A Fuzzy-Statistics-Based Affinity Propagation Technique for Clustering in Multispectral Images

Chen Yang, Lorenzo Bruzzone, *Fellow, IEEE*, Fengyue Sun, Laijun Lu, Renchu Guan, and Yanchun Liang

***Abstract*—Due to a high number of spectral channels and a large information quantity, multispectral remote-sensing images are difficult to be classified with high accuracy and efficiency by conventional classification methods, particularly when training data are not available and when unsupervised clustering techniques should be considered for data analysis. In this paper, we propose a novel image clustering method [called fuzzy-statistics-based affinity propagation (FS-AP)] which is based on a fuzzy statistical similarity measure (FSS) to extract land-cover information in multispectral imagery. AP is a clustering algorithm proposed recently in the literature, which exhibits a fast execution speed and finds clusters with small error, particularly for large datasets. FSS can get objective estimates of how closely two pixel vectors resemble each other. The proposed method simultaneously considers all data points to be equally suitable as initial exemplars, thus reducing the dependence of the final clustering from the initialization. Results obtained on three kinds of multispectral images (Landsat-7 ETM+, Quickbird, and moderate resolution imaging spectroradiometer) by comparing the proposed technique with K-means, fuzzy K-means, and AP based on Euclidean distance (ED-AP) demonstrate the good efficiency and high accuracy of FS-AP.**

***Index Terms*—Affinity propagation (AP), clustering, fuzzy clustering, fuzzy sets, fuzzy statistical similarity measure (FSS), image classification, unsupervised classification.**

I. Introduction

CLUSTERING techniques can be used in unsupervised classification to partition multispectral (and hyperspectral) feature spaces for extracting clusters of patterns that can be associated with land-cover classes [1]–[3]. As, in clustering, training data are not available (unlike supervised classification, clustering is an ill-posed problem), a common approach is to use data to learn a set of centers such that the sum of squared errors between data points and their nearest centers is small. When the centers are selected from actual data points, they are called "exemplars" [4]. In order to mathematically identify clusters in a dataset, it is usually necessary to first define a measure of similarity [5], which establishes a rule for assigning patterns to the domain of a particular cluster center (or exemplar). This similarity measure places similar data close to one another to form a group, thus generating different clusters [6].

One of the most popular clustering algorithms is the K-means, which was developed by MacQueen in 1967 [7]. It exploits the Euclidean distance as similarity measure and begins with an initial set of randomly selected exemplars. Then, it iteratively refines this set to decrease the sum of squared errors. K-means algorithm has many advantages such as simplicity, low computational complexity, etc. [8]. Therefore, it is widely used in remote sensing [9]–[11]. However, K-means is sensitive to initialization [8], [12] and to the choice of the number of clusters, which usually is a critical issue. Different random initializations of the cluster centers result in significantly different clusters at the convergence. Thus, the algorithm is usually run many times with different initializations in an attempt to find a good solution [13]. In addition, K-means is prone to find clusters with spherical shape, and it is sensitive to noisy data [14]. Most important, K-means algorithm is a crisp clustering method which restricts each point of the data to be associated with only one cluster.

In remote-sensing images, depending on both the spatial resolution of the sensor and the considered scene, a pixel can represent a mixture of land covers that cannot be properly described by a single class [15]. A fuzzy approach to data classification is more suitable in managing both the uncertainty intrinsic in the classification problem and the relation one-to-many of a pattern with the related information classes [16], [17]. Since a single class cannot describe these patterns, fuzzy clustering has been developed. Fuzzy K-means algorithm (proposed by Dunn [18] and extended by Bezdek [19]) is the extension of crisp K-means. It has been shown to have a better performance than K-means due to its ability to deal with uncertain situations. One of the most significant advantages of fuzzy K-means is that it more naturally handles situations in which subclasses are formed by mixing (or interpolating) extreme examples. Fuzzy K-means has been widely used in remote sensing, and many algorithms are derived from it [20]–[25]. However, these foregoing algorithms have similar drawbacks when used in remote-sensing imagery analysis. First, the use of Euclidean distance as a measure of similarity is not very suitable for remote-sensing clustering, because the scatter diagram of multispectral remote-sensing data tends to

Manuscript received October 25, 2009. Date of publication March 1, 2010; date of current version May 19, 2010. This work was supported in part by the "211" Project-Evolution of the Lithosphere and Accumulation of Mineral Resources in Northeast Asia, by the 11th Five-Year Plan for National Land and Resources-Middle and Large Scale Mineral Resources Prediction and Programming in Baishan, by the Graduate Innovation Fund Project (20091024) of Jilin University, and by the National Science Foundation of China under Grants 60673023 and 10872077. This work was developed during the stay of C. Yang at the University of Trento, Italy.

C. Yang is with the College of Earth Sciences, Jilin University, Changchun 130061, China, and also with the College of Computer Science and Technology, Changchun Normal University, Changchun 130000, China (e-mail: yangc_616@yahoo.com.cn).

L. Bruzzone is with the Department of Information Engineering and Computer Science, University of Trento, 38123 Trento, Italy (e-mail: lorenzo.bruzzone@ing.unitn.it).

F. Sun and L. Lu are with the College of Earth Sciences, Jilin University, Changchun 130061, China (e-mail: fengyuesuncc@yahoo.com; laijun_lu@yahoo.com.cn).

R. Guan and Y. Liang are with the College of Computer Science and Technology, Jilin University, Changchun 130012, China (e-mail: guan_renchu@yahoo.com.cn; ycliang@jlu.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/TGRS.2010.2040035

hyperellipsoid distributions in the feature space, owing to uncertainty and existence of mixed pixels. This can be mitigated by considering kernelized versions of fuzzy K-means, known as the kernel fuzzy K-means [26]. Second, initial centers of K-means and fuzzy K-means are defined randomly, thus leading to unstable clustering results and requiring multiple trials for obtaining reasonable results. This has a considerable trouble in a noisy environment and inaccuracy with a large number of different sample sized clusters [27]. Gath *et al.* [28] proposed an improved version of fuzzy K-means called unsupervised fuzzy partition-optimal number of classes. This method combines fuzzy techniques with statistical algorithms to obtain reliable clusters. Nasser *et al.* [29] proposed a clustering algorithm that combines fuzzy K-means and expectation-maximization algorithm. Li *et al.* [30] presented an agglomerative fuzzy K-means clustering algorithm by introducing a penalty term to the objective function to make the clustering process insensitive to the initial cluster centers. It is worth noting that clustering precision of the algorithm is affected by its equal partition trend for datasets. Optimized clustering results of the fuzzy K-means algorithm might not be a right partition of the feature space when datasets have a large discrepancy in the number of class samples [23].

Remote-sensing data are general reflection of the spatial characteristics of ground objects. Statistical pattern recognition is the most common approach used in the classification of multispectral and hyperspectral remote-sensing data [31]. When unsupervised clustering methods are used for data analysis, results are affected by uncertainty, and multiple reliable solutions can be obtained. After introducing and developing fuzzy set theory, many studies have been carried out to combine statistical methods and fuzzy set theory. These studies, called fuzzy statistics, have been developed in several branches [32]. Fuzzy set theory is the basis in studying membership relationships from the fuzziness of the phenomena. Fuzzy statistics is used to estimate the degree of membership of a pattern to a class according to the use of a membership function. Multispectral and hyperspectral remote-sensing images consist of multiple channels and a large quantity of data. These data usually have a complex structure, which results in time-consuming and slow convergence rate in the clustering process. Therefore, the development of efficient and fast fuzzy clustering algorithms for multispectral and hyperspectral remote-sensing imagery is an important topic of current research.

Recently, Frey and Dueck [4] published a paper in *Science* where they describe an algorithm, called affinity propagation (AP), that clusters data points based on similarity measures and considers that all data points can be equally suitable as exemplars. This algorithm aims to find several exemplars such that the sum of the similarities between the data points and the corresponding exemplars is maximized. There are two kinds of messages communicated between data points, namely, responsibility and availability, and each takes a different kind of competition into account. In [4], the authors first describe the algorithm and then apply it to some different examples of clustering problems from diverse fields [4], [33]–[35]. They argue that AP finds clusters with much lower error than other methods by requiring less than one-hundredth the amount of time of standard clustering algorithms, which slashes computing times while keeping accuracy. It performs well, particularly in face image recognition, gene finding, etc. However, it has not been used in remote sensing and on multispectral images. In this context, unless a meaningful measure of similarity between pairs of points has been established, no meaningful cluster analysis is possible. Aiming to further solve this knotty problem, in this paper, an improved similarity measure integrating fuzzy statistics with AP is proposed, which is called fuzzy-statistics-based AP (FS-AP). The proposed method considers that all data points can be equally suitable as initial exemplars. First, according to the characteristics of multispectral images, we propose fuzzy mean deviation and then develop a fuzzy statistical similarity measure (FSS) in evaluating the similarity between two pixel vectors. We iteratively merge cluster centers to extract land-cover information by FSS. Experimental results show that our method can improve not only the clustering accuracy but also the computational efficiency, compared with the standard K-means, the fuzzy K-means algorithms, and the AP method presented in [4].

This paper is organized as follows. Section II reviews some important related researches about similarity measures and briefly introduces the AP algorithm and fuzzy schemes used in remote-sensing data clustering. Section III presents the fuzzy mean deviation and the FSS used in this paper. Then, the proposed FS-AP based on the presented similarity measure is introduced. Section IV illustrates the experimental results obtained on three different multispectral images and analyzes the accuracy and efficiency of the proposed method. Finally, Section V draws the conclusion of this paper.

II. Background

A. Similarity Measures for Clustering

For remote-sensing images, clustering implies a grouping of pixels in a multidimensional space. Pixels belonging to a particular cluster are therefore spectrally similar. In order to quantify this relationship, it is necessary to define a similarity measure. Many similarity metrics have been proposed in the literature, but those commonly used in clustering procedures are usually simple distance measures in a multidimensional space [31]. Nowadays, some similarity measures, including Euclidian distance, spectral angle, correlation coefficient, spectral information divergence, encoding and matching, and others, are used, and each has its advantages and disadvantages.

The most frequently encountered similarity measure is the Euclidean distance. By using Euclidean distance, hyperspherical-shaped clusters of equal size are usually detected. This measure is not very useful or even undesirable when clusters tend to develop along principal axes [36]. In addition, Euclidean distance primarily measures overall brightness differences but does not respond to the correlation between two pixels of the spectra [37]. To take care of hyperellipsoidal-shaped clusters, the Mahalanobis distance is one of the most popular choices. One of the major difficulties associated with using the Mahalanobis distance as a similarity measure is that we have to recompute the inverse of the sample covariance matrix each time that a pattern changes its cluster domain, which is computationally expensive.

The correlation coefficient is very responsive to differences in direction (i.e., spectral shape), but it does not respond to brightness differences due to band-independent gain or offset

factors. Spectral angle [38] is closely related mathematically to the correlation coefficient and is primarily responsive to differences in spectral shape. However, spectral angle does respond to brightness differences due to a uniform offset, which confounds the interpretation of the spectral angle value [37].

B. AP

AP [4] was proposed as a new technique for exemplar learning. It takes input measures as similarity between pairs of data points. In contrast to K-means and fuzzy K-means techniques, AP operates by simultaneously considering all data points as potential exemplars and iteratively exchanging messages between data points until a good set of exemplars and clusters emerges.

Let $s(x_i, x_k)$ be the similarity between points x_i and x_k, i.e., the suitability of point x_i to serve as the exemplar for data point x_k (shorten as i and k). In conventional AP, a common choice for similarity is the negative Euclidean distance

$$s(i,k) = -\|i-k\|^2.$$

AP can be applied by using more general notions of similarity, and the similarities may be positive or negative.

The preference of point k, called $s(k,k)$, is the *a priori* suitability of point k to serve as an exemplar. Preferences can be set to a global (shared) value or customized for particular data points. High values of the preferences will cause AP to find many exemplars (clusters), while low values will lead to a small number of exemplars (clusters). A good initial choice for the preference is the minimum similarity or the median similarity.

The AP algorithm is as follows [4], [39]:

Initialization:

$$r(i,k) = s(i,k) - \max\{s(i,k')\}, a(i,k) = 0, \qquad k' \neq k. \quad (1)$$

Responsibility updates:

$$r(i,k) = s(i,k) - \max_{k' s.t. k' \neq k} \{a(i,k') + s(i,k')\} \quad (2)$$

$$r(k,k) = s(k,k) - \max_{k' s.t. k' \neq k} \{a(k,k') + s(k,k')\}. \quad (3)$$

Availability updates:

$$a(i,k) = \min\left(0, r(k,k) + \sum_{i' s.t. i' \notin \{i,k\}} \max\{0, r(i',k)\}\right) \quad (4)$$

$$a(k,k) = \sum_{i' s.t. i' \neq k} \max(0, r(i',k)). \quad (5)$$

Making assignments:

$$c_i^* \leftarrow \arg\max_{1 \le k \le n} r(i,k) + a(i,k).$$

In the processing, two kinds of messages are exchanged among data points, and each takes into account a different kind of competition. The "responsibility" $r(i,k)$, sent from data point i to candidate exemplar point k, indicates how well suited point i would be as a member of the candidate exemplar point k. The "availability" $a(i,k)$, sent from candidate exemplar point k to potential cluster members point i, indicates how the cluster k would represent point i. Responsibilities and availabilities are initialized as (1), and in the whole process, they follow the updating rules (2)–(5). Messages are updated on the basis of simple equations searching for minima of an appropriately chosen energy function. However, computing responsibilities and availabilities according to simple updating rules often lead to oscillations caused by "overshooting" the solution, so the responsibilities and availability messages are "damped" according to the following equation:

$$R^{t+1} = \alpha R^{t-1} + (1-\alpha) R^t$$
$$A^{t+1} = \alpha A^{t-1} + (1-\alpha) A^t \quad (6)$$

where R and A represent the responsibility and availability vectors, respectively; α is the factor of damping, which should be satisfied at $0.5 \le \alpha < 1$; and t is the number of iterations. A higher α will lead to a slower convergence.

At any time during the clustering, the magnitude of each message reflects the current affinity of a data point to choose another data point as its exemplar. For point i, if point k ($k \neq i$) maximizes $r(i,k) + a(i,k)$, then k would be considered as the exemplar of i, whereas $k = i$ means that point i itself is an exemplar.

The message-passing procedure may be terminated after a fixed number of iterations, after changes in the messages are less than a threshold, or after the local decisions stay constant for some iterations.

C. Fuzzy Schemes for Remote-Sensing Image Clustering

Fuzziness is an intrinsic characteristic of remote-sensing imagery. Probabilistic clustering techniques use the concept of memberships to describe the degree by which a vector belongs to a cluster. The use of memberships provides probabilistic methods with more realistic clustering than hard or crisp techniques. In conventional fuzzy classification, pixels can belong to several classes with different degrees of membership, which is the case when class descriptions overlap, e.g., in the presence of mixed pixels. Pixels whose feature values are within these overlapping ranges can be seen as ambiguous pixels. Although fuzzy concepts make it possible to describe these ambiguities, the main aim of each classification is to define classes as unambiguously as possible.

Conventional fuzzy clustering, like the fuzzy K-means, needs the given cluster numbers, and the clustering results strongly depend on the initial sequence of samples. There are two main deficiencies associated with fuzzy K-means, namely, inability to distinguish outliers from nonoutliers by weighing the memberships and attraction of the centroid toward the outliers [40]. Both deficiencies together are referred to as "noise sensitivity." Moreover, conventional fuzzy schemes are based on maximum and minimum paradigms. Most of the cluster analysis results are thus easy to trap in local optimizations which increase randomicity and cause difficulty in getting accurate results.

III. Proposed Clustering Method

A. FSS

Fuzzy statistics is a subject based on the combination of fuzzy set theory and statistical methods. Fuzzy set theory is the basis in studying membership relationships from the fuzziness

of the phenomena [41]. In this section, we will discuss the mathematical definition of the FSS.

Multispectral and hyperspectral remote-sensing images often have extensive interband correlations. As a result, the images may contain similar information and have similar spatial structures [42]. At the same time, multispectral and hyperspectral images have their own special characteristics, namely, the spatial variability of the spectral signature. According to this, we introduce the following statistical characteristics, which are based on fuzzy statistics.

1) Fuzzy Set: Let $X = \{x_1, x_2, \ldots, x_n\} \subset R^p$ be a set of pixels vectors, where X represents all pixels in the dataset, $x_i = [x_{i1}, x_{i2}, \ldots, x_{ip}]$ $(i = 1, 2, \ldots, n)$ is the feature vector of pixel i, n is the total number of pixels in the image, and p is the number of features considered (e.g., bands of the image). Let $Z = \{z_1, z_2, \ldots, z_m\} \subset R^p$ be a set of cluster exemplars, where Z represents all exemplars set, $z_k = [z_{k1}, z_{k2}, \ldots, z_{kp}]$ $(k = 1, 2, \ldots, m)$ is an exemplar vector, and m is the number of clustering exemplars in the image (at initialization $n = m$). $N = \{\mu_1(1), \mu_2(2), \ldots, \mu_n(m)\}$ is the membership degree set, and $\mu_i(k) = [\mu_{i1}(k), \mu_{i2}(k), \ldots, \mu_{ip}(k)]$ $(i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m)$ is a membership degree vector. The fuzzy set is defined by

$$F = \{X, Z, N\} = \{\langle x_i, z_k, \mu_i(k)\rangle\} = \{\langle x_1, z_1, \mu_1(1)\rangle, \langle x_1, z_2, \mu_1(2)\rangle, \ldots, \langle x_n, z_m, \mu_n(m)\rangle\}.$$

2) Fuzzy Mean Distance: Distance is the amount of difference between individual pixel vectors and clustering exemplars vector. Let $Dis = \{dis_1(1), dis_2(2), \ldots, dis_n(m)\}$ be a set of distance vectors values, where $dis_i(k) = [dis_{i1}(k), dis_{i2}(k), \ldots, dis_{ip}(k)]$ $(i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m)$ is a distance vector. The value of the distance is defined as

$$dis_{ij}(k) = |x_{ij} - z_{kj}|, \quad i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, p;\ k = 1, 2, \ldots, m;\ i \neq k \quad (7)$$

in which $dis_{ij}(k)$ represents the value of the distance between the ith pixel value and the kth clustering exemplar of the jth band.

Let $\overline{Dis} = \{\overline{dis}_1(1), \overline{dis}_2(2), \ldots, \overline{dis}_n(m)\}$ be a set of fuzzy mean distance vectors values, where $\overline{dis}_i(k)$ $(i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m)$ represents the fuzzy mean distance computed on all bands between the ith pixel vector and the kth cluster exemplar.

Rather than being assigned to a single class, the unknown measurement vector (pixel) now has membership grade values describing how close the pixel is to the clustering exemplars. Fuzzy mean distance is computed using the membership degree. It represents the distance between two pixel vectors and is defined as

$$\overline{dis}_i(k) = \sum_{j=1}^{p} dis_{ij}(k)\mu_{ij}(k) \Big/ \sum_{j=1}^{p} \mu_{ij}(k), \quad i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, p;\ k = 1, 2, \ldots, m;\ i \neq k \quad (8)$$

where $\mu_{ij}(k)$ $(i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, p;\ k = 1, 2, \ldots, m)$ represents the membership of the fuzzy mean distance.

3) Mean Distance Deviation and Membership Function: Deviation is a measure of difference for interval and ratio variables between the distance value and the mean. Let $Dev = \{dev_1(1), dev_2(2), \ldots, dev_n(m)\}$ be a set of distance deviation values, where $dev_i(k) = [dev_{i1}(k), dev_{i2}(k), \ldots, dev_{ip}(k)]$ $(i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m)$ is a distance deviation vector. The value of distance deviation is defined as

$$dev_{ij}(k) = \left|dis_{ij}(k) - \overline{dis}_i(k)\right|, \quad i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, p;\ k = 1, 2, \ldots, m;\ i \neq k. \quad (9)$$

Let $\overline{Dev} = \{\overline{dev}_1(1), \overline{dev}_2(2), \ldots, \overline{dev}_n(m)\}$ be a set of mean distance deviation vectors values, where $\overline{dev}_i(k)$ $(i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m)$ represents the mean distance deviation computed on all bands between the ith pixel vector and the kth cluster exemplar. The mean distance deviation of the pixels in the space is computed using

$$\overline{dev}_i(k) = \frac{1}{p}\sum_{j=1}^{p}\left|dis_{ij}(k) - \overline{dis}_i(k)\right|, \quad i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, p;\ k = 1, 2, \ldots, m;\ i \neq k. \quad (10)$$

The membership degree of the fuzzy mean distance is computed using

$$\mu_{ij}(k) = \exp\left(-dev_{ij}(k)^{\beta}/\overline{dev}_i(k)^{\beta}\right), \quad i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, p;\ k = 1, 2, \ldots, m;\ i \neq k \quad (11)$$

where β is a parameter determining the scalar of $\mu_{ij}(k)$ and ranges from $(0, \infty)$. It determines the degree of fuzziness of the final solution, which is the degree of overlapping between groups. If $\beta = 0$, the solution is a hard partition. As β becomes close to infinity, the solution approaches its highest degree of fuzziness. β is aimed to accommodate the outliers in a special class to decrease their effect on clustering.

The membership degree depends on distance deviation in the spectral space. Because the scatter diagram distribution of most remote-sensing data tends to hyperellipsoid distribution in the feature space, the membership function adopts an exponential function where the exponential part is the Mahalanobis distance based on the standard variance matrix.

4) FSS: The FSS between pixel vector $x_i = [x_{i1}, x_{i2}, \ldots, x_{ip}]$ and clustering exemplar $z_k = [z_{k1}, z_{k2}, \ldots, z_{kp}]$ can be computed using the following:

$$\text{FSS} = s(i, k) = -\overline{dis}_i(k), \quad i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m;\ i \neq k \quad (12)$$

and the preference is calculated using

$$s(i, i) = \min - \text{CTS}(\max - \min), \quad i = 1, 2, \ldots, n \quad (13)$$

where $\max$ and $\min$ are the maximum and minimum values of all $s(i, k)$ $(i \neq k)$. Cluster threshold scalar (CTS) is used to get the expected number of clusters through setting the appropriate values.

The similarity is a negative value; therefore, a small value is equivalent to a large similarity. By introducing fuzzy mean deviation into similarity measure, we can exploit fuzzy sets in decision making. Thus, FSS can take into account the difference

of the same band between pixels. When working with real remote-sensing data, the actual fuzzy partition of the spectral space is the merger based on FSS.

B. FS-AP

The algorithm proposed in the following is based on fuzzy statistics and AP and is called FS-AP. Compared with the conventional fuzzy clustering methods, it simultaneously considers all data points in the feature space to be initial clustering exemplars and iteratively refines with the mean distance deviation until getting the optimal FSS.

The procedure associated with the FS-AP is as follows.

Step 1) *Set the initial value of the exemplars and parameters.* At the beginning, we simultaneously consider all samples to be initial clustering exemplars. Therefore, all pixel vectors are set as clustering exemplars $(Z = X)$.

Step 2) *Calculate the fuzzy mean distance between the sample vector and the clustering exemplars using* (7)–(11). In practical applications, we need to get $\mu_{ij}(k)$, which can be obtained by using a fuzzy iteration method. Accordingly, the membership degree of the fuzzy mean distance can be obtained as follows:

$$\mu_{ij}^{t}(k) = \begin{cases} 1, & t = 0 \\ \exp\left(-dev_{ij}^{t}(k)^{\beta} / \overline{dev}_{i}^{t}(k)^{\beta}\right), & 0 < t \le c \end{cases}$$
$$i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, p;\ k = 1, 2, \ldots, m;\ i \ne k$$

where t is the number of iterations and $dev_{ij}^{t}(k)$ and $\overline{dev}_{i}^{t}(k)$ are defined as

$$dev_{ij}^{t}(k) = \left| dis_{ij}(k) - \overline{dis}_{i}^{t}(k) \right|$$

$$\overline{dev}_{i}^{t}(k) = \frac{1}{p} \sum_{j=1}^{p} \left| dis_{ij}(k) - \overline{dis}_{i}^{t}(k) \right|, \qquad i = 1, 2, \ldots, n;$$
$$j = 1, 2, \ldots, p;\ k = 1, 2, \ldots, n;\ i \ne k.$$

The initial fuzzy mean distance is $(\mu_{ij}^{0}(k) = 1)$

$$\overline{dis}_{i}^{0}(k) = \sum_{j=1}^{p} dis_{ij}(k) \Big/ p, \qquad i = 1, 2, \ldots, n;$$
$$j = 1, 2, \ldots, p;\ k = 1, 2, \ldots, n;\ i \ne k.$$

If $\overline{dis}_{i}^{t}(k)$ and $\overline{dis}_{i}^{t+1}(k)$ are close enough, i.e., $|\overline{dis}_{i}^{t}(k) - \overline{dis}_{i}^{t+1}(k)| < \varepsilon$, where ε is the iteration accuracy and has a predefined value, then the iteration is stopped, or else the iteration is continued.

Step 3) *Calculate the FSS for all data points according to* (12) *and* (13).

Step 4) *Update the responsibility and availability.* At the beginning, responsibilities and availabilities are initialized according to (1). Then, update responsibility and availability according to (2)–(6).

Step 5) *Identify the fuzzy cluster centers and the number of clusters.* Identify the fuzzy cluster centers by looking at the maximum value of availabilities and responsibilities. For point i, if point k $(k \ne i)$ maximizes $r(i,k) + a(i,k)$, then k is considered as the fuzzy cluster exemplar of i, or else if $k = i$, then the point i itself is considered a fuzzy cluster exemplar.

Step 6) *Convergence.* Repeat steps 4)–5) until the decisions for fuzzy cluster exemplars and cluster boundaries are unchanged for some number of iterations. Then, we can get the fuzzy cluster centers and the number of clusters.

The flowchart for FS-AP is shown in Fig. 1.

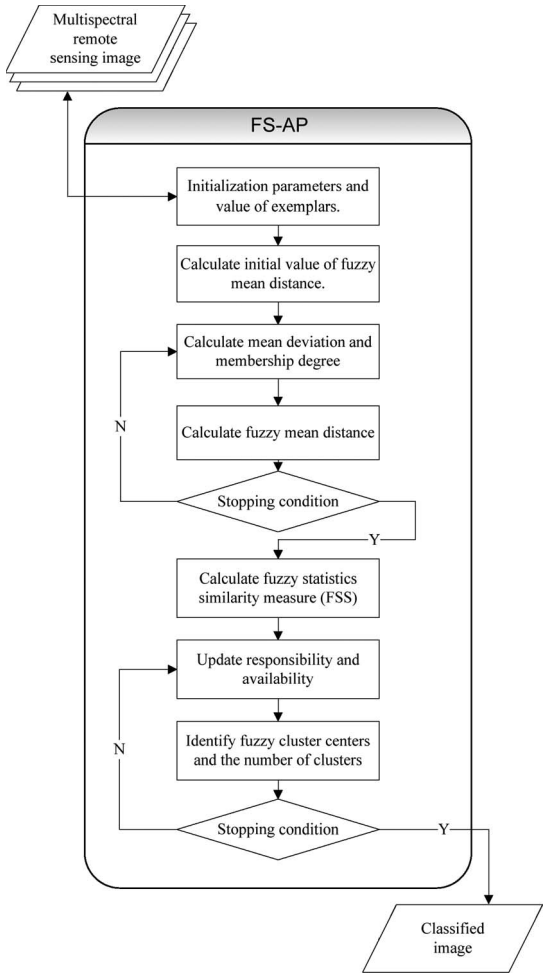

Fig. 1. Flowchart of the proposed FS-AP.

Fig. 2. Color composite (RGB = bands $7, 4, 2$) of Landsat-7 ETM+ images of the west of Haerbin, Heilongjiang, China.

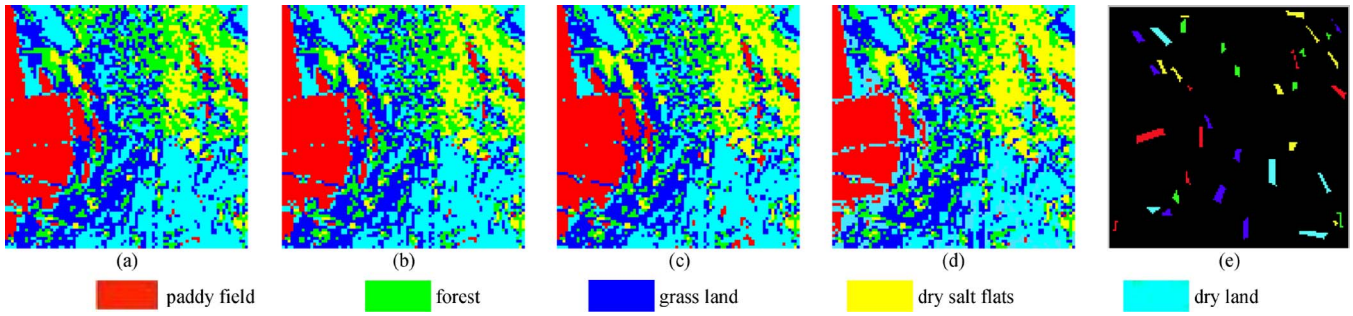

Fig. 3. Unsupervised clustering maps of the west of Haerbin, Heilongjiang, China (Landsat-7 ETM+ images). (a) K-means. (b) Fuzzy K-means. (c) ED-AP. (d) FS-AP. (e) Test field map.

IV. Experimental Results

In order to show the effectiveness of the proposed approach, three different types of multispectral images are considered to test its performance. Consistent comparisons between FS-AP and traditional unsupervised algorithms (K-means and fuzzy K-means) and the standard AP method based on the Euclidean distance (ED-AP) are carried out. The estimations of clustering accuracy provided by these algorithms are given in terms of classification accuracy by manually associating cluster labels with land-cover classes and exploiting (only for numerical validation) available ground truth information. In addition, computational efficiency for these algorithms is provided. The comparison is performed on PC workstation (Intel(R) Pentium(R) Dual E2180 2.0 GHz, 2.0 GHz with 2.0 GB of RAM).

A. Landsat-7 ETM+ Dataset

The first dataset is a portion of a Landsat-7 ETM+ multispectral image (bands 1, 2, 3, 4, 5, and 7) acquired over the west of Haerbin, Heilongjiang, China, on August 11, 2001. This site mainly contains two land-cover types, which are vegetation and exposed land. We use the displayed color composite image (Fig. 2) as a guide to make the comparison and to evaluate in a qualitative way the effectiveness of the proposed FS-AP. Even in a simple three-band image, it is easy to see that there are areas that have similar spectral characteristics. Green areas represent vegetation. Dark green areas represent dry land. Slightly darker green areas on the image usually represent forest land. Bright green areas represent grass land, and deep darker green areas represent paddy field. We can see that the degree of class mixture in the vegetation area is high. Brown areas mainly represent exposed land. Slightly brown areas represent dry salt flats which are blocked by forest land and dry land that are highly mixed too. This points out the difficulty in land-cover clustering.

For both the FS-AP and the ED-AP algorithms, we carried out the experiments with an α that is equal to 0.7 and by considering values of CTS between 2 and 15. The value of β was fixed to one in the FS-AP method. In the following, we report the results obtained by the K-means (best of 50 runs and $k = 5$), the fuzzy K-means (best of 10 runs and $k = 5$), the ED-AP $(\text{CTS} = 14, \alpha = 0.7)$, and the FS-AP $(\text{CTS} = 9, \alpha = 0.7, \beta = 1)$ techniques applied to the aforementioned images. Fig. 3(a)–(d) shows the unsupervised clustering maps provided by the four considered algorithms. To evaluate the classification accuracy, a test field map based on the ground truth data is shown in Fig. 3(e). The ETM+ image is classified into the following five clusters: paddy field, forest, grass land, dry salt flats, and dry land.

From the comparison between Figs. 2 and 3, it can be observed that the FS-AP shows better classification results than the K-means, the fuzzy K-means, and the ED-AP. In these methods, many pixels of the image are wrongly classified to other cover types.

In order to quantitatively compare the clustering performance, we used 670 reference pixels collected for Haerbin, Heilongjiang, China, on the basis of a stratified random sampling. We assigned the obtained cluster labels to land-cover classes manually. The clustering results of K-means, fuzzy K-means, ED-AP, and FS-AP are evaluated using overall accuracy, Kappa value, average of producer's accuracy [43], average of user's accuracy [43], and average of Short's mapping accuracy index [44]. Among them, average of producer's accuracy, average of user's accuracy, overall accuracy, and Kappa value are widely used in the validation of the land use/land cover classification [44]. The average of the Short's mapping accuracy index is the arithmetic mean of the Short's mapping accuracy index [45], [46] (a monotonic function of the harmonic mean of the user's and producer's accuracies) that explicitly combines both user's and producer's accuracies in one measure. This measure is supposed to be supplementary to the average of the user's accuracy and the average of the producer's accuracy.

The error matrices obtained from all the considered methods are shown in Table I. For a more detailed verification of the results, we assess the accuracy of each method using the producer's and user's accuracy measures (see Table I). As an example, for the paddy field, the producer's accuracy of the K-means, fuzzy K-means, ED-AP, and FS-AP are 80.92%, 84.73%, 92.37%, and 93.89%, respectively. FS-AP exhibits the highest producer's accuracy among the four considered methods. In the classification results of K-means, although 80.92% of the paddy field pixels are correctly identified as paddy field, only 67.09% of the areas in the cluster of the paddy field are actually paddy field. A similar situation occurred for the fuzzy K-means and the ED-AP. On the contrary, the user's accuracy of FS-AP for the paddy field is 82.55%. A careful evaluation of the error matrix also reveals that there is confusion when discriminating dry salt flats from forest in the first three methods. In other words, although the user's accuracy of K-means, fuzzy K-means, and ED-AP for dry salt flats is high, the producer's accuracy of these methods shows that some of the dry salt flat pixels are wrongly identified as forest. It can be observed that there is a significant confusion also when discriminating grass land

TABLE I
ERROR MATRICES OF THE CLASSIFICATION MAPS DERIVED FROM THE WEST OF HAERBIN, HEILONGJIANG, CHINA (LANDSAT-7 ETM+ DATA)

Methods	Class	Paddy field	Forest	Grass land	Dry salt flats	Dry land	Row total
K-means	Paddy field	**106**	4	17	6	20	153
	Forest	8	**41**	9	34	3	95
	Grass land	12	19	**99**	13	31	174
	Dry salt flats	0	2	5	**73**	0	80
	Dry land	5	7	25	0	**131**	168
	Column total	131	73	155	126	185	**450**
	Producer's accuracy (%)	80.92	56.16	63.87	57.94	70.81	
	User's accuracy (%)	67.09	43.16	56.90	91.25	77.98	
Fuzzy K-means	Paddy field	**111**	3	18	3	18	153
	Forest	6	**43**	8	29	2	88
	Grass land	10	17	**101**	10	27	165
	Dry salt flats	0	4	7	**84**	0	95
	Dry land	4	6	21	0	**138**	169
	Column total	131	73	155	126	185	**477**
	Producer's accuracy (%)	84.73	58.90	65.16	66.67	74.60	
	User's accuracy (%)	72.55	48.86	61.21	88.42	81.66	
ED-AP	Paddy field	**121**	6	21	7	24	179
	Forest	3	**45**	4	23	1	76
	Grass land	6	14	**105**	6	17	148
	Dry salt flats	0	3	9	**90**	0	102
	Dry land	1	5	16	0	**143**	165
	Column total	131	73	155	126	185	**504**
	Producer's accuracy (%)	92.37	61.64	67.74	71.43	77.30	
	User's accuracy (%)	67.60	59.21	70.95	88.24	86.67	
FS-AP	Paddy field	**123**	3	9	1	13	149
	Forest	2	**52**	5	8	2	69
	Grass land	4	10	**116**	6	15	151
	Dry salt flats	0	6	8	**111**	0	125
	Dry land	2	2	17	0	**155**	176
	Column total	131	73	155	126	185	**557**
	Producer's accuracy (%)	93.89	71.23	74.84	88.10	83.78	
	User's accuracy (%)	82.55	75.36	76.82	88.80	88.07	

TABLE II
PERFORMANCE OF K-MEANS, FUZZY K-MEANS, ED-AP, AND THE PROPOSED FS-AP (LANDSAT-7 ETM+ DATA)

Parameter	K-means	Fuzzy K-means	ED-AP	FS-AP
Execution numbers	50	10	1	1
Execution time (min)	76	113	68	57
Overall accuracy (%)	67.16	71.19	75.22	83.13
Kappa value	0.583	0.634	0.686	0.785
Average of producer's accuracy (%)	65.94	70.01	74.10	82.37
Average of user's accuracy (%)	67.28	70.54	74.53	82.32
Average of Short's mapping accuracy index	0.498	0.542	0.588	0.702

from other land covers, particularly for the K-means, the fuzzy K-means, and the ED-AP. The FS-AP recognizes the grass land better than the other three methods. The producer's and user's accuracies of the K-means and fuzzy K-means for dry land are low because the mixed pixels in this class are not distinguished accurately. Although the producer's and user's accuracies of the ED-AP are higher than those of the K-means and fuzzy K-means, the producer's and user's accuracies provided by the FS-AP for this class are further increased in some degrees.

Table II shows the global performance in terms of execution time and classification accuracy yielded by the K-means, fuzzy K-means, ED-AP, and the proposed FS-AP. From the table, we can observe that FS-AP exhibits the highest overall accuracy and Kappa value, with a gain of the overall accuracy of 15.97%, 11.94%, and 7.91% over the K-means, the fuzzy K-means, and the ED-AP, respectively. This behavior is confirmed by the other quality indices considered.

Fig. 4. Color composite (RGB = bands $1, 2, 3$) of the Quickbird data of the south part of the city of Trento, Italy.

B. Quickbird Dataset

The second dataset used in our experiments is a portion of a multispectral Quickbird image, which covers a small area of the south part of the city of Trento, Italy (see Fig. 4). This image was acquired on July 17, 2006. Seven land-cover classes, i.e., agricultural field, road, tree, soil, roof, shadow, and grass, characterize this image.

For both the FS-AP and the ED-AP algorithms, we carried out the experiments with an α that is equal to 0.9 and by considering values of CTS between 8 and 42. The value of β was fixed to 0.5 in the FS-AP method.

Fig. 5(a)–(d) shows the unsupervised classification maps obtained by using the K-means (best of 80 runs and $k = 7$), the fuzzy K-means (best of 5 runs and $k = 7$), the ED-AP (CTS $= 40, \alpha = 0.9$), and the FS-AP (CTS $= 38, \alpha = 0.9, \beta = 0.5$). To evaluate the classification accuracy, a test field map is provided in Fig. 5(e) based on the ground truth data. A manual

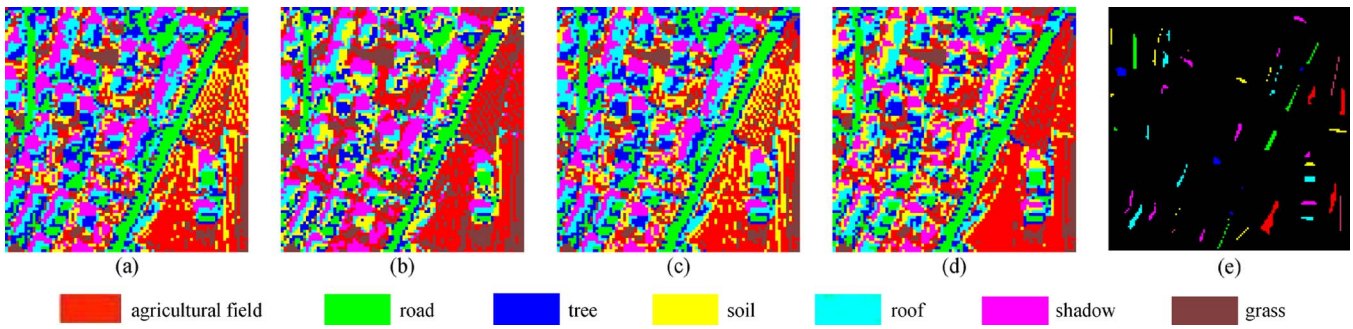

Fig. 5. Unsupervised classification maps of the south part of the city of Trento, Italy (Quickbird data). (a) K-means. (b) Fuzzy K-means. (c) ED-AP. (d) FS-AP. (e) Test field map.

association of the cluster labels to the land-cover classes was carried out.

The error matrices and global performance indices obtained by applying K-means, fuzzy K-means, ED-AP, and FS-AP to this dataset are shown in Tables III and IV. As an example, from Table III, one can see that, for the agricultural field, the user's accuracy of K-means and ED-AP is 79.34% and 87.26%, respectively. However, they get a lot of confusion between agricultural field and soil (only 64.00% and 59.33% of the agricultural field pixels are classified correctly). By contrast, fuzzy K-means and FS-AP exhibit a higher accuracy (i.e., 85.33% and 91.33%, respectively) than the other techniques. On the other hand, there is a confusion when discriminating agricultural field from grass, and many shadow pixels are wrongly identified as soil and roof by the fuzzy K-means. It can also be observed that there is a significant confusion between trees and other land covers, particularly for fuzzy K-means and K-means.

From Table IV, we can observe that, in general, the proposed FS-AP obtains the highest overall accuracy and Kappa value, improving 9.11%, 5.82%, and 6.87% of the overall accuracy yielded by the K-means, the fuzzy K-means, and the ED-AP, respectively. These results are also confirmed by the other quality indices considered, including those related to the computational time. Thus, we can conclude that, on this dataset, the FS-AP technique is superior to all the other three algorithms considered.

C. MODIS Dataset

The third dataset consists of Moderate Resolution Imaging Spectroradiometer (MODIS) data, acquired in the west of Changchun, Jilin, China, on June 12, 2008 (see Fig. 6). The considered level 1B dataset includes 500-m resolution images acquired in channels 3–7 of the sensor. Seven land-cover classes, i.e., wetland, river, cultivated land, open grass, bare soil, dry salt flats, and grass land, characterize this image.

For both the FS-AP and the ED-AP algorithms, we carried out the experiments with an α that is equal to 0.85 and by considering values of CTS between 10 and 25. The value of β was fixed to 1.5 in the FS-AP method. Fig. 7(a)–(d) shows the unsupervised classification maps obtained by using the K-means (best of 50 runs and $k = 7$), the fuzzy K-means (best of 3 runs and $k = 7$), the ED-AP $(\text{CTS} = 35, \alpha = 0.85)$, and the FS-AP $(\text{CTS} = 30, \alpha = 0.85, \beta = 1.5)$. To quantitatively evaluate the classification accuracy, a test field map is provided in Fig. 7(e) based on the ground truth data. Also, in this case, the land-cover classes are assigned to cluster labels manually.

The error matrices and global performance indices obtained by applying K-means, fuzzy K-means, ED-AP, and FS-AP to this dataset are shown in Tables V and VI. As an example, one can see from Table V that, for the wetland class, the producer's accuracy of K-means and fuzzy K-means is equal to 72.1% and 74.78%, respectively. However, this class is significantly confused with river and cultivated land, i.e., only 59.7% and 65.35% of the wetland pixels are classified correctly, respectively. By contrast, ED-AP and FS-AP recognize the wetland better than the aforementioned algorithms, resulting in a user's accuracy of 82.88% and 76.74%, respectively. On the other hand, the ED-AP involves a significant confusion between cultivated land and the other land covers and between grass and bare soil.

From Table VI, we can observe that, in general, the FS-AP obtains the highest overall accuracy and Kappa value, with a gain of 7.07%, 5.38%, and 4.48% over the K-means, the fuzzy K-means, and the ED-AP, respectively. These results are also confirmed by the other quality indices considered, thus pointing out the superiority of the FS-AP over the other three algorithms also for this dataset.

D. Analysis of Computational Efficiency

Based on all the experimental results presented in the previous sections, we can conclude that the proposed FS-AP produces better clustering accuracies than the K-means, the fuzzy K-means, and the ED-AP techniques. In this section, we focus our attention on the execution time taken from the different algorithms.

The K-means and fuzzy K-means are quite sensitive to the initial selection of exemplars. Different initializations cause different evolutions of the algorithm, which affect the number of iterations and the accuracy of clustering. As a result, K-means and fuzzy K-means often need to run many times with different initial exemplars and then require a cluster validation procedure for selecting the best final result (in the experiments reported in this paper, we used classification accuracy as a validation measure). For this reason, K-means and fuzzy K-means executed 50/80/50 and 10/5/3 runs with different initial random choice in the experiments presented in Sections IV-A–C), respectively, for achieving the best results. It can be observed from Tables II, IV, and VI that the FS-AP takes much less time and gets better classification results

TABLE III
ERROR MATRICES OF THE CLASSIFICATION MAPS DERIVED FROM THE SOUTH PART OF THE CITY OF TRENTO, ITALY (QUICKBIRD DATA)

Methods	Class	Agricultural field	Road	Tree	Soil	Roof	Shadow	Grass	Row total
K-means	Agricultural field	**96**	0	5	9	6	5	0	121
	Road	0	**99**	0	0	18	0	0	117
	Tree	4	5	**32**	2	23	2	4	72
	Soil	29	0	0	**60**	3	9	2	103
	Roof	4	0	2	6	**72**	3	0	87
	Shadow	5	3	6	7	12	**89**	1	123
	Grass	12	0	0	2	0	3	**30**	47
	Column total	150	107	45	86	134	111	37	**478**
	Producer's accuracy (%)	64.00	92.52	71.11	69.77	53.73	80.18	81.08	
	User's accuracy (%)	79.34	84.62	44.45	58.25	82.76	72.36	63.83	
Fuzzy K-means	Agricultural field	**128**	0	6	12	3	2	0	151
	Road	0	**101**	0	0	21	0	0	122
	Tree	0	4	**26**	3	23	0	0	56
	Soil	0	0	4	**45**	8	1	0	58
	Roof	0	0	2	4	**64**	0	0	70
	Shadow	3	2	7	20	15	**101**	2	150
	Grass	19	0	0	2	0	7	**35**	63
	Column total	150	107	45	86	134	111	37	**500**
	Producer's accuracy (%)	85.33	94.39	57.78	52.33	47.76	90.99	94.60	
	User's accuracy (%)	84.77	82.79	46.43	77.59	91.43	67.33	55.56	
ED-AP	Agricultural field	**89**	0	1	2	5	5	0	102
	Road	0	**102**	0	0	16	0	0	118
	Tree	1	2	**34**	1	19	2	1	60
	Soil	37	0	0	**73**	6	13	4	133
	Roof	5	0	3	4	**79**	5	0	96
	Shadow	7	3	7	2	9	**84**	0	112
	Grass	11	0	0	4	0	2	**32**	49
	Column total	150	107	45	86	134	111	37	**493**
	Producer's accuracy (%)	59.33	95.33	75.56	84.88	58.96	75.68	86.49	
	User's accuracy (%)	87.26	86.44	56.67	54.89	82.29	75.00	65.31	
FS-AP	Agricultural field	**137**	0	1	6	1	0	0	145
	Road	0	**103**	0	0	18	0	0	121
	Tree	0	1	**32**	2	19	0	0	54
	Soil	2	0	5	**57**	5	5	1	75
	Roof	0	0	4	3	**78**	2	0	87
	Shadow	2	3	3	13	13	**99**	3	136
	Grass	9	0	0	5	0	5	**33**	52
	Column total	150	107	45	86	134	111	37	**539**
	Producer's accuracy (%)	**91.33**	96.26	71.11	66.28	58.21	89.19	89.19	
	User's accuracy (%)	**94.48**	85.12	59.26	76.00	89.66	72.79	63.46	

TABLE IV
PERFORMANCE OF K-MEANS, FUZZY K-MEANS, ED-AP, AND THE PROPOSED FS-AP (QUICKBIRD DATA)

Parameter	K-means	Fuzzy K-means	ED-AP	FS-AP
Execution numbers	80	5	1	1
Execution time (min)	101	122	86	72
Overall accuracy (%)	71.34	74.63	73.58	80.45
Kappa value	0.661	0.698	0.688	0.767
Average of producer's accuracy (%)	73.20	74.74	76.60	80.22
Average of user's accuracy (%)	69.37	72.27	72.55	77.25
Average of Short's mapping accuracy index	0.548	0.567	0.584	0.647

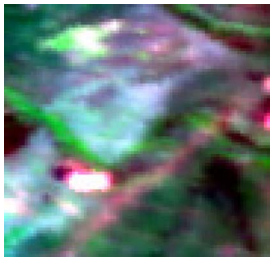

Fig. 6. Color composite (RGB = bands $7, 4, 2$) of the MODIS data of the west of Changchun, Jilin, China.

than K-means and fuzzy K-means. From the same tables, one can observe that the FS-AP algorithm is faster (and also more accurate) than the ED-AP. In order to further compare the execution time when getting different number of clusters, different values of CTS were used in the three considered datasets. By analyzing Figs. 8–10, which report the execution time versus the number of clusters, one can note that FS-AP has a smaller execution time than ED-AP with the same number of clusters. This mainly depends on the effectiveness of the FSS used by the proposed technique. Although FSS similarity computing needs a little longer time, it can make the whole algorithm converge more quickly (i.e., in a smaller number of iterations). From this analysis, we can conclude that the proposed FS-AP clustering technique gets better efficiency than the K-means, the fuzzy K-means, and the ED-AP methods.

E. Parameter Sensitivity Analysis for FS-AP

The following are the two user-defined parameters in FS-AP and ED-AP: 1) α, which affects both the convergence speed and the computational time, and 2) CTS, which affects the number of clusters. We performed experiments by varying these two parameters when running the FS-AP in the three considered datasets.

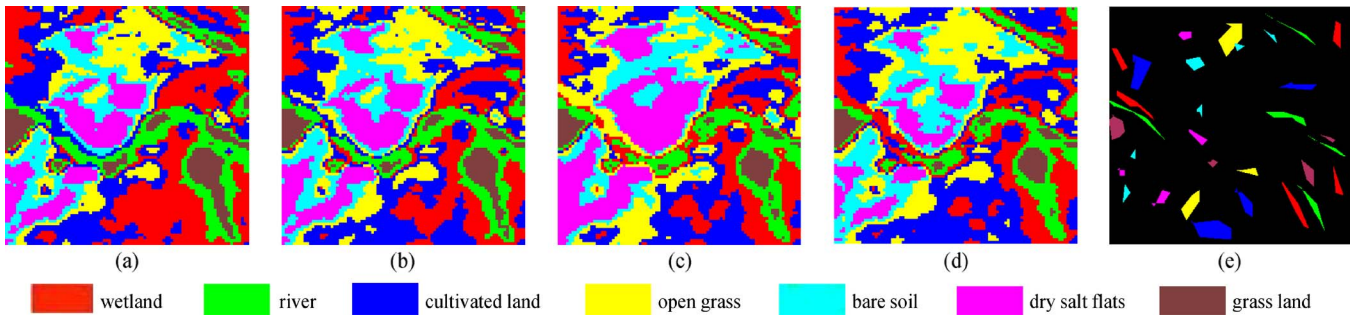

Fig. 7. Unsupervised classification maps of the west of Changchun, Jilin, China (MODIS data). (a) K-means. (b) Fuzzy K-means. (c) ED-AP. (d) FS-AP. (e) Test field map.

TABLE V
ERROR MATRICES OF THE CLASSIFICATION MAPS DERIVED FROM THE WEST OF CHANGCHUN, JILIN, CHINA (MODIS DATA)

Methods	Class	Wetland	River	Cultivated land	Open grass	Bare soil	Dry salt flats	Grass land	Row total
K-means	Wetland	**80**	1	53	0	0	0	0	134
	River	12	**65**	0	0	0	0	5	82
	Cultivated land	19	5	**128**	21	0	0	0	173
	Open grass	0	0	0	**82**	4	0	0	86
	Bare soil	0	0	0	7	**39**	5	0	51
	Dry salt flats	0	0	0	0	13	**46**	0	59
	Grass land	0	17	0	0	0	0	**68**	85
	Column total	111	88	181	110	56	51	73	**508**
	Producer's accuracy (%)	72.10	73.86	70.72	74.55	69.64	90.20	93.15	
	User's accuracy (%)	59.70	79.27	73.99	95.35	76.47	77.97	80.00	
Fuzzy K-means	Wetland	**83**	0	44	0	0	0	0	127
	River	11	**62**	0	0	0	0	6	79
	Cultivated land	9	8	**137**	10	0	0	0	164
	Open grass	8	0	0	**79**	7	0	0	94
	Bare soil	0	0	0	21	**43**	3	0	67
	Dry salt flats	0	0	0	0	6	**48**	0	54
	Grass land	0	18	0	0	0	0	**67**	85
	Column total	111	88	181	110	56	51	73	**519**
	Producer's accuracy (%)	74.78	70.46	75.69	71.82	76.79	94.12	91.78	
	User's accuracy (%)	65.35	78.48	83.54	84.04	64.18	88.89	78.82	
ED-AP	Wetland	**92**	8	11	0	0	0	0	111
	River	2	**66**	0	0	0	0	12	80
	Cultivated land	0	0	**145**	0	0	0	0	145
	Open grass	17	5	13	**76**	0	0	0	111
	Bare soil	0	0	12	34	**34**	0	0	80
	Dry salt flats	0	0	0	0	22	**51**	0	73
	Grass land	0	9	0	0	0	0	**61**	70
	Column total	111	88	181	110	56	51	73	**525**
	Producer's accuracy (%)	82.88	75.00	80.11	69.09	60.71	100	83.56	
	User's accuracy (%)	82.88	82.50	100.00	68.47	42.50	69.86	87.14	
FS-AP	Wetland	**99**	10	20	0	0	0	0	129
	River	0	**67**	0	0	0	0	10	77
	Cultivated land	0	2	**152**	18	0	0	0	172
	Open grass	2	0	0	**85**	7	0	0	94
	Bare soil	0	0	9	7	**46**	8	0	70
	Dry salt flats	0	0	0	0	3	**43**	0	46
	Grass land	0	9	0	0	0	0	**63**	72
	Column total	111	88	181	110	56	51	73	**555**
	Producer's accuracy (%)	89.19	76.14	83.98	77.27	82.14	84.31	86.30	
	User's accuracy (%)	76.74	87.01	88.37	90.43	65.71	93.48	87.50	

- *Parameter* α: The parameter α is related to the convergence of the algorithm. By increasing the α value, we increase the convergence probability, but we also increase the execution time. As an example, Fig. 11 shows the behavior of the execution time versus the value of α for the Landsat ETM+ dataset (CTS $= 9, \beta = 1$). By analyzing the figure, one can conclude that the execution time increases almost linearly with the value of α.
- *Parameter CTS*: Table VII reports the relation between the number of clusters obtained and the values of the CTS parameter for the three considered datasets. One can observe that the number of clusters is close to be monotonically related to the CTS value, and it decreases while CTS increases.

V. CONCLUSION

In this paper, a novel FS-AP clustering method has been presented and implemented. The key concepts contained in the FS-AP include fuzzy mean deviation and FSS. FSS allows the algorithm to assign proper memberships to uncertain data and can get an accurate and objective estimate of how closely two

TABLE VI
PERFORMANCE OF K-MEANS, FUZZY K-MEANS, ED-AP, AND THE PROPOSED FS-AP (MODIS DATA)

Parameter	K-means	Fuzzy K-means	ED-AP	FS-AP
Execution numbers	50	3	1	1
Execution time (min)	52	61	50	46
Overall accuracy (%)	75.82	77.46	78.36	82.84
Kappa value	0.710	0.730	0.743	0.794
Average of producer's accuracy (%)	77.75	79.35	78.76	82.76
Average of user's accuracy (%)	77.54	77.61	76.19	84.18
Average of Short's mapping accuracy index	0.633	0.648	0.637	0.714

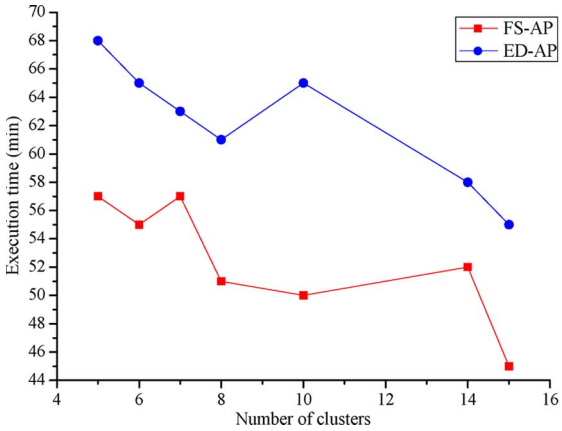

Fig. 8. Comparison between the execution time taken by the ED-AP and the proposed FS-AP algorithms versus the number of clusters (Landsat-7 ETM+ data).

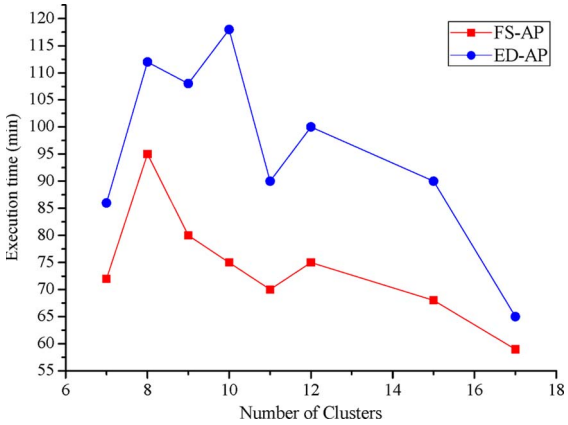

Fig. 9. Comparison between the execution time taken by the ED-AP and the proposed FS-AP algorithms versus the number of clusters (Quickbird data).

pixel vectors resemble each other. This ensures high robustness against noise and involves accurate clustering results in the case of mixed (or complex) pixels. Meanwhile, the proposed method simultaneously considers all the data points as candidate exemplars and passes soft information around until a subset of data points becomes the exemplar. It can avoid poor solutions caused by unlucky initializations and hard decisions and can save a significant amount of execution time for getting optimal results.

The experimental results based on three types of multispectral remote-sensing images (Landsat-7 ETM+, Quickbird, and MODIS) show that the proposed FS-AP clustering method

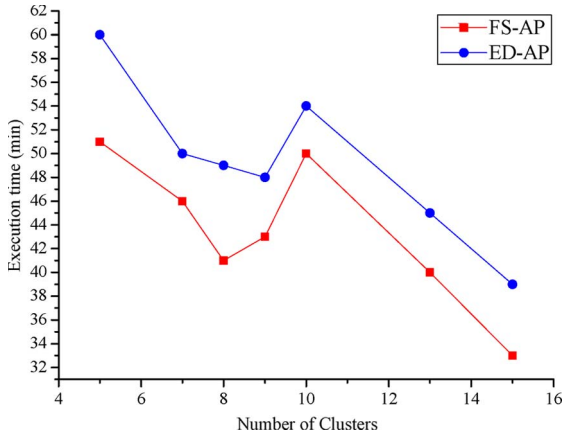

Fig. 10. Comparison between the execution time taken by the ED-AP and the proposed FS-AP algorithms versus the number of clusters (MODIS data).

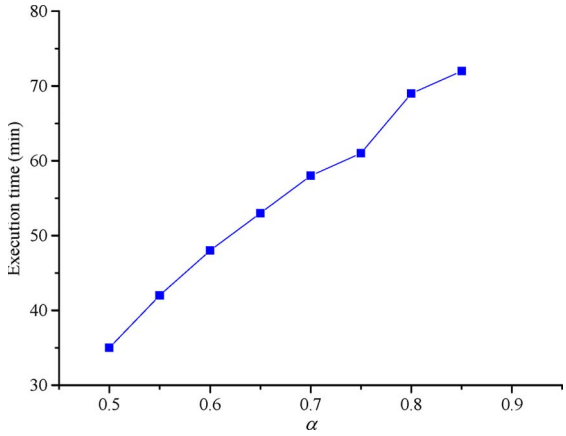

Fig. 11. Execution time versus the α value for the proposed FS-AP technique (Landsat ETM+ data).

TABLE VII
NUMBER OF CLUSTERS VERSUS THE VALUE OF THE CTS PARAMETER FOR THE PROPOSED FS-AP TECHNIQUE

	CTS	1	2	3	4	5	6	7	8	9	10
Landsat 7 ETM+	**Number of clusters**	19	15	14	10	8	7	6	6	5	4
	CTS	8	12	15	18	20	25	30	35	38	40
Quickbird	**Number of clusters**	17	15	12	11	10	9	9	8	7	6
	CTS	5	10	12	13	15	18	20	25	30	35
MODIS	**Number of clusters**	20	15	13	10	10	9	9	8	7	5

exhibits better accuracy indices and higher efficiency than the K-means, the fuzzy K-means, and the ED-AP algorithms. The obtained classification accuracy, Kappa value, and general accuracy indices related to the proposed FS-AP clustering method are always higher than those yielded by the other considered algorithms. Meanwhile, the FS-AP takes much less execution time than the K-means (best of 50/80/50 runs), the fuzzy K-means (best of 10/5/3 runs), and the ED-AP. This confirms that the FS-AP is effective in processing multispectral remote-sensing images.

In the future, we will focus the attention on the definition of alternative similarity measures and on their comparison and integration with other metrics, particularly with reference to the applications to multispectral and hyperspectral remote-sensing image classification.

ACKNOWLEDGMENT

The authors would like to thank Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, for their support by providing some of the multispectral remotely sensed data used in this paper.

REFERENCES

[1] 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.

[2] M. M. Chi, Q. Qian, and J. A. Benediktsson, "Cluster-based ensemble classification for hyperspectral remote sensing," in *Proc. IEEE IGARSS*, 2008, pp. 209–212.

[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] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, no. 5814, pp. 951–972, Feb. 2007.

[5] A. S. Robert, *Remote Sensing: Models and Methods for Image Processing*, 3rd ed. New York: Elsevier, 2007.

[6] M. S. Aldenderfer and R. K. Blashfield, *Cluster Analysis*. Beverly Hills, CA: Sage, 1984.

[7] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probability*, 1967, vol. 1, pp. 281–297.

[8] Z. J. Ding, J. Yu, and Y. Q. Zhang, "A new improved K-means algorithm with penalized term," in *Proc. IEEE ICCC*, 2007, p. 313.

[9] N. Chehata and F. Bretar, "Terrain modeling from lidar data: Hierarchical K-means filtering and Markovian regularization," in *Proc. IEEE ICIP*, 2008, pp. 1900–1903.

[10] J. Zheng, Z. Z. Cui, A. F. Liu, and Y. Jia, "A K-means remote sensing image classification method based on AdaBoost," in *Proc. IEEE ICNC*, 2008, pp. 27–32.

[11] P. Maheshwary and N. Srivastav, "Retrieving similar image using color moment feature detector and K-means clustering of remote sensing images," in *Proc. IEEE Int. Conf. Comput. Elect. Eng.*, Dec. 20–22, 2008, pp. 821–824.

[12] L. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed. New York: Elsevier, 2005.

[13] X. D. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z.-H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg, "Top 10 algorithms in data mining," *Knowl. Inf. Syst.*, vol. 14, no. 1, pp. 1–37, Dec. 2008.

[14] P. Y. Liu, K. B. Jia, and P. Z. Zhang, "An effective method of image retrieval based on modified fuzzy C-means clustering scheme," in *Proc. IEEE ICSP*, 2006, vol. 3.

[15] Y. Wang and J. Mo, "Fuzzy logic applied in remote sensing image classification," in *Proc. Int. Conf. Syst., Man Cybern.*, 2004, pp. 6378–6382.

[16] M. S. Dinesh, K. Chidananda Gowda, and P. Nagabhuehan, "Unsupervised classification for remotely sensed data using fuzzy set theory," in *Proc. IEEE IGARSS—A Scientific Vision for Sustainable Development*, 1997, vol. 1, pp. 521–523.

[17] B. Borasca, L. Bruzzone, L. Carlin, and M. Zusi, "A fuzzy-input fuzzy-output SVM technique for classification of hyperspectral remote sensing images," in *Proc. 7th NORSIG*, 2006, pp. 2–5.

[18] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *J. Cybern.*, vol. 3, no. 3, pp. 32–57, 1974.

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

[20] D. Altman, "Efficient fuzzy clustering of multi-spectral images," in *Proc. IEEE IGARSS*, 1999, vol. 3, pp. 1594–1596.

[21] P. V. Gorsevski, P. E. Gessler, and P. Jankowski, "Integrating a fuzzy K-means classification and a Bayesian approach for spatial prediction of landslide hazard," *J. Geographical Syst.*, vol. 5, no. 3, pp. 223–251, Nov. 2003.

[22] Y. F. Zhong, L. P. Zhang, B. Huang, and P. X. Li, "An unsupervised artificial immune classifier for multi/hyperspectral remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 3957–3966, Feb. 2006.

[23] X. F. Liu, X. W. Li, Y. Zhang, C. J. Yang, W. B. Xu, M. Li, and H. M. Luo, "Remote sensing image classification based on dot density function weighted FCM clustering algorithm," in *Proc. IEEE IGARSS*, 2007, pp. 2010–2013.

[24] C. C. Hung, W. P. Liu, and B. C. Kuo, "A new adaptive fuzzy clustering algorithm for remotely sensed images," in *Proc. IGARSS*, 2008, vol. 2, pp. 863–866.

[25] J. C. Fan, M. Han, and J. Wang, "Single point iterative weighted fuzzy C-means clustering algorithm for remote sensing image segmentation," *Pattern Recognit.*, vol. 42, no. 11, pp. 2527–2540, Nov. 2009.

[26] H. Y. Zhang, Q. T. Wu, and J. X. Pu, "A novel fuzzy kernel clustering algorithm for outlier detection," in *Proc. IEEE ICMA*, 2007, pp. 2378–2382.

[27] K. L. Wu and M. S. Yang, "Alternative C-means clustering algorithms," *Pattern Recognit.*, vol. 35, no. 10, pp. 2267–2278, Oct. 2002.

[28] I. Gath and A. B. Geva, "Unsupervised optimal fuzzy clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 773–781, Jul. 1989.

[29] S. Nasser, R. Alkhaldi, and G. Vert, "A modified fuzzy K-means clustering using expectation maximization," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, 2006, pp. 231–235.

[30] M. J. Li, M. K. Ng, Y. Cheung, and J. Z. Huang, "Agglomerative fuzzy K-means clustering algorithm with selection of number of clusters," *IEEE Trans. Knowl. Data Eng.*, vol. 20, no. 11, pp. 1519–1534, Nov. 2008.

[31] J. A. Richards and X. P. Jia, *Remote Sensing Digital Image Analysis: An Introduction*, 4th ed. Berlin, Germany: Springer-Verlag, 2006.

[32] S. M. Taheri, "Trends in fuzzy statistics," *Austrian J. Stat.*, vol. 32, no. 3, pp. 239–257, 2003.

[33] B. J. Frey and D. Dueck, "Mixture modeling by affinity propagation," in *Advances in Neural Information Processing Systems*, vol. 18. Cambridge, MA: MIT Press, 2006. 2, 3, 7.

[34] B. J. Frey and D. Dueck, "Response to comment on 'Clustering by passing messages between data points'," *Science*, vol. 319, no. 5864, p. 726d, Feb. 2008.

[35] [Online]. Available: http://www.psi.toronto.edu/affinitypropagation/faq.html

[36] M. C. Sun and C. H. Chou, "A modified version of the K-means algorithm with a distance based on cluster symmetry," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 674–680, Jun. 2001.

[37] J. N. Sweet, "The spectral similarity scale and its application to the classification of hyperspectral remote sensing data," in *Proc. IEEE Workshop Adv. Techn. Anal. Remotely Sensed Data*, 2003, pp. 92–99.

[38] R. A. Schowengerdt, *Remote Sensing Models and Methods for Image Processing*, 2nd ed. San Diego, CA: Academic, 1997.

[39] B. J. Frey and D. Dueck, "Non-metric affinity propagation for unsupervised image categorization," in *Proc. IEEE 11th ICCV*, 2007, pp. 1–8.

[40] K. K. Chintalapudi and M. Kam, "A noise-resistant fuzzy C-means algorithm for clustering," in *Proc. IEEE World Congr. Comput. Intell. Fuzzy Syst.*, 1998, vol. 2, pp. 1458–1463.

[41] C. Yang, L. J. Lu, H. P. Lin, R. C. Guan, X. H. Shi, and Y. C. Liang, "A fuzzy-statistics-based principal component analysis (FS-PCA) method for multispectral image enhancement and display," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3937–3947, Nov. 2008.

[42] J. Bryant, "On displaying multispectral imagery," *Photogramm. Eng. Remote Sens.*, vol. 54, no. 12, pp. 1739–1743, Dec. 1988.

[43] T. Fung and E. L. Drew, "The determination of optimal threshold levels for change detection using various accuracy indices," *Photogramm. Eng. Remote Sens.*, vol. 54, no. 10, pp. 1449–1454, 1998.

[44] C. Liu, P. Frazier, and L. Kumar, "Comparative assessment of the measures of thematic classification accuracy," *Photogramm. Eng. Remote Sens.*, vol. 107, no. 4, pp. 606–616, Apr. 2007.

[45] N. M. Short, *The Landsat Tutorial Workbook Basics of Satellite Remote Sensing*. Greenbelt, MD: Goddard Space Flight Center, 1982. NASA Reference Publication 1078.

[46] G. H. Rosenfield and K. Fitzpatrick-Lins, "A coefficient of agreement as a measure of thematic classification accuracy," *Photogramm. Eng. Remote Sens.*, vol. 52, no. 2, pp. 223–227, 1986.

Chen Yang received the M.S. degree in computer applied technology from Northeast Normal University, Changchun, China, in 2007. She is currently working toward the Ph.D. degree in Jilin University, Changchun.

She worked as an Exchange Student at the University of Trento, Trento, Italy, from June 2008 to December 2008. Her research interests include remote-sensing image processing, spatial data mining, and machine learning.

Lorenzo Bruzzone (S'95–M'98–SM'03–F'10) received the Laurea (M.S.) degree in electronic engineering (*summa cum laude*) and the Ph.D. degree in telecommunications from the University of Genoa, Genoa, Italy, in 1993 and 1998, respectively.

He is currently a Full Professor of telecommunications with the University of Trento, Trento, Italy, where he teaches remote sensing, pattern recognition, radar, and electrical communications. He is the Head of the Remote Sensing Laboratory, Department of Information Engineering and Computer Science, University of Trento. His current research interests are in the areas of remote sensing, signal processing, and pattern recognition (analysis of multitemporal images, feature extraction and selection, classification, regression and estimation, data fusion, and machine learning). He conducts and supervises research on these topics within the frameworks of several national and international projects. He is the author (or coauthor) of 82 scientific publications in referred international journals (55 in IEEE journals), more than 140 papers in conference proceedings, and 11 book chapters. He is the Editor/Coeditor of nine books/conference proceedings.

Dr. Bruzzone ranked first place in the Student Prize Paper Competition of the 1998 IEEE International Geoscience and Remote Sensing Symposium, July 1998. He was a recipient of the Recognition of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS) Best Reviewers in 1999 and was a Guest Coeditor of different Special Issues of the IEEE TGRS. He was the General Chair and Cochair of the First and Second IEEE International Workshop on the Analysis of Multi-temporal Remote-Sensing Images and is currently a member of the Permanent Steering Committee of this series of workshops. Since 2003, he has been the Chair of the SPIE Conference on Image and Signal Processing for Remote Sensing. From 2004–2006, he served as an Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS and currently is an Associate Editor for the IEEE TGRS and the *Canadian Journal of Remote Sensing*. In 2008, he was appointed a member of the joint NASA/ESA Science Definition Team for Outer Planet Flagship Missions. He is also a member of the International Association for Pattern Recognition and of the Italian Association for Remote Sensing. He is a Referee for many international journals and has served on the Scientific Committees of several international conferences. He is a member of the Managing Committee of the Italian Inter-University Consortium on Telecommunications and a member of the Scientific Committee of the India–Italy Center for Advanced Research. Since 2009, he has been a member of the Administrative Committee of the IEEE Geoscience and Remote Sensing Society.

Fengyue Sun received the M.S. and Ph.D. degrees from the China University of Geosciences, Wuhan, China, in 1987 and 1995, respectively.

He is currently a Professor with the College of Earth Sciences, Jilin University, Changchun, China. He was a Visiting Scholar with the University of Western Australia, Perth, WA, Australia, from 2005 to 2006. He has published two monographs and over 90 journal and conference papers. His research interests include metallogeny and prediction for hydrothermal deposits.

Dr. Sun was the recipient of several grants from National Natural Science Foundation of China, "985 Project" innovation platform, etc.

Laijun Lu received the M.S. and Ph.D. degrees in mathematical geology from Jilin University, Changchun, China, in 1982 and 1992, respectively.

He is currently a Professor with the College of Earth Sciences and the Laboratory of Digital Geoscience, Jilin University. He was a Postdoctoral Fellow with the Mining and Petroleum of Exploration Postdoctoral Workstation, Northeastern University, Shenyang, China, from 1992 to 1994. He has published six books and over 40 journal and conference papers. His research interests include digital geosciences, resources and environment information system, and application of spatial and time-series-related Markov process.

Renchu Guan received the B.S. degree in computer science and technology and the M.S. degree in computer applied technology from Northeast Normal University, Changchun, China, in 2004 and 2007, respectively. He is currently working toward the Ph.D. degree in Jilin University, Changchun.

He worked as an Exchange Student at the University of Trento, Trento, Italy, from June 2008 to December 2008. His research interests include machine learning, text mining, and bioinformatics.

Yanchun Liang received the Ph.D. degree in applied mathematics from Jilin University, Changchun, China, in 1997.

He is a Professor with the College of Computer Science and Technology, Jilin University. He was a Visiting Scholar with Manchester University, Manchester, U.K., from 1990–1991; a Visiting Professor with the National University of Singapore, Singapore, from 2000–2001; a Guest Professor with the Institute of High Performance Computing of Singapore, Singapore, from 2002–2004; and a Guest Professor with the University of Trento, Trento, Italy, from 2006–2008. He has published over 280 papers. His research has been featured in *Bioinformatics*, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART A: SYSTEMS AND HUMANS, *Journal of Micromechanics and Microengineering*, *Physical Review E*, *Neural Computing and Applications*, *Smart Materials and Structures*, *Artificial Intelligence in Medicine*, *Applied Artificial Intelligence*, etc. His research interests include computational intelligence, machine learning methods, text mining, microelectromechanical system modeling, and bioinformatics.

Dr. Liang was a recipient of several grants from NSFC, European Union, etc.