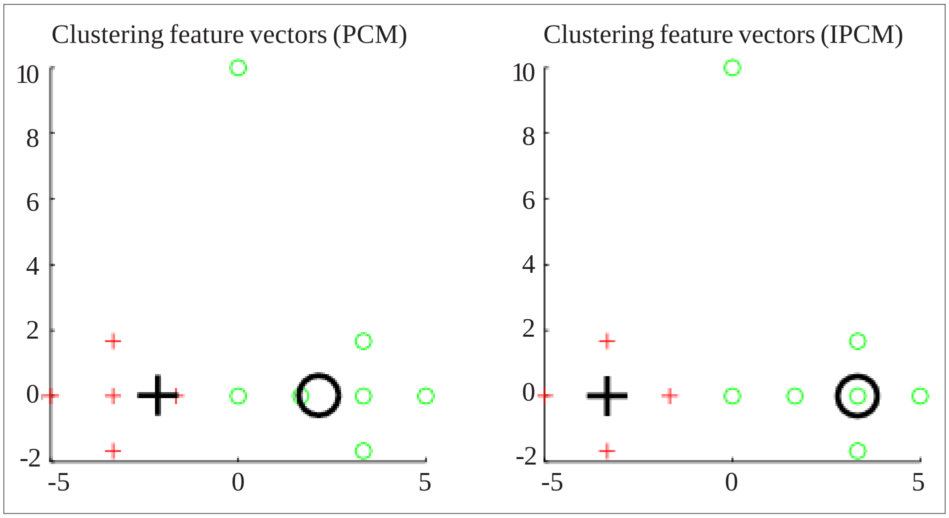

Figure 2. PCM, IPCM clustering results for the two-cluster data set with an outlier, X12

The PCM and IPCM are compared in the two previous experiences, using the following criteria for the cluster centers locations: the mean square error (MSE) of the centers, which its formula is:

$$MSE = \sqrt{||v_c - v_t||^2} \qquad (16)$$

where v_c is the computed center and v_t is the true center [10].

	X11		X12	
	PCM	***IPCM***	***PCM***	***IPCM***
MSE	0.468	7.36 e –5	1.688	0.0061
NI	14	3	26	9

Table 1. MSE And NI Generated By Pcby PCM And IPCM For X11 And X12

Table 1 shows that the cluster centers found by IPCM are closer the true centers, than the centers found by PCM. The number of iterations (*NI*) is much smaller in IPCM than in PCM.

4.2 Results on MR images

For the data set we work some MR images, which are attained by tumors. This choice has been imposed by specialists because the performance of clustering algorithm is judged on the basis of detection tumor in MR images. The number of clusters varies from one image to another.

In the experiment, we make a comparison of the new approach with some algorithms: FCM, S-FCM, MS-FCM and IPCM.

MRI 1: shows an image of the brain of a patient with a bronchial tumor (metastasis well rounded). Number of tested classes: 4 classes [30].

MRI 2: shows a thoracic lung cancer. Number of tested classes: 5 classes [30].

MRI 3: presents an MRI image of a brain of a patient with a tumor called a lipoma. Number of tested classes: 10 classes [30].

An indicator of classification quality index proposed by *XB* Xie and Beni (1991) and modified by Pal and Bezdek (1995) [11]:

$$XB(c) = \frac{\sum_{i=1}^{c}\sum_{j=1}^{n} \mu_{ij} \| x_j - v_i \|^2}{n \times \min_{i \neq j} \| v_i - v_j \|^2} \qquad (17)$$

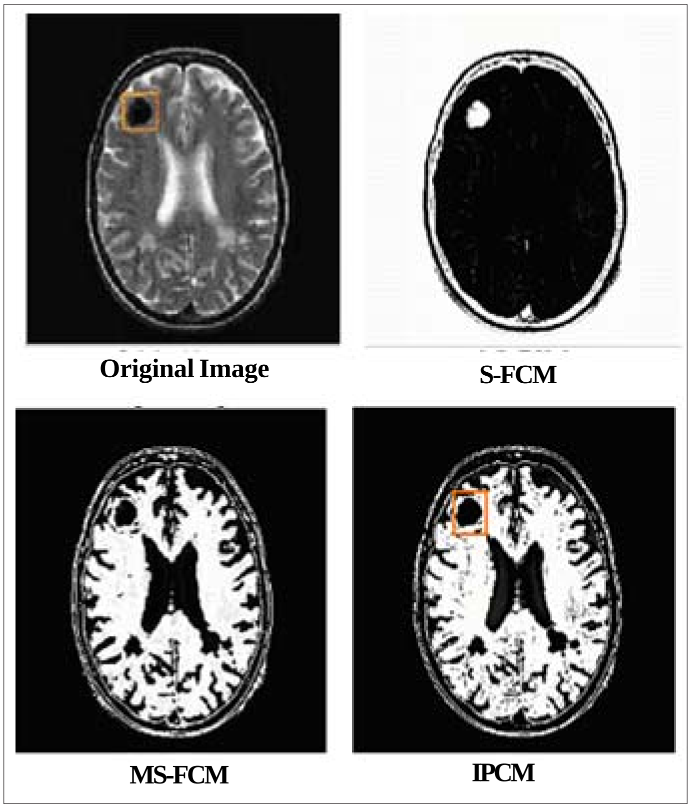

Figure 3. S-FCM; MS-FCM , IPCM clustering results for IRM1

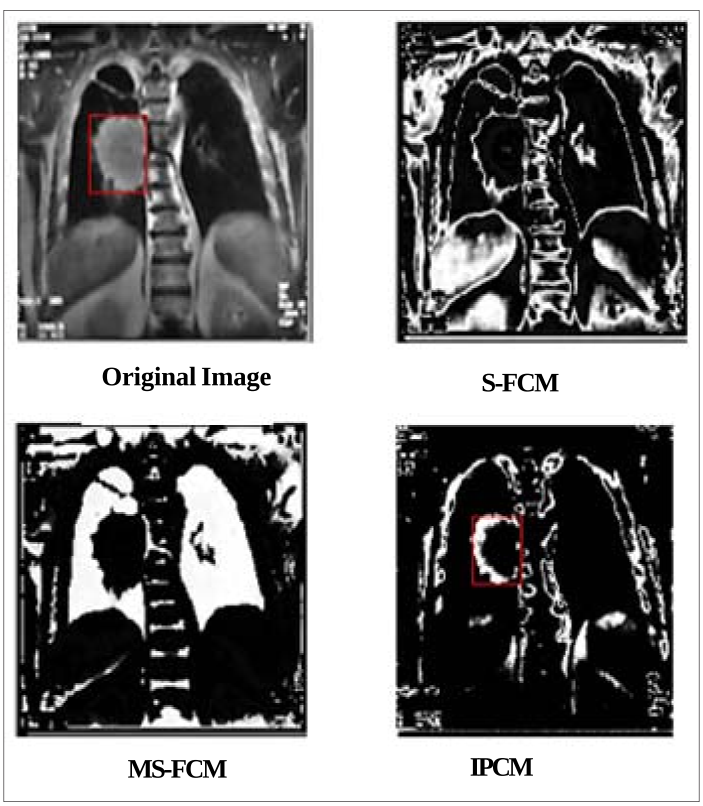

Figure 4. S-FCM; MS-FCM , IPCM clustering results for IRM2

For IRM1 the XB of IPCM is the greatest compared with the indexes of S-FCM and MS-FCM ($XB_{S\text{-}FCM}$ = 0.0392, $XB_{MS\text{-}FCM}$ = 0.1501 and XB_{IPCM} = 0.3512).

For IRM2 the XB of IPCM is the greatest compared with the indexes of S-FCM and MS-FCM ($XB_{S\text{-}FCM}$ = 0.0868, $XB_{MS\text{-}FCM}$ = 0.0958 and XB_{IPCM} = 0.1435).

For IRM3 the XB of IPCM is the greatest compared with the indexes of S-FCM and MS-FCM ($XB_{S\text{-}FCM}$ = 0.0675, $XB_{MS\text{-}FCM}$ = 0:0741 and XB_{IPCM} = 0.1118).

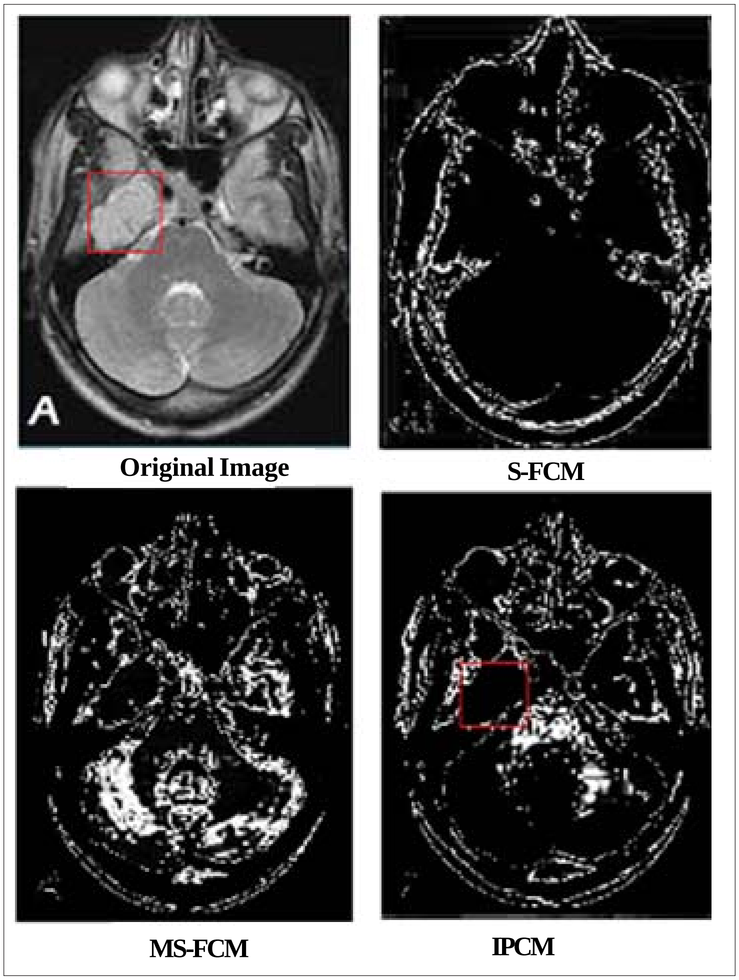

Figure 5. MS-FCM; IMS-FCM , IPCM clustering results for IRM3

According to the results of the three MRI images, the proposed approach gives a maximum value of the index relative to previous approaches, also the gap between the indices is large; which implies that IPCM gives a good classification result with a very clear difference.

5. Conclusions

In this paper, we analyzed a set of algorithms based on FCM and PCM algorithms; the objective was to study the behavior and the specificities of each algorithm in order to propose a new classification method more efficient than existing ones. We focused our work on a particular algorithm called MS-FCM, it has improved enormously the basic algorithm FCM and accelerated its convergence to the optimal. The idea of the proposed approach was to improve PCM using the principle of MS-FCM method.

The Improved Possibilistic C-Means uses a weight β, which is specific to each point relative to a class. To evaluate the new algorithm performance, we make an experimental study based on a comparison between the new approach and some existent algorithms. In our evaluation, we took account of the quality of the image after clustering. According to the results given on synthetic and real images, we remark that the new algorithm is better than the others. But the shortcomings of the proposed method are the complications and the use of more computation, because of the continued use of the new parameter β, which is calculated for each point and each iteration of the algorithm.

References

[1] Yang, X., Song, Q., Wang, Y. (2007). A weighted support vector machine for data classification. *International Journal of Pattern Recognition and Artificial Intelligence*, 21, p. 961–976.

[2] Zadeh, L. A. (1965). Fuzzy Sets. Inf.control, 8, p. 338–352.

[3] Ruspini, E. R. Numerical Methods for Fuzzy Clustering. *Information Science*, 2, p. 319–350.

[4] Dunn, J. C. (1973). A Fuzzy Relative of the ISODATA Process and its Use in Detecting Compact Well-Separated Clusters. J. *Cybernetics*, 3, p.32–57.

[5] Fan, J. L., Zhen, W. Z., Xie, W. X. (2003). Suppressed fuzzy c-means clustering algorithm. *Pattern Recognition Letters*, 24 (9)1607–1612.

[6] Hung, W. L., Yang, M. S., Chen, D. H. (2006). Parameter selection for suppressed fuzzy c-means with an application to MRI segmentation. *Pattern Recognition Letters*, 27 (5) 424–438.

[7] Bezdek, J. C., Harris, J. D. (1978). Fuzzy partitions and relations: axiomatic basic for clustering. Fuzzy Sets and Systems.

[8] Zadeh, L. A., Bellman, R., Kalaba, R. (1996). Abstraction and Pattern classification. *J. Math. Anal*, 13, p. 1–7.

[9] Dovzan, D., Skrjanc, I. (2011). Recursive fuzzy c-means clustering for recursive fuzzy identification of time-varying processes. *ISA Transactions*, 50, p. 159–169.

[10] Zhonghang, Y., Yangang, T., S. Funchun, S., Zengqi, S. (2006). Fuzzy Clustering with Novel Serable Criterion. *Tsinghua Science and Technology*, 11 (1) 50–53.

[11] Wu, K. L., Yang, M. S. (2005). A cluster validity index for fuzzy clustering. *Pattern Recognition Letters*, 26, p. 1275–1291.

[12] Timm, H., Borgelt, C., Doring, C., Kruse, R. (2004). Extension to possibilistic fuzzy cluster analysis. *Fuzzy Sets and Systems*, 147, p. 3–16.

[13] Zadeh, L. A . Fuzzy Sets as a Basis for a Theory of Possibility. Fuzzy Sets and Systems, 1, p. 3–28.

[14] Masulli, F., Schenone, A. (1999). A Fuzzy clustering based segmentation system as support to diagnosis in medical imaging. *Artificial Intelligence in Medicine*, 16, p. 129–147.

[15] Khodja, L. (1997). Contribution `a la classification non supervisé. PhD thesis, University of Savoie, France.

[16] Borgelt, C. (2005). Prototype-based classification and clustering. Habilitation, University of Magdeburg Germany, June.

[17] Bezdek, C. (1973). Fuzzy mathematics in pattern classification, PhD thesis dissertation, Cornell Univ., Ithaca, NY.

[18] Barra, V. (2000). Fusion 3D Images of the Brain: Study of Models and Applications. PhD thesis, University Blaise Pascal and University Auvergne, France, July.

[19] Krishnapuram, R., Keller, J. (1993). A possibilistic Approch to Clustering. *IEEE Trans. on Fuzzy Systems*, 1 (2), May.

[20] Barni, M., Cappellini, V., Mecocci, A. (1996). Comments on A possibilistic approach to clustering. *IEEE Trans. on Fuzzy Systems*, 4, p. 393–396.

[21] Bezdek, J. C. (1980). A convergence theorem for the fuzzu ISODATA clustering. *IEEE Trans. Pattern Anal. Machine Intell.*, 2 (1) 1–8.

[22] Berry, M. W. (2003). Survey of Text Mining. In Springer-Verlag, New York, NY, USA.

[23] Berthold, M R., Hand, D. J. (1999). Intelligent Data Analysis. *In*: Springer-Verlag, Berlin, Germany.

[24] Bezdek, J. C. (1981). Pattern recognition with fuzzy objective function algorithms. *In*: Plenum, New York.

[25] Pal, N. R., Pal, K., Bezdek, J. C. (1997). A mixed c-means clustering model. *In*: Proc. of the Sixth IEEE International Conference on Fuzzy Systems, July.

[26] SAAD, M. F., Alimi, A. M. (2009). Modified Fuzzy Possibilistic C-means. *Internationnal MultiConference of Engineers and Computer Scientists* (IMECS'09), 1, p. 177–182.

[27] SAAD, M. F., Karem, F., Alimi, A. M. (2009). Classification of MRI images par Robust Modified Suppressed Fuzzy C-means. Traitement et analyse de l'information : Méthodes et Applications (TAIMA'09), 1, p. 491– 197, May.

[28] SAAD, M. F., Alimi, A. M. (2010). Improved Modified Suppressed Fuzzy C-means. I*nternational Conference on Image Processing Theory, Tools and Applications* (IPTA'09), p. 313–318 , July.

[29] Batagelj, V., Ferligoj, A. (1998). Constrained Clustering Problems. *Conference of the International federation of classification societes* (IFCS'98), Rome.

[30] Website of Harvard University, www.med.harvard.edu.