Fuzzy c-means clustering based on Gaussian spatial information for brain MR image segmentation

Abbas Biniaz, Ataollah Abbassi
Computational Neuroscience Laboratory, Sahand University of Technology
Tabriz, Iran
CNLab@sut.ac.ir

Mousa Shamsi, Afshin Ebrahimi.
Department of Electrical Engineering, Sahand University of Technology
Tabriz, Iran
{shamsi,aebrahimi}@sut.ac.ir

***Abstract*—Conventional fuzzy c-means (FCM) algorithm is highly vulnerable to noise due to not considering the spatial information in image segmentation. This paper aims to develop a Gaussian spatial FCM (gsFCM) for segmentation of brain magnetic resonance (MR) images. The proposed algorithm uses fuzzy spatial information to update fuzzy membership with a Gaussian function. Proposed method has less sensitivity to noise specifically in tissue boundaries, angles, and borders than spatial FCM (sFCM). Furthermore by the proposed algorithm a pixel which is a distinct tissue from anatomically point of view for example a tumor in preliminary stages of its appearance, has more chance to be a unique cluster. The quantitative assessment of presented FCM techniques is evaluated by conventional validity functions. Experimental results show the efficiency of proposed algorithm in segmentation of MR images.**

Keywords-component; Segmentation; MRI; FCM; spatial information .

I. Introduction

Magnetic resonance imaging (MRI) is most common imaging modalities employed as a diagnostic technique [1]. Segmentation of medical images inferred to partition pixels/voxels in an image into the number of 2D/3D tissues, each with unique features and similar properties. Segmentation process could be based on numerous features of input data. Therefore a variety of edge based techniques has been developed in image segmentation. Here is a list of edge operators which commonly is used in the image segmentation trials: Sobel, Roberts, Prewitt, Canny, Zero-crossing, Laplacian, and Laplacian of Gaussian(LoG) [2, 3]. There are the large number of gray level based approaches for segmentation of medical images using both local and universal image intensity information. Thresholding is one of the image segmentation techniques and has two common types: Global thresholding , and Local thresholding [4].

Region based approaches are popular segmentation procedures. A well-developed region based method is region growing. Based on some predefined criteria, a connected area is portrayed by region growing. Disadvantageous of these methods are creation of holes and disconnectedness in segmented image [5, 6]. Other methods like deformable models and active contours models (ACMs) or level set are applied as numerical methods for tracking boundaries and borders in an image [7].

Fuzzy clustering has many applications in medical image segmentation, because they can preserve more information about original image using fuzziness membership than other methods [8]. However standard FCM doesn't exploit spatial information of neighborhood pixels in image segmentation. In order to develop a modified FCM algorithm compared with sFCM approach [9], this paper presents a modified sFCM algorithm based on Gaussian spatial information as gsFCM. New approach extracts tissue boundaries, borders, angles, and small organisms successfully. The rest of this paper is organized as follows: Section 2 introduces methodology of this paper. Section 3 describes quantitative validity functions; and Section 4 presents experimental results. Section 5 summarizes conclusions of this paper.

II. Methodology

A. Fuzzy c-Means Clustering

Fuzzy c-Means clustering algorithms, developed in 1970s and optimized later [10]. Let $X = \{x_1, x_2, ..., x_n\}$ denotes an input vector with n number of elements to be partitioned into c ($2\leq c\leq n$) clusters, and x_j denotes the feature value. The FCM algorithm is an iterative optimization process that minimizes the following cost function:

$$J_m(U,V) = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m \|x_j - v_i\|^2 \quad (1)$$

Where n is the number of data points and m is the fuzziness value (1 in hard clustering, and will be increased in fuzzy clustering). u_{ij} is membership of pixel x_j in the i-th cluster that v_i is centroid of it; and $\|.\|$ is a norm metric. Cluster centers and membership functions in FCM are updated by the following [8]:

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{\|x_j - v_i\|}{\|x_j - v_k\|}\right)^{2/(m-1)}} \quad (2)$$

And

$$v_i = \frac{\sum_{j=1}^{n} u_{ij}^m x_j}{\sum_{j=1}^{n} u_{ij}^m} \quad (3)$$

B. Spatial FCM (sFCM)

Correlation among neighboring pixels is one of the significant characteristics in the brain MR images. This means that neighborhood pixels have many similarities and analogous feature properties hence with great probability they are members of the unique clusters. To utilize spatial information in FCM algorithm, the spatial function can be impressively represented as [9]:

$$h_{ij} = \sum_{k \in NB(x_j)} u_{ik} \quad (4)$$

The spatial function h_{ij} just like the membership function u_{ij} signifies the probability that pixel x_j belongs to i-th cluster. However, h_{ij} contains spatial information of MR image. $NB_{(x_j)}$ represents a 5×5 lattice window and x_j is a pixel in the lattice window [8].

C. Proposed Gaussian spatial FCM (gsFCM)

The sFCM algorithm with a linear filter on membership function reduces effects of noise in MR images [9]. However this has disadvantages on tissue boundaries, borders, angles, and small organisms. Furthermore, one pixel which is anatomically a distinct tissue for example a tumor or pathological lesion in preliminary of its appearance has less chance to be classified as a unique cluster. In this paper to surmount on mentioned disadvantages a Gaussian function is applied in standard FCM[11]. This function handles Gaussian spatial information on FCM and proposes gsFCM algorithm. The optimized algorithm preserves superiority of sFCM and modifies its disadvantages is as follow:

$$h_{ij} = \sum_{k \in NB'_{(x_j)}} \sum_{l \in NB'_{(x_j)}} \frac{1}{2\pi\sigma^2} e^{-\frac{k^2+l^2}{2\sigma^2}} u_{k,l} \quad (5)$$

The proposed Gaussian spatial function h_{ij} just like the membership function $u_{k,l}$, indicates the possibility that pixel x_j belongs to i-th cluster. The lattice window $NB'_{(x_j)}$ denotes a 5×5 square window with the Gaussian spatial information. Incorporation of the Gaussian spatial function into membership function is as follows [9]:

$$u_{ij}^* = \frac{u_{ij}^p \times h_{ij}^q}{\sum_{k=1}^{c} u_{kj}^p \times h_{kj}^q} \quad (6)$$

Where u_{ij}^* is new membership function, and the parameters p and q signifies the comparative influence of both membership and Gaussian spatial functions u_{ij} and h_{ij} respectively. The improved spatial FCM by Gaussian function with parameters p and q is represented as $gsFCM_{p,q}$. The proposed gsFCM algorithm is summarized as follows:

Step 1: Select the number of clusters c and fuzziness grade m; let ε be a small positive constant, and initialize $V^{(0)}$ matrix by randomly small values.

Step 2: Update the membership matrix U^*, using (6) .

Step 3: Update cluster center matrix V, using (3).

Step 4: Repeat steps 2-3 until termination. The termination criterion in two successive iteration is as follows $\left\|v_i^{(t+1)} - v_i^{(t)}\right\| < \varepsilon$, where $\|.\|$ is norm metric.

III. CLUSTER VALIDITY FUNCTIONS

Mostly two types of validity functions are used to evaluate the performance of clustering: fuzzy partition and geometric structure. Partition coefficient V_{pc} and partition entropy V_{pe} are fuzzy partition functions defined as following [12, 13]:

$$V_{PC} = \frac{\sum_{j=1}^{n}\sum_{i=1}^{C} {u_{ij}}^2}{N} \quad (8)$$

$$V_{pe} = (-1) \times \frac{\sum_{j=1}^{n}\sum_{i=1}^{c} u_{ij} log(u_{ij})}{n} \quad (9)$$

The best clustering is achieved when the V_{pc} has maximum value (close to 1) or V_{pe} has minimum value (close to 0). However, fuzzy partition functions can only measure the fuzzy partition and don’t have a direct access to feature vector. To quantify the ratio of total variation within clusters using geometric structure, V_{fs} and V_{xb} are defined as follow [12, 13]:

$$V_{fs} = \sum_{j=1}^{n}\sum_{i=1}^{C} u_{ij}^m \left(\|x_j - v_i\|^2 - \|v_i - \bar{v}\|^2\right) \quad (10)$$

$$V_{xb} = \frac{\sum_{j=1}^{n}\sum_{i=1}^{c} {u_{ij}}^m \left(\|x_j - v_i\|^2\right)}{n * \left(min_{i,k}\{\|v_k - v_i\|^2\}\right)} \quad (11)$$

Where $v_i \neq v_k$, and minimized V_{fs} or V_{xb} lead to optimal clustering.

IV. RESULTS AND DISCUSSION

The synthetic and real MR images with various white Gaussian noise values have been used in the experiments. Fig. 1(a) depict four-level synthetic T1 weighted image [9] corrupted by additive Gaussian noise ($m=0$, σ=0.003). Fig. 1 (b)-(f) show the clustering results with FCM techniques respectively. The $sFCM_{0,2}$ and proposed $gsFCM_{0,2}$ techniques as qualitative are superior to other FCM techniques. However $gsFCM_{0,2}$ is superior in the inner boundary than $sFCM_{0,2}$. In all experiments a symmetric 5×5 lattice window by a Gaussian spatial filter and standard deviation 0.8 in gsFCM technique is used.

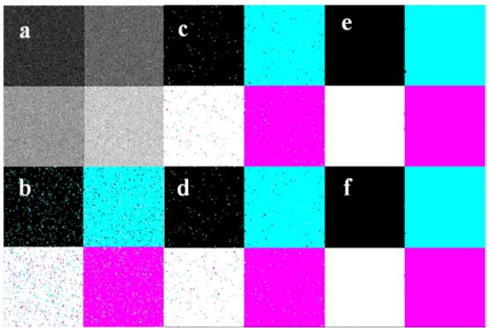

Figure 1. (a) Simulated MR image corrupted by additive Gaussian noise (m=0, σ=0.003). The gray levels are 50 (UL), 100 (UR), 150 (LL), and 200 (LR). Clustering results using (b) FCM, (c) $sFCM_{1,1}$, (d) $gsFCM_{1,1}$, (e) $sFCM_{0,2}$, and (f) $gsFCM_{0,2}$.

Segmentation results on simulated T1 weighted image of humane brain is portrayed on Fig. 2. This image tainted by additive Gaussian noise (m=0, σ=0.001). Performances of FCM, sFCM, and proposed gsFCM techniques are observed in this image. As can be seen gsFCM due to influence of Gaussian spatial function on neighborhood pixels, efficiently manages tissue boundaries, borders, angles, and small tissues.

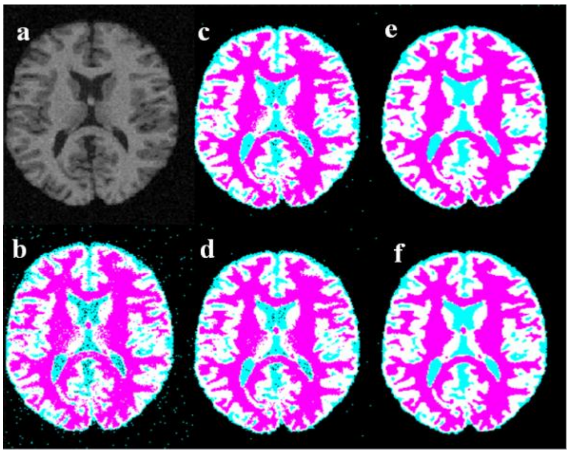

Figure 2. (a) Simulated MR image tainted by Gaussian noise (m=0, σ=0.001). Segmentation results using (b) FCM, (c) $sFCM_{1,1}$, (d) $gsFCM_{1,1}$, (e)$sFCM_{0,2}$, and (f) $gsFCM_{0,2}$.

To scrutinize between $sFCM_{0,2}$ and $gsFCM_{0,2}$ in Fig. 3 fuzzy and hard clustering results are portrayed. Columns from left to right are background (BGND), CSF, GM, and WM respectively. 1st and 2nd rows are results of fuzzy clustering by $sFCM_{0,2}$ and $gsFCM_{0,2}$ respectively. By fuzzy clustering, fuzzy membership of each cluster is portrayed. As can be seen tissue boundaries are correctly segmented by proposed gsFCM. Hard clustering is represented in 3rd and 4th rows for $sFCM_{0,2}$ and $gsFCM_{0,2}$ respectively. Assigning rigid membership in hard clustering to each cluster, tissue boundaries are exactly distinguished by two approaches and their differences are comparable.

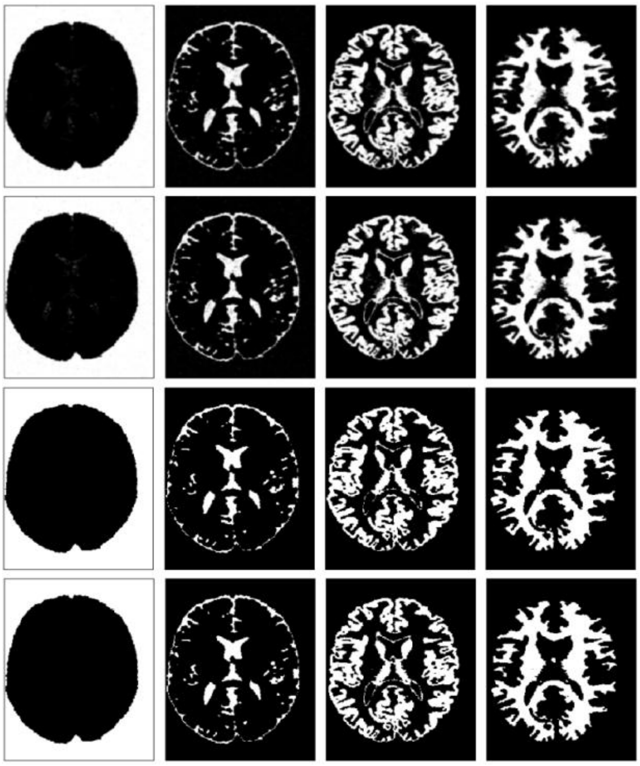

Figure 3. Fuzzy and hard clustering on simulated MR image. Columns from left to right are BGND, CSF, GM, and WM respectively. 1st and 2nd rows are results of fuzzy clustering respectively by $sFCM_{0,2}$ and $gsFCM_{0,2}$. In 3rd and 4th rows results of hard clustering by $sFCM_{0,2}$ and $gsFCM_{0,2}$ are represented respectively.

In addition to more investigate on FCM techniques, simulations were done on real MR images. Fig. 4 (a) depict T1 weighted image; and Fig. 4 (b)-(f) show clustering results with the FCM techniques.

Moreover in Fig. 5, fuzzy and hard clustering results on real MR image is portrayed. Columns from left to right are BGND, WM, GM, and CSF successively. 1st and 2nd rows are result of fuzzy clustering by $sFCM_{0,2}$ and $gsFCM_{0,2}$ respectively. Fuzzy membership in fuzzy clustering is assigned to each pixel then each tissue is portrayed by its real intensity. Proposed gsFCM efficiently reduces noise effects like sFCM however gsFCM has better performances in tissue boundaries. Hard clustering results is represented in 3rd and 4th rows for $sFCM_{0,2}$, and $gsFCM_{0,2}$ techniques respectively. Then by assigning rigid membership to each pixel tissue boundaries are depicted exactly by two algorithms.

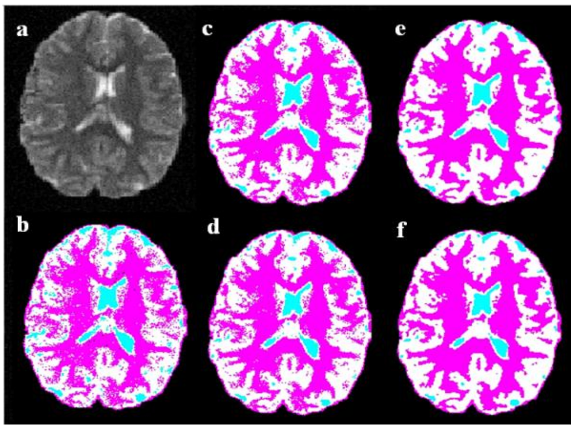

Figure 4. (a) Real T1 weighted brain image corrupted by additive Gaussian noise (m=0, σ=0.001). Segmented images using (b) FCM; (c) $sFCM_{1,1}$,(d) $gsFCM_{1,1}$, (e)$sFCM_{0,2}$, and (f) $gsFCM_{0,2}$.

TABLE.I and figure 6 show fuzzy validity function results to evaluate performance of FCM techniques on various MR images. High negative values of V_{fs} represents the high performance of algorithm; V_{pc} when is closer to one reveals that algorithm performance is closer to optimum. Most close to zero in V_{pe} and V_{xb} , reflects the highest quality of segmentation.

V. Conclusion

Standard FCM has been applied efficiently to brain MR image segmentation. These images have high homogeneity in spatial domain however these spatial relationships among neighborhood pixels are seldom employed in standard FCM. In this paper, spatial information was applied in two different linear and nonlinear modes. In linear mode (sFCM), spatial information was incorporated by equal weigh coefficients; these equal coefficients caused to misclassification in tissue boundaries, borders, angles, and small organisms. Therefore by the proposed gsFCM algorithm a weighed summation of spatial information by a Gaussian function was assigned to neighborhood pixels in MR images. By the proposed approach, sFCM disadvantages were modified. Furthermore quantitative assessment of sFCM and proposed gsFCM techniques were evaluated by conventional fuzzy validity functions.

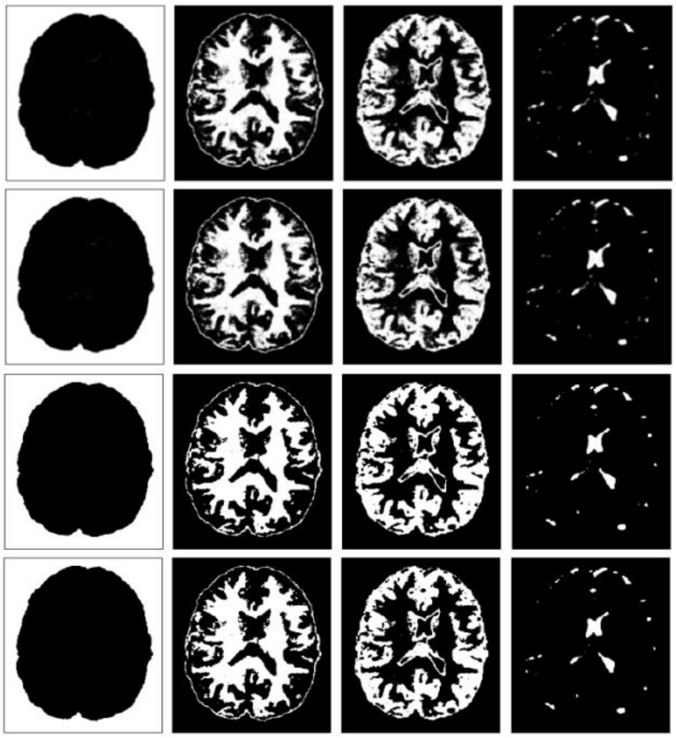

Figure 5. Fuzzy and hard clustering on real MR image. Columns from left to right are BGND, WM, GM and CSF respectively. 1st and 2nd rows are result of fuzzy clustering by $sFCM_{0,2}$ and $gsFCM_{0,2}$ successively. 3rd and 4th rows show result of hard clustering by $sFCM_{0,2}$ and $gsFCM_{0,2}$ respectively.

TABLE I. VALIDATION FUNCTIONS FOR DIFFERENT SIMULATED AND REAL MR IMAGES.

	Images	*Standard FCM*	*$SFCM_{1,1}$*	*$gsFCM_{1,1}$*	*$sFCM_{0,2}$*	*$gsFCM_{0,2}$*
V_{xb}	simulated 4 level MRI (σ=0.003)	0.047	0.062	0.061	0.076	0.071
	simulated MRI (σ=0.001)	0.048	0.053	0.053	0.071	0.064
	Real MRI(0.001)	0.080	0.079	0.078	0.106	0.091
*V_{fs} *(-10^6)*	simulated 4 level MRI (σ=0.003)	115	129	129	125	126
	simulated MRI (σ=0.001)	92	100	100	96	97
	Real MRI(σ=0.001)	128	146	147	143	146

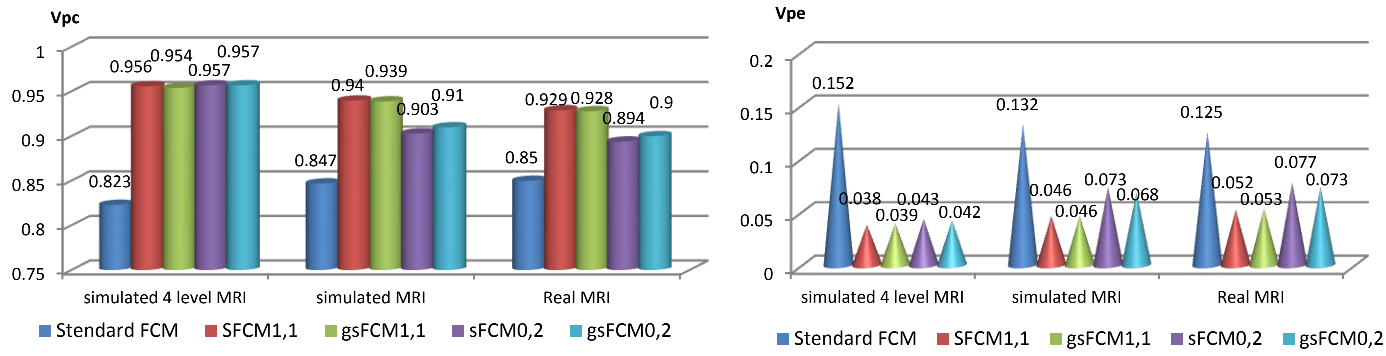

Figure 6. Vpc and Vpe for various MR images (better segmentation is most close to one for Vpc and most close to zero for Vpe).

REFERENCES

[1] J. Bezdek, L. Hall, and L. Clarke, "Review of MR image segmentation techniques using pattern recognition," *MEDICAL PHYSICS-LANCASTER PA-*, vol. 20, pp. 1033-1033, 1993.

[2] S. Rujikietgumjorn, "Segmentation methods for multiple body parts," *Project in lieu of Thesis, University of Tennessee, Knoxville*, 2008.

[3] D. Lu, "Edge detection improvement by ant colony optimization," *Pattern Recognition Letters*, vol. 29, pp. 416-425, 2008.

[4] O. Wirjadi, *Survey of 3d image segmentation methods*: ITWM, 2007.

[5] Y. Lu, T. Jiang, and Y. Zang, "Region growing method for the analysis of functional MRI data," *NeuroImage*, vol. 20, pp. 455-465, 2003.

[6] R. Pohle and K. D. Toennies, "Segmentation of medical images using adaptive region growing," 2001, pp. 1337-1346.

[7] Y. T. Chen, "A level set method based on the Bayesian risk for medical image segmentation," *Pattern Recognition*, vol. 43, pp. 3699-3711, 2010.

[8] Y. Li and Y. Shen, "Fuzzy c-means clustering based on spatial neighborhood information for image segmentation," *Systems Engineering and Electronics, Journal of*, vol. 21, pp. 323-328, 2010.

[9] K. Chuang, H. Tzeng, S. Chen, J. Wu, and T. Chen, "Fuzzy c-means clustering with spatial information for image segmentation," *Computerized Medical Imaging and Graphics*, vol. 30, pp. 9-15, 2006.

[10] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," 1973.

[11] M. Yang and H. Tsai, "A Gaussian kernel-based fuzzy c-means algorithm with a spatial bias correction," *Pattern Recognition Letters*, vol. 29, pp. 1713-1725, 2008.

[12] K. Xiao, S. H. Ho, and A. Bargiela, "Automatic brain MRI segmentation scheme based on feature weighting factors selection on fuzzy c-means clustering algorithms with Gaussian smoothing," *International Journal of Computational Intelligence in Bioinformatics and Systems Biology*, vol. 1, pp. 316-331, 2010.

[13] W. Wang and Y. Zhang, "On fuzzy cluster validity indices," *Fuzzy Sets and Systems*, vol. 158, pp. 2095-2117, 2007.