An Adaptive Memetic Fuzzy Clustering Algorithm With Spatial Information for Remote Sensing Imagery

Yanfei Zhong, *Member*, *IEEE*, Ailong Ma, and Liangpei Zhang, *Senior Member*, *IEEE*

***Abstract*—Due to its inherent complexity, remote sensing image clustering is a challenging task. Recently, some spatial-based clustering approaches have been proposed; however, one crucial factor with regard to their clustering quality is that there is usually one parameter that controls their spatial information weight, which is difficult to determine. Meanwhile, the traditional optimization methods of the objective functions for these clustering approaches often cannot function well because they cannot simultaneously possess both a local search capability and a global search capability. Furthermore, these methods only use a single optimization method rather than hybridizing and combining the existing algorithmic structures. In this paper, an adaptive fuzzy clustering algorithm with spatial information for remote sensing imagery (AFCM_S1) is proposed, which defines a new objective function with an adaptive spatial information weight by using the concept of entropy. In order to further enhance the capability of the optimization, an adaptive memetic fuzzy clustering algorithm with spatial information for remote sensing imagery (AMASFC) is also proposed. In AMASFC, the clustering problem is transformed into an optimization problem. A memetic algorithm is then utilized to optimize the proposed objective function, combining the global search ability of a differential evolution algorithm with a local search method using Gaussian local search (GLS). The optimal value of the specific parameter in GLS, which determines the local search efficiency, can be obtained by comparing the objective function increment for different values of the parameter. The experimental results using three remote sensing images show that the two proposed algorithms are effective when compared with the traditional clustering algorithms.**

***Index Terms*—Fuzzy clustering, memetic algorithm, remote sensing, spatial information.**

I. INTRODUCTION

CLUSTERING is one of the most important techniques in remote sensing image processing. The aim of remote sensing clustering is to partition a given image into groups such that pixels in the same group are as similar to each other as possible, while those assigned to different groups are dissimilar [1]–[3]. Among the clustering methods, fuzzy clustering is popular and has been widely used in remote sensing image clustering [3]–[6]. The fuzzy clustering approach can retain more information from the original image than the crisp or hard clustering methods such as K-means [7] and ISODATA [8], which usually do not perform well when the mixed pixel problem appears.

Fuzzy c-means (FCM) [9] is one of the most widely used fuzzy clustering methods in remote sensing image clustering. In some cases, the original FCM-based clustering algorithms do function well, to a certain extent; however, due to the characteristics of remote sensing imagery and the influence of external conditions, there are still some problems in remote sensing image clustering. For example, some isolated pixels may appear in the clustering image due to the existence of noise, outliers, or mixed pixels. This may be a result of not taking the spatial information in the image into account. Ahmed *et al.* [10] proposed FCM_S, with the aim of incorporating the spatial information by modifying the objective function of FCM. However, FCM_S is time-consuming because of the computation of the spatial neighborhood term in each iteration step. Chen and Zhang [11] reduced the computational complexity of FCM_S by introducing a mean-filtered image named FCM_S1. However, one common drawback of the above methods is that they both need a parameter α to control the trade-off between robustness to outliers and the effectiveness of the detail preservation [12]. Moreover, this parameter is often selected empirically, which is time-consuming and unreliable. Second, the traditional clustering algorithms, such as K-means and FCM, belong, in essence, to mountain-climbing methods. That is, it is easy for them to get stuck in a local optimum, especially when considering the complexity of remote sensing processing [13]. Hence, global optimization methods such as the genetic algorithm, differential evolution algorithm, clonal selection algorithm [14], and particle swarm optimization have been used to optimize the corresponding objective functions [15]–[19]. Although these global optimization methods can locate the promising solutions of the search space, it is difficult for them to refine the solutions in the space. As a result, the optimization performance is usually unsatisfactory if only one optimization method is utilized to optimize the objective function.

In this paper, to overcome the problems mentioned above, an adaptive fuzzy clustering algorithm with spatial information for remote sensing imagery, namely AFCM_S1, is proposed. In AFCM_S1, a new objective function with an adaptive spatial information weight is constructed. Inspired by the physical meaning of entropy, an increased spatial constraint is assigned to pixels with greater entropy, due to the fact that the pixels with greater entropy are much more uncertain. In AFCM_S1, the adaptive spatial information weight is set intuitively as normalized entropy. In addition, an adaptive memetic fuzzy clustering algorithm with spatial information for remote sensing imagery, namely

Manuscript received July 14, 2013; revised October 10, 2013 and January 07, 2014; accepted January 20, 2014. Date of publication February 13, 2014; date of current version April 18, 2014. This work was supported in part by the National Natural Science Foundation of China under Grant 41371344, in part by A Foundation for the Author of National Excellent Doctoral Dissertation of China (FANEDD) under Grant 201052, and in part by a Program for Changjiang Scholars and Innovative Research Team in University under Grant IRT1278. *(Corresponding author: Y. Zhong.)*

The authors are with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: zhongyanfei@whu.edu.cn; maailong007@whu.edu.cn; zlp62@whu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2014.2303634

AMASFC, is also proposed, in which the clustering problem is transformed into an optimization problem. A memetic algorithm is introduced to optimize the proposed objective function. Memetic algorithms, as first proposed by Moscato [20], can be seen as a population-based search method that is coupled with one or more refinement methods. Evolutionary algorithms perform well for global searching because they are capable of quickly finding and exploiting the promising regions of the search space, but their capability of converging to a local optimum is limited. Local search methods can quickly find the local optimum of a small region of the search space, but they have poor global search capability. As a result, hybrid algorithms have been proposed, which can combine the excellent global exploration characteristics of an evolution algorithm with the efficient refinement capabilities of a local search algorithm. These hybrid algorithms are known as memetic algorithms [21], [22]. Memetic algorithms have been successfully applied to many problems, including combinatorial optimization [23], multi-objective optimization [24], gene features [25], and feature selection [26], [27]. They have also been used in some real-world applications in remote sensing image processing, such as image segmentation [28], feature selection [29], and sub-pixel mapping [30]. In this paper, the global search in the memetic algorithm is set to be a differential evolution algorithm [31] because of its powerful global search capability, which has been proved in many applications [32]–[34]. Gaussian local search (GLS) is used as the local search method. The optimal value of parameter δ in GLS, which determines the local search efficiency, can be obtained by comparing the Jm increment for different values of δ.

The rest of the paper is organized as follows. Section II introduces some related background, including fuzzy clustering with spatial constraints (FCM_S), and the framework of the memetic algorithm. Section III describes the proposed algorithm in detail. The experimental results are shown in Section IV, and Section V provides the conclusion.

II. The Fuzzy Clustering Algorithm With Spatial Information

A. The FCM Clustering Algorithm

The FCM clustering algorithm performs clustering by minimizing the objective function (1)

$$Jm = \sum_{i=1}^{C}\sum_{k=1}^{N} u_{ik}^{m}\|x_k - v_i\|^2 \tag{1}$$

$$U_{C\times N} = \left\{ u_{ik} \in [0,1] \middle| \sum_{i=1}^{C} u_{ik} = 1 \ and \ 0 < \sum_{k=1}^{N} u_{ik} < N \right\} \tag{2}$$

$$v_i = \frac{\sum_{k=1}^{N} u_{ik}^{m} x_k}{\sum_{k=1}^{N} u_{ik}^{m}} \tag{3}$$

$$u_{ik} = \frac{\|x_k - v_i\|^{-2/m-1}}{\sum_{j=1}^{C} \|x_k - v_j\|^{-2/m-1}} \tag{4}$$

where x_k is the gray-level value of the kth pixel; N is the total number of pixels; C is the number of clusters; v_i is the value of the ith cluster center; and u_{ik} represents the fuzzy membership of the kth pixel, with respect to the ith cluster center, which satisfies (2). The parameter m is a fuzzy weighting exponent on each fuzzy membership u_{ik}, and it controls the fuzziness of the membership $\mathrm{m} \in (1, \infty)$. When the parameter m approaches 1, the FCM algorithm tends to be a crisp clustering algorithm, the same as K-means. When the parameter m approaches positive infinity, the entire data tend to be classified into one class.

By minimizing the objective function (1) using the Lagrange multiplier method, the update equations of membership u_{ik} and cluster centers v_i are as (3) and (4), respectively. The FCM can be implemented in the following steps:

1) Initialize the membership matrix $U_{C\times N}$ by randomly selecting $C \times N$ values between 0 and 1. Then, the constraint $\sum_{i=1}^{C} u_{ik} = 1$ in (2) is conducted on the initialized membership matrix $U_{C\times N}$.
2) Calculate the cluster centers by (3).
3) Calculate the membership matrix U by (4).
4) Repeat steps 2) and 3) until the difference between the current membership matrix and the previous membership matrix is under the specified threshold value, which is 0.0001 in this paper.

A major disadvantage of FCM is that its clustering quality is sensitive to the initial value, and it is also heavily influenced by noise.

B. FCM_S and Its Improved Variant (FCM_S1)

In a remote sensing clustering image, there are often some isolated pixels, which can result for the following reasons: 1) there is noise, and/or outliers in the remote sensing image; 2) because of the low resolution of the remote sensing image, some mixed pixels may exist; and 3) different objects may have the same spectral characteristic, while similar objects may have different spectral characteristics. Hence, in the above situation, the standard FCM cannot function well because it does not apply any spatial information in its objective function, which makes it sensitive to noise or other isolated pixels. To overcome this problem, clustering algorithms incorporating spatial information have been proposed.

FCM_S, as introduced by Ahmed *et al.* [10], is a modification of FCM that introduces a term that allows the labeling of a pixel to be influenced by the labels in its neighborhood. The neighborhood effect acts as a regularizer and biases the solution toward homogeneous labeling. The objective function of FCM_S is defined as follows:

$$Jm = \sum_{i=1}^{C}\sum_{k=1}^{N} u_{ik}^{m}\|x_k - v_i\|^2 + \frac{\alpha}{N_R}\sum_{i=1}^{C}\sum_{k=1}^{N} u_{ik}^{m} \times \sum_{\mathrm{r}\in N_k} \|x_r - v_i\|^2 \tag{5}$$

where x_k is the gray-level value of the kth pixel; N is the total number of pixels; v_i is the value of the ith cluster center; u_{ik} represents the fuzzy membership of the kth pixel, with respect to the ith cluster center; and NR is its cardinality. x_r represents the

Input: an instance, size of population, probability *Pls*, and some other parameters
Output: a feasible individual

Initialization: Generate an initial population
While stopping criteria are not satisfied do
  Evaluate all individuals in the population
  Evolve a new population using evolutionary operator
  for each individual do
    Perform local search around it with probability *Pls*
  end
end

Fig. 1. General framework of the memetic algorithm.

neighbor of x_k and N_k stands for the set of neighbors falling into a window around pixel x_k. The parameter α is used to control the effect of the neighbor term.

FCM_S is time-consuming because of the calculation of the neighbor term at each iteration step. As a result, FCM_S1 [11] was proposed as a variant of FCM_S. The objective function of FCM_S1 is written as follows:

$$Jm = \sum_{i=1}^{C}\sum_{k=1}^{N} u_{ik}^{m}\|x_k - v_i\|^2 + \alpha \sum_{i=1}^{C}\sum_{k=1}^{N} u_{ik}^{m}\|\overline{x_k} - v_i\|^2 \quad (6)$$

where $\overline{x_k}$ is the mean of the neighboring pixels lying within a window around x_k. Unlike FCM_S, $\overline{x_k}$ can be calculated in advance, reducing the calculation time.

The two methods mentioned above (FCM_S and FCM_S1) can function well in certain situations. However, the clustering quality largely depends on the parameter α, which controls the trade-off between robustness to outliers and the effectiveness of detail preservation, and parameter α is often selected by trial and error. In Section III, we formulate the proposed method of adaptively determining the trade-off parameter by introducing the concept of entropy.

C. The General Framework of the Memetic Algorithm

The traditional clustering algorithms, such as K-means and FCM, belong, in essence, to mountain-climbing methods. That is, it is easy for them to get stuck in a local optimum. Some global optimization methods such as the genetic algorithm, differential evolution algorithm, and clonal selection algorithm have been used to optimize the corresponding objective functions. Although these global optimization methods can locate the promising solutions of the search space, it is difficult for them to refine the solutions in the space. Hence, the optimization performance is usually unsatisfactory if only one optimization method is utilized to optimize the objective function. As the "no free lunch" theorem shows, there is no universal optimizer which performs well on all classes of problems. Hence, a memetic algorithm is needed, which can be seen as a population-based search method that is coupled with one or more local search methods; the framework of which is summarized in Fig. 1 [23].

As can be seen from Fig. 1, the general framework of the memetic algorithm is the same as a traditional evolution algorithm such as the genetic algorithm or differential evolution algorithm, except for the addition of a local search procedure that refines some individuals of the population. The success of the memetic algorithm is, therefore, largely dependent on the selection of the local search method, which often incorporates domain knowledge of the specific problem. For example, the most commonly used local search methods in combinatorial domains are hill climbing, simulated annealing, and tabu search. In continuous domains, downhill, gradient, quasi-Newton, and trust-region strategies are often used [35].

III. An Adaptive Memetic Fuzzy Clustering Algorithm With Spatial Information for Remote Sensing Imagery

In this paper, to adaptively determine the spatial information weight in the process of fuzzy clustering and to enhance the capability of the traditional optimization methods, an adaptive memetic fuzzy clustering algorithm with spatial information for remote sensing imagery, namely AMASFC, is proposed. AMASFC consists of two main processes: 1) the construction of the objective function and 2) the optimization of the objective function. In the process of the construction of the objective function, a new objective function is proposed with an adaptive spatial information weight by introducing the concept of entropy. In the process of the optimization of the objective function, a memetic algorithm is utilized to optimize the proposed objective function, combining a differential evolution algorithm with a GLS method. The flowchart of the proposed algorithm (AMASFC) is shown in Fig. 2. The two processes of AMASFC are described in detail in the following sections.

A. Construction of the Objective Function and Its Traditional Optimization Method

The objective function of the proposed algorithm is as follows:

$$\begin{aligned} Jm = &\sum_{i=1}^{C}\sum_{k=1}^{N} u_{ik}^{m}(1-\alpha_k)\|x_k - v_i\|^2 \\ &+ \sum_{i=1}^{C}\sum_{k=1}^{N} u_{ik}^{m}\alpha_k\|\overline{x_k} - v_i\|^2. \end{aligned} \quad (7)$$

As with the objective function of FCM (1), the objective function (7) can also be minimized by updating (8) and (9) through iteration. In the rest of the paper, we refer to the above optimization method as AFCM_S1

$$v_i = \frac{\sum_{k=1}^{N} u_{ik}^{m}(1-\alpha_k)x_k + \sum_{k=1}^{N} u_{ik}^{m}\alpha_k\overline{x_k}}{\sum_{k=1}^{N} u_{ik}^{m}} \quad (8)$$

$$u_{ik} = \frac{\{(1-\alpha_k)\|x_k - v_i\|^2 + \alpha_k\|\overline{x_k} - v_i\|^2\}^{-1/m-1}}{\sum_{j=1}^{C}\{(1-\alpha_k)\|x_k - v_j\|^2 + \alpha_k\|\overline{x_k} - v_j\|^2\}^{-1/m-1}} \quad (9)$$

$$U_{C\times N} = \left\{ u_{ik} \in [0,1] | \sum_{i=1}^{C} u_{ik} = 1 \; and \; 0 < \sum_{k=1}^{N} u_{ik} < N \right\} \quad (10)$$

where x_k is a vector representing the kth pixel for a multispectral remote sensing image, N is the total number of pixels,

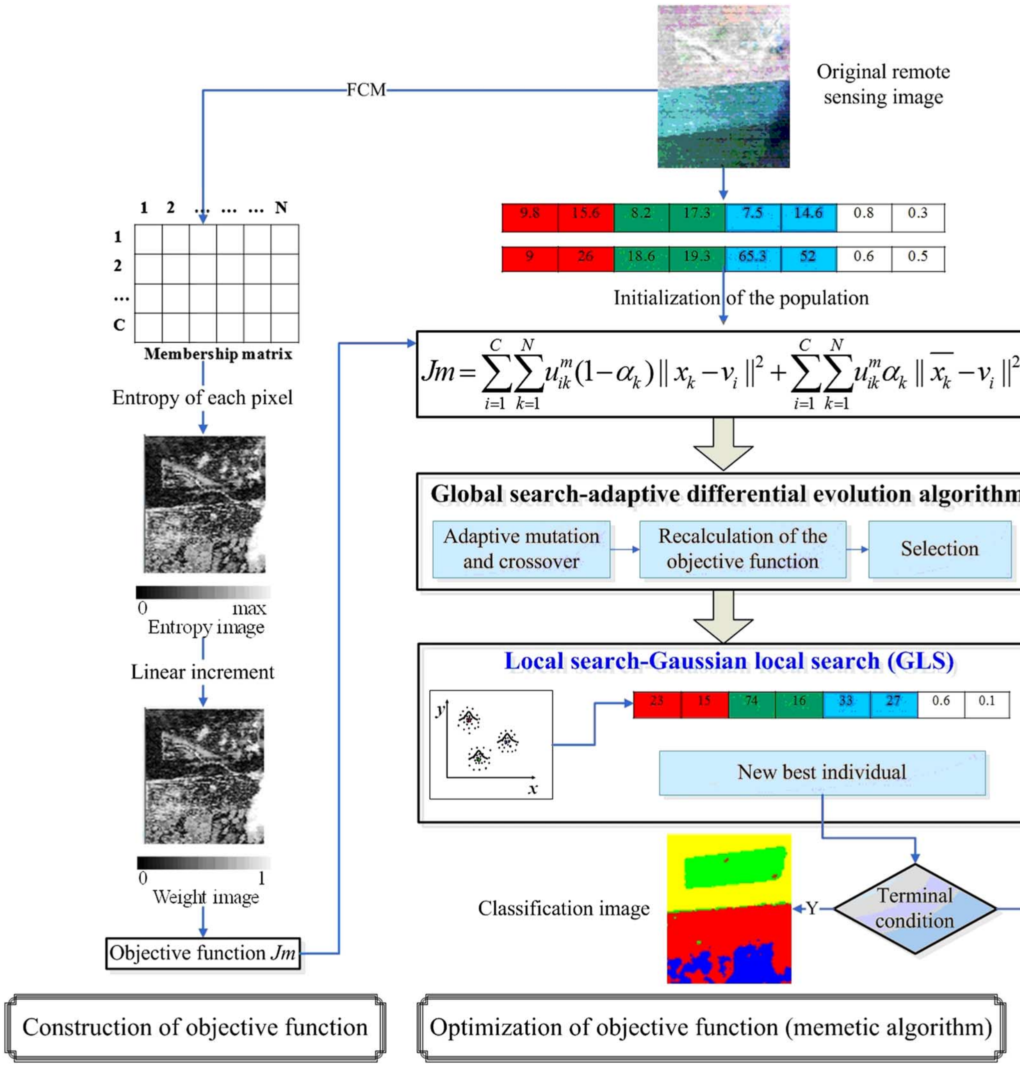


Fig. 2. Flowchart of the adaptive memetic fuzzy clustering algorithm with spatial information (AMASFC).

and C is the number of clusters, and $\{\nu_i\}_{i=1}^{C}$ are the centroids of the clusters. The parameter m is a fuzzy weighting exponent on each fuzzy membership and the function of the parameter m is similar to FCM. The array $U_{C\times N}$ is a fuzzy membership matrix satisfying (10). $\overline{x_k}$ represents the mean of the pixels falling into a window around x_k, which can be calculated in advance. The parameter α_k is used to control the effect of the neighbor term determined by the entropy of the kth pixel. The formula for the entropy [36] is as follows:

$$E_k = -\sum_{i=1}^{C} u_{ik}\log_2 u_{ik} \tag{11}$$

where E_k is the entropy of the kth pixel. As can be seen from (11), the greater the entropy of the pixel, the more uncertain the pixel is. On the other hand, for isolated pixels, their fuzzy membership to each class is comparable to the FCM-based clustering methods. That is, the labels of these pixels are uncertain. Hence, our motivation for introducing entropy is to put an increased spatial constraint onto the pixels with more uncertainty.

In order to obtain the entropy of each pixel, the FCM algorithm is first applied to the remote sensing image, and the membership matrix can then be acquired. The entropy of each pixel is then calculated by (11). Because the range of the entropy is not equal to the range of α_k, namely [0,1], it needs to be mapped linearly to [0,1]. The linear mapping schedule in our study is as follows:

$$\alpha_k = \frac{E_k - E_{\min}}{E_{\max} - E_{\min}} \tag{12}$$

where E_k is the entropy of the kth pixel. E_{max} and E_{min} are the maximum and minimum entropies of all the pixels, respectively, and α_k is the trade-off parameter of the kth pixel.

Fig. 3(a) is one-dimensional simulated data with a size of 4×7 pixels. After applying FCM to the data, the cluster centers v_1 and v_2 are 86.71 and 171.00, respectively. Moreover, the trade-off parameter matrix can be calculated as shown in Fig. 3(b). As can be seen from Fig. 3(a), the pixels with values of 255 and 10 are very different from the other pixels. In Fig. 3(b), these pixels have a

89	91	87	162	163	158	162
90	86	92	88	160	10	160
89	255	255	90	90	161	159
90	87	86	91	89	160	158

(a)

0.013999	0.041825	0	0.147130	0.119080	0.283360	0.147130
0.026332	0.000031	0.060308	0.005076	0.210642	0.957160	0.210642
0.013999	1.000000	1.000000	0.026332	0.026332	0.177683	0.245904
0.026332	0	0.000031	0.041825	0.013999	0.210642	0.283360

(b)

Fig. 3. Example of the determination of parameter α: (a) simulated data ($\nu_1 = 86.71$, $\nu_2 = 171.00$) and (b) corresponding trade-off parameter matrix of the data in (a).

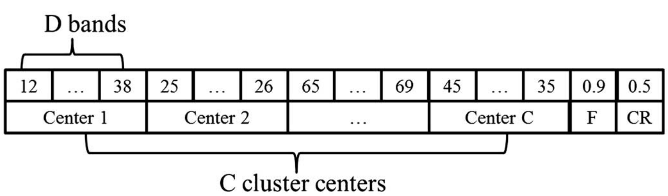


Fig. 4. Example of individual encoding.

much greater spatial information weight than the other pixels, meaning that an increased spatial constraint is put on these pixels.

B. The Proposed Optimization Method of the Objective Function (Memetic Algorithm)

A memetic algorithm is used to optimize the objective function (7) and the fuzzy clustering problem is transformed into an optimization problem. An adaptive differential evolution algorithm [37] jDE is used as the global search method, while GLS is used to refine the solution. The whole optimization process can be implemented according to the following steps:

Step 1) Initialization of the population. In jDE, considering the adaptive mutation and cross-over in Step 3), the mutation scale factor F and the crossover constant CR need to be encoded into the individual. We assume that there are C cluster centers and D bands in the remote sensing image (see Fig. 4). Hence, each individual contains $(C \times D + 2)$ dimensions. C is the number of cluster centers. D is the number of bands of the image. The number of parameters in DE is 2. The individual in the population is initialized by randomly selecting pixels from the whole image as the cluster centers.

Step 2) Calculation of the fitness of each individual, using the objective function (7).

Step 3) Adaptive mutation and crossover. In DE, the mutation operator amounts to creating a donor vector $\boldsymbol{V_i}(t) = [v_{i,1}(t), v_{i,2}(t), \ldots, v_{i,D}(t)]$ for changing each individual of the population. The mutation process can be expressed as follows:

$$\boldsymbol{V_i}(t) = \boldsymbol{X}_{r_1^i}(t) + F_i(\boldsymbol{X}_{r_2^i}(t) - \boldsymbol{X}_{r_3^i}(t)) \tag{13}$$

where $\boldsymbol{X}_{r_1^i}(t)$, $\boldsymbol{X}_{r_2^i}(t)$, $and \boldsymbol{X}_{r_3^i}(t)$ are picked up randomly from the population.

Input: individual with the best fitness obtained by the global search Output: new individual with the best fitness
For each dimension of each cluster center ➢Perform the Gaussian mutation to produce a trial individual. ➢Evaluate the trial solution by (7). ➢Update the best individual with the best fitness by selecting the winner between the old individual and the trial individual. end

Fig. 5. Update of the best individual in the process of local search.

After the mutation operator, crossover is undertaken between the donor vector $\boldsymbol{V}_i(t)$ and the target vector $\boldsymbol{X}_i(t)$, generating a trial vector $\boldsymbol{U}_i(t) = [u_{i,1}(t), u_{i,2}(t), \ldots, u_{i,D}(t)]$. The crossover operator can be implemented as follows:

$$u_{i,j}(t) = \begin{cases} v_{i,j}(t), & if\ (rand_{i,j}(0,1) \le CR_i\ or\ j = j_{rand}) \\ x_{i,j}(t), & otherwise. \end{cases} \tag{14}$$

There are two main parameters F and CR in DE. As shown in Fig. 4, each individual not only encodes the cluster centers but also the parameters F and CR, enabling their update in the process of evolution. F_i and CR_i can be updated according to

$$F_i' = \begin{cases} 0.1 + 0.9 \times rand(0,1), & if\ rand(0,1) < 0.1 \\ F_i, & otherwise \end{cases} \tag{15}$$

$$CR_i' = \begin{cases} rand(0,1), & if\ rand(0,1) \le 0.1 \\ CR_i, & otherwise \end{cases} \tag{16}$$

where F_i' and CR_i are the updated values of the corresponding individual.

Step 4) Recalculation of the fitness of the offspring, using the objective function (7).

Step 5) Selection. The selection operator is used to decide if the target vector $\boldsymbol{X}_i(t)$ or the trial vector $\boldsymbol{U}_i(t)$ is the winner. The vector with the better fitness can then be selected for the next generation. The target vector of the next generation is generated by the selection operator, as follows:

$$\boldsymbol{X}_i(t+1) = \begin{cases} \boldsymbol{U}_i(t), & if\ f(\boldsymbol{U}_i(t)) \le f(\boldsymbol{X}_i(t)) \\ \boldsymbol{X}_i(t), & otherwise \end{cases} \tag{17}$$

where $f(\boldsymbol{X})$ is the objective function and an individual with a lower value of $f(\boldsymbol{X})$ has the better fitness, assuming that this is a minimization problem. $\boldsymbol{X}_i(t+1)$ is the individual that is selected to the $t+1$ generation.

Step 6) Elitist strategy. In order to speed up the convergence of the iteration and enhance the efficiency of the optimization, an elitist strategy is applied, which preserves the individual with the best fitness found so far.

Step 7) Local search. If the search stagnates for consecutive generations, i.e., the candidate best individual does not update for consecutive generations, then local search

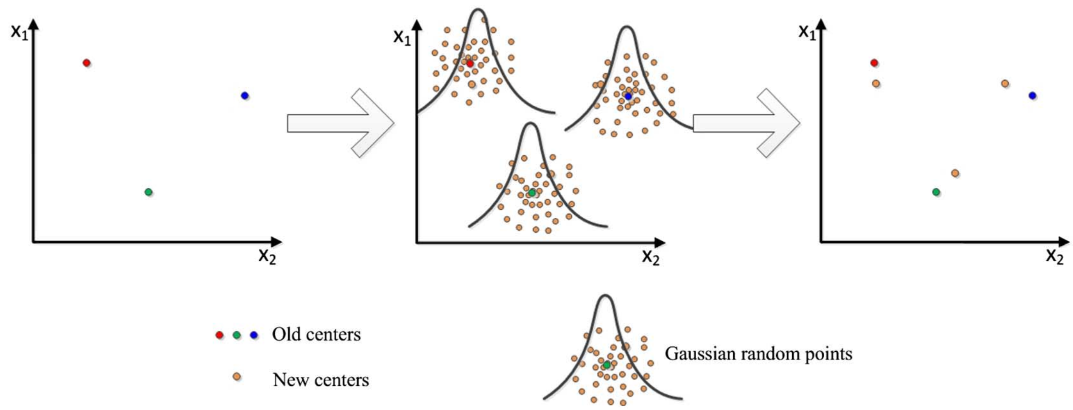


Fig. 6. Process of GLS.

	FLC	Wuhan TM	AVIRIS_Salinas
Size of population	60	30	50
Maximum generations	100	100	100
The parameter δ in GLS	0.1	1	1
Exponent weighting *m*	2	2	2

Fig. 7. Parameter settings.

can be executed. The local search method is described in detail in Section III-C.

Step 8) Terminal condition. Repeat Steps 3)–7) until the terminal condition is met. The terminal condition is to either reach the maximum number of iterations or the beginning of stagnation for the update of the best individual.

Step 9) After the optimization layer, the optimal cluster centers can be acquired. However, the optimal cluster centers cannot be used to cluster the original remote sensing image directly because of the physical meaning of the objective function (7). Before clustering, (9) is used to generate the corresponding fuzzy membership matrix, and each pixel should be assigned to the class with the largest membership.

C. Local Search

Local search is an important part of the memetic algorithm. The role of the local search is fundamental, and the selection of its search rule and its balance with the global search scheme determines the success of the memetic framework [35]. The local search method used in this paper is as follows.

GLS: Suppose that $x(k) = \{x_1(k), x_2(k), \ldots, x_N(k)\}$ is a vector with N dimensions that represents a cluster center. The Gaussian mutation can be represented as

$$x'_i(k) = N(x_i(k-1), \delta^2) \tag{18}$$

where $i \in [1, N]$ and $N(x_i(k-1), \delta^2)$ is a normal distribution with a mean of $x_i(k-1)$ and standard deviation δ.

The Gaussian mutation is performed on each dimension of the vector because of the sparsity of the efficient solutions. For example, for the individual in Fig. 4, it needs a six-times Gaussian mutation and six-times evaluation to perform the local search. As for the minimization problem, the individual with the best fitness obtained by the global search can be updated as follows (Fig. 5):

$$individual\_new = \begin{cases} individual\_trial, f(individual\_trial) \\ \qquad\qquad < f(individual\_old) \\ individual\_old, otherwise \end{cases} \tag{19}$$

where $individual\_old$, $individual\_trial$, and $individual\_new$ are the old best individual, the trial individual, and the new best individual, respectively.

Fig. 6 shows the process of GLS. Trial solutions are generated by performing Gaussian mutation on the individual with the best fitness. If the fitness of the trial solution is higher than the old one, then the trial solution will replace the old one.

As can be seen from (18), the parameter δ is crucial to the GLS and is discussed further in Section IV.

IV. Experimental Results

A. Competing Methods and Parameter Settings

The proposed algorithm AMASFC is compared with several other clustering algorithms: the FCM clustering algorithm, FCM_S1 [11], AFCM_S1, and automatic fuzzy clustering using an improved differential evolution algorithm (FCIDE) [16]. It should be noted that the AFCM_S1 method only uses the proposed adaptive spatial information weight instead of introducing a memetic algorithm, as in AMASFC. FCIDE involves the application of DE to the automatic clustering of large unlabeled datasets and aims to automatically determine the optimal class number of the unlabeled dataset. It can give a promising result with a modification of the traditional DE chromosome representation scheme. However, in some cases, FCIDE cannot determine the correct class number of the unlabeled dataset. Hence, for the different remote sensing datasets of FCIDE, the best clustering results with the correct class numbers are reported, instead of the mean and standard deviation of the clustering accuracy.

Fig. 7 lists the parameter settings of the experiments. For the FLC image, the Wuhan TM image, and the Salinas Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) image, the size

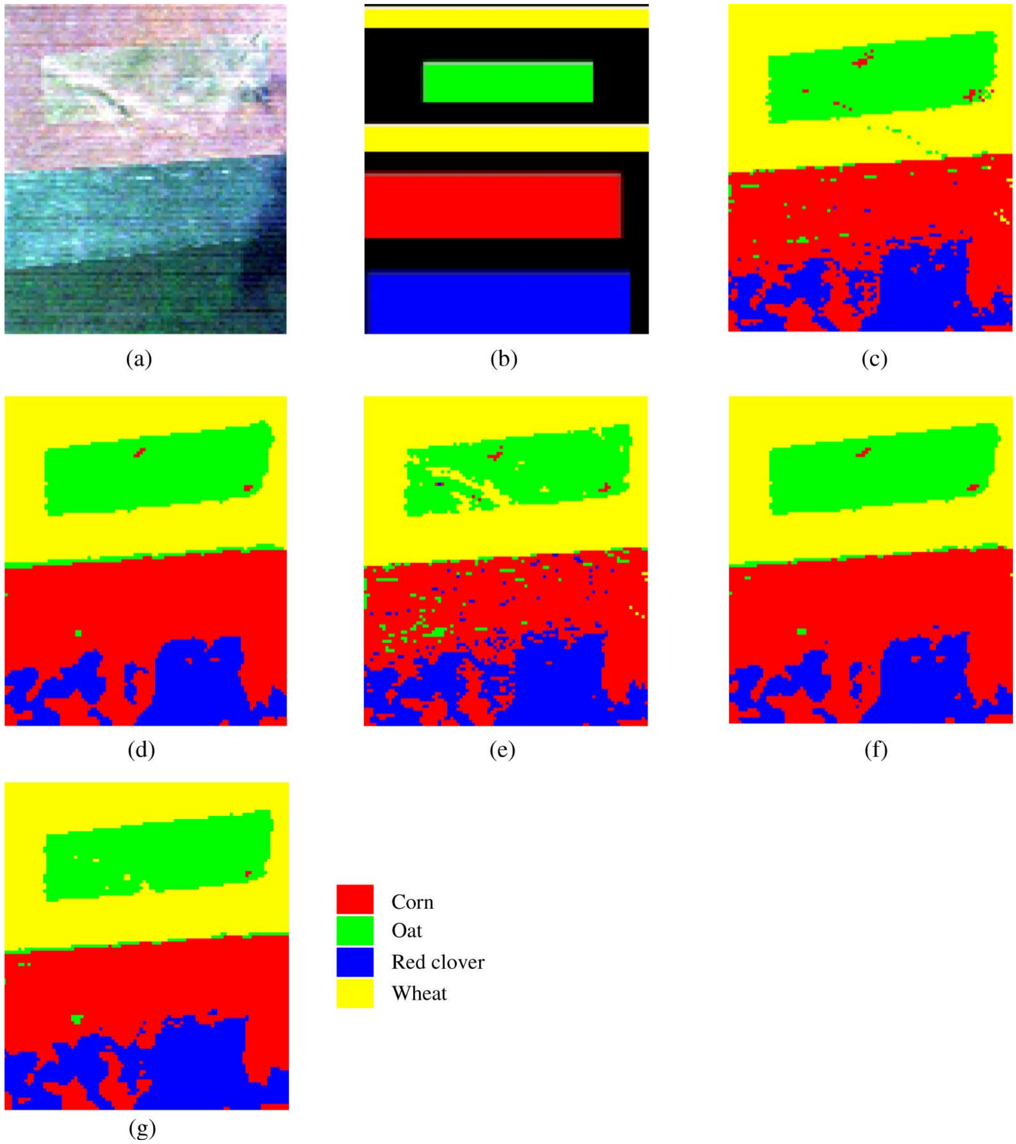


Fig. 8. FLC image and the clustering results: (a) FLC image, (b) ground truth, (c) FCM, (d) FCM_S1($\alpha = 4.5$), (e) FCIDE, (f) AFCM_S1, and (g) AMASFC.

of the population in FCIDE and AMASFC is $5 \times D$ [38], where D is the dimension of the data. The number of maximum generations is 100. The standard deviation, namely the parameter δ in GLS, is 0.1, 1, and 1 in experiments 1, 2, and 3, respectively. The exponent weighting parameter m is 2 in all the experiments.

B. Experiment 1—FLC Multispectral Image

In experiment 1, a Flightline C1 image of Tippecanoe County, Indiana, US, is used, which was acquired from the M7 scanner, at a resolution of $36.25\ \mathrm{m} \times 36.25\ \mathrm{m}$ and a size of 97×102 pixels, in June 1966. Twelve bands are contained in this image. This image contains four classes: corn, oat, red clover, and wheat. The original image and the ground truth image are shown in Fig. 8(a)–(b).

Fig. 8(c)–(g) illustrates the clustering results of the FLC image using FCM, FCM_S1, FCIDE, AFCM_S1, and AMASFC, respectively. It should be noted that Fig. 8(g) shows the best clustering result of AMASFC, and the overall accuracy (OA) of which is 92.17%. Visually, the clustering results of FCM_S1, AFCM_S1, and AMASFC are different from FCM and FCIDE due to the application of the spatial information. Although the red clover class is largely misclassified into the corn class for all five results, the clustering result of the red clover class for AMASFC is better because of the reduced misclassification in the middle-lower part of Fig. 8(g) when compared with the other results. The result of FCM_S1 looks smoother, compared with AFCM_S1 and AMASFC, especially in some parts of the corn and red clover classes, and the reason for this is that the trade-off parameter can be tuned manually for FCM_S1.

To compare the above algorithms quantitatively, the OA and kappa coefficient [39], [40] for the image are listed in Table I. It should be noted that in order to compare AMASFC with the competing methods more fairly and to demonstrate the stability of AMASFC, the experiments are repeated 10 times. The mean and the standard deviation of the 10 runs are listed in Table I. As can be seen from Table I, AMASFC obtains the best OA, 89.62%, with gains of 2.53%, 0.32%, 1.64%, and 1.10% over FCM, FCM_S1, FCIDE, and AFCM_S1, respectively. It can be

TABLE I
COMPARISON OF THE RESULTS FOR THE FLC IMAGE

Classes	FCM		FCM_S1		FCIDE		AFCM_S1		AMASFC	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Corn	97.90	71.40	**99.93**	73.87	92.90	**77.20**	99.78	72.52	99.25 ± 0.27	75.15 ± 3.17
Oat	96.90	94.70	98.77	**99.69**	88.50	89.70	98.61	99.38	**99.52 ± 0.34**	98.41 ± 0.33
Red clover	60.40	99.40	66.04	100.00	**72.20**	96.00	63.72	**100.00**	66.34 ± 5.34	99.78 ± 0.08
Wheat	99.50	**100.0**	99.91	99.83	**100.0**	95.20	99.91	99.83	99.97 ± 0.06	99.95 ± 0.15
Overall accuracy	87.09		89.30		87.98		88.52		**89.62** ± **1.59**	
Kappa	0.8228		0.8539		0.8349		0.8433		**0.8462** ± **0.0306**	

PA, producer's accuracy; UA, user's accuracy.

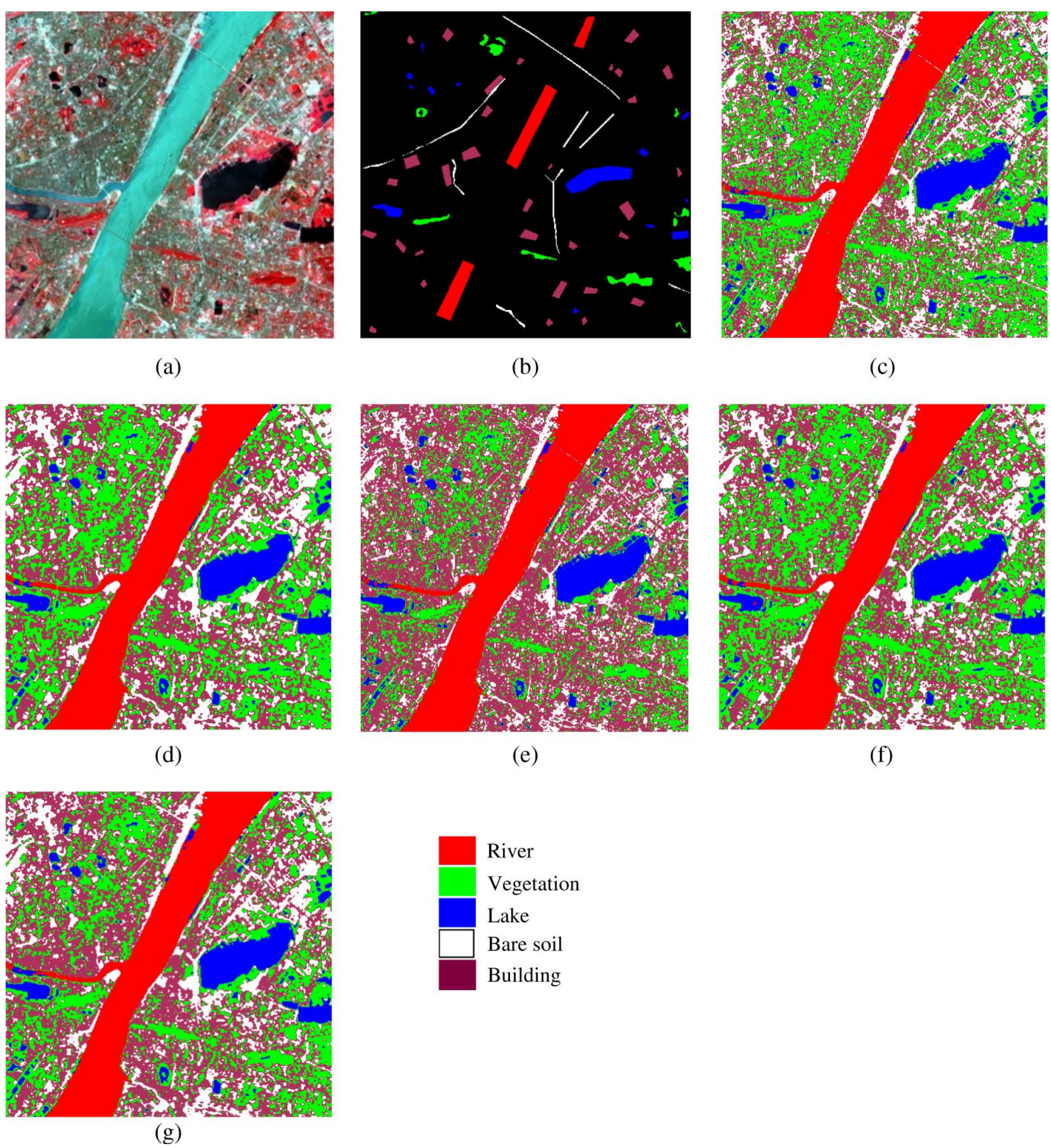


Fig. 9. Wuhan TM image and the clustering results: (a) Wuhan TM image, (b) ground truth, (c) FCM, (d) FCM_S1($\alpha = 5.7$), (e) FCIDE, (f) AFCM_S1, and (g) AMASFC.

TABLE II
COMPARISON OF THE RESULTS FOR THE WUHAN TM IMAGE

Classes	FCM		FCM_S1		FCIDE		AFCM_S1		AMASFC	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
River	100	100	100	100	100	100	100	100	**100 ± 0**	**100 ± 0**
Vegetation	82.50	75.10	**89.18**	82.20	66.60	**87.60**	88.42	79.47	88.76 ± 1.87	83.56 ± 3.92
Lake	99.80	91.90	**100.00**	96.34	99.90	**98.40**	99.94	97.15	99.94 ± 0	95.72 ± 2.50
Bare soil	71.80	85.10	67.93	84.41	69.50	87.30	65.93	**88.85**	**72.22 ± 1.10**	85.93 ± 1.82
Building	61.20	67.40	71.58	72.86	**84.30**	59.00	67.99	68.79	73.57 ± 6.10	**75.62 ± 1.88**
Overall accuracy	81.66		85.69		81.75		84.35		**86.55 ± 1.73**	
Kappa	0.7619		0.8154		0.7652		0.7976		**0.8276 ± 0.0228**	

PA, producer's accuracy; UA, user's accuracy.

seen that the producer's accuracy is markedly different from the user's accuracy for the red clover class of the five results. This is because, for many pixels, the red clover class is misclassified into the corn class. Overall, the quantitative comparison of the five algorithms is consistent with the above qualitative finding.

AMASFC achieves the best performance both visually and quantitatively. The reason for this may be that there is no spatial information in the objective function of FCM and FCIDE, which could result in some isolated pixels. For FCM_S1 and AFCM_S1, they can easily get stuck in a locally optimal solution, due to the lack of a global search capability and their sensitivity to the initial values. Moreover, by the optimization of the memetic algorithm, the clustering result of AFCM_S1 is dramatically enhanced. Hence, for AMASFC, not only is the spatial information weight determined adaptively but also the clustering result is the best among all the clustering algorithms mentioned above due to the optimization of the memetic algorithm. Although the clustering result of AMASFC is not overwhelmingly impressive when compared with FCM_S1, the spatial information weight of FCM_S1 needs to be determined by trial and error, whereas it is determined adaptively for AMASFC. Furthermore, because the FLC image is simple, it is inclined to be oversmoothed. In the next part, more complicated images are tested with the proposed methods.

C. Experiment 2—Wuhan TM Image

In order to further test the validity of the proposed algorithm, another image is used, which is a 30-m resolution multispectral Landsat TM image of Wuhan City, China, with a size of 400×400 pixels and six bands. This region of the image was expected to contain five classes: river, vegetation, lake, bare soil, and building. The original Wuhan TM image and the ground truth image are shown in Fig. 9(a)–(b).

Fig. 9(c)–(g) illustrates the clustering results for the Wuhan TM image using FCM, FCM_S1, FCIDE, AFCM_S1, and AMASFC, respectively. Again, it should be noted that Fig. 9(g) shows the best clustering result of AMASFC, and the OA of which is 88.94%. First, there are more isolated pixels in the result of FCM and FCIDE than for the three other classifiers, due to it not taking the spatial information into account. Second, as can be seen from Fig. 9, because of the simple characteristics of the river class and lake class, the five classifiers all achieve similar clustering results. For FCM, FCM_S1, and AFCM_S1, the building class is largely misclassified into the vegetation class, especially in the left-middle part of the image, compared with AMASFC, which gives better visual results. On the other hand, for FCIDE, the vegetation class is largely misclassified into the building class, when compared with AMASFC. Lastly, the bare soil class is largely misclassified into other classes such as the river class (in the top part of the image) and the building class (in the left-middle of the image). A possible reason for this is that the shape of the bare soil class in the ground truth image is linear, which could easily be smoothed by neighboring pixels. On the whole, AMASFC achieves the best visual accuracy.

To compare the above algorithms quantitatively, the OA and kappa coefficient for the image are listed in Table II. It should be noted that in order to compare AMASFC with the competing methods more fairly and to demonstrate the stability of AMASFC, the experiments are repeated 10 times. The mean and the standard deviation of the 10 runs are listed in Table II. AMASFC obtains the best OA, 86.55%, with gains of 4.89%, 0.86%, 4.80%, and 2.20% over FCM, FCM_S1, FCIDE, and AFCM_S1, respectively. The quantitative comparison of the five algorithms is consistent with the above qualitative finding: based on the above analysis, AMASFC outperforms the four other classifiers.

D. Experiment 3—Salinas AVIRIS Image

In order to test the performance of the proposed algorithms when clustering a remote sensing image with a large number of classes, another remote sensing image dataset is used, which was acquired by the 224-band AVIRIS sensor over the Salinas Valley, CA, USA [41]. The size of the image we use is 245 lines by 217 samples. A total of 20 water absorption bands

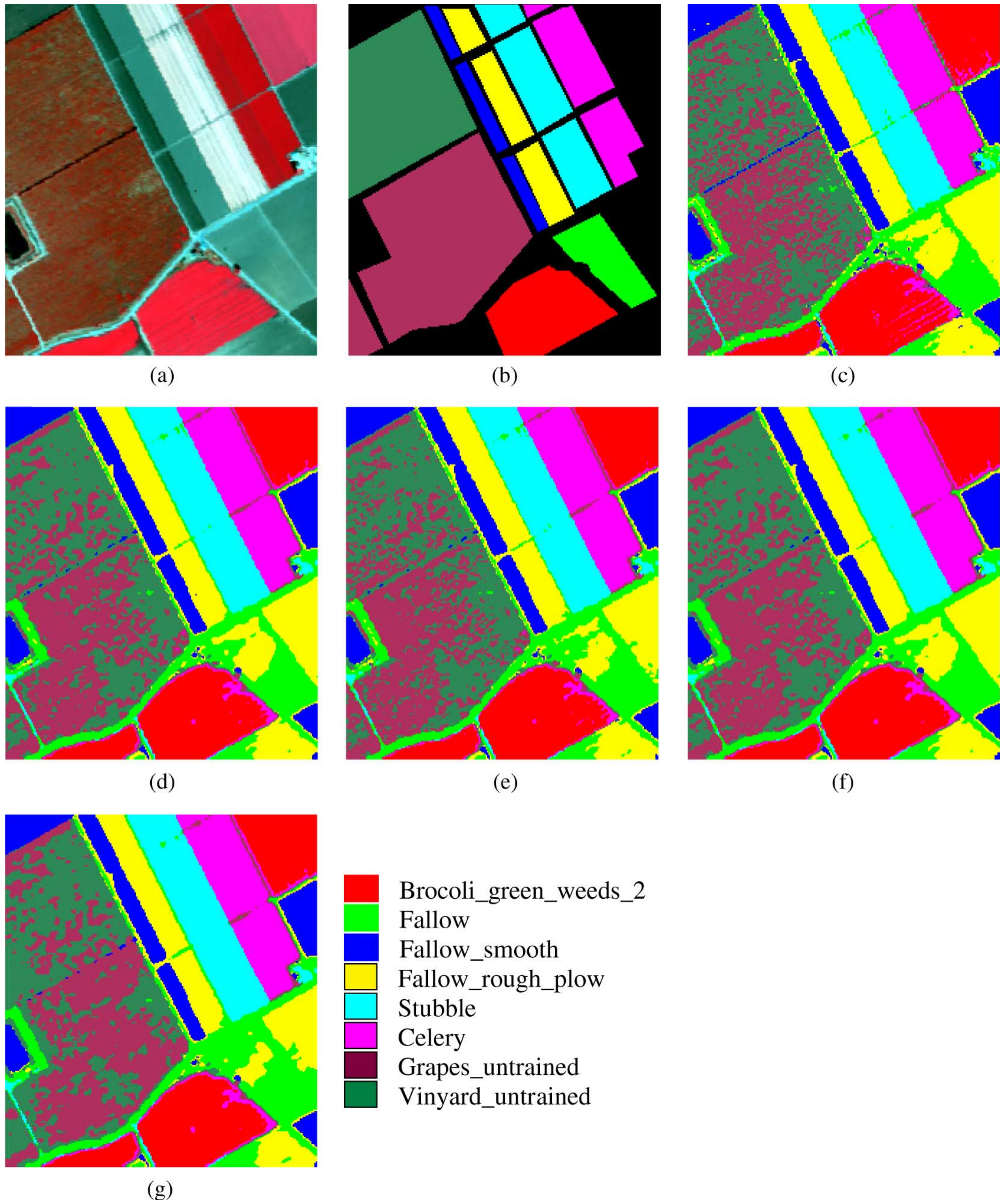


Fig. 10. Salinas AVIRIS image and the clustering results: (a) Salinas image (RGB 70, 27, 17), (b) ground truth, (c) FCM, (d) FCM_S1 ($\alpha = 2.1$), (e) FCIDE, (f) AFCM_S1, and (g) AMASFC.

(108–112, 154–167, 224) were removed. The image contains eight classes: Brocoli_green_weeds_2, Fallow, Fallow_smooth, Fallow_rough_plow, Stubble, Celery, Grapes_untrained, and Vinyard_untrained. Fig. 10(a) shows the Salinas dataset. The ground truth of the Salinas dataset is provided in Fig. 10(b). In order to enhance the efficiency, PCA feature reduction is conducted and the first 10 features are used for the clustering.

Fig. 10(c)–(g) illustrates the clustering results of the Salinas AVIRIS image using FCM, FCM_S1, FCIDE, AFCM_S1, and AMASFC, respectively. Here, it should be noted that the class number is fixed for FCIDE because of the fact that FCIDE cannot determine the correct number of classes for the Salinas AVIRIS image with a large number of classes. As can be seen from Fig. 10, there are more isolated pixels in the results of FCM and FCIDE, especially in the Grapes_untrained class and the Vinyard_untrained class, than for the three other classifiers due to these methods not taking the spatial information into account. In addition, visually, AMASFC achieves the best clustering performance for the Fallow class.

The quantitative comparisons for the image, the OA and kappa coefficient, are listed in Table III, in which the mean and the standard deviation of the 10 runs of AMASFC are listed. The results indicate that the methods using local spatial information, such as FCM_S1, AFCM_S1, and AMASFC, achieve better clustering results, compared with the methods that do not use local spatial information, such as FCM and FCIDE. In addition, for the Fallow class, AMASFC achieves the best producer's accuracy, 83.44%, and the smallest difference between the

TABLE III
COMPARISON OF THE RESULTS FOR THE SALINAS AVIRIS IMAGE

Classes	FCM		FCM_S1		FCIDE		AFCM_S1		AMASFC	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Brocoli_green_weeds 2	92.17	100	95.48	100	95.87	100	**95.97**	100	95.72 ± 0.43	**100 ± 0**
Fallow	52.53	78.88	55.52	80.78	66.85	**85.23**	55.36	81.1	**83.44 ± 9.11**	80.00 ± 3.77
Fallow_smooth	**99.57**	96.93	99.43	96.79	99.35	**97.40**	99.35	97.06	97.62 ± 2.53	95.57 ± 1.05
Fallow_rough_plow	**97.61**	72.81	95.29	73.5	96.68	78.88	95.44	73.51	88.84 ± 5.68	**86.94 ± 4.45**
Stubble	98.08	99.85	99.6	99.92	99.55	99.95	99.75	**99.97**	**99.85 ± 0**	99.91 ± 0.01
Celery	**99.11**	92.44	98.97	95.6	99.05	95.97	98.8	95.83	98.83 ± 0.27	**95.94 ± 0.31**
Grapes_untrained	56.95	73.87	60.79	77.21	58.94	74.95	61.05	77.34	**60.39 ± 9.57**	**78.61 ± 3.92**
Vinyard_untrained	68.53	51.66	72.03	55.16	69.30	52.98	72.25	55.36	**73.05 ± 10.43**	**55.53 ± 2.59**
Overall accuracy	76.11		78.51		78.14		78.64		**79.72 ± 0.87**	
Kappa	0.7104		0.7392		0.7348		0.7408		**0.7586 ± 0.0081**	

PA, producer's accuracy; UA, user's accuracy.

producer's accuracy and the user's accuracy is 3.44%. AMASFC obtains the best OA, 79.72%, with gains of 3.61%, 1.21%, 1.58%, and 1.08% over FCM, FCM_S1, FCIDE, and AFCM_S1, respectively. Overall, the quantitative comparison of the five algorithms is consistent with the above qualitative finding in that AMASFC outperforms the four other classifiers. As can be seen from the clustering results in Fig. 10, roads that show up in the Fallow class suggest that the image could be classified into more clusters. However, it is still a problem to acquire the ground truth and to assess the clustering results. In [42], remote sensing image with many classes is clustered. In [1], CONNindex, based on CONN matrix concept, is specially proposed to assess complex cluster structures. In our future work, more focus will be put on the above problems.

E. The Sensitivity of the Parameters

1) The Sensitivity of the Weight of the Spatial Information in FCM_S1, AFCM_S1, and AMASFC: In order to further test the validity of the adaptive spatial information weight, Fig. 11 shows the OA of FCM_S1 corresponding to different α values for the above three images (the FLC image, the Wuhan TM image, and the Salinas AVIRIS image). It should be noted that both AFCM_S1 and AMASFC are the proposed methods in this paper. However, AFCM_S1 only uses the proposed adaptive trade-off parameter, instead of introducing a memetic algorithm, as in AMASFC. As can be seen from the above discussion, the value of α for the two proposed methods, namely AFCM_S1 and AMASFC, is determined adaptively. However, the value of α for FCM_S1 is tuned and determined manually. In Fig. 11, the horizontal axis denotes the variation range of α for FCM_S1. The two lines represent the constant OA for the two proposed methods, and the constant clustering accuracy of the two proposed methods and the varying clustering accuracy of FCM_S1 are compared. As shown in Fig. 11, although the OA of AFCM_S1 is a little lower than the best OA of FCM_S1, the spatial information weight of AFCM_S1 is determined adaptively. For the Salinas AVIRIS image, in particular, AFCM_S1 can get nearly the same OA as FCM_S1. Moreover, by the optimization of the memetic algorithm, the clustering result of AFCM_S1 is dramatically enhanced for the above three dataset images. Again, AMASFC achieves the best clustering result.

2) The Sensitivity of Parameter δ in the GLS: The parameter δ determines the efficiency of the local search in the GLS because of the sparsity of the solutions with the increment in the dimension. Figs. 12(a) and 13(a) show the variation of Jm for different values of δ for the FLC image and the Wuhan TM image, respectively. The range of generation is from 8 to 100, considering that the local search plays a much more important role in the later stage than in the initial stage. Hence, the variation of Jm caused by the local search can be more distinct. As can be seen from Figs. 12(b) and 13(b), the biggest Jm increment appears when $\delta = 0.1$ and $\delta = 1$ for the FLC image and Wuhan TM image, respectively, meaning that the GLS functions much better in the optimization progress when $\delta = 0.1$ and $\delta = 1$, compared with the other δ values. For example, when a larger value of δ is chosen, the chance of high-quality solutions becomes less due to the sparsity of the

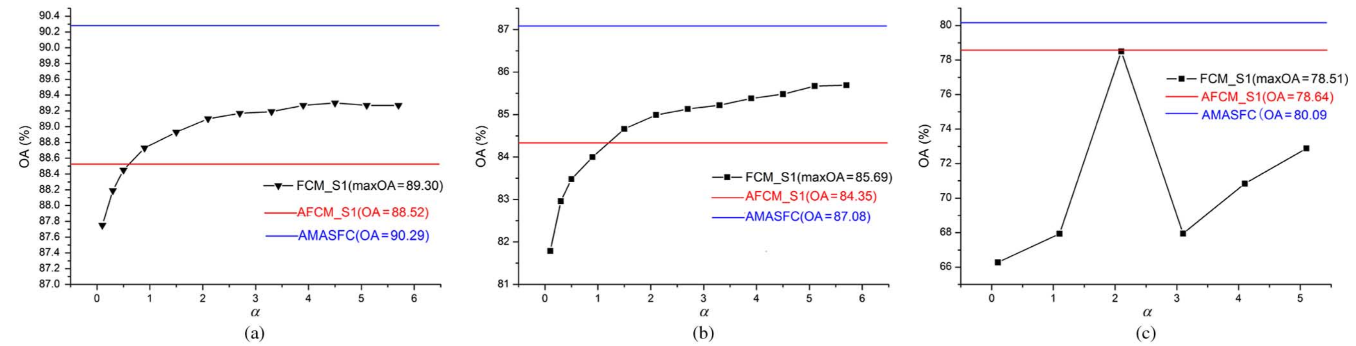


Fig. 11. Influence of α for FCM_S: (a) FLC image, (b) Wuhan TM image, and (c) Salinas AVIRIS image.

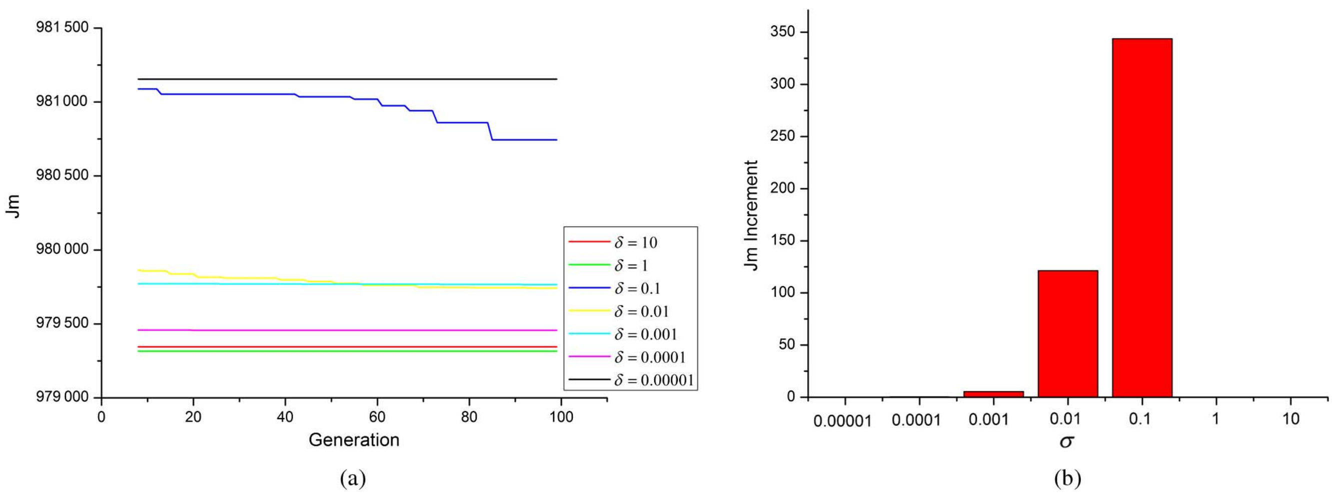


Fig. 12. Sensitivity of parameter δ in the GLS for the FLC image: (a) the variation of Jm for different values of δ and (b) the Jm increment for different values of δ.

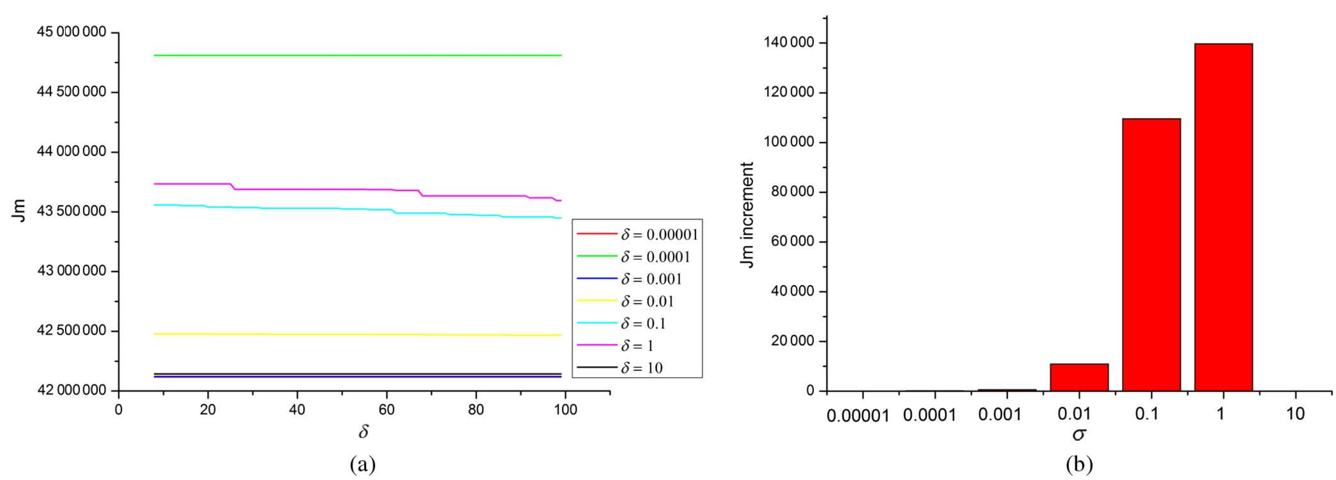


Fig. 13. Sensitivity of parameter δ in the GLS for the Wuhan TM image: (a) the variation of Jm for different values of δ and (b) the Jm increment for different values of δ.

solutions. On the other hand, if a smaller value of δ is chosen, the performance improvement of the solutions is limited. Hence, when $\delta = 0.1$ and $\delta = 1$ for the FLC image and the Wuhan TM image, respectively, the largest Jm increment can be achieved. In the same way, the optimal value of δ can be obtained, which is $\delta = 1$ for the Salinas AVIRIS image. As can be seen from the results, in most cases, the different images have different values of the parameter δ.

F. The Impact of Different Types of Areas

Fig. 14(f) shows the trade-off parameter map after density slicing. After the density slicing, four levels of trade-off parameter value are generated, which are marked with different colors. As can be seen from Fig. 14(f), the homogeneous areas such as river and lake have smaller trade-off parameter values, indicating that the pixels in these areas do not need to consider much spatial

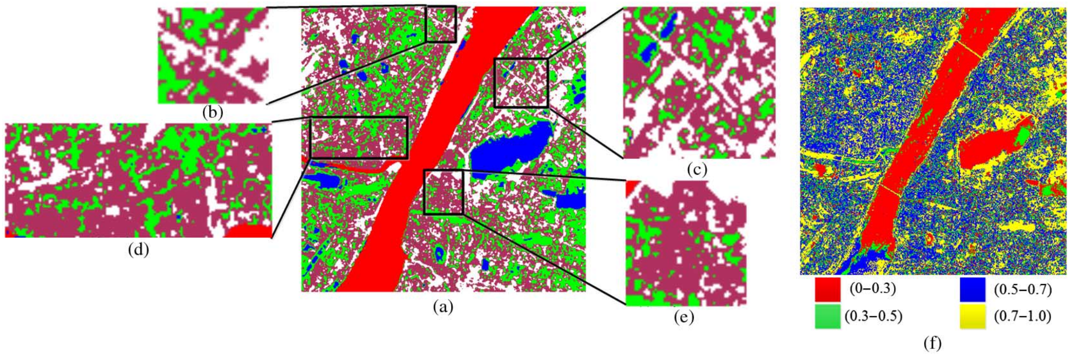


Fig. 14. Influence of smoothing: (a) the clustering result of Wuhan TM image, (b) road 1, (c) road 2, (d) urban area 1, (e) urban area 2, and (f) trade-off parameter after density slice.

information. The heterogeneous areas such as the urban areas are constrained with more spatial information. Furthermore, as can be seen from Fig. 14(b) and (c), the basic structure of the roads can be preserved. On the other hand, due to more consideration of the spatial constraint, the urban areas in Fig. 14(d) and (e) have fewer isolated pixels, which is more in line with the real situation when compared with the results of FCM.

V. Conclusion

This paper proposes an adaptive fuzzy clustering algorithm with spatial information for remote sensing imagery, namely AFCM_S1, which defines an objective function to adaptively determine the trade-off parameter. An adaptive memetic fuzzy clustering algorithm with spatial information for remote sensing imagery, namely AMASFC, is also proposed. In AMASFC, a memetic algorithm is used to further enhance the clustering performance by introducing a local search method, namely GLS. The parameter δ in GLS, which determines the local search efficiency, is obtained by comparing the Jm increment for different values of δ. The experimental results confirm the efficiency of the proposed method. Although the performance improvements in clustering precision are not so impressive, the proposed methods can adaptively determine the weight of contribution from the spatial neighborhood. The proposed methods are, therefore, more convenient and automated than the traditional clustering algorithms.

The experimental data in this paper are relatively simple. Therefore, in our future work, more complex scenarios will be considered, which can be found in many real remote sensing situations such as those in [42] and [43], in which complex, multiple-class clustering with subtle spectral differences is addressed. In addition, the acquisition of the ground truth of the remote sensing image with many classes is still a problem, which results in the difficult subject in the process of clustering results assessment. It is an alternative to specially develop a new index suitable for the proposed method such as in [1], which will be one of our future works.

Acknowledgments

The authors would like to thank the editor, associate editor, and anonymous reviewers for their helpful comments and suggestions, and also thank the research group supervised by Prof. D. Landgrebe, Purdue University, West Lafayette, IN, USA, and Computational Intelligence Group, Basque University (UPV/EHU) for providing the free downloads of the FLC and AVIRIS images.

References

[1] K. Ta demir and E. Merényi, "A validity index for prototype-based clustering of data sets with complex cluster structures," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 41, no. 4, pp. 1039–1053, Aug. 2011.

[2] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognit. Lett.*, vol. 31, no. 8, pp. 651–666, Aug. 2010.

[3] U. Maulik and I. Saha, "Modified differential evolution based fuzzy clustering for pixel classification in remote sensing imagery," *Pattern Recognit.*, vol. 42, no. 9, pp. 2135–2149, Sep. 2009.

[4] S. Bandyopadhyay, U. Maulik, and A. Mukhopadhyay, "Multiobjective genetic clustering for pixel classification in remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1506–1511, May 2007.

[5] A. Ghosh, N. S. Mishra, and S. Ghosh, "Fuzzy clustering algorithms for unsupervised change detection in remote sensing images," *Inf. Sci.*, vol. 181, no. 4, pp. 699–715, Feb. 2011.

[6] S. Niazmardi, S. Homayouni, and A. Safari, "An improved FCM algorithm based on the SVDD for unsupervised hyperspectral data classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 1939–1404, Feb. 2013.

[7] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.

[8] G. H. Ball and D. J. Hall, "ISODATA, a novel method of data analysis and pattern classification," Stanford Res. Inst., Menlo Park, CA, USA, Tech. Rep. AD0699616, 1965.

[9] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms.* New York, NY, USA: Plenum, 1981.

[10] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty, "A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data," *IEEE Trans. Med. Imag.*, vol. 21, no. 3, pp. 193–199, Mar. 2002.

[11] S. Chen and D. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 34, no. 4, pp. 1907–1916, Aug. 2004.

[12] S. Krinidis and V. Chatzis, "A robust fuzzy local information c-means clustering algorithm," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1328–1337, May 2010.

[13] L. Groll and J. Jakel, "A new convergence proof of fuzzy c-means," *IEEE Trans. Fuzzy Syst.*, vol. 13, no. 5, pp. 717–720, Oct. 2005.

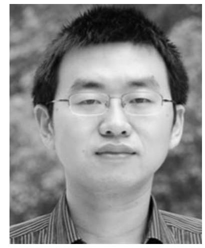
[14] L. N. De Castro and F. J. Von Zuben, "Learning and optimization using the clonal selection principle," *IEEE Trans. Evol. Comput.*, vol. 6, no. 3, pp. 239–251, Jun. 2002.

[15] S. Bandyopadhyay, "Genetic algorithms for clustering and fuzzy clustering," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 1, no. 6, pp. 524–531, Nov./Dec. 2011.

[16] S. Das, A. Abraham, and A. Konar, "Automatic clustering using an improved differential evolution algorithm," *IEEE Trans. Syst. Man, Cybern. A*, vol. 38, no. 1, pp. 218–237, Jan. 2008.

[17] U. Maulik and I. Saha, "Automatic fuzzy clustering using modified differential evolution for image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3503–3510, Sep. 2010.

[18] Y. Zhong and L. Zhang, "A new fuzzy clustering algorithm based on clonal selection for land cover classification," *Math. Probl. Eng.*, vol. 2011, pp. 1–21, 2011, article ID 708459.
[19] J. Senthilnath, S. Omkar, V. Mani, N. Tejovanth, P. Diwakar, and B. Shenoy, "Hierarchical clustering algorithm for land cover mapping using satellite images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 762–768, Jun. 2012.
[20] P. Moscato, "On evolution, search, optimization, genetic algorithms and martial arts: Towards memetic algorithms," Caltech Concurrent Computation Program, California Inst. Technol., Pasadena, CA, Tech. Rep. 826, 1989.
[21] Y. S. Ong, M. Lim, and X. Chen, "Memetic computation-past, present & future," *IEEE Comput. Intell. Mag.*, vol. 5, no. 2, pp. 24–31, May 2010.
[22] X. Chen, Y. S. Ong, M. H. Lim, and K. C. Tan, "A multi-facet survey on memetic computation," *IEEE Trans. Evol. Comput.*, vol. 15, no. 5, pp. 591–607, Oct. 2011.
[23] K. Tang, Y. Mei, and X. Yao, "Memetic algorithm with extended neighborhood search for capacitated arc routing problems," *IEEE Trans. Evol. Comput.*, vol. 13, no. 5, pp. 1151–1166, Oct. 2009.
[24] S. Goudos, K. Gotsis, K. Siakavara, E. Vafiadis, and J. Sahalos, "A multi-objective approach to subarrayed linear antenna arrays design based on memetic differential evolution," *IEEE Trans. Antennas Propag.*, vol. 61, no. 6, pp. 3042–3052, Jun. 2013.
[25] Z. Zhu, Y. S. Ong, and M. Dash, "Markov blanket-embedded genetic algorithm for gene selection," *Pattern Recognit.*, vol. 40, no. 11, pp. 3236–3248, Nov. 2007.
[26] Z. Zhu, Y. S. Ong, and M. Dash, "Wrapper–filter feature selection algorithm using a memetic framework," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 37, no. 1, pp. 70–76, Feb. 2007.
[27] Z. Zhu, S. Jia, and Z. Ji, "Towards a memetic feature selection paradigm," *IEEE Comput. Intell. Mag.*, vol. 5, no. 2, pp. 41–53, May 2010.
[28] L. Jiao, M. Gong, S. Wang, B. Hou, Z. Zheng, and Q. Wu, "Natural and remote sensing image segmentation using memetic computing," *IEEE Comput. Intell. Mag.*, vol. 5, no. 2, pp. 78–91, May 2010.
[29] Z. Zhu, Z. Ji, and S. Jia, "Memetic ant colony optimization for band selection of hyperspectral imagery classification," in *Proc. 2010 Chin. Conf. Pattern Recognit.* (CCPR), 2010, pp. 1–6.
[30] Y. Zhong and L. Zhang, "Adaptive multi-objective sub-pixel mapping framework based on memetic algorithm for hyperspectral remote sensing imagery," in *Proc. 22th Congr. Int. Soc. Photogramm. Remote Sens.* (ISPRS), Melbourne, Australia, 2012, pp. 1–6.
[31] R. Storn and K. Price, "Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, Dec. 1997.
[32] Y. Zhong and L. Zhang, "Remote sensing image subpixel mapping based on adaptive differential evolution," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 42, no. 5, pp. 1306–1329, Oct. 2012.
[33] S. Das and S. Sil, "Kernel-induced fuzzy clustering of image pixels with an improved differential evolution algorithm," *Inf. Sci.*, vol. 180, no. 8, pp. 1237–1256, Apr. 2010.
[34] A. Ghosh, A. Datta, and S. Ghosh, "Self-adaptive differential evolution for feature selection in hyperspectral image data," *Appl. Soft Comput.*, vol. 13, no. 4, pp. 1969–1977, Apr. 2013.
[35] F. Neri, C. Cotta, and P. Moscato, *Handbook of Memetic Algorithms.* New York, NY, USA: Springer, 2011.
[36] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.
[37] J. Brest, S. Greiner, B. Boskovic, M. Mernik, and V. Zumer, "Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems," *IEEE Trans. Evol. Comput.*, vol. 10, no. 6, pp. 646–657, Dec. 2006.
[38] R. Gämperle, S. D. Müller, and P. Koumoutsakos, "A parameter study for differential evolution," *Adv. Intell. Syst. Fuzzy Syst. Evol. Comput.*, vol. 10, pp. 293–298, 2002.
[39] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, 1960.
[40] J. L. Fleiss, in *Statistical Methods for Rates and Proportions*, 2nd ed. New York, NY, USA: Wiley, 1981.
[41] [Online]. Available: http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes.
[42] K. Ta demir and E. Merényi, "Cluster analysis in remote sensing spectral imagery through graph representation and advanced SOM visualization," in *Proc. 11th Int. Conf. Discovery Sci.* (DS-2008), vol. 5255, (Lect. Notes Comput. Sci.), Budapest, Hungary, Oct. 13–16, 2008, pp. 272–283.
[43] E. Merényi, B. Csathó, and K. Ta demir, "Knowledge discovery in urban environments from fused multi-dimensional imagery," in *Proc. 4th IEEE GRSS/ISPRS Joint Workshop Remote Sens. Data Fusion over Urban Areas* (URBAN 2007), Paris, France, Apr. 11–13, 2007, pp. 1–13.



Yanfei Zhong (M'11) received the B.S. degree in information engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

He has been with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University since 2007 and is currently a Professor. His research interests include multi- and hyperspectral remote sensing data processing, high resolution image processing and scene analysis, and computational intelligence. He has published more than 60 research papers, including more than 25 peer-reviewed articles in international journals such as IEEE Transactions on Geoscience and Remote Sensing and IEEE Transactions on Systems, Man and Cybernetics Part B, and Pattern Recognition.

Dr. Zhong was the recipient of the National Excellent Doctoral Dissertation Award of China (2009) and New Century Excellent Talents in University of China (2009). He was a Referee of IEEE Transactions on Cybernetics, IEEE Transactions on Geoscience and Remote Sensing, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Geoscience and Remote Sensing Letters, and Pattern Recognition.

Ailong Ma received the B.S. degree from China University of Petroleum, Qingdao, China, in 2010. He is currently working toward the Ph.D. degree in photogrammetry and remote sensing with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China.

His major research interests include remote sensing image processing and evolutionary computing.

Liangpei Zhang (M'06–SM'08) received the B.S. degree in physics from Hunan Normal University, ChangSha, China, in 1982, the M.S. degree in optics from Xi'an Institute of Optics and Precision Mechanics of Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He is currently the Head of the Remote Sensing Division with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University. He is also a "Chang-Jiang Scholar" Chair Professor appointed by the Ministry of Education, China. He is currently the Principal Scientist for the China State Key Basic Research Project (2011–2016) appointed by the Ministry of National Science and Technology of China to lead the remote sensing program in China. He has more than 300 research papers and is the holder of five patents. His research interests include hyperspectral remote sensing, high resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang is an Executive Member (Board of Governor) of the China National Committee of International Geosphere–Biosphere Programme. He also serves as an Associate Editor of *International Journal of Ambient Computing and Intelligence*, *International Journal of Image and Graphics*, *International Journal of Digital Multimedia Broadcasting*, *Journal of Geo-spatial Information Science*, and the *Journal of Remote Sensing*. He regularly serves as a Co-chair of the series SPIE Conferences on multispectral image processing and pattern recognition, Conference on Asia Remote Sensing, and many other conferences. He edits several conference proceedings, issues, and the Geoinformatics Symposiums. He is currently serving as an Associate Editor for the IEEE Transactions on Geoscience and Remote Sensing. He is a fellow of the Institution of Electrical Engineers, an executive Member for the China Society of Image and Graphics, and others.