Computerized Medical Imaging and Graphics 30 (2006) 9–15

Computerized Medical Imaging and Graphics

www.elsevier.com/locate/compmedimag

Fuzzy c-means clustering with spatial information for image segmentation

Keh-Shih Chuang [a,*], Hong-Long Tzeng [a,b], Sharon Chen [a], Jay Wu [a,b], Tzong-Jer Chen [c]

[a] Department of Nuclear Science, National Tsing-Hua University, Hsinchu 30013 Taiwan
[b] Health Physics Division, Institute of Nuclear Energy Research, Atomic Energy Council, Taiwan
[c] Department of Medical Imaging Technology, Shu-Zen College of Medicine and Management, Kaohsiung, Taiwan

Received 20 August 2004; revised 26 August 2005; accepted 6 September 2005

Abstract

A conventional FCM algorithm does not fully utilize the spatial information in the image. In this paper, we present a fuzzy c-means (FCM) algorithm that incorporates spatial information into the membership function for clustering. The spatial function is the summation of the membership function in the neighborhood of each pixel under consideration. The advantages of the new method are the following: (1) it yields regions more homogeneous than those of other methods, (2) it reduces the spurious blobs, (3) it removes noisy spots, and (4) it is less sensitive to noise than other techniques. This technique is a powerful method for noisy image segmentation and works for both single and multiple-feature data with spatial information.

Keywords: Fuzzy c-means; Spatial information; Image segmentation; Clustering

1. Introduction

Many neurological conditions alter the shape, volume, and distribution of brain tissue; magnetic resonance imaging (MRI) is the preferred imaging modality for examining these conditions. Reliable measurement of these alterations can be performed by using image segmentation. Several investigators have developed methods to automate such measurements by segmentation [1–4]. However, some of these methods do not exploit the multispectral information of the MRI signal.

Fuzzy c-means (FCM) clustering [1,5,6] is an unsupervised technique that has been successfully applied to feature analysis, clustering, and classifier designs in fields such as astronomy, geology, medical imaging, target recognition, and image segmentation. An image can be represented in various feature spaces, and the FCM algorithm classifies the image by grouping similar data points in the feature space into clusters. This clustering is achieved by iteratively minimizing a cost function that is dependent on the distance of the pixels to the cluster centers in the feature domain.

The pixels on an image are highly correlated, i.e. the pixels in the immediate neighborhood possess nearly the same feature data. Therefore, the spatial relationship of neighboring pixels is an important characteristic that can be of great aid in imaging segmentation. General boundary-detection techniques have taken advantage of this spatial information for image segmentation. However, the conventional FCM algorithm does not fully utilize this spatial information. Pedrycz and Waletzky [7] took advantage of the available classified information and actively applied it as part of their optimization procedures. Ahmed et al. [8] modified the objective function of the standard FCM algorithm to allow the labels in the immediate neighborhood of a pixel to influence its labeling. The modified FCM algorithm improved the results of conventional FCM methods on noisy images. However, the way in which they incorporate the neighboring information limits their application to single-feature inputs.

The aim of this study is to introduce a new segmentation method for FCM clustering. In a standard FCM technique, a noisy pixel is wrongly classified because of its abnormal feature data. Our new method incorporates spatial information, and the membership weighting of each cluster is altered after the cluster distribution in the neighborhood is considered. This scheme greatly reduces the effect of noise and biases the algorithm toward homogeneous clustering.

* Corresponding author. Tel.: +886 3 574 2681; fax: +886 3 571 8649.
E-mail address: kschuang@mx.nthu.edu.tw (K.-S. Chuang).

0895-6111/$ - see front matter © 2005 Published by Elsevier Ltd.
doi:10.1016/j.compmedimag.2005.10.001

2. Method

2.1. FCM clustering

The FCM algorithm assigns pixels to each category by using fuzzy memberships. Let $X=(x_1, x_2,\ldots,x_N)$ denotes an image with N pixels to be partitioned into c clusters, where x_i represents multispectral (features) data. The algorithm is an iterative optimization that minimizes the cost function defined as follows:

$$J = \sum_{j=1}^{N} \sum_{i=1}^{c} u_{ij}^{m} \|x_j - v_i\|^2, \tag{1}$$

where u_{ij} represents the membership of pixel x_j in the ith cluster, v_i is the ith cluster center, $\|\cdot\|$ is a norm metric, and m is a constant. The parameter m controls the fuzziness of the resulting partition, and $m=2$ is used in this study.

The cost function is minimized when pixels close to the centroid of their clusters are assigned high membership values, and low membership values are assigned to pixels with data far from the centroid. The membership function represents the probability that a pixel belongs to a specific cluster. In the FCM algorithm, the probability is dependent solely on the distance between the pixel and each individual cluster center in the feature domain. The membership functions and cluster centers are updated by the following:

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

and

$$v_i = \frac{\sum_{j=1}^{N} u_{ij}^{m} x_j}{\sum_{j=1}^{N} u_{ij}^{m}}. \tag{3}$$

Starting with an initial guess for each cluster center, the FCM converges to a solution for v_i representing the local minimum or a saddle point of the cost function. Convergence can be detected by comparing the changes in the membership function or the cluster center at two successive iteration steps.

2.2. Spatial FCM

One of the important characteristics of an image is that neighboring pixels are highly correlated. In other words, these neighboring pixels possess similar feature values, and

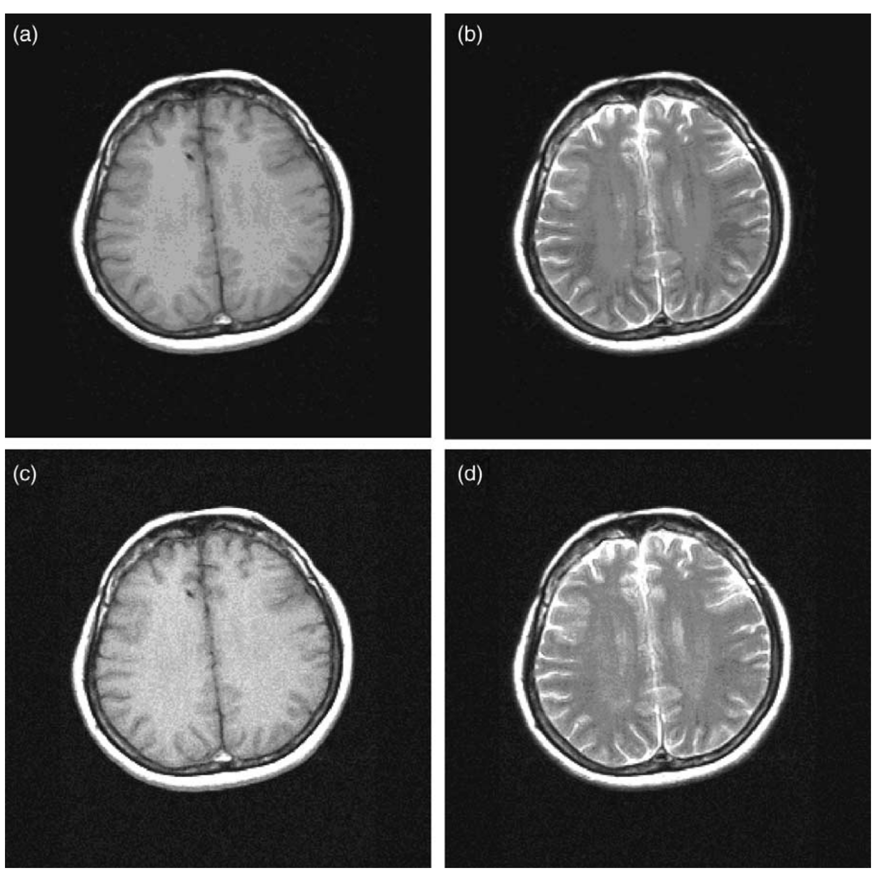

Fig. 1. (a) T1 and (b) T2 images used for the study. (c) T1 and (d) T2 images added with uniform random noise.

the probability that they belong to the same cluster is great. This spatial relationship is important in clustering, but it is not utilized in a standard FCM algorithm. To exploit the spatial information, a spatial function is defined as

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

where $\mathrm{NB}(x_j)$ represents a square window centered on pixel x_j in the spatial domain. A 5×5 window was used throughout this work. Just like the membership function, the spatial function h_{ij} represents the probability that pixel x_j belongs to ith cluster. The spatial function of a pixel for a cluster is large if the majority of its neighborhood belongs to the same clusters. The spatial function is incorporated into membership function as follows:

$$u'_{ij} = \frac{u_{ij}^p h_{ij}^q}{\sum_{k=1}^{c} u_{kj}^p h_{kj}^q}, \quad (5)$$

where p and q are parameters to control the relative importance of both functions. In a homogenous region, the spatial functions simply fortify the original membership, and the clustering result remains unchanged. However, for a noisy pixel, this formula reduces the weighting of a noisy cluster by the labels of its neighboring pixels. As a result, misclassified pixels from noisy regions or spurious blobs can easily be corrected. The spatial FCM with parameter p and q is denoted $\mathrm{sFCM}_{p,q}$. Note that $\mathrm{sFCM}_{1,0}$ is identical to the conventional FCM.

The clustering is a two-pass process at each iteration. The first pass is the same as that in standard FCM to calculate the membership function in the spectral domain. In the second pass, the membership information of each pixel is mapped to the spatial domain, and the spatial function is computed from that. The FCM iteration proceeds with the new membership that is incorporated with the spatial function. The iteration is stopped when the maximum difference between two cluster centers at two successive iterations is less than a threshold ($=0.02$). After the convergence, defuzzification is applied to assign each pixel to a specific cluster for which the membership is maximal.

2.3. Cluster validity functions

Two types of cluster validity functions, fuzzy partition and feature structure, are often used to evaluate the performance of clustering in different clustering methods. The representative functions for the fuzzy partition are partition coefficient V_{pc} [9] and partition entropy V_{pe} [10].

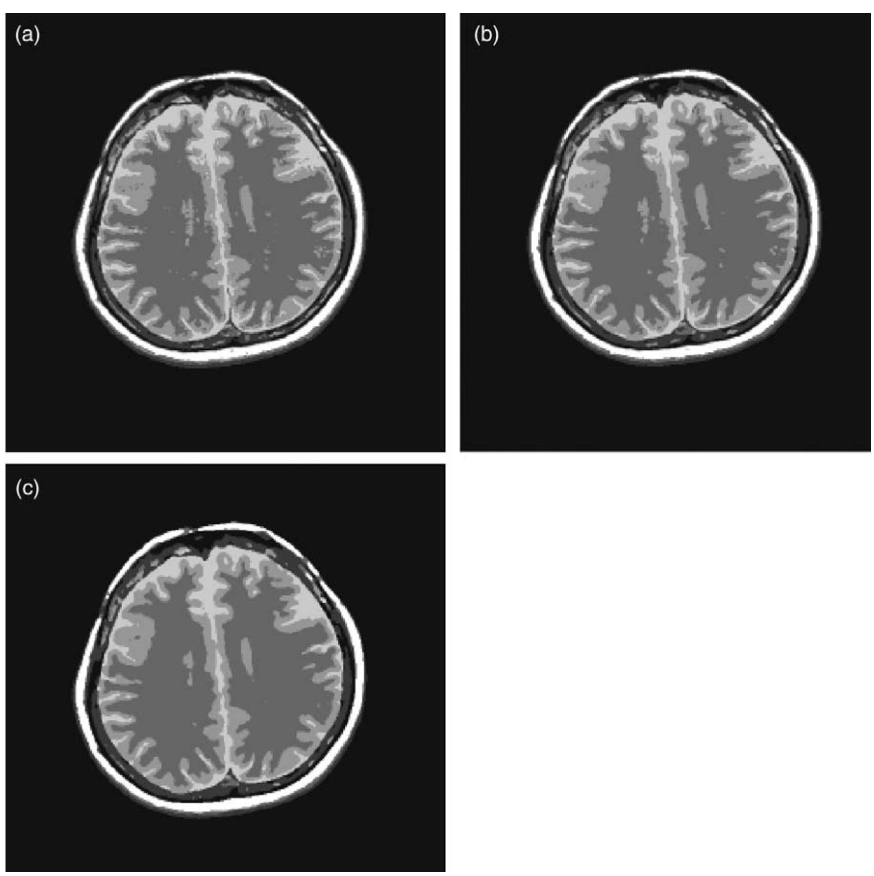

Fig. 2. Segmented images of an MRI image using (a) FCM; (b) $\mathrm{sFCM}_{1,1}$; and (c) $\mathrm{sFCM}_{0,2}$.

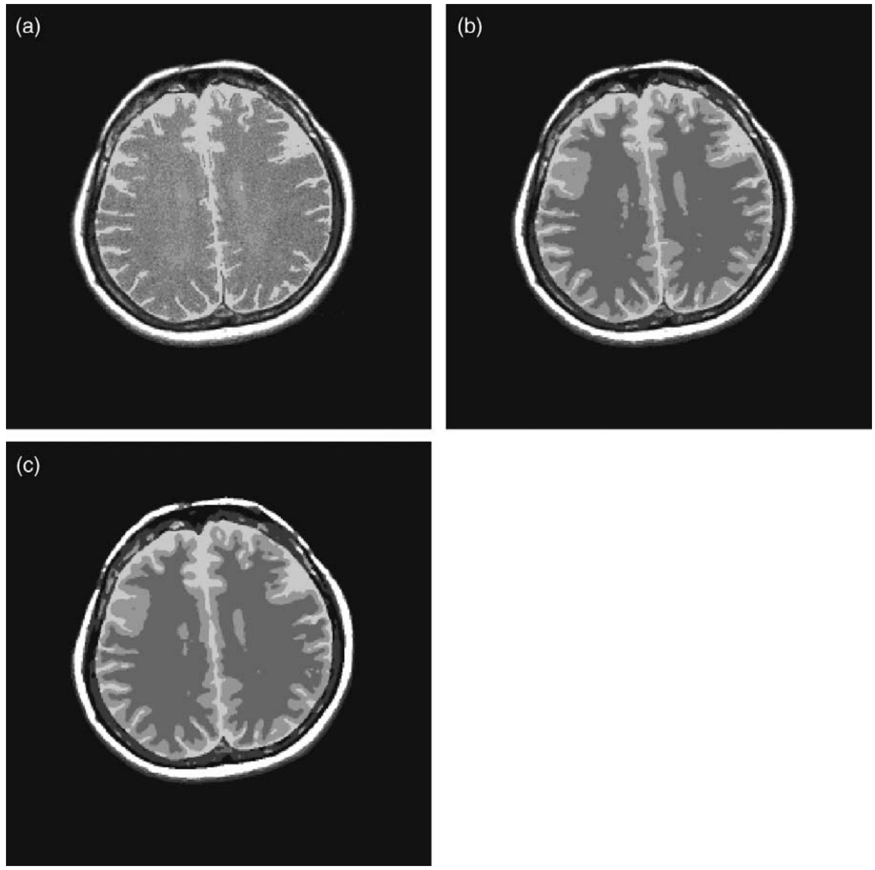

Fig. 3. Segmented images of a noisy MRI image using (a) standard FCM; (b) $\mathrm{sFCM}_{1,1}$; (c) $\mathrm{sFCM}_{0,2}$.

They are defined as follows:

$$V_{\mathrm{pc}} = \frac{\sum_{j}^{N}\sum_{i}^{c} u_{ij}^{2}}{N} \tag{6}$$

and

$$V_{\mathrm{pe}} = \frac{-\sum_{j}^{N}\sum_{i}^{c}[u_{ij}\log u_{ij}]}{N} \tag{7}$$

The idea of these validity functions is that the partition with less fuzziness means better performance. As a result, the best clustering is achieved when the value V_{pc} is maximal or V_{pe} is minimal.

Disadvantages of V_{pc} and V_{pe} are that they measure only the fuzzy partition and lack a direct connection to the featuring property. Other validity functions based on the feature structure are available [11,12]. For example, Xie and Beni [11] defined the validity function as

$$V_{\mathrm{xb}} = \frac{-\sum_{j}^{N}\sum_{i}^{c} u_{ij}\|x_j - v_i\|^2}{N * \left(\min_{i\neq k}\left\{\|v_k - v_i\|^2\right\}\right)}. \tag{8}$$

A good clustering result generates samples that are compacted within one cluster and samples that are separated between different clusters. Minimizing V_{xb} is expected to lead to a good clustering.

2.4. Image data

In this study, we used one T1-weighted image and one T2-weighted MRI image from the same patient. To demonstrate the effect of noise on the segmentation, both images were added with a uniform random noise of magnitude between (−50, 50). The images were divided into six clusters: gray matter (GM), white matter (WM), cerebrospinal fluid, fat, bone, and air. In addition, two synthetic images with four gray levels are generated to serve as ‘ground truth’ images for segmentation evaluation. The images were then corrupted by using a uniform noise between $(-n, n)$, where $n = 3.5 \times (\text{pixel value})^{1/2}$.

3. Results and discussion

Fig. 1(a) and (b) show the T1- and T2-weighted images used for the study, respectively. Fig. 2(a) shows the segmentation results obtained by using a standard FCM algorithm and Fig. 2(b) and (c) show the results of the FCM incorporated into the spatial information with parameters $(p=1,\ q=1)$ and $(p=0,\ q=2)$, respectively.

The conventional FCM successfully classifies the MRI images into six clusters. However, spurious blobs of GM appear inside the WM cluster. The spatial function modifies the membership function of a pixel according to the membership statistics of its neighborhood. Such neighboring effect biases the solution toward piecewise-homogeneous labeling. Both sFCM techniques reduce the number of spurious blobs, and the segmented images are more homogeneous. The sFCM algorithm with a higher q parameter shows a better smoothing effect. The possible disadvantages of using higher spatial weighting are the blurring of some of the finer details. However, this is difficult to judge from the results.

Fig. 1(c) and 1(d) show the T1 and T2 images added with a uniform noise. Fig. 3 shows the segmented results of a noisy MRI image by using the three FCM techniques. As can be seen, the standard FCM technique misclassifies GM and WM at numerous places because the added noise changes the location of GM and WM pixels in the feature space and causes the misclassification of these noisy pixels. Because no similar cluster is present in the neighborhood, the weight of the noisy cluster is greatly reduced with sFCM. Furthermore, the membership of the correct cluster is enhanced by the cluster distribution in the neighboring pixels. As a result, both sFCM techniques effectively correct the misclassification caused by the noise.

Fig. 4 shows the segmentation results of MRI images in the feature domain for all FCM techniques. The cluster centered on the noisy image was different from that of the original noiseless image with standard FCM clustering. The sFCM techniques successfully corrected the misclassified pixels and kept the cluster centers unaffected by noise.

Fig. 5(a) and (b) depict the corrupted four-level simulated T1 and T2 images, and Fig. 5(c)–(e) show the clustering results of the FCM techniques. As can be seen, the clustering results of our sFCM algorithms were superior to those obtained by using conventional FCM. The classification error of the FCM was mostly the salt-and-pepper type. These spots were corrected with the sFCM by using the spatial contextual information.

Table 1 tabulates the validity functions used to evaluate the performance of FCM clustering for six images. In most cases, the validity functions based on the fuzzy partition were better for the sFCM than the conventional FCM. The $\mathrm{sFCM}_{1,1}$ technique showed the best clustering results. The improvement was even better for noisy images. For

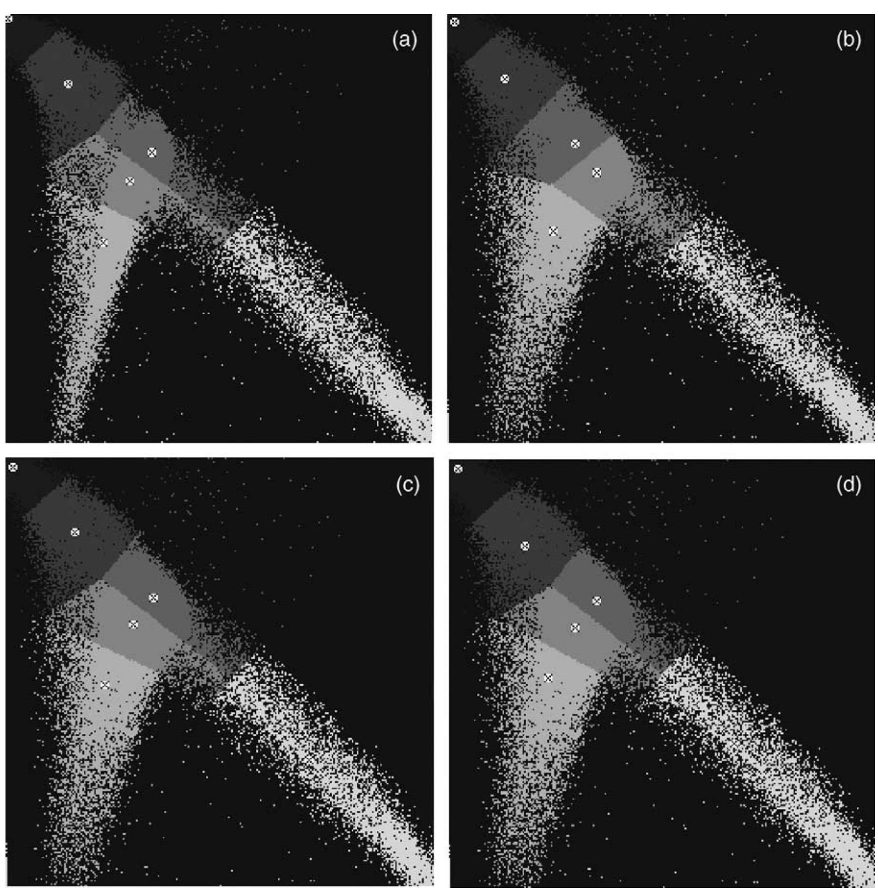

Fig. 4. The calculated centroids (⊗) of cluster in the feature domain (where T1 and T2 values are coordinates) for (a) FCM on the original image; (b) FCM on the noisy image; (c) $\mathrm{sFCM}_{1,1}$ on the noisy image; and (d) $\mathrm{sFCM}_{0,2}$ on the noisy image.

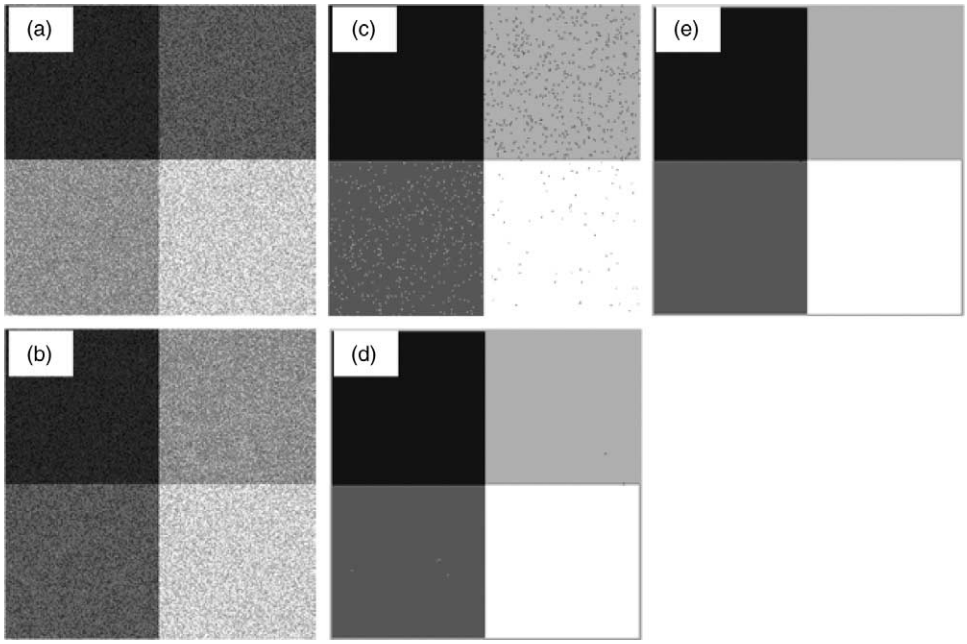

Fig. 5. Simulated (a) T1 and (b) T2 image corrupted by additive uniform noise. The gray levels are 50 (UL), 100 (UR), 150 (LL), and 200 (LR) in the T1 image and 50 (UL), 150 (UR), 100 (LL), and 200 (LR) in the T2 image. Clustering resulted using (c) FCM; (d) $\mathrm{sFCM}_{1,1}$; and (e) $\mathrm{sFCM}_{0,2}$.

example, the reduction in V_{pe} with the $\mathrm{sFCM}_{1,1}$, as compared with the conventional technique, was 50% for original images and 64.6% for noisy data.

Paired-sample t tests were performed on these data sets to test the significance of any difference in the measures. Table 2 lists statistical results. For V_{pe} (V_{pc}), the $\mathrm{sFCM}_{1,1}$ is significantly greater (smaller) than conventional FCM and $\mathrm{sFCM}_{0,2}$. However, the validity function based on feature structure showed different results: V_{xb} increased for the proposed FCM, which was understandable. V_{xb} measured the compactness in the feature domain. Conventional FCM achieves partition by minimizing the metric difference in the feature domain, and its V_{xb} is minimized. The sFCM modifies the partition on the basis of spatial distribution and causes deterioration of the compactness in the feature domain and subsequently the increase in V_{xb}.

Other variations of spatial function are available. For example, it can be expressed as

$$h_{ij} = \sum_{k \in NB(x_j)} g_{ik} \tag{9}$$

where

$$g_{ik} = \begin{cases} 1, & \text{if } u_{ik} \geq u_{lk} \text{ for } l = 1, c \\ 0, & \text{otherwise} \end{cases}.$$

The calculation of this membership has less fuzziness than others and causes a loss in the details.

This technique is applied only to image data with spatial information. The spatial function can be

Table 1
The clustering results of six images using various FCM techniques

Images	Techniques	V_{pc}	V_{pe}	$V_{\mathrm{xb}}(\times 10^{-3})$
Original MRI images	FCM	0.888	0.234	0.918
	$\mathrm{sFCM}_{1,1}$	0.924	0.117	2.46
	$\mathrm{sFCM}_{0,2}$	0.899	0.160	3.38
Noise added MRI images	FCM	0.779	0.414	1.22
	$\mathrm{sFCM}_{1,1}$	0.909	0.147	3.14
	$\mathrm{sFCM}_{0,2}$	0.884	0.191	4.14
Synthesized image uniform noise (−35,35)	FCM	0.718	0.571	0.620
	$\mathrm{sFCM}_{1,1}$	0.915	0.133	0.748
	$\mathrm{sFCM}_{0,2}$	0.911	0.146	0.789
Synthesized image uniform noise ($-n$, n)	FCM	0.740	0.529	0.563
	$\mathrm{sFCM}_{1,1}$	0.920	0.122	0.688
	$\mathrm{sFCM}_{0,2}$	0.916	0.133	0.728
Synthesized image SNR = 5	FCM	0.685	0.615	0.793
	$\mathrm{sFCM}_{1,1}$	0.840	0.271	1.032
	$\mathrm{sFCM}_{0,2}$	0.832	0.301	1.182
Synthesized image SNR = 10	FCM	0.857	0.317	0.292
	$\mathrm{sFCM}_{1,1}$	0.959	0.032	0.333
	$\mathrm{sFCM}_{0,2}$	0.952	0.044	0.355

The MRI images and the simulated four-level T1 and T2 images are the same as what is shown in Figs. 1 and 5.

Table 2
Statistical results of the paired-sample t-test on the differences in the three performance index

	$(X{:}Y)$	$\bar{W}$	S_W	T	p-value
V_{pc}	($sFCM_{1,1}$: FCM)	0.133	0.0586	5.57	$0.001 < p < 0.0025$
	($sFCM_{1,1}$: $sFCM_{0,2}$)	0.0122	0.01	2.99	$0.025 < p < 0.01$
V_{pe}	($sFCM_{1,1}$: FCM)	−0.309	0.115	−6.58	$p < 0.001$
	($sFCM_{1,1}$: $sFCM_{0,2}$)	−0.0255	0.0156	−4.00	$p \cong 0.005$
V_{xb}	($sFCM_{1,1}$: FCM)	0.666	0.836	1.95	$p > 0.05$
	($sFCM_{1,1}$: $sFCM_{0,2}$)	−0.362	0.466	−1.90	$p > 0.05$

$\left(H_0{:}\ \mu_x = \mu_y,\ H_1{:}\ \mu_x > \mu_y,\ W_i = X_i - Y_i,\ T = \bar{W}/(S_W/\sqrt{n})\right)$ (Note that in testing V_{pe} and V_{xb}, the alternative hypothesis H_1: $\mu_x < \mu_y$ is used).

incorporated with other FCM techniques without a great deal of modifications. For example, the feature-weight learning FCM technique [13] assigns various weights to different features to improve the performance of clustering. The spatial function can be estimated at each iteration and incorporated into the membership function. Owing to the incorporation of spatial information, the new FCM technique is less sensitive than other methods to noise.

4. Summary

FCM clustering is an unsupervised clustering technique applied to segment images into clusters with similar spectral properties. It utilizes the distance between pixels and cluster centers in the spectral domain to compute the membership function. The pixels on an image are highly correlated, and this spatial information is an important characteristic that can be used to aid their labeling. However, the spatial relationship between pixels is seldom utilized in FCM.

In this paper, we proposed a spatial FCM that incorporates the spatial information into the membership function to improve the segmentation results. The membership functions of the neighbors centered on a pixel in the spatial domain are enumerated to obtain the cluster distribution statistics. These statistics are transformed into a weighting function and incorporated into the membership function. This neighboring effect reduces the number of spurious blobs and biases the solution toward piecewise homogeneous labeling. The new method was tested on MRI images and evaluated by using various cluster validity functions. Preliminary results showed that the effect of noise in segmentation was considerably less with the new algorithm than with the conventional FCM.

Acknowledgements

This work is supported in part by a research grant of NSC92-2218-E007-012 from the National Science Council, Taiwan.

References

[1] Bezdek J, Hall L, Clarke L. Review of MR image segmentation using pattern recognition. Med Phys 1993;20:1033–48.
[2] Brandt ME, Bohan TP, Kramer LA, Fletcher JM. Estimation of CSF, white matter and gray matter volumes in hydrocephalic children using fuzzy clustering of MR images. Comput Med Imaging Graph 1994; 18:25–34.
[3] Clark MC, Hall LO, Goldgof DB, Clarke LP, Velthuizen RP, Silbiger MS. MRI segmentation using fuzzy clustering techniques. IEEE Eng Med Biol 1994;13:730–42.
[4] Pham DL, Prince JL. Adaptive fuzzy segmentation of magnetic resonance images. IEEE Trans Med Imaging 1999;18:737–52.
[5] Lyer NS, Kandel A, Schneider M. Feature-based fuzzy classification for interpretation of mammograms. Fuzzy Sets Syst 2002;114: 271–80.
[6] Yang MS, Hu YJ, Lin KCR, Lin CCL. Segmentation techniques for tissue differentiation in MRI of Ophthalmology using fuzzy clustering algorithms. Magn Reson Imaging 2002;20:173–9.
[7] Pedrycz W, Waletzky J. Fuzzy clustering with partial supervision. IEEE Trans Syst Man Cybern Part B Cybern 1997;27:787–95.
[8] Ahmed MN, Yamany SM, Mohamed N, Farag AA, Moriarty T. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. IEEE Trans Med Imaging 2002;21:193–9.
[9] Bezdek JC. Cluster validity with fuzzy sets. J Cybern 1974;3:58–73.
[10] Bezdek JC. Mathematical models for systematic and taxonomy. In: proceedings of eigth international conference on numerical taxonomy, San Francisco; 1975, p. 143–66.
[11] Xie XL, Beni GA. Validity measure for fuzzy clustering. IEEE Trans Pattern Anal Mach Intell 1991;3:841–6.
[12] Fukuyama Y, Sugeno M. A new method of choosing the number of clusters for the fuzzy c-means method. In: proceedings of fifth fuzzy system symposium; 1989, p. 247–50.
[13] Wang X, Wang Y, Wang L. Improving fuzzy c-means clustering based on feature-weight learning. Pattern Recognit Lett 2004;25: 1123–32.