

## Research Article

# Generalized Fuzzy C-Means Clustering with Improved Fuzzy Partitions and Shadowed Sets

Seyed Mohsen Zabihi<sup>1</sup> and Mohammad-R Akbarzadeh-T<sup>2</sup>

<sup>1</sup> Machine Vision Laboratory, Computer Engineering Department, Ferdowsi University of Mashhad, Mashhad 9177948944, Iran

<sup>2</sup> Departments of Electrical Engineering and Computer Engineering, Ferdowsi University of Mashhad, Mashhad 9177948944, Iran

Correspondence should be addressed to Seyed Mohsen Zabihi, mohsen.zabihi@hotmail.com

Received 13 September 2011; Accepted 18 October 2011

Academic Editor: I. Buci

Copyright © 2012 S. M. Zabihi and M.-R. Akbarzadeh-T. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Clustering involves grouping data points together according to some measure of similarity. Clustering is one of the most significant unsupervised learning problems and do not need any labeled data. There are many clustering algorithms, among which fuzzy c-means (FCM) is one of the most popular approaches. FCM has an objective function based on Euclidean distance. Some improved versions of FCM with rather different objective functions are proposed in recent years. Generalized Improved fuzzy partitions FCM (GIFP-FCM) is one of them, which uses  $L_p$  norm distance measure and competitive learning and outperforms the previous algorithms in this field. In this paper, we present a novel FCM clustering method with improved fuzzy partitions that utilizes shadowed sets and try to improve GIFP-FCM in noisy data sets. It enhances the efficiency of GIFP-FCM and improves the clustering results by correctly eliminating most outliers during steps of clustering. We name the novel fuzzy clustering method shadowed set-based GIFP-FCM (SGIFP-FCM). Several experiments on vessel segmentation in retinal images of DRIVE database illustrate the efficiency of the proposed method.

## 1. Introduction

One of the first mental activities of humans is clustering, where the goal is to find data structures and assign similar data to a group (cluster). Actually it tries to put unlabeled data into groups so that data points in a group are more similar to each other compared to those in other groups. It means that the goal is to maximize the intraclass likeness while minimizing the intercluster similarity.

Clustering has been used in diverse areas like machine vision and pattern recognition as well as in medical applications. There is a wide array of clustering algorithms [1–8]. But they can be generally classified to some groups as listed below:

- (i) exclusive clustering: one data point to belong to one cluster. No Overlapping (K-means, any linear classifier belongs to this kind),
- (ii) overlapping clustering: data can be in two or more clusters (fuzzy C-means),
- (iii) hierarchical clustering,
- (iv) probabilistic clustering (model based).

Among the above, overlapping clustering or partition-based clustering methods that group objects with some membership degrees are often used to handle noisy and uncertain data and hence can have a good utility in many practical applications.

Probabilistic and fuzzy clustering are two kinds of overlapping clustering methods when we consider the uncertainty in data. Actually, crisp clustering can be considered as a special type of these two clustering methods. FCM [9] is the most popular one among the other presented fuzzy clustering methods. It goes without saying that the later proposed fuzzy clustering methods are based on FCM in

general. FCM and its derived methods cluster data according to an objective function and several constraints. For instance, the summation of all the memberships of each data point to all clusters must be one.

A novel presentation of constraining the membership functions can be seen in [10]. In addition, a new FCM algorithm is also presented in [10], which is based on improved fuzzy partitions, named IFP-FCM. This algorithm is driven by crisp membership degrees, so it may seem resistant to noise and outliers. However, the fuzziness index  $m$  [10] is fixed and cannot be changed. Since the clusters in a data set have different densities, the performance of FCM may significantly depend on the choice of fuzziness index. Therefore, a good value of this parameter should be adopted to be consequent on the data distribution in data set [11]. So, another objective function should be introduced so that adjusts  $m$ . In [12] GIFF-FCM is proposed that does not restrict  $m$  to a fixed value. There is a parameter in this algorithm ( $\alpha$ ) that connects it to FCM and IFP-FCM. We can say briefly that this algorithm has the benefit of IFP-FCM and also generalizes it because of the various values that  $m$  can get.

In the paper, we present a novel Shadowed set-based Generalized Improved fuzzy partitions FCM (SGIFP-FCM). This new algorithm uses shadowed sets and the most important idea of the proposed method is to improve performance of GIFF-FCM in the stage of determining new cluster centers at each iteration. This is accomplished by removing the outliers and unsuitable data that have negative effects on structure of clusters. We will see that it improves the results of clustering in noisy systems and also it decreases the time of clustering in comparison with GIFF-FCM and the previous methods.

Clustering methods are often used in image processing applications such as in image segmentation [7]. To investigate the proficiency and utility of the presented method, the algorithm is applied here to vessel extraction of retinal images. These images are often used in medical applications to diagnosis of some diseases such as diabetes. The proposed approach is compared against other competing clustering algorithms on this database as well as an artificial dataset.

Some fuzzy clustering algorithms are briefly reviewed in Section 2. Section 3 describes the proposed algorithm (SGIFP-FCM) in detail. Section 4 shows experimental results of our work. Finally, conclusion is brought in Section 5.

## 2. Previous Fuzzy Clustering Algorithms

Many areas such as pattern recognition and machine vision utilize fuzzy clustering in solving their problems. A great number of fuzzy clustering algorithms are there, and it is noticeable that most of them use distance criteria. One of these algorithms that is so common and popular is FCM [13]. Reverse distance is used in FCM to fuzzy memberships.

In FCM, each feature vector can belong to every cluster with a coefficient between zero and one. Finally, algorithms label each data point (feature vector) based on the maximum coefficient of this data point over all clusters.

The fuzzy membership matrix and the cluster centers are computed by minimizing the following partition formula:

$$J_f(C, m) = \sum_{i=1}^c \sum_{k=1}^n (u_{i,k})^m d_{i,k} \quad (1)$$

subject to  $\sum_{i=1}^c u_{i,k} = 1.$

In this equation,  $n$  denotes the number of data,  $c$  the number of clusters,  $u_{i,k}$  the fuzzy membership of the  $k$ th data point to the  $i$ th cluster,  $d_{i,k}$  the Euclidean distance between the data point and the cluster center, and  $m \in (1, \infty)$  a fuzzy weighting factor that defines the degree of fuzziness of the results. The data class becomes fuzzier and less discriminating with increasing  $m$ . In general,  $m = 2$  is chosen (it is mentioned that this value of  $m$  does not produce optimal solution for all problems).

The constraint in (1) implies that each point must entirely distribute its membership among all the clusters. The cluster centers (centroids) are determined as the fuzzy-weighted center of gravity of the data  $X$ ,

$$v_i = \frac{\sum_{k=1}^n (u_{i,k})^m x_k}{\sum_{k=1}^n (u_{i,k})^m}, \quad i = 1, 2, \dots, c. \quad (2)$$

Since  $u_{i,k}$  affects the computation of the cluster center  $v_i$ , the data with a high membership will influence the prototype location more than points with a low membership. For the fuzzy C-means algorithm, distance  $d_{i,k}$  is defined as follows:

$$(d_{i,k})^2 = \|x_k - v_i\|^2. \quad (3)$$

The cluster centers  $v_i$  represent the typical values of that cluster, whereas the  $u_{i,k}$  component of the membership matrix denotes the extent to which the data point  $x_k$  is similar to its prototype. The minimization of the partition functional (1) will give the following expression for the membership:

$$u_{i,k} = \frac{1}{\sum_{j=1}^c (d_{i,k}/d_{j,k})^{1/(m-1)}}. \quad (4)$$

Equation (4) is determined in an iterative way since the distance  $d_{i,k}$  depends on membership  $u_{i,k}$ .

The procedure to calculate the FCM is as follows.

- (1) Opt for the number of clusters  $c$ ,  $2 \leq c < n$ ; Choose  $m$ ,  $1 \leq m < \infty$ . Initialize  $U^{(0)}$ .
- (2) Calculate the cluster centers  $v_i$  using (2).
- (3) Calculate new partition matrix  $U^{(1)}$  using (4).
- (4) Compare  $U^{(j)}$  and  $U^{(j+1)}$ . If the variation of the membership degree  $u_{k,i}$ , calculated with an appropriate norm, is smaller than a given threshold, terminate the algorithm, in other respects go back to step (2).

FCM clustering has a shortcoming in producing membership functions. For instance, these functions are not bounded and do not have to decay rapidly, so cannot be

understood locally. Distances to the Voronoi cell of the cluster instead of using distances to the cluster prototypes are applied in IFP-FCM method [10] by changing FCM's objective function. Objective function of IFP-FCM can be seen below:

$$J_{\text{IFP-FCM}} = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^2 d^2(x_i, v_i) - \sum_{j=1}^n a_j \sum_{i=1}^c \left(u_{ij} - \frac{1}{2}\right)^2. \quad (5)$$

In this equation  $a_j$  is a rewarding parameter. IFP-FCM seems more resistance to outliers and even noise.

Fuzzy index  $m$  is equal to 2 in many cases for FCM and IFP-FCM. Despite the fact that this parameter is required to be various or in different values for optimal or near-to-optimal results, fuzziness index  $m$  should be flexible and generalized. To do that, another objective function is necessary accordingly.

For that reason, GIFF-FCM clustering approach was presented in [12]. It is based on the rival-penalized competitive learning (RPCL) concept [14]:

$$J_{\text{GIFF-FCM}} = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d^2(x_i, v_i) + \sum_{j=1}^n a_j \sum_{i=1}^c u_{ij} (1 - u_{ij}^{m-1}). \quad (6)$$

It is clear that the new objective function uses the opinion of RPCL that for minimizing, only a specified  $u_{ij}$  gets the maximum reward and the other  $u_{kj}$  (rivals) get the minimum reward.

Authors in [12] showed that GIFF-FCM can convert to FCM or IFP-FCM with choosing proper values for some parameters and also it converges quicker rather than the other two clustering algorithms that we discussed before, but computational complexity of GIFF-FCM is the same as theirs. Fuzziness index  $m$  in IFP-FCM was equal to 2, whereas this parameter can be changed in GIFF-FCM properly.

### 3. The Proposed Method

In this section, we briefly describe shadowed sets and then try to propose our new clustering algorithm, SGIFF-FCM.

**3.1. Shadowed Sets.** Suppose that  $A$  is a fuzzy set in  $X$  which is an interval-valued one and maps element of  $X$  into  $[0, 1]$ , and the unit interval  $[0, 1]$ . Shadowed set  $B$  can be defined by this fuzzy set  $A$  so that  $B$  is a mapping  $B : X \rightarrow \{0, 1, [0, 1]\}$ , where  $0, 1$ , and  $[0, 1]$  illustrate complete exclusion from  $B$ , complete inclusion in  $B$ , and complete ignorance, respectively. Shadowed sets have some characteristics. For example, they are isomorphic with a three-valued logic. Shadowed sets and logic have similar operations. (For more details you can refer to [15, 16]).

Figure 1 depicts a fuzzy set and also its shadowed set.

**3.2. Creating a Shadowed Set.** Pedrycz in [17] presented the way of creating a shadowed set from a fuzzy set. Two threshold of  $\alpha$  and  $\beta$  should be defined at first as follows:  $\alpha \in [0, 0.5]$  and  $\beta \in [0.5, 1]$ ; next, low membership grades

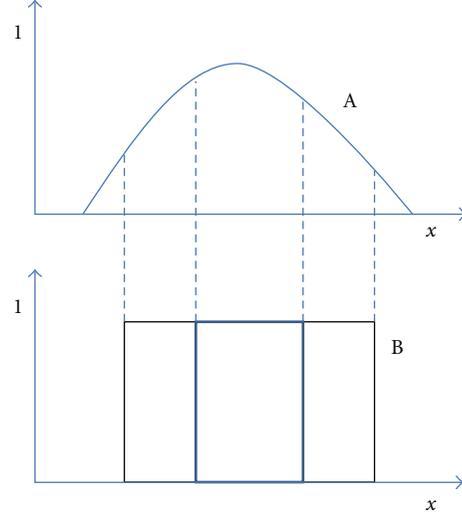


FIGURE 1: Fuzzy set A and its shadowed set (B).

and high membership grades should be changed to 0 and 1, respectively. At last, memberships are converted to some grades between  $\alpha$  and  $\beta$  in  $[0, 1]$ . Figure 2 shows this trend for a unimodal fuzzy set.

In Fact, the thresholds  $\alpha$  and  $1 - \alpha$  are proper choices, since the threshold  $\alpha$  is an essential part in process of constructing a shadowed set. So, finding such an optimal threshold is of utmost importance. Pedrycz considers the balance of uncertainty for this issue. The balance of uncertainty can be preserved by recompensing the changes of membership grades to zero and one for by creating the shadowed set that “absorb” the former elimination of partial membership at low and high ranges of membership.

According to Figure 2, the balance equation for fuzzy set with discrete membership function is defined as

$$\Omega_1 + \Omega_2 = \Omega_3,$$

$$\Omega_1 = \sum_{u_i \leq \alpha} u_i,$$

$$\Omega_2 = \sum_{u_i \geq 1 - \alpha} (1 - u_i),$$

$$\Omega_3 = \text{card}\{u_i \mid \alpha < u_i < 1 - \alpha\}.$$

It is better to change the above problem to (8) so that find the optimal value of parameter  $\alpha$

$$\text{Minimize } Q(\alpha) = |\Omega_1 + \Omega_2 - \Omega_3|. \quad (8)$$

**3.3. Applying Shadowed Sets to GIFF-FCM.** As we saw, clusters centers are modified in each iteration of the clustering algorithm and in addition to this, all data are participated in computation of each cluster center that increases time complexity of the algorithm. For decreasing the time of algorithm and improve the efficiency, we use shadowed sets to remove unsuitable data, outliers, and also data in borders before determining centers of clusters. We show partitions matrix with  $U = (u_{ij})_{c \times n}$  and each row of this matrix

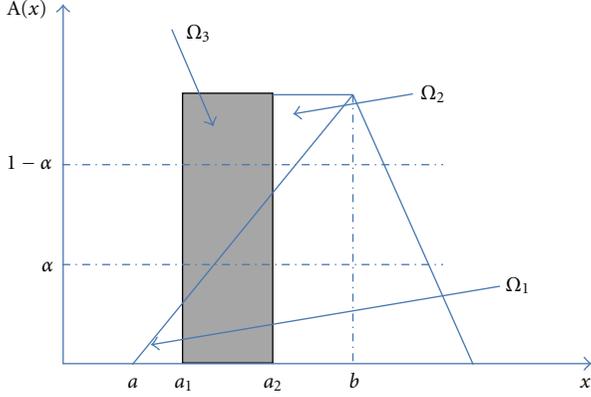


FIGURE 2: Constructing of a shadowed set B.

describes a cluster. We should compute optimal threshold  $\alpha_i$  ( $i = 1, 2, \dots, c$ ) for each row and then can remove outliers.  $\alpha_i$  ( $i = 1, 2, \dots, c$ ) can be computed with solving the below optimization problem:

$$\text{Minimize } Q(\alpha) = |\Omega_{i1} + \Omega_{i2} + \Omega_{i3}|, \quad (9)$$

where

$$\begin{aligned} \Omega_{i1} &= \sum_{j:u_{ij} \leq \alpha_i} u_{ij}, \\ \Omega_{i2} &= \sum_{j:u_{ij} \geq 1-\alpha_i} (1 - u_{ij}), \end{aligned} \quad (10)$$

$$\Omega_{i3} = \text{card}\{u_{ij} \mid \alpha_i < u_{ij} < 1 - \alpha_i, j = 1, 2, \dots, n\}.$$

**3.4. SGIFP-FCM Algorithm.** Now we can summarize our proposed algorithm (S FIFP-FCM) in some steps as follows.

*Step 1.* Set the number of clusters ( $c$  is between 1 and  $n$ ) and also choose proper values for threshold of stopping algorithm ( $\epsilon$ ), parameter  $\alpha'$  in  $[0, 1)$  that determines the rate of rewarding, number of iterations and fuzziness index  $m$ , and initial values of  $u_{ij}$ .

*Step 2.* Compute the optimum value  $\alpha$  for each row of partition matrix using (9). If  $u_{ij} < \alpha$ , it does not have any role in determining cluster center.

*Step 3.* Compute clusters centers ( $v_i^{(l+1)}$ ) using

$$v_i^{(l+1)} = \frac{\sum_{j:u_{ij} \geq \alpha} u_{ij}^m x_j}{\sum_{j:u_{ij} \geq \alpha} u_{ij}^m} \quad (i = 1, 2, \dots, c). \quad (11)$$

*Step 4.* Compute membership functions  $u_{ij}^{(l+1)}$  using

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left( \frac{d^2(x_j, v_i) - \alpha' \min_{1 \leq a \leq c} d^2(x_j, v_a)}{d^2(x_j, v_k) - \alpha' \min_{1 \leq a \leq c} d^2(x_j, v_a)} \right)^{1/(m-1)}}. \quad (12)$$

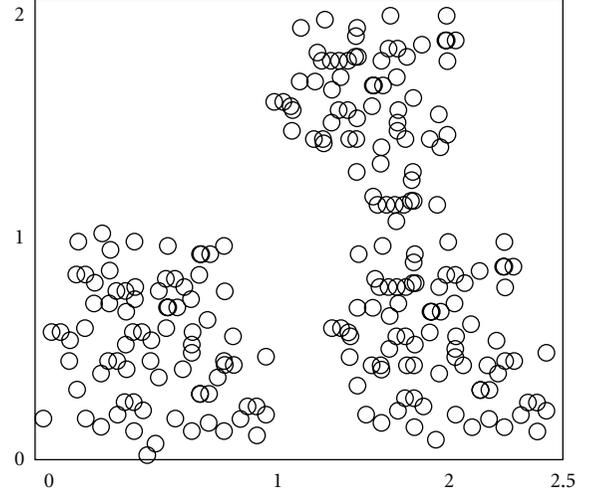


FIGURE 3: Data set of the first experiment with three clusters.

*Step 5.* If the terminating condition is satisfied, the algorithm is finished, otherwise increase iteration number and go back to Step 2.

## 4. Results

In the first experiment, we create a random and artificial data set with three clusters or groups of data as it is shown in Figure 3. We can observe the results of these data clustering with three algorithms of FCM, GIFP-FCM, and SGIFP-FCM in Figure 4. Almost FCM and SGIFP-FCM depict similar results and detect the three clusters correctly, but in GIFP-FCM contrary to SGIFP-FCM all data are participated in determining centers and then could not cluster data exactly. This is because of the effects of outliers and data in borders of clusters on determining centers, which are removed in SGIFP-FCM.

In Figure 4, clusters are distinguished with different colors and centers of clusters are specified with "+". Values of parameters  $\alpha$  and  $m$  are supposed to be 2 and 0.9, respectively.

In the second experiment, we use retinal images of DRIVE database [18]. We want to detect vessels in these images using clustering methods. For this purpose, we first extract some features for each pixel of the processed image and then we get these feature vectors to clustering algorithm to label the pixels as vessels or nonvessels or clearly divide the pixels into two clusters of vessels and nonvessels. There are many feature extraction methods. We use LBP (Local Binary Patterns) method for feature extraction [19]. The features vector of each pixel is composed of gray level intensity, variance, and LBP value in three local windows around the central pixel [20].

Sample of a segmented retina image using the proposed algorithm (SGIFP-FCM) is presented in Figure 5 along with manually segmented version of the image.

To evaluate the algorithm, we calculated true positive ratio (TPR) and False Positive Ratio (FPR). TPR is number

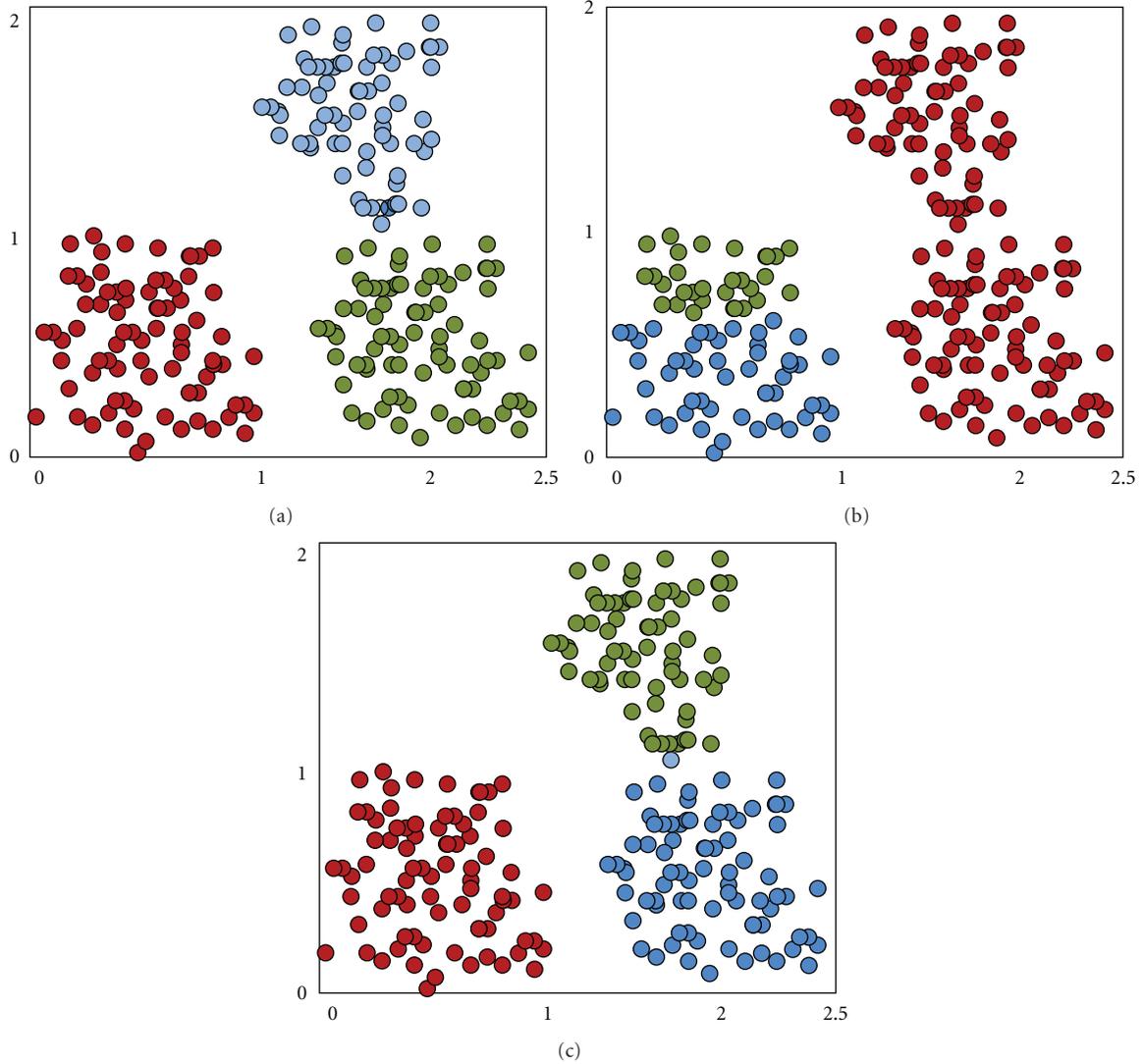


FIGURE 4: Clustering results of data set in Figure 3. with methods of (a) FCM, (b) GIFP-FCM, and (c) SGIFP-FCM.

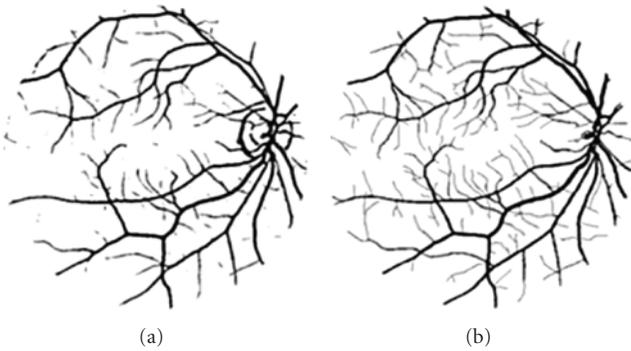


FIGURE 5: Clustering result, (a) segmentation using SGIFP-FCM, and (b) manually segmented image.

of pixels of resulting image, which are correctly clustered as vessel (according image generated by human expert) to total number of pixels of human-expert generated image

TABLE 1: Comparison of SGIFP-FCM and in retina vessel extraction.

| Algorithm          | TPR    | FPR    | Time(S) |
|--------------------|--------|--------|---------|
| 2nd human observer | 0.7761 | 0.0275 | —       |
| SGIFP-FCM          | 0.6944 | 0.029  | 2.76    |
| GIFP-FCM           | 0.6821 | 0.032  | 3.41    |

that are labeled as vessel. FPR is the number of pixel of resulting image, which are incorrectly clustered as vessel to total number of pixel of human-expert generated image that are labeled as background.

Table 1 compares performance of retina vessel extraction with SGIFP-FCM and GIFP-FCM versus 2nd human observer with regard to the TPR, FPR, and time of clustering.

The results in Table 1 are average results of 20 test images of DRIVE database.

As we see in Table 1, performance of SGIFP-FCM is better than GIFP not only in extracted vessels but also in time consuming.

## 5. Conclusion

In this paper, we present a Generalized Fuzzy C-Means Clustering With Improved Fuzzy Partitions and Shadowed sets (SGIFP-FCM). As is illustrated, this new algorithm improves the previous clustering algorithms such as GIFP-FCM in noisy data sets. In the proposed algorithm, the effects of outliers and noisy data are reduced in determining clusters centers using shadowed sets. Finally this new algorithm is tested on several experiments. Performance of this method is evaluated on image segmentation application (retina images) and because there are many noises in these images this algorithm presented better results than GIFP-FCM algorithm.

## References

- [1] R. Ng and J. Han, "Efficient and effective clustering methods for spatial data mining," in *Proceedings of the 20th International Conference on Very Large Data Bases (VLDB '94)*, pp. 144–155, Santiago, Chile, 1994.
- [2] J. C. Bezdek, R. J. Hathaway, M. J. Sabin, and W. T. Tucker, "Convergence theory for fuzzy c-means: counterexamples and repairs," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 17, no. 5, pp. 873–877, 1987.
- [3] G. Karypis, E. H. Han, and V. Kumar, "Chameleon: hierarchical clustering using dynamic modeling," *Computer*, vol. 32, no. 8, pp. 68–75, 1999.
- [4] T. Zhang, R. Ramakrishnan, and M. Livny, "BIRCH: a new data clustering algorithm and its applications," *Data Mining and Knowledge Discovery*, vol. 1, no. 2, pp. 141–182, 1997.
- [5] G. Sheikholeslami, S. Chatterjee, and A. Zhang, "WaveCluster: a multiresolution clustering approach for very large spatial databases," in *Proceedings of the 24th International Conference on Very Large Data Bases (VLDB '98)*, pp. 428–439, New York, NY, USA, 1998.
- [6] Y. Xiu, S. T. Wang, X. Wu, and D. Hu, "The directional similarity-based clustering method DSCM," *Journal of Computer Research and Development*, vol. 43, no. 8, pp. 1425–1431, 2006 (Chinese).
- [7] F. L. Chung, S. T. Wang, M. Xu, D. Hu, and Q. Lin, "Possibility theoretic clustering and its preliminary application to large image segmentation," *Soft Computing*, vol. 11, no. 2, pp. 103–113, 2007.
- [8] J. L. Fan, W. Z. Zhen, and W. X. Xie, "Suppressed fuzzy c-means clustering algorithm," *Pattern Recognition Letters*, vol. 24, no. 9–10, pp. 1607–1612, 2003.
- [9] J. C. Bezdek, *Pattern Recognition With Fuzzy Objective Function Algorithms*, Plenum, New York, NY, USA, 1981.
- [10] F. Höppner and F. Klawonn, "Improved fuzzy partitions for fuzzy regression models," *International Journal of Approximate Reasoning*, vol. 32, no. 2, pp. 85–102, 2003.
- [11] C. Hwang and F. Rhee, "Uncertain fuzzy clustering: interval type-2 fuzzy approach to C-means," *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 1, pp. 107–120, 2007.
- [12] L. Zhu, F.-L. Chung, and S. Wang, "Generalized fuzzy C-means clustering algorithm with improved fuzzy partitions," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 39, no. 3, pp. 578–591, 2009.
- [13] F. Höppner and F. Klawonn, "A contribution to convergence theory of fuzzy c-means and derivatives," *IEEE Transactions on Fuzzy Systems*, vol. 11, no. 5, pp. 682–694, 2003.
- [14] L. Xu, A. Krzyzak, and E. Oja, "Rival penalized competitive learning for clustering analysis, RBF net, and curve detection," *IEEE Transactions on Neural Networks*, vol. 4, no. 4, pp. 636–649, 1993.
- [15] W. Pedrycz, "Shadowed sets: representing and processing fuzzy sets," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 28, no. 1, pp. 103–109, 1998.
- [16] W. Pedrycz and G. Vukovich, "Granular computing with shadowed sets," *International Journal of Intelligent Systems*, vol. 17, no. 2, pp. 173–197, 2002.
- [17] W. Pedrycz, "Shadow sets: bridging fuzzy and rough sets," in *Rough Fuzzy Hybridization: A New Trend in Decision-Making*, S. K. Pal and A. Skowron, Eds., Springer, Singapore, 1999.
- [18] M. Niemeijer and B. van Ginneken, <http://www.isi.uu.nl/Research/Databases/DRIVE/results.php>, 2002.
- [19] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [20] S. H. Rezatofghi, A. Roodaki, and H. A. Noubari, "An enhanced segmentation of blood vessels in retinal images using Contourlet," in *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '08)*, pp. 3530–3533, August 2008.