We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

12

Application of Particle Swarm Optimization in Accurate Segmentation of Brain MR Images

Nosratallah Forghani[1], Mohamad Forouzanfar[1,2], Armin Eftekhari[1], Shahab Mohammad-Moradi[1] and Mohammad Teshnehlab[1]
[1]K.N. Toosi University of Technology, [2]University of Tehran
[1,2]Iran

1. Introduction

Medical imaging refers to the techniques and processes used to obtain images of the human body for clinical purposes or medical science. Common medical imaging modalities include ultrasound (US), computerized tomography (CT), and magnetic resonance imaging (MRI). Medical imaging analysis is usually applied in one of two capacities: i) to gain scientific knowledge of diseases and their effect on anatomical structure in vivo, and ii) as a component for diagnostics and treatment planning (Kannan, 2008).

Medical US uses high frequency broadband sound waves that are reflected by tissue to varying degrees to produce 2D or 3D images. This is often used to visualize the fetus in pregnant women. Other important uses include imaging the abdominal organs, heart, male genitalia, and the veins of the leg. US has several advantages which make it ideal in numerous situations. It studies the function of moving structures in real-time, emits no ionizing radiation, and contains speckle that can be used in elastography. It is very safe to use and does not appear to cause any adverse effects. It is also relatively cheap and quick to perform. US scanners can be taken to critically ill patients in intensive care units, avoiding the danger caused while moving the patient to the radiology department. The real time moving image obtained can be used to guide drainage and biopsy procedures. Doppler capabilities on modern scanners allow the blood flow in arteries and veins to be assessed. However, US images provides less anatomical detail than CT and MRI (Macovski, 1983).

CT is a medical imaging method employing tomography (Slone et al., 1999). Digital geometry processing is used to generate a three-dimensional image of the inside of an object from a large series of two-dimensional X-ray images taken around a single axis of rotation. CT produces a volume of data which can be manipulated, through a process known as windowing, in order to demonstrate various structures based on their ability to block the X-ray beam. Although historically the images generated were in the axial or transverse plane (orthogonal to the long axis of the body), modern scanners allow this volume of data to be reformatted in various planes or even as volumetric (3D) representations of structures. CT was the first imaging modality to provide in vivo evidence of gross brain morphological abnormalities in schizophrenia, with many CT reports of increase in cerebrospinal fluid (CSF)-filled spaces, both centrally (ventricles), and peripherally (sulci) in a variety of psychiatric patients.

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. MRI is based on the phenomenon of nuclear magnetic resonance (NMR). Nuclei with an odd number of nucleons, exposed to a uniform static magnetic field, can be excited with a radio frequency (RF) pulse with the proper frequency and energy. After the excitation pulse, NMR signal can be recorded. The return to equilibrium is characterized by relaxation times T1 and T2, which depend on the nuclei imaged and on the molecular environment. Mainly hydrogen nuclei (proton) are imaged in clinical applications of MRI, because they are most NMR-sensitive nuclei (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 is becoming 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 tumours 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, most notably thresholding (Suzuki & Toriwaki, 1991), edge detection (Canny, 1986), region growing (Pohle & Toennies, 2001), and clustering (Dubes & Jain, 1988). Thresholding is the simplest segmentation method, where the classification of each pixel depends on its own information such as intensity and colour. 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 in the edge image. These drawbacks of early image segmentation methods, has led to region growing algorithms. Region growing extends thresholding by combining it with connectivity conditions or region homogeneity 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 (HCM) and its fuzzy equivalent, fuzzy c-means (FCM) algorithms (Li et al., 1993) are the typical methods of clustering. A common disadvantage of EM algorithms is that the intensity distribution of brain images is modeled as a normal distribution, which is untrue, 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. Unfortunately, the greatest shortcoming of FCM is its over-sensitivity to noise, which is also a drawback of many other intensity-based segmentation methods. 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 extensions of the FCM algorithm have been reported in the literature to overcome the effects of noise, such as noisy clustering (NC) (Dave, 1991), possibilistic c-means (PCM)

(Krishnapuram & Keller, 1993), robust fuzzy c-means algorithm (RFCM) (Pham, 2001), bias-corrected FCM (BCFCM) (Ahmed et al., 2002), spatially constrained kernelized FCM (SKFCM) (Zhang & Chen, 2004), 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.

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 neighbourhood attraction. The first parameter is the feature difference between neighbouring pixels in the image and the second one is the relative location of the neighbouring pixels. Therefore, segmentation is decided not only by the pixel's intensity but also by neighbouring 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 (Haykin, 1998).

In this paper we 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 both GAs and PSO are superior to ANN in segmentation of noisy MR images; however, PSO obtains the best results. The achieved improvements are 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, common clustering algorithms, including EM, FCM, and different extensions of FCM, are introduced. Section 3 presents new parameter optimization methods based on GAs and PSO for determination of optimum degree of attraction in IFCM algorithm. Section 4 is dedicated to a comprehensive comparison of the proposed segmentation algorithms based on GAs and PSO with related recent techniques. The paper in concluded in Section 5 with some remarks.

2. Clustering Algorithms

According to the limitation of conventional segmentation methods such as thresholding, edge detection, and region growing, clustering methods are utilized for complex pathology. Clustering is an unsupervised classification of data samples with similar characteristics into larger units of analysis (clusters). While classifiers are trained on pre-labeled data and tested on unlabeled data, clustering algorithms take as input a set of unlabeled samples and organize them into clusters based on similarity criteria. The algorithm alternates between dividing the data into clusters and learning the characteristics of each cluster using the current division. In image segmentation, a clustering algorithm iteratively computes the characteristics of each cluster (e.g. mean, standard deviation) and segments the image by classifying each pixel in the closest cluster according to a distance metric. The algorithm alternates between the two steps until convergence is achieved or a maximum number of

iterations is reached (Lauric & Frisken, 2007). In this Section, typical methods of clustering including EM algorithm, FCM algorithm, and extensions of FCM are described.

2.1 EM Algorithm

The EM algorithm is an estimation method used in statistics for finding maximum likelihood estimates of parameters in probabilistic models, where the model depends on unobserved latent variables (Wells et al., 1994). In image segmentation, the observed data are the feature vectors x_j associated with pixels j, while the hidden variables are the expectations E_{ji} for each pixel j that it belongs to each of the given clusters i.

The algorithm starts with an initial guess at the model parameters of the clusters and then re-estimates the expectations for each pixel in an iterative manner. Each iteration consists of two steps: the expectation (E) step and the maximization (M) step. In the E-step, the probability distribution of each hidden variable is computed from the observed values and the current estimate of the model parameters (e.g. mean, covariance). In the M-step, the model parameters are re-estimated assuming the probability distributions computed in the E-step are correct. The parameters found in the M step are then used to begin another E step, and the process is repeated.

Assuming Gaussian distributions for all clusters, the hidden variables are the expectations E_{ji} that pixel j belongs to cluster i. The model parameters to estimate are the mean, the covariance and the mixing weight corresponding to each cluster. The mixing weight is a measure of a cluster's strength, representing how prevalent the cluster is in the data. The E and M step of the EM algorithm are as follows.

E-step:

$$E_{ji}^{t} = P\left(i\middle|x_j, \theta_i^t\right) = \frac{P\left(x_j\middle|i, \theta_i^t\right)\pi_i^t}{\sum_{k=1}^{c} P\left(x_j\middle|k, \theta_k^t\right)\pi_k^t} \tag{1}$$

M-step:

$$\pi_i^{t+1} = \frac{1}{n}\sum_{j=1}^{n} E_{ji}^{t} \tag{2}$$

$$v_i^{t+1} = \frac{1}{n\pi_i^{t+1}}\sum_{j=1}^{n} E_{ji}^{t} x_j \tag{3}$$

$$\Sigma_i^{t+1} = \frac{1}{n\pi_i^{t+1}}\sum_{j=1}^{n} E_{ji}^{t}\left(x_j - v_i^{t+1}\right)\left(x_j - v_i^{t+1}\right)^T \tag{4}$$

where θ_i^t are the model parameters of class i at time t and π_i^t is the mixing weight of class i at time t. Note that $\sum_{i=1}^{c}\pi_i^t = 1, \forall t$. $P\left(i\middle|x_j\right)$ is the a posteriori conditional probability that pixel j is a member of class i, given its feature vector x_j. $P\left(i\middle|x_j\right)$ gives the membership value of pixel j to class i, where i takes values between 1 and c (the number of classes), while j takes values between 1 and n (the number of pixels in the image). Note that $\sum_{k=1}^{c} P\left(k\middle|x_j\right) = 1$. $P\left(x_j\middle|i\right)$ is the conditional density distribution, i.e., the probability that pixel j has feature

vector x_j given that it belongs to class i. If the feature vectors of class i have a Gaussian distribution, the conditional density function has the form:

$$P(x_j|i) = \frac{1}{(2\pi)^{\frac{D}{2}} det^{\frac{1}{2}}(\Sigma_i)} e^{-\frac{1}{2}(x_j-v_i)^T \Sigma_i (x_j-v_i)} \quad (5)$$

where v_i and Σ_i are the mean feature vector and the covariance matrix of class i. The mean and the covariance of each class are estimated from training data. D is the dimension of the feature vector. The prior probability of class i is:

$$P(i) = \frac{|w_i|}{\sum_{k=1}^{c}|w_k|} \quad (6)$$

where $|w_i|$ is a measure of the frequency of occurrence of class i and $\sum_{k=1}^{c}|w_k|$ is a measure of the total occurrence of all classes. In image segmentation, $|w_i|$ is usually set to the number of pixels which belong to class i in the training data, and $\sum_{k=1}^{c}|w_k|$ to the total number of pixels in the training data.

The algorithm iterates between the two steps until the log likelihood increases by less than some threshold or a maximum number of iterations is reached. EM algorithm can be summarized as follows (Lauric & Frisken, 2007):

1. Initialize the means v_i^0, the covariance matrices Σ_i^0 and the mixing weights π_i^0. Typically, the means are initialized to random values, the covariance matrices to the identity matrix and the mixing weights to $1/c$.
2. E-step: Estimate E_{ji} for each pixel j and class i, using (1).
3. M-step: Estimate the model parameters for class i, using (2)-(4).
4. Stop if convergence condition $\left(\log\prod_{j=1}^{n} E_{ji}^{t+1} - \log\prod_{j=1}^{n} E_{ji}^{t}\right) \leq \epsilon$ is achieved. Otherwise, repeat steps 2 to 4.

A common disadvantage of EM algorithms is that the intensity distribution of brain images is modeled as a normal distribution, which is untrue, especially for noisy images.

2.2 FCM Algorithm

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 $X_1, \ldots, X_c$ by hard c-partition $\{\mu_1, \ldots, \mu_c\}$. 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):

$$J_{FCM}(\mu, v) = \sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m} d^2(x_j, v_i) \quad (7)$$

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:

$$v_i = \frac{\sum_{j=1}^{n} \mu_{ij}^m x_j}{\sum_{j=1}^{n} \mu_{ij}^m} \quad (i = 1, \dots, c) \tag{8}$$

and

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d(x_j, v_i)}{d(x_j, v_k)} \right)^{2/(m-1)}} \tag{9}$$

The FCM algorithm iteratively optimizes $J_{FCM}(\mu, v)$ with the continuous update of μ and v, until $\left|\mu^{(l+1)} - \mu^l\right| \leq \varepsilon$, where l is the number of iterations.

From (7), it is clear that the objective function of FCM does not take into consideration any spatial dependence among X and deals with each image pixel as a separate point. Also, the membership function in (9) is mostly decided by $d^2(x_j, v_i)$, which measures the similarity between the pixel intensity and the cluster center. Higher membership depends on closer intensity values to the cluster center. It therefore increases the sensitivity of the membership function to noise. If an MR image contains noise or is affected by artifacts, their presence can change the pixel intensities, which will result in an incorrect membership and improper segmentation.

There are several approaches to reduce sensitivity of FCM algorithm to noise. The most direct way is the use of low pass filters in order to smooth the image and then applying the FCM algorithm. However low pass filtering, may lead to lose some important details. Different extensions of FCM algorithm were proposed by researchers in order to solve sensitivity to noise. In the following Subsections we will introduce some of these extensions.

2.2.1 NC algorithm

The most popular approach for increasing the robustness of FCM to noise is to modify the objective function directly. Dáve (1991) proposed the idea of a noise cluster to deal with noisy clustering data in the approach known as NC. Noise is effectively clustered into a separate cluster which is unique from signal clusters. However, it is not suitable for image segmentation, since noisy pixels should not be separated from other pixels, but assigned to the most appropriate clusters in order to reduce the effect of noise.

2.2.2 PCM algorithm

Another similar method, developed by Krishnapuram and Keller (1993), is called PCM, which interprets clustering as a possibilistic partition. Instead of having one term in the objective function, a second term is included, forcing the membership to be as high as possible without a maximum limit constraint of one. However, it caused clustering being stuck in one or two clusters. The objective function of PCM is defined as follows:

$$J_{PCM}(\mu, v) = \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij}^m d^2(x_j, v_i) + \sum_{i=1}^{c} \eta_i \sum_{j=1}^{n} (1 - \mu_{ij})^m \tag{10}$$

where η_i are suitable positive numbers. The first term demands that the distances from the feature vectors to the prototypes be as low as possible, whereas the second term forces the μ_{ij} to be as large as possible, thus avoiding the trivial solution.

2.2.3 RFCM algorithm

Pham presented a new approach of FCM, named the robust RFCM (Pham, 2001). A modified objective function was proposed for incorporating spatial context into FCM. A parameter controls the tradeoff between the conventional FCM objective function and the smooth membership functions. However, the modification of the objective function results in the complex variation of the membership function. The objective function of RFCM is defined as follows:

$$J_{RFCM}(\mu, v) = \sum_{j\in\Omega}\sum_{i=1}^{c} \mu_{ij}^{m} d^2(x_j, v_i) + \frac{\beta}{2}\sum_{j\in\Omega}\sum_{i=1}^{c} \mu_{ij}^{m} \sum_{r\in N_j}\sum_{s\in M_i} \mu_{rs}^{m} \tag{11}$$

where N_j is the set of neighbors of pixel j in Ω, and $M_i = \{1, \ldots, c\}\backslash\{i\}$. The parameter β controls the trade-off between minimizing the standard FCM objective function and obtaining smooth membership functions.

2.2.4 BCFCM algorithm

Another improved version of FCM by the modification of the objective function was introduced by Ahmed et al. (2002). They proposed a modification of the objective function by introducing a term that allows the labeling of a pixel to be influenced by the labels in its immediate neighborhood. The neighborhood effect acts as a regularizer and biases the solution toward piecewise-homogeneous labeling. Such regularization is useful in segmenting scans corrupted by salt and pepper noise. The modified objective function is given by:

$$J_{PCM}(\mu, v) = \sum_{i=1}^{c}\sum_{j=1}^{n} \mu_{ij}^{m} d^2(x_j, v_i) + \frac{\alpha}{N_R}\sum_{i=1}^{c}\sum_{j=1}^{n} \mu_{ij}^{m} \left(\sum_{x_r\in N_j} d^2(x_r, v_i) \right) \tag{12}$$

where N_j stands for the set of neighbors that exist in a window around x_j and N_R is the cardinality of N_j. The effect of the neighbors term is controlled by the parameter α. The relative importance of the regularizing term is inversely proportional to the signal-to-noise ratio (SNR) of the MRI signal. Lower SNR would require a higher value of the parameter α.

2.2.5 SKFCM algorithm

The SKFCM uses a different penalty term containing spatial neighborhood information in the objective function, and simultaneously the similarity measurement in the FCM, is replaced by a kernel-induced distance (Zhang & Chen, 2004). We know every algorithm that only uses inner products can implicitly be executed in the feature space F. This trick can also be used in clustering, as shown in support vector clustering (Hur et al., 2001) and kernel FCM (KFCM) algorithms (Girolami, 2002; Zhang & Chen, 2002). A common ground of these algorithms is to represent the clustering center as a linearly-combined sum of all $\emptyset(x_j)$, i.e. the clustering centers lie in feature space. The KFCM objective function is as follows:

$$J_{KFCM}(\mu, v) = \sum_{i=1}^{c}\sum_{j=1}^{n} \mu_{ij}^{m} d^2\left(\emptyset(x_j), \emptyset(v_i)\right) \tag{13}$$

where $\emptyset$ is an implicit nonlinear map. Same as BCFCM, the KFCM-based methods inevitably introduce computation issues, by modifying most equations along with the modification of the objective function, and have to lose the continuity from FCM, which is well-realized with many types of software, such as MATLAB.

2.2.6 IFCM algorithm

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, an attraction entitled 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 λ, and the spatial position of the neighbors or distance attraction ξ, which also depends on the neighborhood structure. Considering this neighborhood attraction, they defined the similarity function as below:

$$d^2(x_j, v_i) = \|x_j - v_i\|^2 (1 - \lambda H_{ij} - \xi F_{ij}) \tag{14}$$

where H_{ij} represents the feature attraction and F_{ij} represents the distance attraction. Magnitudes of two parameters λ and ζ are between 0 and 1; adjust the degree of the two neighborhood attractions. H_{ij} and F_{ij} are computed in a neighborhood containing S pixels as follow:

$$H_{ij} = \frac{\sum_{k=1}^{S} \mu_{jk} g_{jk}}{\sum_{k=1}^{S} \mu_{jk}} \tag{15}$$

$$F_{ij} = \frac{\sum_{k=1}^{S} \mu_{ik}^2 q_{jk}^2}{\sum_{k=1}^{S} \mu_{ik}^2} \tag{16}$$

with

$$g_{jk} = |x_j - x_k|, \; q_{jk} = (a_j - a_k)^2 + (b_j - b_k)^2 \tag{17}$$

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 leads to stronger feature attraction and a higher value of ξ leads to stronger distance attraction. Optimized values of these parameters enable the best segmentation results to be achieved. However, inappropriate values can be detrimental. Therefore, parameter optimization is an important issue in IFCM algorithm that can significantly affect the segmentation results.

3. Parameter Optimization of IFCM Algorithm

Optimization algorithms are search methods, where the goal is to find a solution to an optimization problem, such that a given quantity is optimized, possibly subject to a set of constrains. Although this definition is simple, it hides a number of complex issues. For

example, the solution may consist of a combination of different data types, nonlinear constrains may restrict the search area, the search space can be convoluted with many candidate solutions, the characteristics of the problem may change over time, or the quantity being optimized may have conflicting objectives (Engelbrecht, 2006).
As mentioned earlier, the problem of determining optimum attraction parameters constitutes an important part of implementing the IFCM algorithm. Shen et al. (2005) computed these two parameters using an ANN through an optimization problem. However, designing the neural network architecture and setting its parameters are always complicated tasks which slow down the algorithm and may lead to inappropriate attraction parameters and consequently degrade the partitioning performance. In this Section we introduce two new algorithms, namely GAs and PSO, for optimum determination of the attraction parameters. The performance evaluation of the proposed algorithms is carried out in the next Section.

3.1. Structure of 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 (1992) and have been widely successful in optimization problems. 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:
    a. [Selection] Select two parent chromosomes from a population according to their fitness (the better fitness, the bigger chance to be selected).
    b. [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.
    c. [Mutation] With a mutation probability, mutate new offspring at each locus (position in chromosome).
    d. [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 & Cheng, 1997):

1. Generate a random number r from the rang $[0,\ 1]$.
2. If $r \leq P_1$, then select the first chromosome, otherwise select the k-th 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.

3.2. Structure of PSO

Team formation has been observed in many animal species. For some animal species, teams or groups are controlled by a leader, for example a pride of lions, a troop of baboon, and a troop of wild buck. In these societies the behavior of individuals is strongly dictated by social hierarchy. More interesting is the self-organizing behavior of species living in groups where no leader can be identified, for example, a flock of birds, a school of fish, or a herd of sheep. Within these social groups, individuals have no knowledge of the global behavior of the entire group, nor they have any global information about the environment. Despite this, they have the ability to gather and move together, based on local interactions between individuals. From the simple, local interaction between individuals, more complex collective behavior emerges, such as flocking behavior, homing behavior, exploration and herding. Studies of the collective behavior of social animals include (Engelbrecht, 2006):

1. Bird flocking behavior;
2. Fish schooling behavior;
3. The hunting behavior of humpback whales;
4. The foraging behavior of wild monkeys; and
5. The courtship-like and foraging behavior of the basking shark.

PSO, introduced by Kennedy and Eberhart (1995), is a member of wide category of swarm intelligence methods (Kennedy & Eberhart, 2001). Kennedy originally proposed PSO as a simulation of social behavior and it was initially introduced as an optimization method. 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 & 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 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 are moved toward the best solution by repeating the following steps:
   a. If $f(\bar{x}_i(t)) < pbest_i$ 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 i-th particle and $\bar{x}_{pbest_i}$ is the corresponding coordinate.
   b. If $f(\bar{x}_i(t)) < lbest_i$ then $lbest = f(\bar{x}_i(t))$, where $lbest_i$ is the best fitness over the topological neighbors.
   c. Change the velocity v_i of each particle:

$$\bar{v}_i(t) = \bar{v}_i(t-1) + \rho_1 \left(\bar{x}_{pbest_i} - \bar{x}_i(t)\right) + \rho_2 \left(\bar{x}_{lbest_i} - \bar{x}_i(t)\right) \tag{18}$$

   where ρ_1 and ρ_2 are random accelerate constants between 0 and 1.
   d. 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.

4. 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 & 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[1]; and 3) real MR images acquired from IBSR[2]. 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 (14) 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).

4.1 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, respectively. In order to investigate the sensitivity of the algorithms to noise, a uniformly distributed noise in the

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

interval (0, 120) is added to the image. The reference noise-free image and the noisy one are illustrated in Figures 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:

$$UnS = \frac{N_{fp}}{N_n} \times 100 \tag{19}$$

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

$$OvS = \frac{N_{fn}}{N_p} \times 100 \tag{20}$$

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

$$InC = \frac{N_{fp} + N_{fn}}{N} \times 100 \tag{21}$$

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, the PSO-based is superior to the others. For better comparison, the segmentation results of IFCM-based methods are illustrated in Figures 1(c)-(e); where the segmented classes are demonstrated in red, green, blue and black colors.

Evaluation parameters	FCM	PCM	RFCM	ANN-IFCM	GAs-IFCM	PSO-IFCM
UnS(%)	9.560	25.20	6.420	0.0230	0.0210	0.0110
OvS(%)	23.79	75.00	16.22	0.0530	0.0468	0.0358
InC(%)	14.24	43.75	9.88	0.0260	0.0220	0.0143

Table 1. Segmentation evaluation of synthetic square image

Since the segmentation results of IFCM-based algorithms are too closed to each other, we define another metric 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:

$$SI = 2 \times \frac{A \cap B}{A + B} \times 100 \tag{22}$$

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. Figure 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 PSO algorithm.

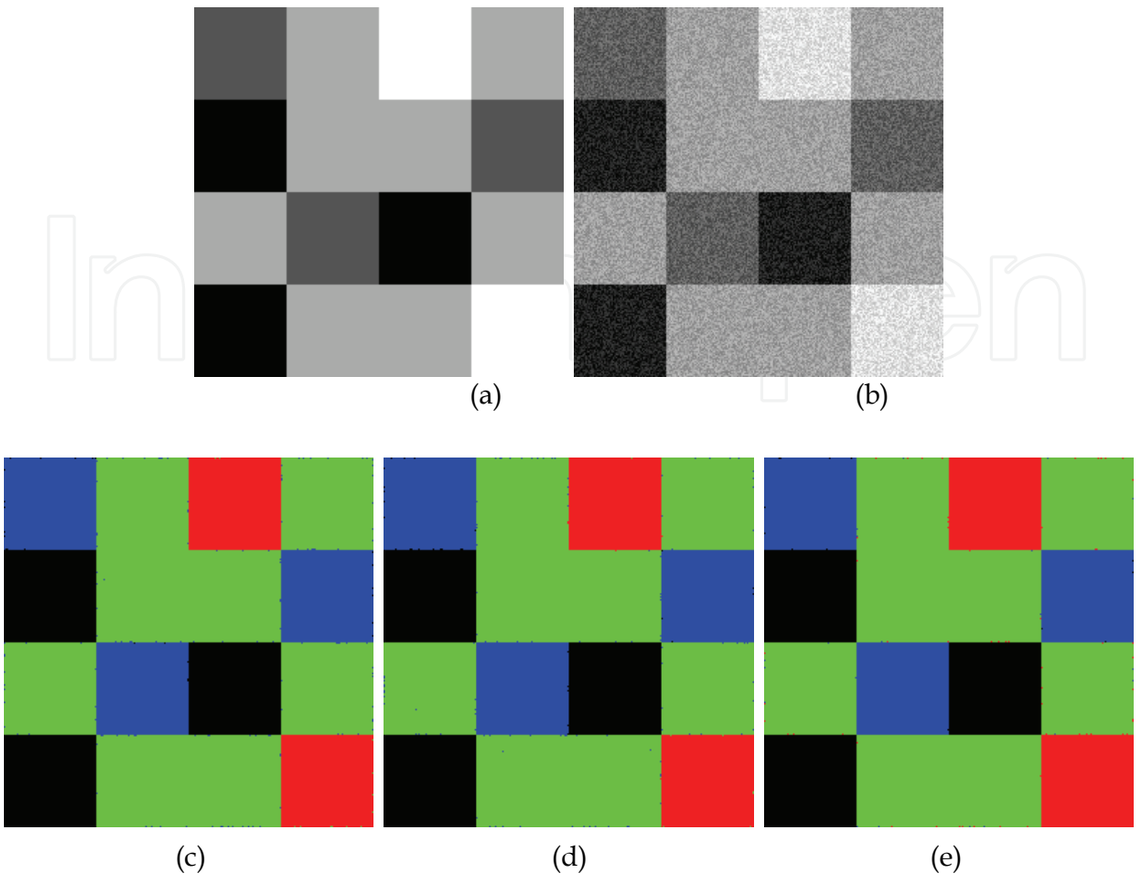


Figure 1. Segmentation results on a synthetic square image with a uniformly distributed noise in the interval (0, 120). (a) Noise-free reference image, (b) Noisy image, (c) ANN-IFCM, (d) GAs-IFCM, (e) PSO-IFCM

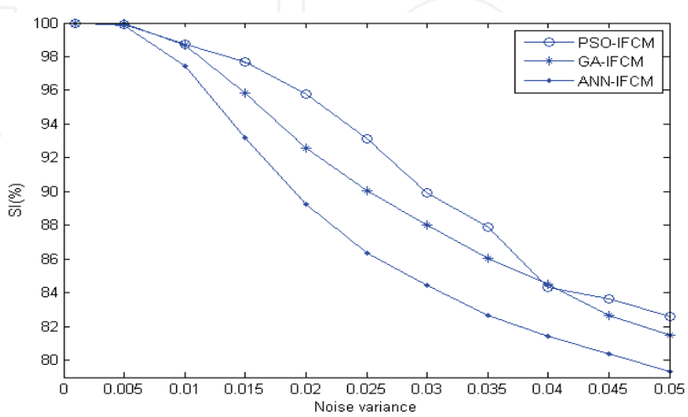


Figure 2. Performance comparison of IFCM-based methods using the SI metric at different noise levels

4.2 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 × 217 × 181) was downloaded from Brainweb. 7% noise was applied to each slice of the simulated image. The 100th brain region slice of the simulated image is shown in Figure 3(a) and its discrete anatomical structure consisting of cerebral spinal fluid (CSF), white matter, and gray matter is shown in Figure 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 Figures 3(c)-(e). Also, the performance evaluation parameters of FCM, PCM, RFCM, and IFCMs are compared in Table 2. Again, it is obvious that the PSO-IFCM has achieved the best segmentation results. These observations are consistent with the simulation results obtained in the previous Section.

4.3 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.

Figure 4(a) shows a slice of the image with 5% Gaussian noise and Figure 4(b) shows the manual segmentation result provided by the IBSR. For comparison with the manual segmentation result, 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 Figure 4(c), while segmentation of IFCM-based methods are shown in Figures 4(d)-(f). Table 3 lists the evaluation parameters for all methods. PSO-IFCM showed a significant improvement over other IFCMs both visually and parametrically, and eliminated the effect of noise, considerably. These results nominate the PSO-IFCM algorithm as a good technique for segmentation of noisy brain MR images in real application.

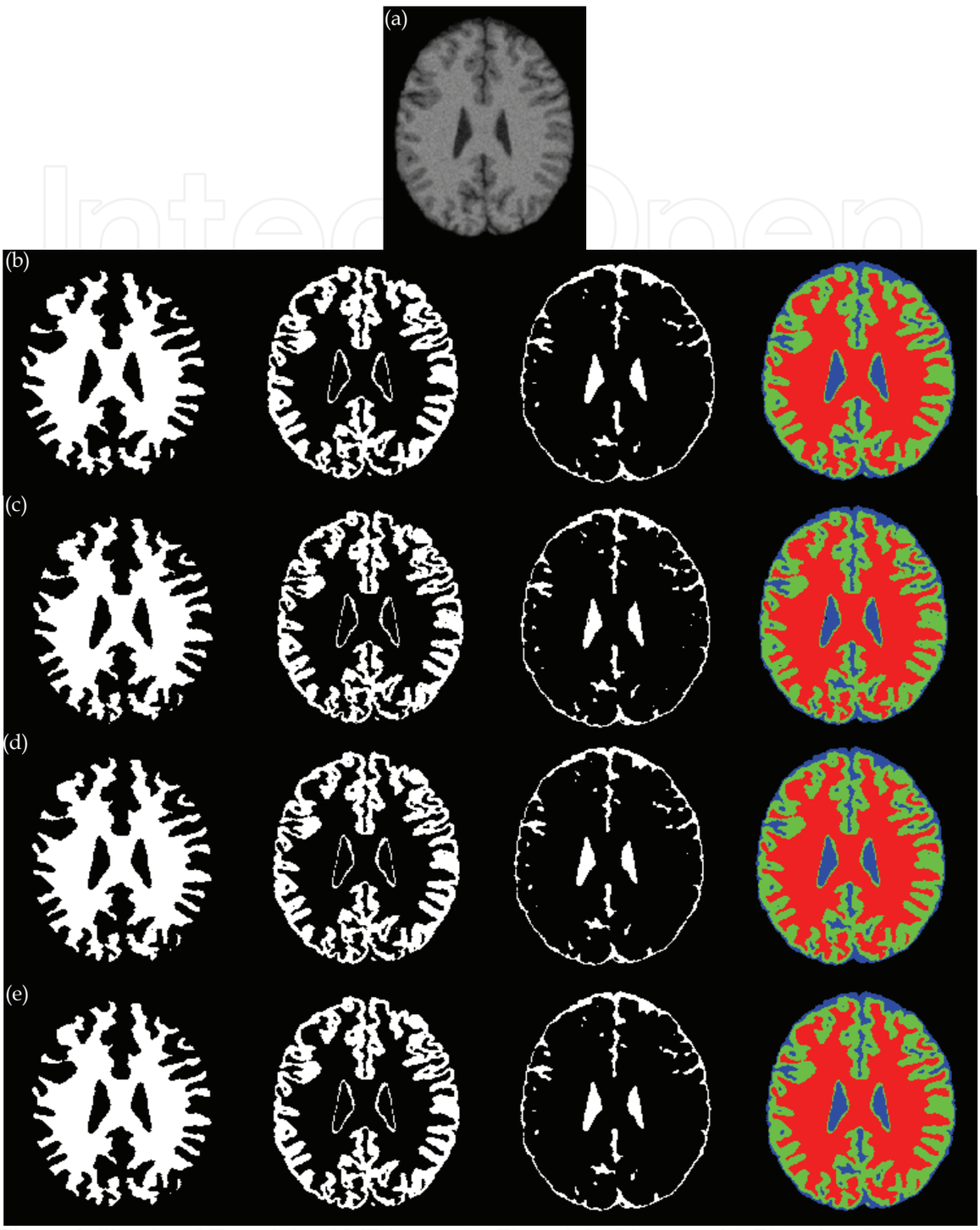


Figure 3. Simulated T_1-weighted MR image. (a) The original image with 7% noise, (b) Discrete anatomical model (from left to right) white matter, gray matter, CSF, and the total segmentation, (c) Segmentation result of ANN-IFCM, (d) Segmentation result of GAs-IFCM, (e) Segmentation result of PSO-IFCM

class	Evaluation parameters	FCM	PCM	RFCM	ANN-IFCM	GAs-IFCM	PSO-IFCM
CSF	UnS(%)	0.50	0	0.47	0.20	0.16	0.11
	OvS(%)	7.98	100	7.98	6.82	5.91	4.36
	InC(%)	0.76	34.0	0.73	0.57	0.45	0.31
White matter	UnS(%)	1.35	0	1.11	0.95	0.91	0.78
	OvS(%)	11.08	100	10.92	7.31	7.02	5.56
	InC(%)	2.33	10.16	2.11	1.59	1.39	1.06
Gray matter	UnS(%)	0.75	15.86	0.76	0.54	0.48	0.29
	OvS(%)	7.23	0	5.72	2.65	2.61	2.13
	InC(%)	1.68	13.57	1.47	0.93	0.87	0.71
Average	UnS(%)	0.87	5.29	0.78	0.56	0.52	0.39
	OvS(%)	8.76	66.67	8.21	5.59	5.18	4.02
	InC(%)	1.59	19.24	1.44	1.03	0.90	0.69

Table 2. Segmentation evaluation on simulated T_1-weighted MR

class	Evaluation parameters	FCM	ANN-IFCM	GAs-IFCM	PSO-IFCM
CFS	UnS(%)	11.1732	11.1142	10.6406	10.1619
	OvS(%)	44.4444	45.1356	41.4939	40.9091
	InC(%)	12.4009	12.7177	11.7791	11.2965
	SI(%)	87.5991	87.6305	88.2209	88.7035
White matter	UnS(%)	3.3556	2.7622	0.9783	1.5532
	OvS(%)	14.8345	9.6177	3.1523	9.0279
	InC(%)	6.2951	4.5178	1.5350	3.4673
	SI(%)	93.7049	95.4822	98.4650	96.5327
Gray matter	UnS(%)	5.9200	3.8073	3.8469	3.5824
	OvS(%)	36.9655	35.9035	31.4066	30.5603
	InC(%)	16.9305	15.1905	13.6211	13.1503
	SI(%)	83.0695	87.6305	86.3789	86.8497
Average	UnS(%)	5.7635	5.1014	4.5606	4.4094
	OvS(%)	24.2163	22.7491	20.7562	20.2043
	InC(%)	9.3831	8.4900	7.6465	7.3856
	SI(%)	90.6169	91.5100	92.3535	92.6144

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

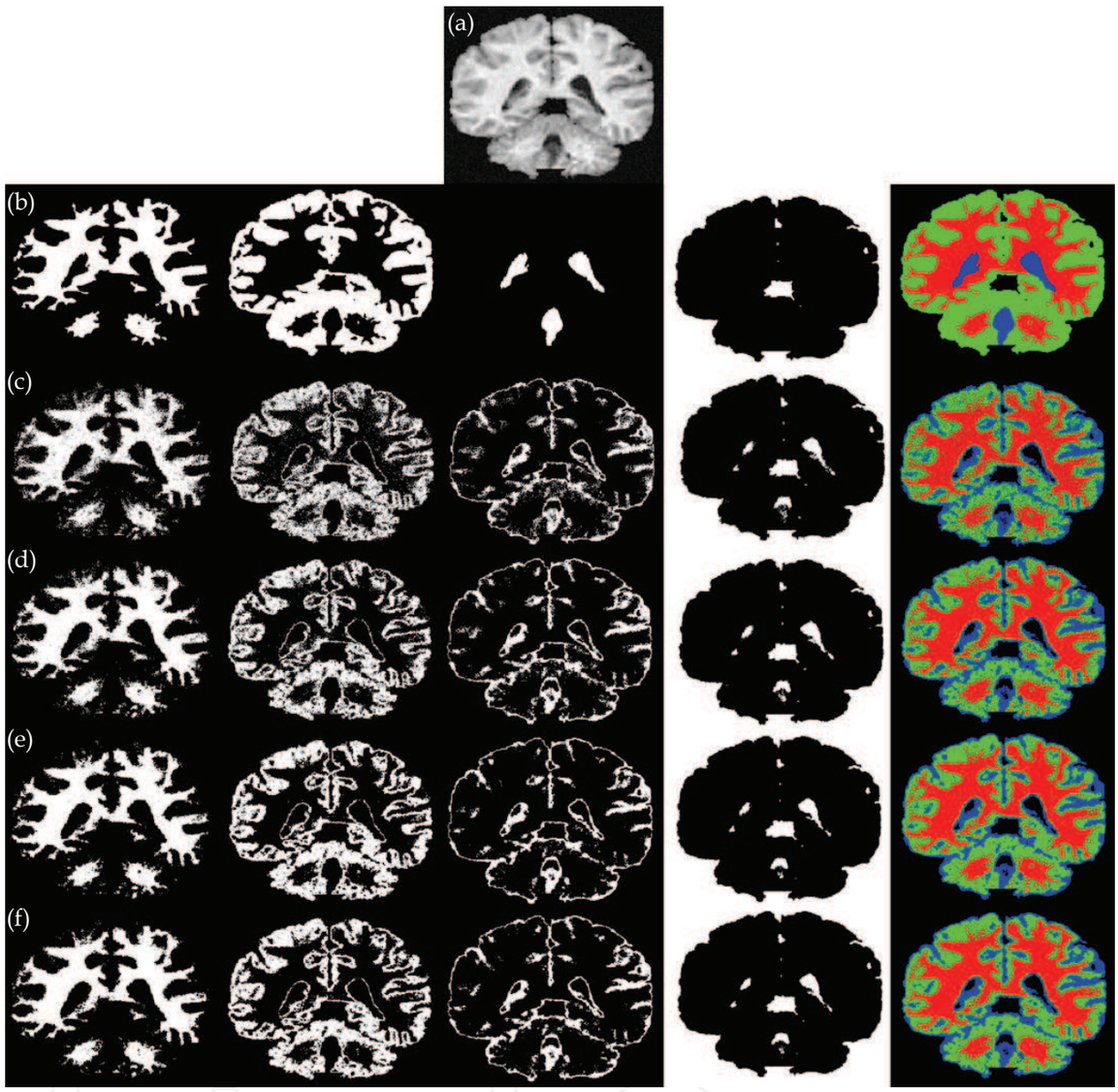


Figure 4. Real T_1-weighted MR image. (a) The original image with 5% noise, (b) Discrete anatomical model (from left to right) white matter, gray matter, CSF, others, and the total segmentation, (c) Segmentation results of FCM, (d) Segmentation result of ANN-IFCM, (e) Segmentation result of GAs-IFCM,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

5. Conclusion and Future Work

Brain MRI segmentation is becoming an increasingly important image processing step in many applications including automatic or semiautomatic delineation of areas to be treated prior to radiosurgery, delineation of tumors before and after surgical or radiosurgical intervention for response assessment, and tissue classification. A traditional approach to segmentation of MR images is the 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 the degree of attraction is a challenging task which can considerably affect the segmentation results.

In this context, we introduced new optimized IFCM-based algorithms for segmentation of noisy brain MR images. We utilized 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 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, best segmentation results were achieved using the PSO algorithm. These results nominate the PSO-IFCM algorithm as a good technique for segmentation of noisy brain MR images. It is expected that a hybrid method combining the strengths of PSO with GAs, simultaneously, would result to significant improvements that will be addressed in a future work.

6. Acknowledgement

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

7. References

Ahmed, M.N.; Yamany, S.M.; Mohamed, N.; Farag, A.A. & Moriarty, T. (2002). A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data, *IEEE Trans. Med. Imag*., vol. 21, pp. 193–199, Mar. 2002

Angeline, P.J. (1998). Evolutionary Optimization Versus Particle Swarm Optimization: Philosophy and Performance Differences, *Evolutionary Programming VII, Lecture Notes in Computer Science 1447*, pp. 601-610, Springer, 1998

Bezdek, J.C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981

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*, vol. 47, pp. 47-51, Jan. 1990

Canny, J.A. (1986). Computational approach to edge detection, *IEEE Trans. Patt. Anal. Machine Intell*., vol. 8, pp. 679-698, 1986

Clarke, L.P. et al. (1995). MRI segmentation: Methods and applications, *Magn. Reson. Imag*., vol. 13, pp. 343-368, 1995

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

Dave, R.N. (1991). Characterization and detection of noise in clustering, Pattern Recogn. Lett., vol. 12, pp. 657–664, 1991

Dubes, R. & Jain, A. (1988). *Algorithms That Cluster Data*, Prentice Hall, Englewood Cliffs, 1988

Eberhart, R.C.; Dobbins R.W. & Simpson, P. (1996). *Computation Intelligence PC Tools*, Academic Press, 1996

Eberhart R.C. & Shi, Y. (1998). Comparison between Genetic Algorithms and Particle Swarm Optimization, *Evolutionary Programming VII*, *Lecture Notes in Computer Science 1447*, pp. 611-616, Springer, 1998

Engelbrecht, A.P. (2006). *Fundamentals of Computational Swarm Intelligence*, John Wiley, 2006.

Gen, M. & Cheng, R. (1997). *Genetic Algorithms and Engineer Design*, John Wiley, 1997

Girolami, M. (2002). Mercer kernel-based clustering in feature space. *IEEE Trans. Neural Networks,* vol. 13, pp. 780-784, 2002

Haacke, E. M.; Brown, R.W.; Thompson, M.L. & Venkatesan, R. (1999). *Magnetic Resonance Imaging: Physical Principles and Sequence Design*, John Wiley, 1999

Haykin, S. (1998). *Neural Networks: A Comprehensive Foundation,* Prentice Hall, 1998

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

Hur, A.B.; Horn, D.; Siegelmann, H.T. & Vapnik, V. (2001). Support vector clustering. *J. Mach. Learn. Res.,* vol. 2, pp. 125—37, 2001

Kannan, S.R. (2008). A new segmentation system for brain MR images based on fuzzy techniques, *Applied Soft Computing*, in press

Kennedy, J. & Eberhart, R.C. (1995). Particle Swarm Optimisation, *Proceedings of IEEE International Conference on Neural Networks, IV*, pp. 1942-1948, 1995

Kennedy, J. & Eberhart, R.C. (2001). *Swarm Intelligence*, Morgan Kaufman Publishers ,2001

Krishnapuram R.R. & Keller, J.M. (1993). A possibilistic approach to clustering, *IEEE Trans. Fuzzy Syst.*, vol. 1, pp. 98–110, May 1993

Lauric, A. & Frisken, S. (2007). Soft Segmentation of CT Brain Data, Tufts University, Medford, MA, USA, Tech. Rep. 2007

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.*, vol. 12, pp. 740–750, Apr. 1993

Macovski, A. (1983). *Medical Imaging Systems*, Prentice-Hall, New Jersey, 1983

Mitchel, M. (1999). *An Introduction to Genetic Algorithms*, MIT Press, 1999

Pham, D.L. (2001). Spatial models for fuzzy clustering, *Comput. Vis. Imag. Understand.*, vol. 84, pp. 285–297, 2001

Pohle, R. & Toennies, K.D. (2001). Segmentation of medical images using adaptive region growing, *Proc. SPIE Med. Imag.*, vol. 4322, pp. 1337–1346, 2001

Shen, S.; Sandham, W.; Granat, M. & Sterr, A. (2005). MRI fuzzuy segmentation of brain tissue using neighborhood attraction with neural-network optimization, *IEEE Trans. Information Technology in Biomedicine*, vol. 9; pp 459-467, Sep. 2005

Slone, R.M.; Fisher, A.J.; Pickhardt, P.J.; Guitierrez, F. & Balfe D.M. (1999). *Body CT : a practical approach*, McGraw-Hill, Health Professions Division, New York, 1983

Suzuki, H. & Toriwaki, J. (1992). Automatic segmentation of head MRI images by knowledge guided thresholding, *Comput. Med. Imag. Graph.*, vol. 15, pp. 233–240, 1991

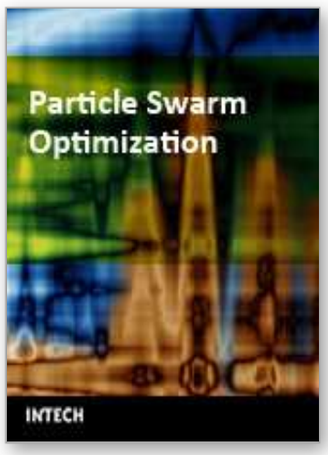
Wells, W.; Kikinis, R.; Grimson, F. & Jolesz, EA. (1994). Statistical intensity correction and segmentation of magnetic resonance image data, *Proc. SPIE Vis. Biomed. Comput.,* 1994

Wells, W.M., III; Grimson, W.E.L.; Kikinis, R. & Jolesz, F.A. (1996). Adaptive segmentation of MRI data, *IEEE Trans. Med. Imag.*, vol. 15, pp. 429–442, Aug. 1996

Zadeh, L.A. (1965). Fuzzy sets, *Inform. Control*, vol. 8, pp. 338–353, Jun. 1965

Zhang, D.Q. & Chen, S.C. (2002). Fuzzy clustering using kernel methods. *Proceedings of the International Conference on Control and Automation*, Xiamen, China, June, 2002

Zhang, D.Q. & Chen, S.C. (2004). A novel kernelized fuzzy c-means algorithm with application in medical image segmentation, *Artif. Intell. Med.*, vol. 32, pp. 37–52, 2004



Particle Swarm Optimization

Edited by Aleksandar Lazinica

ISBN 978-953-7619-48-0
Hard cover, 476 pages
Publisher InTech
Published online 01, January, 2009
Published in print edition January, 2009

Particle swarm optimization (PSO) is a population based stochastic optimization technique influenced by the social behavior of bird flocking or fish schooling.PSO shares many similarities with evolutionary computation techniques such as Genetic Algorithms (GA). The system is initialized with a population of random solutions and searches for optima by updating generations. However, unlike GA, PSO has no evolution operators such as crossover and mutation. In PSO, the potential solutions, called particles, fly through the problem space by following the current optimum particles. This book represents the contributions of the top researchers in this field and will serve as a valuable tool for professionals in this interdisciplinary field.

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen