Engineering Applications of Artificial Intelligence 23 (2010) 160–168

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Parameter optimization of improved fuzzy c-means clustering algorithm for brain MR image segmentation

Mohamad Forouzanfar [a,b,1], Nosratallah Forghani [b,*,1], Mohammad Teshnehlab [c]

[a] School of Information Technology and Engineering (SITE), University of Ottawa, 800 King Edward Avenue, Ottawa, Ontario, Canada K1N 6N5
[b] Department of Biomedical Engineering, Faculty of Electrical Engineering, K.N. Toosi University of Technology, P.O. Box 16315-1355, Tehran, Iran
[c] Department of Control Engineering, Faculty of Electrical Engineering, K.N. Toosi University of Technology, P.O. Box 16315-1355, Tehran, Iran

ARTICLE INFO

Article history:
Received 19 June 2008
Received in revised form
11 October 2009
Accepted 12 October 2009
Available online 13 November 2009

Keywords:
Magnetic resonance imaging (MRI)
Image segmentation
Genetic algorithms (GAs)
Particle swarm optimization (PSO)
Breeding swarm (BS)

ABSTRACT

A traditional approach to segmentation of magnetic resonance (MR) images is the fuzzy c-means (FCM) clustering algorithm. The efficacy of FCM algorithm considerably reduces in the case of noisy data. In order to improve the performance of FCM algorithm, researchers have introduced a neighborhood attraction, which is dependent on the relative location and features of neighboring pixels. However, determination of degree of attraction is a challenging task which can considerably affect the segmentation results.

This paper presents a study investigating the potential of genetic algorithms (GAs) and particle swarm optimization (PSO) to determine the optimum value of degree of attraction. The GAs are best at reaching a near optimal solution but have trouble finding an exact solution, while PSO's-group interactions enhances the search for an optimal solution. Therefore, significant improvements are expected using a hybrid method combining the strengths of PSO with GAs, simultaneously. In this context, a hybrid GAs/PSO (breeding swarms) method is employed for determination of optimum degree of attraction. The quantitative and qualitative comparisons performed on simulated and real brain MR images with different noise levels demonstrate unprecedented improvements in segmentation results compared to other FCM-based methods.

1. Introduction

Magnetic resonance imaging (MRI) is a technique that uses a magnetic field and radio waves to create cross-sectional images of organs, soft tissues, bone, and virtually all other internal body structures (Haacke et al., 1999). MRI possesses good contrast resolution for different tissues and has advantages over computerized tomography (CT) for brain studies due to its superior contrast properties. In this context, brain MRI segmentation has become an increasingly important image processing step in many applications, including: (i) automatic or semiautomatic delineation of areas to be treated prior to radiosurgery, (ii) delineation of tumors before and after surgical or radiosurgical intervention for response assessment, and (iii) tissue classification (Bondareff et al., 1990).

Several techniques have been developed for brain MR image segmentation among which thresholding (Suzuki and Toriwaki, 1991), edge detection (Canny, 1986), region growing (Pohle and Toennies, 2001), and clustering (Clarke et al., 1995) are the most well-known ones. Thresholding is the simplest segmentation method, where the classification of each pixel depends on its own information such as intensity and color. Thresholding methods are efficient when the histograms of objects and background are clearly separated. Since the distribution of tissue intensities in brain MR images is often very complex, these methods fail to achieve acceptable segmentation results. Edge-based segmentation methods are based on detection of boundaries in the image. These techniques suffer from incorrect detection of boundaries due to noise, over- and under-segmentation, and variability in threshold selection for the edge image. These drawbacks of early image segmentation methods, has led to region growing algorithms. Region growing is the extension of thresholding by considering the homogeneity and connectivity criteria. However, only well-defined regions can be robustly identified by region growing algorithms (Clarke et al., 1995). Since the above mentioned methods are generally limited to relatively simple structures, clustering methods are utilized for complex pathology. Clustering is a method of grouping data with similar characteristics into larger units of analysis. Expectation–maximization (EM) (Wells et al., 1996), hard c-means and its fuzzy equivalent, fuzzy c-means (FCM) algorithms (Li et al., 1993) are the typical

* Corresponding author. Tel.: +98 21 88469084; fax: +98 21 88462066.
E-mail addresses: mforo040@site.uottawa.ca (M. Forouzanfar), n_forghani@ee.kntu.ac.ir (N. Forghani), teshnehlab@eetd.kntu.ac.ir (M. Teshnehlab).
[1] The first two authors contributed equally to this work.

0952-1976/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engappai.2009.10.002

methods of clustering. A main drawback of the EM algorithm is that it is based on a Gaussian distribution model for the intensity distribution of brain images, which is not true, especially for noisy images. Since Zadeh (1965) first introduced fuzzy set theory which gave rise to the concept of partial membership, fuzziness has received increasing attention. Fuzzy clustering algorithms have been widely studied and applied in various areas. Among fuzzy clustering techniques, FCM is the best known and most powerful method used in image segmentation. FCM was first conceived in 1973 by Dunn (1973) and further generalized by Bezdek (1981). It is based on minimization of an objective function and is frequently used in pattern recognition. Unfortunately, FCM does not consider the spatial information in the image space and is highly sensitive to noise and imaging artifacts. Since medical images contain significant amount of noise caused by operator, equipment, and the environment, there is an essential need for development of less noise-sensitive algorithms.

Many modifications of the FCM algorithm have been proposed to alleviate the effects of noise, such as noisy clustering (NC) (Dave, 1991), possibilistic c-means (PCM) (Krishnapuram and Keller, 1993), robust fuzzy c-means (RFCM) algorithm (Pham, 2001), and so on. These methods generally modify most equations along with modification of the objective function. Therefore, they lose the continuity from FCM, which inevitably introduce computation issues.

Yu and Yang (2005) proposed a generalized FCM (GFCM) model to unify some variations of FCM and then studied its optimality test with parameter selection. However, the variations of the FCM in this method may not have two kinds of optimality test, i.e., one based on the cluster prototypes and another one based on membership functions. It was shown in Yu and Yang (2005) that the GFCM has only the optimality test with the cluster prototype. In Yu and Yang (2007), an alternative model of GFCM, called a generalized fuzzy clustering regularization (GFCR) was proposed that can have the optimality test with membership functions. Recently, Shen et al. (2005) introduced a new extension of FCM algorithm, called improved FCM (IFCM). They introduced two influential factors in segmentation that address the neighborhood attraction. The first parameter is the feature difference between neighboring pixels in the image and the second one is the relative location of the neighboring pixels. Therefore, segmentation is decided not only by the pixel's intensity but also by neighboring pixel's intensities and their locations. However, the problem of determining optimum parameters constitutes an important part of implementing the IFCM algorithm for real applications. The implementation performance of IFCM may be significantly degraded if the attraction parameters are not properly selected. It is therefore important to select suitable parameters such that the IFCM algorithm achieves superior partition performance compared to the FCM. In Shen et al. (2005), an artificial neural network (ANN) was employed for computation of these two parameters. However, designing the neural network architecture and setting its parameters are always complicated which slow down the algorithm and may also lead to inappropriate attraction parameters and consequently degrade the partitioning performance.

In this paper, we extend the IFCM algorithm to overcome the mentioned drawbacks in segmentation of the intensity MR images. Same as in Shen et al., (2005), a neighborhood attraction is considered to exist between neighboring pixels of the intensity image. The degree of attraction depends on pixel intensities and the spatial position of the neighbors. Two parameters $\lambda(0<\lambda<1)$ and $\xi(0<\xi<1)$ will adjust the degree of the neighborhood attractions. We will then investigate the potential of genetic algorithms (GAs) and particle swarm optimization (PSO) to determine the optimum values of the neighborhood attraction parameters. We will show that both GAs and PSO are superior to the ANN algorithm especially in segmentation of noisy MR images. However, unprecedented improvements are achieved using a hybrid method combining the strengths of PSO with GAs, simultaneously. The achieved improvements of the hybrid GAs/PSO, breeding swarm (BS), method is validated both quantitatively and qualitatively on simulated and real brain MR images at different noise levels.

This paper is organized as follows. In Section 2, the traditional FCM algorithm and its improved extension called IFCM are introduced. Section 3 presents three new parameter optimization methods based on GAs, PSO, and BS. Section 3 compares our proposed algorithms with other published techniques. Section 4 contains conclusions and addresses future work.

FCM and IFCM clustering algorithms

Let $X=\{x_1, \ldots, x_n\}$ be a data set and let c be a positive integer greater than one. A partition of X into c clusters is represented by mutually disjoint sets $X_1, \ldots, X_c$ such that $X_1 \cup \cdots \cup X_c = X$ or equivalently by indicator function $\mu_1, \ldots, \mu_c$ such that $\mu_i(x)=1$ if x is in X_i and $\mu_i(x)=0$ if x is not in X_i for all $i=1, \ldots, c$. This is known as clustering X into c clusters $X_1, \ldots, X_c$ using $\{\mu_1, \ldots, \mu_c\}$. A fuzzy extension allows $\mu_i(x)$ taking values in the interval $[0, 1]$ such that $\sum_{i=1}^{c} \mu_i(x) = 1$ for all x in X. In this case, $\{\mu_1, \ldots, \mu_c\}$ is called a fuzzy c-partition of X. Thus, the FCM objective function J_{FCM} is defined as (Bezdek, 1981)

$$J_{FCM}(\mu,v) = \sum_{i=1}^{c}\sum_{j=1}^{n} \mu_{ij}^{m} d^2(x_j, v_i), \tag{1}$$

where $\mu=\{\mu_1, \ldots, \mu_c\}$ is a fuzzy c-partition with $\mu_{ij}=\mu_i(x_j)$, the weighted exponent m is a fixed number greater than one establishing the degree of fuzziness, $v=\{v_1, \ldots, v_c\}$ is the c cluster centers, and $d^2(x_j,v_i)=||x_j-v_i||^2$ represents the Euclidean distance or its generalization such as the Mahalanobis distance. The FCM algorithm is an iteration through the necessary conditions for minimizing J_{FCM} with the following update equations:

$$v_i = \frac{\sum_{j=1}^{n} \mu_{ij}^{m} x_j}{\sum_{j=1}^{n} \mu_{ij}^{m}} \quad (i=1,\ldots,c) \tag{2}$$

and

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d(x_j,v_i)}{d(x_j,v_k)}\right)^{2(m-1)}}. \tag{3}$$

At each iteration, QUOTE μ and v are updated using (2) and (3). The FCM algorithm iteratively optimizes $J_{FCM}(\mu, v)$ until $|\mu(l+1)-\mu^l| \le \varepsilon$ is the number of iterations.

From (1), it is clear that the objective function of FCM does not take into account any spatial dependence among X and consider each image pixel as an individual point. Also, the membership function in (3) is determined by $d^2(x_j, v_i)$, which measures the similarity between the pixel intensity and the cluster center. The closer the intensity values to the cluster center the higher the value of the membership. Therefore, the membership function is highly sensitive to noise. If an MR image is affected by noise or other artifacts, the intensity of the pixels would change which results in an incorrect membership and improper segmentation.

There are several approaches to reduce sensitivity of FCM algorithm to noise. The most direct technique is low pass filtering of the image and then applying the FCM algorithm. However, low pass filtering may lead to loss of some important details. Different extensions of FCM algorithm have been proposed by researchers in order to solve sensitivity to noise. Dave clustered the noise into a separate cluster which is unique from signal clusters

(Dave, 1991). This method is not proper for image segmentation since noisy pixels are separated from other pixels while they should be assigned to the most appropriate cluster. Pham (2001) modified the FCM objective function by including a spatial penalty on the membership functions. The penalty term leads to an iterative algorithm called as RFCM that allows the estimation of spatially smooth membership functions. However, this method is only slightly different from the original FCM and the new objective function results in the complex variation of the membership function. Krishnapuram and Keller (1993) considered clustering as a possibilistic partition and called their new approach PCM. The drawback of PCM is that it limits the clustering to only one or two classes (Shen et al., 2005).

To overcome these drawbacks, Shen et al., 2005 presented an improved algorithm. They found that the similarity function $d^2(x_j, v_i)$ is the key to segmentation success. In their approach, a kind of relationship named neighborhood attraction is considered to exist between neighboring pixels. During clustering, each pixel attempts to attract its neighboring pixels toward its own cluster. This neighborhood attraction depends on two factors; the pixel intensities or feature attraction $\lambda(0<\lambda<1)$, and the spatial position of the neighbors or distance attraction $\xi(0<\xi<1)$, which also depends on the neighborhood structure. Considering this neighborhood attraction, they defined the similarity function as below

$$d^2(x_j,v_j) = ||x_j - v_i||^2(1 - \lambda H_{ij} - \xi F_{ij}) \tag{4}$$

where H_{ij} represents the feature attraction and F_{ij} represents the distance attraction. The parameters λ and ζ adjust the degree of the two neighborhood attractions. H_{ij} and F_{ij} computed in a neighborhood containing S pixels as follows:

$$H_{ij} = \frac{\sum_{k=1}^{S} \mu_{ik} g_{jk}}{\sum_{k=1}^{S} g_{jk}} \tag{5}$$

$$F_{ij} = \frac{\sum_{k=1}^{S} \mu_{ik}{}^2 q_{jk}{}^2}{\sum_{k=1}^{S} g_{jk}{}^2} \tag{6}$$

with

$$g_{jk} = |x_j - x_k|, \quad q_{jk} = (a_j - a_k)^2 + (b_j - b_k)^2.$$

where (a_j, b_j), and (a_k, b_k) denote the coordinate of pixel j and k, respectively. It should be noted that a higher value of λ results in a stronger feature attraction and a higher value of ξ results in a stronger distance attraction. Optimized values of these parameters lead to the best segmentation results while inappropriate values degrade the results. Therefore, parameter optimization is an important issue in IFCM algorithm that can significantly affect the segmentation results.

2. Parameter optimization of IFCM algorithm

As mentioned earlier, the problem of determining optimum attraction parameters constitutes an important part of implementing the IFCM algorithm. Shen et al. computed these two parameters using an ANN through an optimization problem (Shen et al., 2005). However, designing the ANN architecture and setting its parameters are always complicated tasks that slow down the algorithm and may lead to inappropriate attraction parameters. This consequently degrades the partitioning performance. In this Section, we introduce three new algorithms, namely GAs, PSO, and BS for optimum determination of the attraction parameters. The performance evaluation of the proposed algorithms is carried out in the next Section.

2.1. Structure of genetic algorithms (GAs)

Like neural networks, GAs are based on a biological metaphor, however, instead of the biological brain, GAs view learning in terms of competition among a population of evolving candidate problem solutions. GAs were first introduced by Holland in the early 1970s (Holland, 1992) and have been widely successful in optimization problems. The algorithm is started with a set of solutions (represented by chromosomes) called population. Solutions from one population are taken and used to form a new population. This is motivated by a hope, that the new population will be better than the old one. Solutions which are selected to form new solutions (offspring) are selected according to their fitness; the more suitable they are the more chances they have to reproduce. This is repeated until some condition is satisfied. The GAs can be outlined as follows:

1. [Start] Generate random population of P chromosomes (suitable solutions for the problem).
2. [Fitness] Evaluate the fitness of each chromosome in the population with respect to the cost function J.
3. [New population] Create a new population by repeating following steps until the new population is complete:
   3.1. [Selection] Select two parent chromosomes from a population according to their fitness (the better fitness, the bigger chance to be selected).
   3.2. [Crossover] With a crossover probability, cross over the parents to form a new offspring (children). If no crossover was performed, offspring is an exact copy of parents.
   3.3. [Mutation] With a mutation probability, mutate new offspring at each locus (position in chromosome).
   3.4. [Accepting] Place new offspring in a new population.
4. [Loop] Go to step 2 until convergence.

For selection stage a roulette wheel approach is adopted. Construction of roulette wheel is as follows (Mitchel, 1999):

1. Arrange the chromosomes according to their fitness.
2. Compute summations of all fitness values and calculate the total fitness.
3. Divide each fitness value to total fitness and compute the selection probability (p_k) for each chromosome.
4. Calculate cumulative probability (p_k) for each chromosome.

In selection process, roulette wheel spins equal to the number population size. Each time a single chromosome is selected for a new population in the following manner (Gen and Cheng, 1997):

1. Generate a random number r from the rang [0, 1].
2. If $r \le P_1$, then select the first chromosome, otherwise select the kth chromosome such that $q_{k-1} < r < q_k$.

The mentioned algorithm is iterated until a certain criterion is met. At this point, the most fitted chromosome represents the corresponding optimum values. The specific parameters of the introduced structure are described in Section 4.

2.2. Structure of particle swarm optimization (PSO)

The PSO is a member of wide category of swarm intelligence methods (Kennedy and Eberhart, 2001). Kennedy originally proposed PSO as a simulation of social behavior and it was initially introduced as an optimization method (Kennedy and Eberhart, 1995). The PSO algorithm is conceptually simple and can be implemented in a few lines of code. A PSO individual also

retains the knowledge of where in search space it performed the best, while in GAs if an individual is not selected for crossover or mutation, the information contained by that individual is lost. Comparisons between PSO and GAs are done analytically in Eberhart and Shi (1998) and also with regards to performance in Angeline (1998). In PSO, a swarm consists of individuals, called particles, which change their position $\bar{x}_i(t)$ with time t. Each particle represents a potential solution to the problem and flies around in a multidimensional search space. During flight each particle adjusts its position according to its own experience, and according to the experience of neighboring particles, making use of the best position encountered by itself and its neighbors. The effect is that particles move towards the best solution. The performance of each particle is measured according to a pre-defined fitness function, which is related to the problem being solved.

To implement the PSO algorithm, we have to define a neighborhood in the corresponding population and then describe the relations between particles that fall in that neighborhood. In this context, we have many topologies such as: star, ring, and wheel. Here we use the ring topology. In ring topology, each particle is related with two neighbors and attempts to imitate its best neighbor by moving closer to the best solution found within the neighborhood. The local best

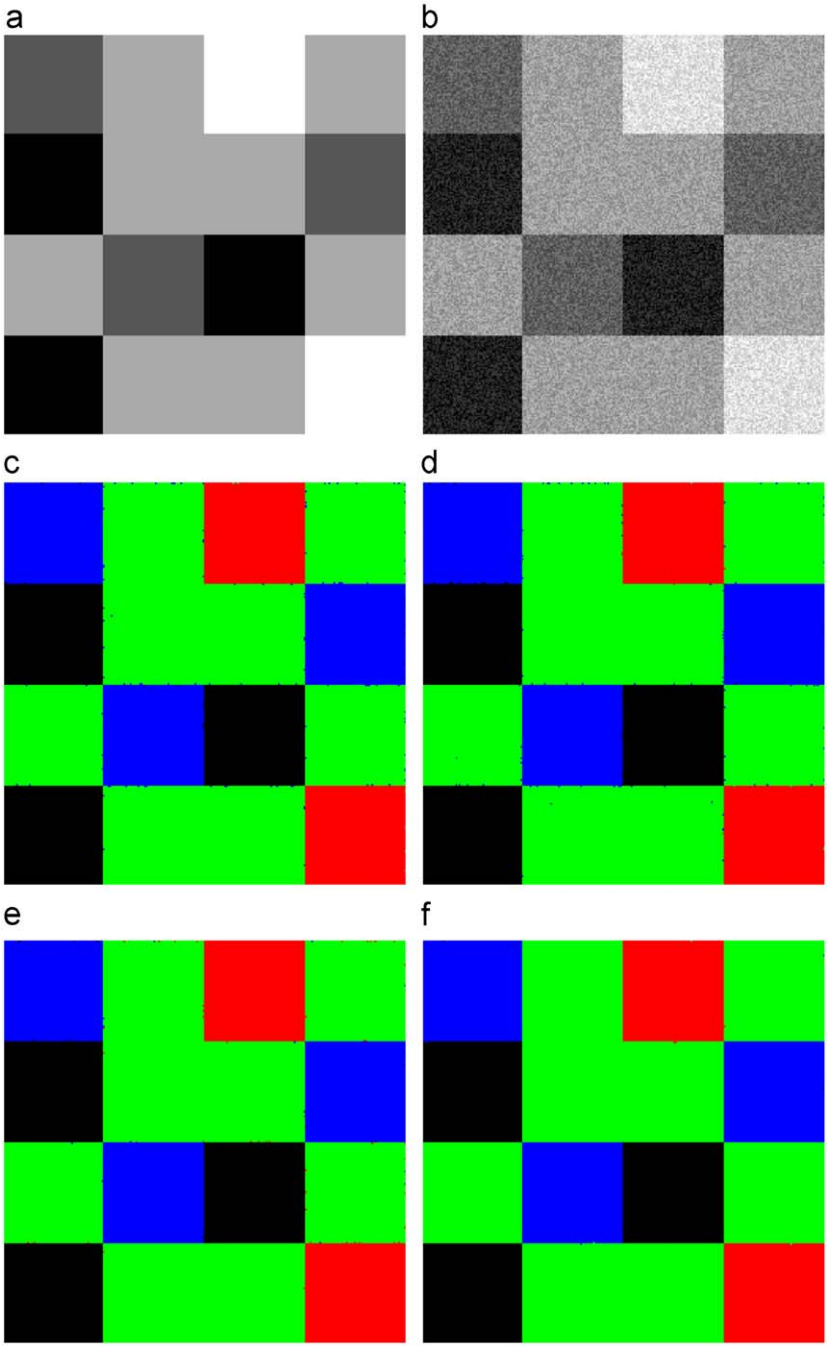

Fig. 1. Segmentation results on a synthetic square image with a uniformly distributed noise in the interval (0, 120). (a) Noise-free reference image. (b) Noisy image. (c) ANN-IFCM. (d) GAs-IFCM. (e) PSO-IFCM. (f) BS-IFCM.

algorithm is associated with this topology (Eberhart et al., 1996; Corne et al., 1999):

1. [Start] Generate a random swarm of P particles in D-dimensional space, where D represents the number of variables (here D=2).
2. [Fitness] Evaluate the fitness $f(\bar{x}_i(t))$ of each particle with respect to the cost function J.
3. [Update] particles or moved toward the best solution by repeating the following steps:
   If $f(\bar{x}_i(t))$ QUOTE then $pbest_i = f(\bar{x}_i(t))$ and $\bar{x}_{pbest_i} = \bar{x}_i(t)$, where $pbest_i$ is the current best fitness achieved by the ith particle and $\bar{x}_{pbest_i}$ is the corresponding coordinate.
   3.1. If $f(\bar{x}_i(t)) < lbest_i$ QUOTE then $lbest_i = f(\bar{x}_i(t))$, where $lbest_i$ is the best fitness over the topological neighbors.
   3.2. Change the velocity v_i of each particle:
   $$\bar{v}_i(t) = \bar{v}_i(t-1) + \rho_1(\bar{x}_{pbest_i} - \bar{x}_i(t) + \rho_2(\bar{x}_{lbesti} - \bar{x}_i(t)) \quad (7)$$
   where ρ_1 and ρ_2 are random accelerate constants between 0 and 1.
   3.3. Fly each particle to its new position $\bar{x}_i(t) + \bar{v}_i(t)$.
4. [Loop] Go to step 2 until convergence.

The above procedures are iterated until a certain criterion is met. At this point, the most fitted particle represents the corresponding optimum values. The specific parameters of the introduced structure are described in Section 4.

2.3. Structure of breeding swarm (BS) optimization

Both GAs and PSO have strengths and weaknesses. The PSO algorithm is theoretically simple and can be implemented in a few lines of code. A PSO individual also keeps the knowledge of where in the search space it performed the best (a memory of the past experience). In GAs, if an individual is not selected for crossover or mutation, the information by that individual is lost. However, without a selection operator PSO may waste resources on poor individuals. A PSO's group interaction enhances the search for an optimal solution, whereas GAs have trouble finding an exact solution (Settles and Soule, 2005).

In this context, our goal is to introduce a hybrid GAs/PSO method, combining the strengths of PSO with GAs, simultaneously. The hybrid algorithm combines the standard velocity and position update rules of PSO with the ideas of selection, crossover and mutation from GAs. The algorithm is designed so that the GAs facilitate a global search and the PSO performs the local search. The structure of BS algorithms can be summarized as follows:

1. [Start] Generate a random population of size P.
2. [Fitness] Evaluate the fitness of each particle with respect to the cost function J.
3. [Selection] Select P best particles using the roulette wheel algorithm (in the first iteration this step is needless).
4. [New population] Perform step 3 of the GAs and PSO in parallel and create a new population gathering the output of both GAs and PSO.
5. [Loop] Go to step 2 until convergence.

Although the BS algorithm seems to be more complicated than GAs and PSO, it is able to locate an optimal, or near optimal, solution significantly faster than either GAs or PSO. This is the result of combining the strengths of PSO with GAs, simultaneously. The GAs facilitate a global search to reach a near optimal solution and the PSO's group interactions enhances the search for the optimal local solution. The specific parameters of the BS structure are described in Section 4.

Table 1
Segmentation evaluation of synthetic square image.

Metrics	Methods						
	FCM	PCM	RFCM	ANN-IFCM	GAs-IFCM	PSO-IFCM	BS-IFCM
UnS(%)	9.56	25.2	6.42	0.0230	0.0210	0.0110	0.0059
OvS(%)	23.79	75.00	16.22	0.0530	0.0468	0.0358	0.0181
InC(%)	14.24	43.75	9.88	0.0260	0.0220	0.0143	0.0084

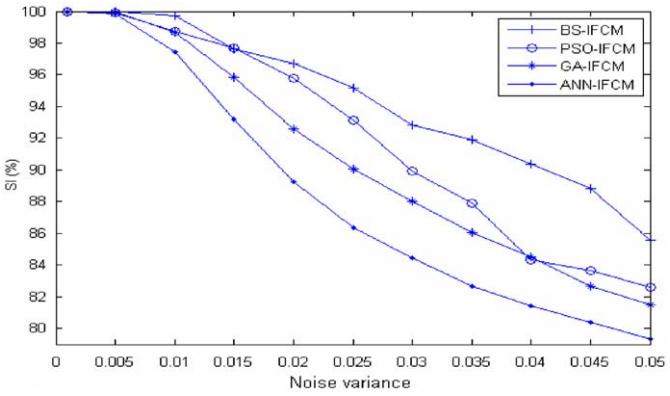

Fig. 2. Performance comparison of IFCM-based methods using the SI metric at different noise levels

3. Experimental results

This section is dedicated to a comprehensive investigation on the proposed methods' performance. To this end, we will compare the proposed algorithms with FCM, PCM (Krishnapuram and Keller, 1993), RFCM (Pham, 2001), and an implementation of IFCM algorithm based on ANN (ANN-IFCM) (Shen et al., 2005).

Our experiments were performed on three types of images: (1) a synthetic square image; (2) simulated brain images obtained from Brainweb[2]; and (3) real MR images acquired from IBSR.[3] In all experiment the size of the population (P) is set to 20 and the cost function J_{FCM} with the similarity index defined in (4) is employed as a measure of fitness. Also, a single point crossover with probability of 0.2 and an order changing mutation with probability of 0.01 are applied. The weighting exponent m in all fuzzy clustering methods was set to 2. It has been observed that this value of weighting exponent yields the best results in most brain MR images (Shen et al., 2005).

3.1. Square image

A synthetic square image consisting of 16 squares of size 64×64 is generated. This square image consists of 4 classes with intensity values of 0, 100, 200, and 300. In order to investigate the sensitivity of the algorithms to noise, a uniformly distributed noise in the interval (0, 120) is added to the image. The reference

[2] www.bic.mni.mcgill.ca/brainweb/
[3] http://www.cma.mgh.harvard.edu/ibsr/

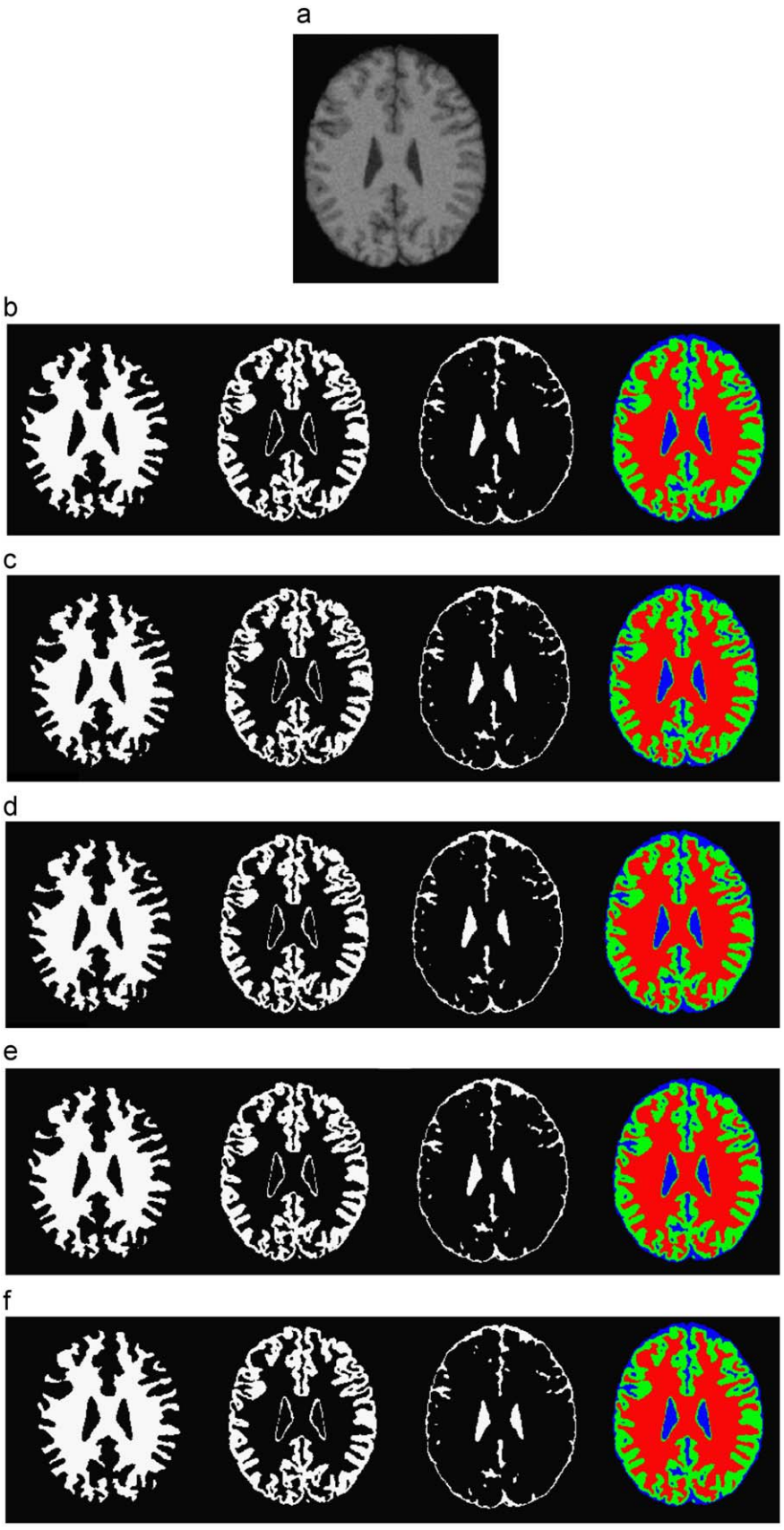

Fig. 3. Simulated T_1-weighted MR image. (a) The original image with 7% noise. (b) Discrete anatomical model (from left to right) white matter, gray matter, CSF, and the total segmentation. (c) Segmentation result of ANN-IFCM. (d) Segmentation result of GAs-IFCM. (e) Segmentation result of PSO-IFCM. (f) Segmentation result of BS-IFCM.

noise-free image and the noisy one are illustrated in Fig. 1(a) and (b), respectively.

In order to evaluate the segmentation performance quantitatively, some metrics are defined as follows:

(1) Under segmentation (Uns), representing the percentage of negative false segmentation:

$$\mathrm{UnS} = \frac{N_{fp}}{N_n} \times 100 \tag{8}$$

(2) Over segmentation (OvS), representing the percentage of positive false segmentation:

$$\mathrm{OvS} = \frac{N_{fn}}{N_p} \times 100 \tag{9}$$

(3) Incorrect segmentation (InC), representing the total percentage of false segmentation:

$$\mathrm{InC} = \frac{N_{fp}+N_{fn}}{N} \times 100 \tag{10}$$

where N_{fp} is the number of pixels that do not belong to a cluster and are segmented into the cluster. N_{fn} is the number of pixels that belong to a cluster and are not segmented into the cluster. N_p is the number of all pixels that belong to a cluster, and N_n is the total number of pixels that do not belong to a cluster.

Table 1 lists the above metrics calculated for the seven tested methods. It is clear that FCM, PCM, and RFCM cannot overcome the degradation caused by noise and their segmentation performance is very poor compared to IFCM-based algorithms. Among IFCM-based algorithms, GAs- and PSO-based methods are superior to the ANN-based method. However, unprecedented improvement in segmentation results is achieved by the BS-based method. This is the result of combining advantages of GAs and PSO, simultaneously. For better comparison, the segmentation results of IFCM-based methods are illustrated in Fig. 1(c)–(f); where the segmented classes are demonstrated in red, green, blue, and black colors.

Since the segmentation results of IFCM-based algorithms are too closed to each other, another metric is defined for better comparison of these methods. The new metric is the similarity index (SI) used for comparing the similarity of two samples defined as follows:

$$\mathrm{SI} = 2 \times \frac{A \cap B}{A+B} \times 100 \tag{11}$$

where A and B are the reference and the segmented images, respectively. We compute this metric on the squared segmented image at different noise levels. The results are averaged over 10 runs of the algorithms. Fig. 2 illustrates the performance comparison of different IFCM-based methods. The comparison clearly indicates that both GAs and PSO are superior to ANN in optimized estimation of λ and ξ. However, best results are obtained using the combination of GAs and PSO, i.e., the BS algorithm.

Simulated MR images

Generally, it is impossible to quantitatively evaluate the segmentation performance of an algorithm on real MR images, since the ground truth of segmentation for real images is not available. Therefore, only visual comparison is possible. However, Brainweb provides a simulated brain database including a set of realistic MRI data volumes produced by an MRI simulator. These data enable us to evaluate the performance of various image analysis methods in a setting where the truth is known.

In this experiment, a simulated T_1-weighted MR image ($181 \times 217 \times 181$) was downloaded from Brainweb and 7% noise was applied to each slice of the simulated image.

The 100th brain region slice of the simulated image is shown in Fig. 3(a) and its discrete anatomical structure consisting of cerebral spinal fluid (CSF), white matter, and gray matter is shown in Fig. 3(b). The noisy slice was segmented into four clusters: background, CSF, white matter, and gray matter (the background was neglected from the viewing results) using FCM, PCM, RFCM, and the IFCM-based methods. The segmentation results after applying IFCM-based methods are shown in Fig. 3(c)–(f). Also, the performance evaluation parameters of FCM, PCM, RFCM, and IFCMs are compared in Table 2.

Again, it is obvious that the BS-IFCM has achieved the best segmentation results. These observations are consistent with the simulation results obtained in the previous section.

Real MR images

Finally, an evaluation was performed on real MR images. A real MR image (coronal T_1-weighted image with a matrix of 256×256) was obtained from IBSR: the Center of Morphometric Analysis at Massachusetts General Hospital. IBSR provides manually guided expert segmentation results along with brain MRI data to support the evaluation and development of segmentation methods.

Fig. 4(a) shows an slice of the image with 5% Gaussian noise and Fig. 4(b) shows the manual segmentation results provided by the IBSR. For comparison with the manual segmentation results,

Table 2
Segmentation evaluation on simulated T_1-weighted MR.

Class	Evaluation parameters	FCM	PCM	RFCM	ANN-IFCM	GAs-IFCM	PSO-IFCM	BS-IFCM
CSF	UnS(%)	0.5	0	0.47	0.2	0.16	0.11	0.08
	OvS(%)	7.98	100	7.98	6.82	5.91	4.36	3.26
	InC(%)	0.76	34	0.73	0.57	0.45	0.31	0.22
White matter	UnS(%)	1.35	0	1.11	0.95	0.91	0.78	0.55
	OvS(%)	11.08	100	10.92	7.31	7.02	5.56	4.69
	InC(%)	2.33	10.16	2.11	1.59	1.39	1.06	0.84
Gray matter	UnS(%)	0.75	15.86	0.76	0.54	0.48	0.29	0.17
	OvS(%)	7.23	0	5.72	2.65	2.61	2.13	1.34
	InC(%)	1.68	13.57	1.47	0.93	0.87	0.71	0.36
Average	UnS(%)	0.87	5.29	0.78	0.56	0.52	0.39	0.27
	OvS(%)	8.76	66.67	8.21	5.59	5.18	4.02	3.1
	InC(%)	1.59	19.24	1.44	1.03	0.9	0.69	0.44

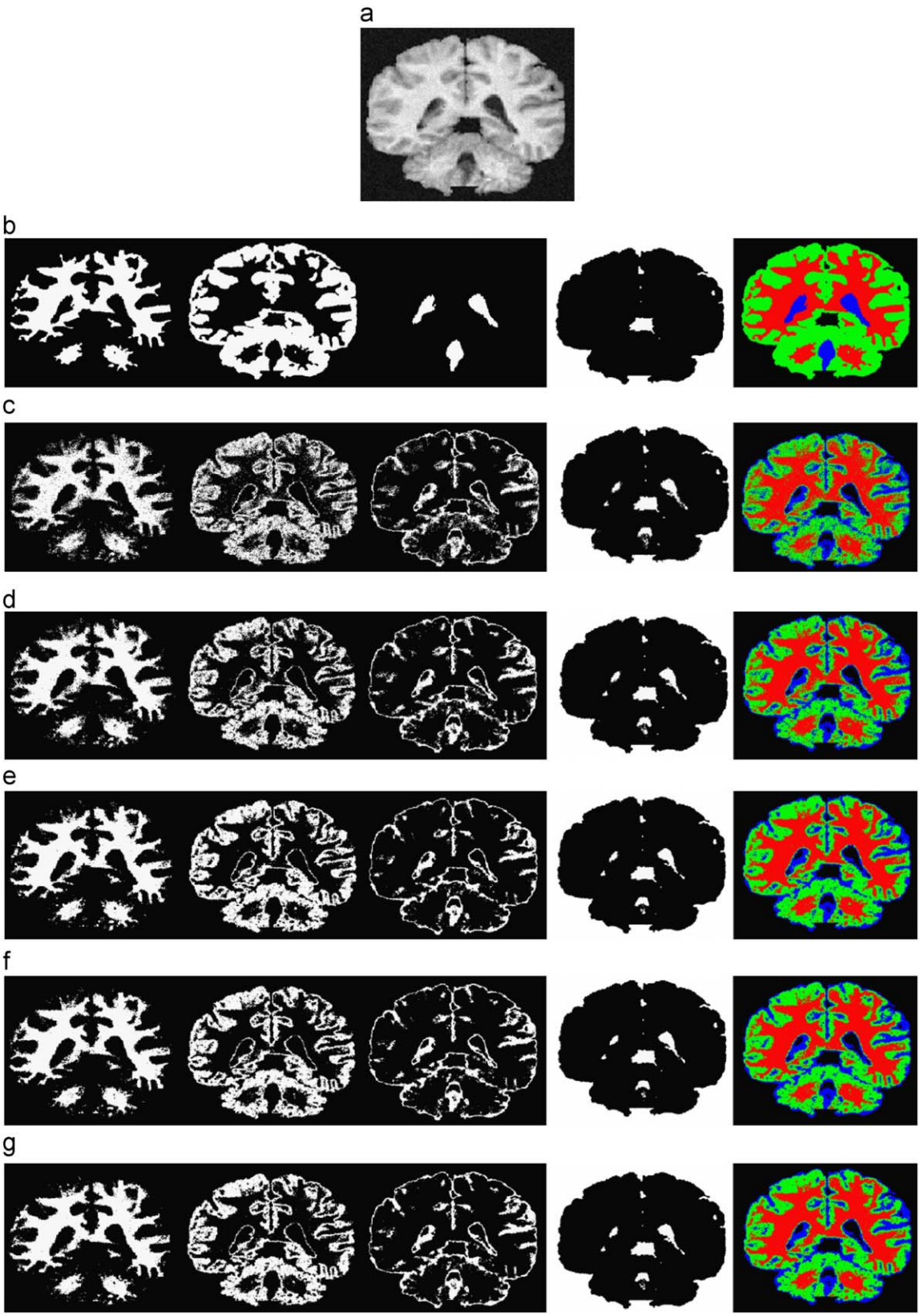

Fig. 4. Real T_1-weighted MR image. (a) The original image with 5% noise. (b) Discrete anatomical model (from left to right) white matter, gray matter, CSF, others, and the total segmentation. (c) Segmentation results of FCM (d) Segmentation result of ANN-IFCM. (e) Segmentation result of GAs-IFCM. (f) Segmentation result of PSO-IFCM. (g) Segmentation result of BS-IFCM.

which included four classes (CSF, gray matter, white matter, and others), the cluster number was set to 4. The segmentation results of FCM algorithm is shown in Fig. 4(c), while segmentation of IFCM-based methods are shown in Fig. 4(d)–(g). Table 3 lists the evaluation parameters for all methods. BS-IFCM showed a significant improvement over other IFCMs both visually and

Table 3
Segmentation evaluation on Real T_1-weighted MR image.

Class	Evaluation parameters	FCM	ANN-IFCM	GA-IFCM	PSO-IFCM	BS-IFCM
CFS	UnS(%)	11.1732	11.1142	10.6406	10.1619	9.5794
	OvS(%)	44.4444	45.1356	41.4939	40.9091	41.8926
	InC(%)	12.4009	12.7177	11.7791	11.2965	10.7717
	SI(%)	87.5991	87.6305	88.2209	88.7035	89.2283
White matter	UnS(%)	3.3556	2.7622	0.9783	1.5532	1.6639
	OvS(%)	14.8345	9.6177	3.1523	9.0279	8.8211
	InC(%)	6.2951	4.5178	1.5350	3.4673	3.4967
	SI(%)	93.7049	95.4822	98.4650	96.5327	96.5033
Gray matter	UnS(%)	5.9200	3.8073	3.8469	3.5824	3.5353
	OvS(%)	36.9655	35.9035	31.4066	30.5603	29.133
	InC(%)	16.9305	15.1905	13.6211	13.1503	12.6138
	SI(%)	83.0695	87.6305	86.3789	86.8497	87.3862
Average	UnS(%)	5.7635	5.1014	4.5606	4.4094	4.2800
	OvS(%)	24.2163	22.7491	20.7562	20.2043	20.0522
	InC(%)	9.3831	8.4900	7.6465	7.3856	7.1315
	SI(%)	90.6169	91.5100	92.3535	92.6144	92.8685

parametrically, and completely eliminated the effect of noise. These results nominate the BS-IFCM algorithm as a good technique for segmentation of noisy brain MR images in real application.

4. Conclusion

In this paper, we proposed new algorithms, namely GAs and PSO, to estimate the optimized values of neighborhood attraction parameters in IFCM clustering algorithm. GAs are best at reaching a near optimal solution but have trouble finding an exact solution, while PSO's group interactions enhances the search for an optimal solution. Therefore, a combined GAs/PSO, the BS, algorithm was employed for further improvements. Although the BS algorithm seems to be more complicated than GAs and PSO, it is able to locate an optimal solution significantly faster than either GAs or PSO. This is the result of combining the strengths of PSO with GAs, simultaneously. The GAs facilitate a global search to reach a near optimal solution and the PSO's group interactions enhances the search for the optimal local solution. We tested the proposed methods on three kinds of images; a square image, simulated brain MR images, and real brain MR images. Both quantitative and quantitative comparisons at different noise levels demonstrated that both GAs and PSO are superior to the previously proposed ANN method in optimizing the attraction parameters. However, significant improvements in segmentation results were achieved using the BS algorithm. These results nominate the BS-IFCM algorithm as a good technique for segmentation of noisy brain MR images.

Acknowledgment

The authors would like to thank Mr. Youness Aliyari-Ghassabeh for the constructive discussions and useful suggestions.

References

Angeline, P.J., 1998. Evolutionary optimization versus particle swarm optimization: philosophy and performance differences. Evolutionary Programming VII (1998), Lecture Notes in Computer Science, 1447. Springer.

Bezdek, J.C., 1981. In: Pattern Recognition with Fuzzy Objective Function Algoritms. Plenum Press, New York.

Bondareff, W., Raval, J., Woo, B., Hauser, D.L., Colletti, P.M., 1990. Magnetic resonance and severity of dementia in older adults. Arch. Gen. Psychiatry 47, 47–51.

Canny, J.A., 1986. Computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. 8, 679–698.

Clarke, L.P., Velthuizen, R.P., Camacho, M.A., et al., 1995. MRI segmentation: methods and applications. Magn. Reson. Imag. 13, 343–368.

Corne, D., Dorigo, M., Glover, F. (Eds.), 1999. New Ideas in Optimization. McGraw Hill.

Dave, R.N.R.N., 1991. Characterization and detection of noise in clustering. Pattern Recognition Lett. 12, 657–664.

Dunn, J.C., 1973. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. J. Cybern. 3, 32–57.

Eberhart, R.C., Shi, Y., 1998. Comparison between genetic algorithms and particle swarm optimization. Evolutionary Programming VII (1998), Lecture Notes in Computer Science, 1447. Springer.

Eberhart, R.C., Dobbins, R.W., Simpson, P., 1996. In: Computation Intelligence PC Tools. Academic Press.

Gen, M, Cheng, R., 1997. In: Genetic Algorithms and Engineer Design. John Wiley.

Haacke, E.M., Brown, R.W., Thompson, M.L., Venkatesan., R., 1999. In: Magnetic Resonance Imaging: Physical Principles and Sequence Design. John Wiley.

Holland, J.H., 1992. Adaptation in Natural and Artificial Systems. MIT Press, Cambridge, MA, USA.

Kennedy, J., Eberhart, R.C., 1995. Particle swarm optimisationProceedings IEEE International Conference on Neural Networks, vol. IV; .

Kennedy, J., Eberhart, R.C., 2001. Swarm Intelligence. Morgan Kaufman Publishers.

Krishnapuram, Keller, J.M., 1993. A possibilistic approach to clustering. IEEE Trans. Fuzzy Syst. 1 (2), 98–110.

Li, C.L., Goldgof, D.B., Hall, L.O., 1993. Knowledge-based classification and tissue labeling of MR images of human brain. IEEE Trans. Med. Imag. 12, 740–750.

Mitchel, M., 1999. In: An Introduction to Genetic Algorithms, fifth printing. MIT Press.

Pham, D.L., 2001. Spatial models for fuzzy clustering. Comput. Vis. Imag. Understand. 84, 285–297.

Pohle, R., Toennies, K.D., 2001. Segmentation of medical images using adaptive region growing. Proc. SPIE Med. Imag., vol. 4322; .

Settles, M., Soule, T., 2005. Breeding swarms: a GA/PSO hybrid. In: Beyer, H.G., O'Reilly, U.-M. (Eds.), Proceedings of the 2005 conference on Genetic and evolutionary computation, GECCO 2005. ACM Press.

Shen, S., Sandham, W., Granat, M., Sterr, A., 2005. MRI fuzzuy segmentation of brain tissue using neighborhood attraction with neural-network optimization. IEEE Trans. Inf. Technol. Biomed. 9, 459–467.

Suzuki, H., Toriwaki, J., 1991. Automatic segmentation of head MRI images by knowledge guided thresholding. Comput. Med. Imag. Graph. 15 233–240

Wells III, W.M., Grimson, W.E.L., Kikinis, R., Jolesz, F.A., 1996. Adaptive segmentation of MRI data. IEEE Trans. Med. Imag. 15, 429–442.

Yu, J., Yang, M.S., 2005. Optimality test for generalized FCM and its application to parameter selection. IEEE Trans. Fuzzy Syst. 13 (2), 164–176.

Yu, J., Yang, M.-S., 2007. A generalized fuzzy clustering regularization model with optimality tests and model complexity analysis. IEEE Trans. Fuzzy Syst. 15 (5), 904–915.

Zadeh, L.A., 1965. Fuzzy sets. Inf. Control 8, 338–353.