Image Clustering Using Improved Particle Swarm Optimization

Thuy Xuan Pham(✉), Patrick Siarry, and Hamouche Oulhadj

Laboratory Images, Signals, and Intelligent Systems, University Paris-Est Creteil, 94400 Vitry sur Seine, France
thuy.phxuan@gmail.com, {siarry,oulhadj}@u-pec.fr

Abstract. In this paper, we propose an improvement method for image segmentation problem using particle swarm optimization (PSO) with a new objective function based on kernelization of improved fuzzy entropy clustering algorithm with spatial local information, called PSO-KFECS. The main objective of our proposed algorithm is to segment accurately images by utilizing the state-of-the-art development of PSO in optimization with a novel fitness function. The proposed PSO-KFECS was evaluated on several benchmark test images including synthetic images (http://pages.upf.pf/Sebastien.Chabrier/ressources.php), and simulated brain MRI images from the McConnell Brain Imaging Center (BrainWeb (http://brainweb.bic.mni.mcgill.ca/brainweb/)). Experimental results show that our proposed PSO-KFECS algorithm can perform better than the competing algorithms.

Keywords: Image segmentation · Particle swarm optimization
Entropy based fuzzy clustering · Fitness function

1 Introduction

Image segmentation is one of the most important and challenging problems in image analysis. Image segmentation is defined as the partitioning of an image into non-overlapped consistent regions or objects that satisfy pre-defined conditions. The level of details depends on the problem being solved or criteria being reached [1]. Image segmentation can be widely applied in many fields, such as medical imaging [2], data mining [3], pattern recognition [4], computer vision [5], etc. Due to the complexity of image content and the variety of image sizes, characteristics and other properties, there is no global solution today able to solve completely the image segmentation problem. Up to now, various methods for image segmentation have been proposed [6–9]. They can be divided into four groups: region based image segmentation [10], clustering based image segmentation [11], thresholding based image segmentation [12], and edge based image segmentation [13]. In this work, we are focused on the clustering based approach, using the state-of-the-art particle swarm algorithm equipped with a novel objective function.

Clustering is considered as the process of dividing data points into homogeneous classes or clusters, so the samples are similar within each group. Many clustering methods have been proposed in the literature [14, 15]. Among these clustering

Y. Chen and T. Q. Duong (Eds.): INISCOM 2017, LNICST 221, pp. 359–373, 2018.
https://doi.org/10.1007/978-3-319-74176-5_31

approaches, fuzzy entropy clustering (FEC) algorithm proposed by Tran and Wagner [16] and fuzzy C-means (FCM) algorithm proposed by Bezdek et al. [17], which are based on minimization of an objective function, are widely used in data clustering and image segmentation for their simplicity and applicability. However, the major drawbacks of both methods are: (i) they are very sensitive to noise and imaging artifacts, since no local spatial information in the image is considered; (ii) the results provided depend not only on the choice of the initial clustering centroids but also on the high vulnerability of the algorithms to be trapped at the local minima. Many approaches have been proposed by researchers to overcome these drawbacks.

To overcome the first problem, many enhanced versions have been proposed by introducing local spatial information to their objective functions. It can be done by adding penalty terms, modifying elements used weighted concepts, using kernel metric, or combining with other methods. For instance, Cheng and Zhang [18] have proposed two algorithms, called KFCM_S1 and KFCM_S2, in which a neighborhood term and the kernel-induced distance replaced the Euclidean distance have been incorporated into the objective function. Krinidis and Chatzis [19] proposed a fuzzy local information c-means (FLICM) algorithm that added a penalty term using both spatial and gray-level local information with free of parameter tuning. Verma et al. [20] proposed an improved fuzzy entropy clustering (IFEC) algorithm. In the IFEC algorithm, the Euclidean distance is weighted by a new fuzzy factor, which determines both local gray-level and spatial information.

To address the second problem, many researchers proposed to use metaheuristic algorithms to perform clustering and image segmentation so as to achieve optimized performance. For example, Maulik [21], Ouadfel and Batouche [22], Hassanzadeh et al. [23], Nandy et al. [24], Omran et al. [11] proposed methods based on genetic algorithm (GA), ant colony optimization (ACO), firefly algorithm (FA), cuckoo search (CS), and particle swarm optimization (PSO), respectively, for image segmentation problem. PSO has become one of the most popular metaheuristics and has been used for clustering problems widely [25] due to simplicity and versatility. Using PSO for image segmentation problem can be in standalone form [11] or combination with other methods, such as FCM [26]. However, in PSO-based clustering, to provide good quality of image clustering, one needs both to tune a range of parameters and to design of good fitness function. Concerning these problems, Filho et al. [27] proposed a method that utilized the state-of-the-art PSO algorithm proposed by Zhang et al. [28] to improve the results of clustering data. Orman et al. [11], Wong et al. [29] proposed new fitness functions for image segmentation. Experimental results showed that their proposed methods can provide better performance when compared to classical methods.

In this work, we introduce a new method for image segmentation problem. The method uses low-discrepancy sequence initialized particle swarm optimization with high-order nonlinear time-varying inertia weight (LHNPSO) algorithm proposed by Yang et al. [30] with a new fitness function. The reason for choosing LHNPSO lies in the ability of this variant to outperform the other variants of PSO with two main advantages, namely, (i) the method can converge faster and give a much more accurate final solution [31], (ii) it is very easy to be implemented. The new fitness function takes advantages from kernelized fuzzy clustering approach; in addition, it also takes into account the spatial local information. Thus, the proposed method may tackle simultaneously the two main problems mentioned above. Experiments using synthetic images, simulated Magnetic Resonance Imaging (MRI) images are reported and the results are compared with recent fuzzy clustering and PSO-based image clustering methods.

The rest of the paper is organized as follows. Section 2 presents the related algorithms on which this work is based. Section 3 introduces the proposed method: LHNPSO algorithm with a new fitness function based on fuzzy entropy clustering for image clustering. The performance of the proposed method is evaluated and the comparison with a set of algorithms from the literature is reported in Sect. 4. Finally, in Sect. 5, we make conclusions.

2 Related Works

The algorithms on which this work is based are described in this section. They are divided in two groups: fuzzy entropy clustering algorithms and particle swarm optimization.

2.1 Fuzzy Entropy Clustering

The FEC algorithm proposed by Tran and Wagner [16] is an alternative generalization of hard c-means (HCM) clustering algorithm, which takes the advantage of fuzzy entropy. For a set of N data patterns $\mathbf{X} = \{x_1, x_2, \ldots, x_N\}$, the algorithm allows partitioning the data space, by minimizing an objective function J_{FEC} with respect to the membership matrix $\mathbf{U} = \{u_{ij}\}$ and the set C of cluster prototypes $\mathbf{C} = \{c_i\}$:

$$J_{FEC}(\mathbf{C}, \mathbf{X}, \mathbf{U}) = \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij} d^2(x_j, c_i) + n \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij} \log u_{ij} \tag{1}$$

Subject to:

$$\begin{array}{ll} 0 \le u_{ij} \le 1 & i = 1, \ldots, C \quad j = 1, \ldots, N \\ \sum_{i=1}^{C} u_{ij} = 1 \quad \forall j & 0 < \sum_{j=1}^{N} u_{ij} < N \quad \forall i \end{array} \tag{2}$$

Where x_j, a data pattern, is j^{th} pixel in the image, which can be intensity value or gray value; $d^2(x_j, c_i)$ is the Euclidean distance between the pixel x_j and centroid c_i; and n is the degree of fuzzy entropy.

The FEC objective function, J_{FEC}, can be minimized by iteratively using the following update equations:

$$u_{ij} = \left\{ \sum_{r=1}^{C} \left[\exp\left(d^2(x_j, c_i) - d^2(x_j, c_r) \right) \right]^{1/n} \right\}^{-1} \tag{3}$$

$$c_i = \frac{\sum_{j=1}^{N} u_{ij} \cdot x_j}{\sum_{j=1}^{N} u_{ij}} \tag{4}$$

The FEC algorithm can uniformly be summarized in the following steps:

Algorithm 1. FEC

Input: Fix the number C of centroids; randomly initialize $\mathbf{U} = \{u_{ij}\}$ satisfying (2); set terminate criterion ε and compute the centers $\mathbf{C} = \{c_i\}$ using **U** and equation (4)
Repeat
  Update the partition matrix using (3)
  Update the centroids using (4)
Until $\max\{||\mathbf{U}_{new} - \mathbf{U}_{old}||\} < \varepsilon$
Return the cluster centroids **C** and partition matrix **U**

After FEC clustering, each pixel will be associated with a membership value for each cluster. To get the image segmentation, the simplest way is assigning pixels to the clusters with the maximum membership values.

From Eq. (1), it is clear that the objective function of FEC does not take into account any local spatial information. Also, the metric is the Euclidean distance, $d^2(x_j, c_i)$, which measures the similarity between the pixel intensity and the cluster centroids. This metric assumes that each feature of data points is equally important and independent from others. This assumption may not be always satisfied in real applications [26]. We make the necessary corrections to this problem as explained in Sect. 2.2.

2.2 Kernelized Fuzzy Entropy Clustering with Spatial Information

2.2.1 Kernel Representation

Recently, ‘kernel method’ was proposed and was proved successful for various applications, such as in pattern recognition and function approximation. The idea of the method of kernel functions (called kernel trick) uses a ‘nonlinear’ transformation $\Phi : X^p \to H$, where X^p is input data space with low dimension and H is a high dimensional feature space. The structure of input data points may be inadequate for the analysis in the original space, but they are capable of analysis in space H [32]. In H, the Euclidean distance $||x_j - c_i||^2$ is substituted by $||\Phi(x_j) - \Phi(c_i)||^2$, which is defined by using kernel function K as

$$\|\Phi(x_j) - \Phi(c_i)\| = K(x_j, x_j) + K(c_i, c_i) - 2K(x_j, c_i) \tag{5}$$

Where $K(x_j, c_i)$ is a kernel function, which calculates the inner product. In this paper, Gaussian kernel is adopted as

$$K(x_j, c_i) = \langle \Phi(x_j), \Phi(c_i) \rangle = \exp\left(-\|x_j - c_i\|^2 / \sigma^2\right) \tag{6}$$

It is clear that the parameter σ of Gaussian kernel can greatly affect the result of kernel methods. For instance, when σ tends to zero, $K(x_j, c_i)$ turns into an impulse function with the value of 1 only at $x_j = c_i$ and 0 elsewhere. In this case, any two points in the feature space have a common value close to 1, in other words, it will be very hard to cluster. On the other hand, when σ towards to infinity, any two points in the feature space toward to zero, which makes them difficult to separate. In this work, the estimation of σ^2 is based on Yang's work [33]. It is defined as:

$$\sigma^2 = \lambda \sum_{j=1}^{N} \left\|x_j - \bar{x}\right\|^2 / N, \ \bar{x} = \sum_{j=1}^{N} x_j \tag{7}$$

λ is a constant defined by experiments.

2.2.2 Kernel Fuzzy Entropy Clustering with Spatial Information Algorithm

In this section, a new objective function based on fuzzy entropy clustering algorithm and kernel method is proposed. The objective function is given as follows:

$$\begin{aligned} J_{KFEC\_S}(\mathbf{C}, \mathbf{U}, \mathbf{X}) = \sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}\left(1 - K\left(x_j, c_i\right)\right) + n\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}\log u_{ij} \\ + \eta\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}\left(1 - K\left(\bar{x}_j, c_i\right)\right) \end{aligned} \tag{8}$$

Where the constraints in Eq. (2) must be satisfied. The third term is the spatial constraint term, in which the parameter η controls the penalty effect of the spatial constraint. The median of the neighbors within a window around the pixel x_j is used to represent $\bar{x}_j$.

By using the Lagrange multiplier method, the necessary conditions for the minimization of the objective function in Eq. (8), with the constraints in Eq. (2), can be found. Specifically, taking the first derivatives of J_{KFEC_S} with respect to u_{ij} and c_i, and zeroing them, respectively, two necessary but not sufficient conditions for J_{KFEC_S} to be at its local optimal solution will be obtained as follows:

$$u_{ij}^{-1} = \sum_{r=1}^{C} \frac{\exp\left(n^{-1}\cdot\left(\left(1 - K\left(x_j, c_i\right)\right) + \eta\cdot\left(1 - K\left(\bar{x}_j, c_i\right)\right)\right)\right)}{\exp\left(n^{-1}\cdot\left(\left(1 - K\left(x_j, c_r\right)\right) + \eta\cdot\left(1 - K\left(\bar{x}_j, c_r\right)\right)\right)\right)} \tag{9}$$

$$c_i = \frac{\sum_{j=1}^{N} u_{ij}\cdot\left(K\left(x_j, c_i\right)\cdot x_j + \eta\cdot K\left(\bar{x}_j, c_i\right)\cdot \bar{x}_j\right)}{\sum_{j=1}^{N} u_{ij}\cdot\left(K\left(x_j, c_i\right) + \eta\cdot K\left(\bar{x}_j, c_i\right)\right)} \tag{10}$$

The proposed kernelized fuzzy entropy clustering with spatial local information (KFEC_S) is described in Algorithm 2.

Algorithm 2. KFEC_S

Input: Fix the number C of centroids; randomly initialize $\mathbf{U} = \{u_{ij}\}$ satisfying (2); set terminate criterion ε, the maximum number of iterations T, the iterative index k to 1, and the parameters: n, λ, η; compute the centers $C = \{c_i\}$ using U and equation (10)
Repeat
  Update the partition matrix using (9)
  Update the centroids using (10)
  k = k + 1
Until $\max\{||\mathbf{U}_{new} - \mathbf{U}_{old}||\} < \varepsilon$ or $k > T$
Return the cluster centroids **C** and partition matrix **U**

As stated above, the Eqs. (9) and (10) are only necessary conditions for minimizing the objective function J_{KFEC_S}; as a result, obtained clustering solutions may be local optima. This is also the drawback of fuzzy clustering and its variants when solving image segmentation problem. To overcome this drawback, we use a metaheuristic optimization-based approach, as explained in the Sect. 2.3.

2.3 Particle Swarm Optimization

PSO is a population based stochastic optimization technique and is regarded as global search strategy, which was originally introduced by Kennedy and Eberhart [34]. In PSO, each individual of the population called a particle represents a potential solution to the optimization problem; and the population of individuals (P) or swarm is evolved through successive iterations. The quality of a candidate solution is evaluated by the fitness value, associated to each particle. Each particle, denoted i, has a position vector ($\mathbf{X}_i = \{x_{ir}\}$), a velocity vector ($\mathbf{V}_i = \{v_{ir}\}$), its own best position (P_{best}) found so far, and interacts with neighboring particles through the best position (G_{best}) discovered in the neighborhood so far. At each iteration k, each particle is moved according to Eq. 11 [35]:

$$V_i(k+1) = wV_i(k) + c_1 r_1 (P_{best}(k) - X_i(k)) + c_2 r_2 (G_{best}(k) - X_i(k)) \tag{11}$$

$$X_i(k+1) = X_i(k) + V_i(k+1) \tag{12}$$

Where, c_1 and c_2 are acceleration coefficients that scale the influence of the 'cognitive' and 'social' components; r_1 and r_2 are two random values, uniformly distributed in [0, 1]; w is inertia weight. The higher the w is, the higher the ability of searching in the global solution space is, and the smaller w is, the higher the ability of searching for the local solutions.

There are two basic criteria for assessing the performance of PSO algorithm, namely, convergence speed and ability to find global optima. To achieve these goals, the balance between global exploration and local exploitation is crucial. From Eqs. (11) and (12), it is clear that the performance is not only dependent on the controlling parameters (w, c_1, c_2), but also dependent on the size and the structure of neighborhood. In this work, initializing the population and tuning parameters (w, c_1, c_2) are adopted from Yang et al. [30], who designed LHNPSO algorithm. In LHNPSO, the initial population of particles is generated by using the Halton sequence to cover the search space efficiently, (c_1, c_2) are set to constants, equal to 2, and inertia weight is updated as follows:

$$w(k+1) = w_{\max} - (w_{\max} - w_{\min})\left(\frac{k}{k_{\max}}\right)^{1/\pi^2} \tag{13}$$

Where (w_{min}, w_{max}) is the range of inertia weight, with $w_{min} = 0.4$ and $w_{max} = 0.9$, k and k_{max} are the iteration number starting from iteration one and the maximum number of allowable iterations, respectively. This law of variation of w increases the exploration of the search space in the first iterations of the algorithm, and the exploitation of the best solutions found so far towards the end of the algorithm. This constitutes a very good balance between the phases of exploration of the search space and the phases of exploitation of the solutions.

The procedures of LHNPSO algorithm can be summarized as:

Algorithm 3. LHNPSO

Input: Set k = 1; initialize particles using the Halton sequence; set 'cognitive' and 'social' coefficients and the range of inertia weight
Repeat
    Calculate fitness value
    Determine the personal best (P_{best}) and the global best (G_{best})
    Update the inertia weight according to (13)
    Calculate the velocity and position according to (11) and (12)
Until the stopping criteria are met or the number of iterations is $k > k_{max}$
Return the optimal solution (G_{best})

The performance of LHNPSO is only tested with the unconstrained optimization problems. For more complex problems, the need of further development is still required.

3 The Proposed Algorithm, Kernelized Fuzzy Entropy Clustering with Spatial Information Using LHNPSO

As discussed in Sect. 2, the kernelized fuzzy entropy clustering with spatial information model is a complex nonlinear model. The KFEC_S algorithm can be trapped into local minima because local search method (gradient method) is used to solve the model. By contrast, the algorithm has the advantage of converging quickly. On the other hand, LHNPSO is a global optimization method, which can provide global optimum solution, but with a relatively longer convergence time. So, we use LHNPSO with minor improvements to optimize KFEC_S, trying to get the global optimum solution by escaping the trap of local optimums, and we use at the same time KFEC_S to guide the LHNPSO research process in order to converge more quickly and anticipate more accurate solution. Thus, a new image clustering algorithm based on KFEC_S and LHNPSO, named PSO-KFECS, is proposed. The details of this algorithm are given in the next section.

3.1 Particle Representations

When LHNPSO is applied to optimize the kernelized fuzzy entropy clustering with spatial information objective function J_{KFEC_S}, cluster prototypes **C** = $\{c_i\}_C$ are chosen to be optimization variables and encoded as positions of particles. For P particles or solutions, there are in total C * P optimization variables needing to be encoded. The position of i^{th} particle can be described as: $\mathbf{X}_i = [x_{i1}, x_{i2}, \ldots, x_{iC}]$. Here, x_{ij} represents the j^{th} cluster centroid among C centroids of the i^{th} solution. In this way, cluster centroids are encoded in position vector $\mathbf{X}_i$ and **C** can be obtained by decoding $\mathbf{X}_i$.

3.2 Fitness Function

After decoding $\mathbf{X}_i$ to get cluster centroids **C** and calculating fuzzy partition matrix **U** according to Eq. (9), the value of the fitness function can be calculated by evaluating J_{KFEC_S} according to Eq. (8).

$$f_i = J_{KFEC\_S} \tag{14}$$

Minimization of f_i is the same as minimization of the objective function J_{KFEC_S}, which leads to the optimal partitioning of the image.

3.3 The Proposed Algorithm

The proposed algorithm, PSO-KFECS, takes both advantage of the excellent feature of LHNPSO, in the aspect of optimizing the objective function of kernelized fuzzy entropy clustering with spatial information and the kernel, and the KFEC_S gradient method, in aspect of speeding up convergence.

To make sure that all particles are moving within the search space and avoiding divergent behavior, the position and the velocity are limited as follows:

$$v_{ij}(k+1) = \begin{cases} rand() \cdot v_{\max} & if \quad v_{ij}(k+1) > v_{\max} \\ -rand() \cdot v_{\max} & if \quad v_{ij}(k+1) < -v_{\max} \\ v_{ij}(k+1) & otherwise \end{cases} \tag{15}$$

$$x_{ij}(k+1) = \begin{cases} rand() \cdot \frac{1}{2} \cdot (x_{\max} - x_{\min}) & if \quad x < x_{\min} \quad or \quad x > x_{\max} \\ x_{ij}(k+1) & otherwise \end{cases} \tag{16}$$

Where v_{max} is the largest allowable step size in any dimension; and, $[x_{min}, x_{max}]$ are the bounds of the search space in each dimension. In image clustering, commonly, v_{max} is set to 5 and $[x_{min}, x_{max}]$ are the minimum and maximum of the feature (intensity or gray value) of the image.

Putting all discussions above together, the PSO-KFECS algorithm is shown as follows:

Algorithm 4. PSO-KFECS

Input: Set $k = 1$; Initialize all parameters for LHNPSO algorithm ($c_1, c_2, k_{max}, P, \ldots$); Initialize all parameters for the fitness function in Eq. (8) (C, n, η); Set constraints for the position [x_{min}, x_{max}] and the velocity [$-v_{max}, v_{max}$]; Generate the initial position $\mathbf{X}_i$ using the Halton sequence, the velocity $\mathbf{V}_i$ randomly; Generate randomly the partition matrix **U** for each solution that satisfies Eq. (2); Evaluate f_i according to Eq. (8); Initialize P_{best} with a copy of X_i; and $G_{best} = P_{best}$ while f_s is the minimum
Repeat
  For each particle **do**
    Calculate kernel distance K using Eq. (6)
    Calculate the partition matrix **U** using Eq. (9)
    Calculate the fitness function f_i using Eq. (8)
  Update the P_{best}, G_{best}
  Update the inertia weight w using Eq. (13)
  Update the position $\mathbf{X}_i$ and the velocity $\mathbf{V}_i$ using Eq. (11-12)
  Map position and velocity into search space using Eq. (15-16)
Until the stopping criteria are met or the number of iterations is $k > k_{max}$
Return the cluster centroids **C** and the partition matrix **U**

In this work, the number of non-significant improvements of the fitness function and the maximum number of iterations are used as the stopping criteria of the algorithm. Thus, if the condition $\left\|f_i^{(k+1)} - f_i^{(k)}\right\| < \varepsilon$ is completed k_{stop} times or the condition $k > k_{max}$ is reached, the algorithm is immediately stopped.

4 Experimental Results

This section presents the results of the experimental evaluation of our algorithm. With this aim in view, our algorithm is compared with those from five well-known image clustering algorithms in the literature: FEC [16], KFCM_S2 [18], FLICM [19], PSO-based image clustering algorithms proposed by Orman et al. [11] (PSO_V1) and Wong et al. [29] (PSO_V2). Two image datasets: synthetic images [36], simulated MRI brain images from BrainWeb [37], with different numbers of cluster centroids and levels of corrupting noises, are used to evaluate the performance of the proposed algorithm. All the algorithms are implemented in MATLAB 2014b and executed with a computer with Intel Core i3 1.5 GHz CPU, 4G RAM under Microsoft Windows 7. The values of the parameters of our method are based on conclusions in the relative literature and try-and-error technique. Parameter settings are given in Table 1.

4.1 Quantitative Evaluation

In order to compare the results of different segmentation methods, supervised evaluation methods are used. In supervised evaluation methods, we use Jaccard index [38] and Hausdorff distance [39], because current research [40] reports that they are suitable

Table 1. Setting of specific parameters in PSO-KFECS

Parameters	Values
Population size	40
$c_1 = c_2$	2.0
Degree of fuzzy entropy, n	20
The spatial constraint parameter, η	2.6
The controlling Gaussian kernel parameter, λ	1/π
The number of non-significant improvement, k_{stop}	10
The terminate criterion parameter, ε	0.00001
Maximum number od iterations, k_{max}	100

metrics for the evaluation when there exists outliers with or without small segments, complex boundaries, low densities in the image.

4.1.1 Jaccard Index

This index is an overlap index, which directly compares the segmented image (I_s) and the ground truth image (I_t) by measuring similarity between them. The higher value indicates the better result. Jaccard index is defined as the intersection between segmentations divided by the size of their union, that is

$$JAC(I_s, I_t) = \frac{TP}{TP + FP + FN} \tag{17}$$

Where *TP*, *FP*, *FN* are basic cardinalities of the confusion matrix, namely, the true positives, the false positives, and the false negatives, respectively.

4.1.2 Hausdorff Distance

This is a spatial distance based metric, which measures the dissimilarity between the segmented image and the ground truth image. The lower value indicates the better result. It is defined as follows:

$$HD(I_s, I_t) = \max(h(I_s, I_t), h(I_t, I_s)) \tag{18}$$

Where $h(I_s, I_t)$ is called the directed Hausdorff distance and given by:

$$h(I_s, I_t) = \max_{p_s \in I_s} \min_{p_t \in I_t} \|p_s - p_t\| \tag{19}$$

Where $\|p_s - p_t\|$ is the Euclidean distance between the gray values of pixel p_s in the segmented image and pixel p_t in the ground truth image.

4.2 Experiments on Synthetic Images

To compare the sensitivity of the algorithms to noise, we apply them to 256 × 256 synthetic images [36], which have 4 or 5 regions and corrupted with salt and pepper or

Gaussian noises with different levels. For the salt and pepper noise, the variance of the noise varies from 0.04 to 0.1 with a step of 0.02; for Gaussian noise, the noise variances are in range of [0.004, 0.01] with steps of 0.002.

Figure 1 describes an example of the segmentation results of a synthetic image, including 5 regions corresponding to gray values taken as 1, 61, 121, 181, and 241, corrupted with 0.1 variance of salt and pepper and 0.01 Gaussian noise. From Fig. 1, one can see, visually, that even though the six competitive methods give a coherent segmentation result, the proposed method produces better results than the others.

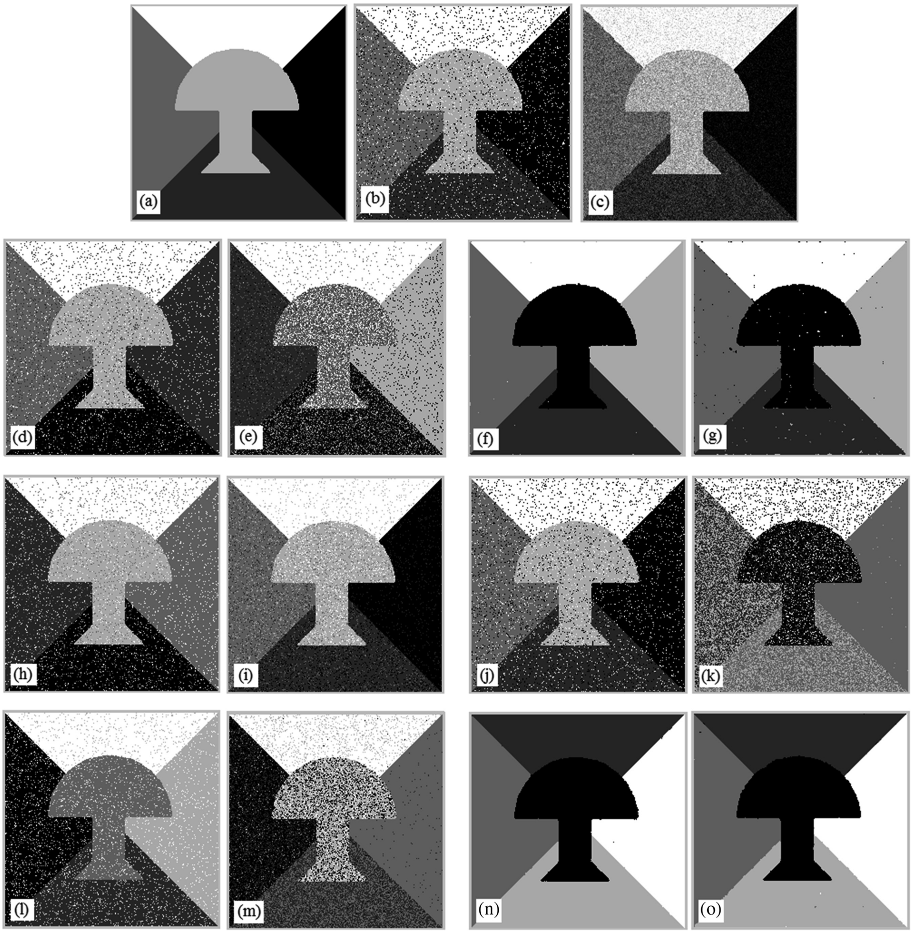

Fig. 1. Segmentation results on synthetic image corrupted by salt and pepper noise (0.1) and Gaussian noise (0.01): (a) original image; (b) salt and pepper noise image; (c) Gaussian noise image; (d, e) FEC results; (f, g) KFCM_S2 results; (h, i) FLICM results; (j, k) PSO_V1 results; (l, m) PSO_V2 results; (n, o) PSO-KFECS results.

Table 2 provides the numerical scores (Eqs. 17 and 18) of the six competitive methods, they are the average of 10 successful program runs. The table shows clearly that the proposed PSO-KFECS algorithm provides superior results. Indeed, the contours of the original images are better reconstructed and the regions are more homogeneous.

Table 2. Segmentation evaluation scores of different methods

Noise type	Metrics	FEC	KFCM_S2	FLICM	PSO_V1	PSO_V2	PSO KFECS
Salt and pepper	*JAC*	0.7318	0.9931	0.7309	0.7325	0.7300	**0.9954**
(0.1)	*HD*	4187	3462	4324	4263	4187	**1660**
Gaussian	*JAC*	0.5105	0.9746	0.6429	0.4765	0.5150	**0.9912**
(0.01)	*HD*	4296	3698	4223	4183	4200	**2980**

4.3 Experiments on Simulated MRI Images

BrainWeb provides a simulated brain database (SBD) that contains a set of realistic MRI data volumes. We have applied the proposed PSO-KFECS and 5 other algorithms on normal brain images, which have characteristics of Mobility T1, slice thickness 1 mm, with different levels of noise (0%, 1%, 3%, 5%, 7%, 9%), and different levels of non-uniformity (0%, 20%, 40%). These images were segmented with 4 cluster centroids: background, cerebral spinal fluid (CSF), gray matter (GM), and white matter (WM).

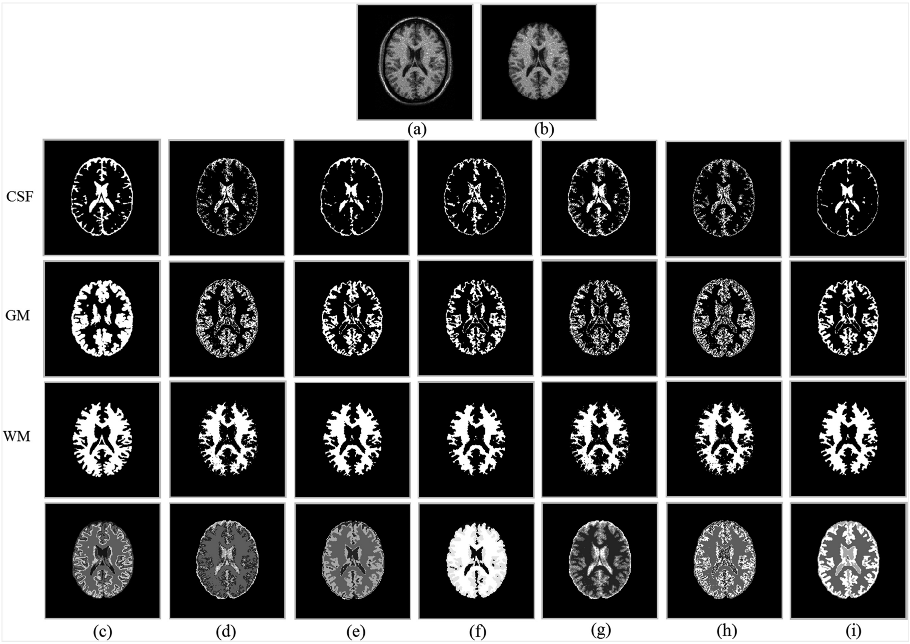

Fig. 2. Original and segmented images by different algorithms: (a) original T1-weighted axial image with 7% noise and 20% intensity of non-uniformity; (b) original image after skull stripping; (c) the ground truth images; (d) FEC results; (e) KFCM_S2 results; (f) FLICM results; (g) PSO_V1 results; (h) PSO_V2 results; (i) PSO-KFECS results.

Figure 2 shows an example of segmentation of simulated brain image with 7% of noise and 20% intensity of non-uniformity. Visually, the comparison of the performance of the same six competitive algorithms is a little bit vague. The quantitative performance evaluation is given in Table 3. In this table, it can be seen that the scores of the proposed algorithm that PSO-KFECS outperforms all competing, only FLICM reveals better for GM with Jaccard index.

Table 3. Segmentation evaluation of the simulated MRI image

Regions	Metrics	FEC	KFCM_S2	FLICM	PSO_V1	PSO_V2	PSO KFECS
CSF	*JAC*	0.8613	0.8978	0.9189	0.8707	0.8957	**0.9297**
	HD	332.6	413.6	229	304.	329.1	**212.6**
GM	*JAC*	0.8112	0.8365	**0.8645**	0.8258	0.8229	0.8549
	HD	264.9	219.4	**205**	215	249.5	**205**
WM	*JAC*	0.8746	0.8869	0.9059	0.8334	0.8584	**0.9210**
	HD	127.7	118.9	113	192.1	171.8	**107.2**
Whole image	*JAC*	0.9388	0.9326	0.9277	0.9114	0.9154	**0.9425**
	HD	263.4	**205**	**205**	259.1	209.3	**205**

5 Conclusion

In this paper, an improved particle swarm optimization with a new fitness function - the kernelized fuzzy entropy clustering with spatial information, for image segmentation - is proposed. Two drawbacks of fuzzy clustering algorithms, which are the trapping of the solution into local minima and the sensitivity to noise and imaging artifacts, have been partially overcome. The experimental results show that our method is more effective in comparison with five competitive methods of the literature. However, there is still much work for us to go further. Specifically, the determination of the parameters: n (degree of fuzzy entropy) and η (controlling parameter of the spatial information) is an open question. In addition, only one criterion (J_{KFECS}) is used to guide the search process of the solution, this may lead to the situation that the solution is a global optimum of the criterion, but it may not be the optimum of the segmentation problem. In addition to the mentioned improvements, we also intend to accelerate the convergence times of our algorithm, by implementing it on multicore computer architecture in order to parallelize the data processing. Indeed, although they are more precise, metaheuristics based algorithms have the drawback of providing more convergence times because they cover more widely the solution search space. Moreover, we plan to apply a multi-objective optimization approach, in order to take benefits of other criteria.

References

1. Gonzalez, R.C., Woods, R.: Digital Image Processing, 3rd edn. Prentice Hall, Englewood Cliffs (2008)
2. Sharma, N., Aggarwal, L.M.: Automated medical image segmentation techniques. J. Med. Phys. **35**(1), 3–14 (2010)

3. Hsu, W., Lee, M.L., Zhang, J.: Image mining: trends and developments. J. Intell. Inf. Syst. **19**(1), 7–23 (2002)
4. Lakshmi, H.C.V., PatiKulakarni, S.: Segmentation algorithm for multiple face detection in color images with skin tone regions using color spaces and edge detection techniques. IJCTE **2**(4), 1793–8210 (2010)
5. Szeliski, R.: Computer Vision: Algorithms and Applications. Springer, London (2011). https://doi.org/10.1007/978-1-84882-935-0
6. Pal, N., Pal, S.: A review on image segmentation techniques. Pattern Recognit. **26**(9), 1277–1294 (1993)
7. Pham, D.L., Xu, C.Y., Prince, J.L.: A survey of current methods in medical image segmentation. Annu. Rev. Biomed. Eng. **2**, 315–337 (2000)
8. Khan, W.: Image segmentation techniques: a survey. J. Image Graph. **1**(4), 166–170 (2013)
9. Taneja, A., Ranjan, P., Ujjlayan, A.: A performance study of image segmentation techniques. In: 4th ICRITO, pp. 1–6 (2015)
10. Freixenet, J., Muñoz, X., Raba, D., Martí, J., Cufí, X.: Yet another survey on image segmentation: region and boundary information integration. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002. LNCS, vol. 2352, pp. 408–422. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-47977-5_27
11. Orman, A.P., Engelbrecht, M., Salman, A.: Particle swarm optimization method for image clustering. Int. J. Pattern Recogn. **19**, 297–321 (2005)
12. Sezgin, M., Sankur, B.: Survey over image thresholding techniques and quantitative performance evaluation. J. Electron. Imaging **13**, 146–148 (2004)
13. Senthilkumaran, N., Rajesh, R.: A study on edge detection methods for image segmentation. In: Proceedings of the International Conferences on Mathematics and Computer Science (ICMCS-2009), vol. 1, pp. 255–259 (2009)
14. Jain, A.K., Murty, M.N., Flynn, P.J.: Data clustering: a review. ACM Comput. Surv. **31**(3), 264–323 (1999)
15. Xu, R., Wunsch, D.: Survey of clustering algorithms. IEEE Trans. Neural. Netw. **16**, 645–678 (2005)
16. Tran, D., Wagner, M.: Fuzzy entropy clustering. In: IEEE International Conference Fuzzy Systems, vol. 1, pp. 152–157 (2000)
17. Bezdek, J.C., Ehrlich, R., Full, W.: FCM: the fuzzy C-means clustering algorithm. Comput. Geosci. **10**(2), 191–203 (1984)
18. Chen, S.C., Zhang, D.Q.: Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure. IEEE Trans. Syst. Man. Cybern. Part B Cybern. **34**(4), 1907–1916 (2004)
19. Krinidis, S., Chatzis, V.: A robust fuzzy local information c-means clustering algorithm. IEEE Trans. Image Process. **19**, 1328–1337 (2010)
20. Verma, H., Agrawal, R.K., Kumar, N.: Improved fuzzy entropy clustering algorithm for MRI brain image segmentation. Int. J. Imaging Syst. Technol. **24**(4), 277–283 (2014)
21. Maulik, U.: Medical image segmentation using genetic algorithms. IEEE Trans. Inf. Technol. Biomed. **13**(2), 166–173 (2009)
22. Ouadfel, S., Batouche, M.: An efficient ant algorithm for swarm-based image clustering. J. Comput. Sci. **3**(3), 162–167 (2007)
23. Hassanzadeh, T., Vojodi, H., Moghadam, A.M.E.: An image segmentation approach based on maximum variance intra-cluster method and firefly algorithm. In: 7th International Conference on Natural Computation, pp. 1817–1821 (2011)
24. Nandy, S., Yang, X.S., Sarkar, P.P., Das, A.: Color image segmentation by cuckoo search. Intell. Autom. Soft Comput. **21**(4), 673–685 (2015)

25. Alam, S., Dobbie, G., Koh, Y.S., Riddle, P., Rehman, S.U.: Research on particle swarm optimization based clustering: a systematic review of literature and techniques. Swarm Evol. Comput. **17**, 1–13 (2014)
26. Benaichouche, A.N., Oulhadj, H., Siarry, P.: Improved spatial fuzzy c-means clustering for image segmentation using PSO initialization, Mahalanobis distance and post-segmentation correction. Digit Signal Process. **23**, 1390–1400 (2013)
27. Filho, T.M.S., Pimentel, B.A., Souza, R.M.C.R., Oliveira, A.L.I.: Hybrid methods for fuzzy clustering based on fuzzy c-means and improved particle swarm optimization. Expert Syst. Appl. **42**(17), 6315–6328 (2015)
28. Zhang, Y., Xiong, X., Zhang, Q.: An improved self-adaptive PSO algorithm with detection function for multimodel function optimization problems. Math Probl. Eng. **2013**, 8 (2013). Article ID 716952. http://dx.doi.org/10.1155/2013/716952
29. Wong, M.T., He, X., Yeh, W.C.: Image clustering using particle swarm optimization. In: 2011 IEEE Congress on Evolutionary Computation (CEC), pp. 262–268, June 2011
30. Yang, C., Gao, W., Liu, N., Song, C.: Low-discrepancy sequence initialized particle swarm optimization algorithm with high-order nonlinear time-varying inertia weight. Appl. Soft Comput. **29**, 386–394 (2015)
31. Harrison, K.R., Engelbrecht, A.P., Ombuki-Berman, B.M.: Inertia weight control strategies for particle swarm optimization. Swarm Intell. **10**(4), 267–305 (2016)
32. Zhao, F., Jiao, L., Liu, H.: Kernel generalized fuzzy c-means clustering with spatial information for image segmentation. Digit Signal Process. **23**, 184–199 (2013)
33. Yang, M., Tsai, H.: A Gaussian kernel-based fuzzy c-means algorithm with a spatial bias correction. Pattern Recogn. Lett. **29**(12), 1713–1725 (2008)
34. Kennedy, J., Eberhart, R.C.: Particle swarm optimization. In: IEEE International Conference on Neural Networks, pp. 1942–1948 (1995)
35. Eberhart, R.C., Shi, Y.: Comparing inertia weights and constriction factors in particle swarm optimization. In: IEEE Congress on Evolutionary Computation, San Diego, pp. 84–88 (2000)
36. http://pages.upf.pf/Sebastien.Chabrier/ressources.php
37. BrainWeb: Simulated Brain database (2016). http://brainweb.bic.mni.mcgill.ca/brainweb/
38. Jaccard, P.: The distribution of the flora in the alpine zone. New Phytol. **11**(2), 37–50 (1912)
39. Beauchemin, M., Thomson, K.P.B., Edwards, G.: On the Hausdorff distance used for the evaluation of segmentation results. CJRS **24**(1), 3–8 (1998)
40. Taha, A.A., Hanbury, A.: Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Med. Imaging. **15**, 29 (2015)