Spectral Clustering Ensemble Applied to SAR Image Segmentation

Xiangrong Zhang, *Member, IEEE*, Licheng Jiao, *Senior Member, IEEE*, Fang Liu, *Senior Member, IEEE*, Liefeng Bo, and Maoguo Gong, *Member, IEEE*

Abstract—**Spectral clustering (SC) has been used with success in the field of computer vision for data clustering. In this paper, a new algorithm named SC ensemble (SCE) is proposed for the segmentation of synthetic aperture radar (SAR) images. The gray-level cooccurrence matrix-based statistic features and the energy features from the undecimated wavelet decomposition extracted for each pixel being the input, our algorithm performs segmentation by combining multiple SC results as opposed to using outcomes of a single clustering process in the existing literature. The random subspace, random scaling parameter, and Nyström approximation for component SC are applied to construct the SCE. This technique provides necessary diversity as well as high quality of component learners for an efficient ensemble. It also overcomes the shortcomings faced by the SC, such as the selection of scaling parameter, and the instability resulted from the Nyström approximation method in image segmentation. Experimental results show that the proposed method is effective for SAR image segmentation and insensitive to the scaling parameter.**

Index Terms—**Image segmentation, spectral clustering (SC), synthetic aperture radar (SAR), unsupervised ensemble.**

I. INTRODUCTION

SYNTHETIC aperture radar (SAR) images find increasingly wide applications because SAR sensors can penetrate clouds and work in bad weather conditions and in nighttime when optical sensors are inoperable. An important problem in SAR image applications is correct segmentation. It is the basis of the understanding of SAR images, such as the change detection of regions for maps updating [1], the recognition of targets [2], and so on. The purpose of SAR image segmentation is to partition an image into regions of different characteristics. There are many approaches available for SAR image segmentation including threshold methods [3], [4], clustering algorithms [5], [6], statistic model-based methods [7]–[9], and morphologic methods [10], [11].

This paper reexamines the approach of clustering-based SAR image segmentation. Among the existing clustering algorithms,

k-means algorithm is the most popular and simplest one. The k-means algorithm is efficient for clusters with spherical shapes in input space. However, many data sets in practice do not have this shape. Recently, a family of spectral clustering (SC) algorithms [12]–[17] were proposed and have shown great promise. Compared with traditional clustering algorithms, SC has some obvious advantages. It can recognize the clusters of unusual shapes and obtain the globally optimal solutions in a relaxed continuous domain by eigendecomposition. This property makes SC more suitable for many applications, such as speech separation [18], image segmentation [12], [13], very large scale integration design [19], and so on. However, several problems are still not solved satisfactorily [20], [21]. For example, the method is computationally expensive because it uses an affinity matrix constructed by the similarity of each pair of pixels and needs to compute the eigenvectors of the affinity matrix. Furthermore, SC requires setting a parameter, namely, the scaling parameter σ in the Gaussian radial basis function (RBF). Appropriate setting of σ is crucial for obtaining good segmentation results in SC. Unfortunately, it is difficult to choose the appropriate σ value, and it is always set manually. The incorrect value of σ can degrade the performance as the SC is highly sensitive to σ, and different values of σ may lead to drastically different results. The proper setting of the scaling parameter still remains an open issue, and there is no known effective method.

In this paper, we develop an SC ensemble (SCE) algorithm to avoid the selection of the appropriate scaling parameter. It was proved that accuracy and diversity of component learners are necessary in the construction of strong ensemble learning [22], [23]. In this paper, the samples are first partitioned by multiple SCs with random scaling parameters, which not only provides the necessary diversity for ensemble, but also avoids the scaling parameter selection for each individual SC.

If SAR images are segmented directly with SC, we face the problem of high computational cost. To make the proposed ensemble method efficient for large-scale problems, an SC algorithm using Nyström method [12] is applied as the base-level learner of the ensemble for its lower computational complexity. In the Nyström method, a small set of randomly sampled data points from all pixels is used to perform the approximation. However, this method is instable for image segmentation. As we know, injecting randomness into base-level learner could be helpful in generating diverse component learners, which has been employed in various ensemble learning [24]–[26]. Therefore, the native instability of the SC using Nyström method can be used in the ensemble construction conveniently.

Furthermore, we can get a robust and stable segmentation result by the SCE.

The remainder of this paper is organized as follows. Two types of features for SAR image segmentation are presented in Section II. In Section III, a classic SC algorithm is reviewed. In Section IV, the SCE construction and an associated aggregation strategy are presented. Experimental results in comparison with other methods are provided in Section V. Conclusions are drawn in Section VI.

II. FEATURES FOR SAR IMAGE SEGMENTATION

The purpose of image segmentation based on region partition is to assign a specific class label for each pixel. Therefore, it comes down to a large-scale clustering problem. At the same time, proper features should be extracted to differentiate the land covers effectively. Texture can provide abundant useful information for SAR image analysis [27], and so far, many different texture features have been proposed. In this paper, we focus on investigating the performance of the statistic features which are derived from the gray-level cooccurrence matrix (GLCM) [28] and the energy features from the undecimated wavelet decomposition [29] for SAR image segmentation.

1) GLCM-Based Statistic Features: GLCM method was frequently used in texture analysis and extraction for SAR images [27], [30]. Texture features are demonstrated by the statistics over the GLCM. Four usually used statistics are angular second moment (also called energy), correlation, entropy, and inverse difference moment (also called inertia). In this paper, the 14 statistics, including previous four, suggested by Haralick *et al.* [28] are exploited because more features will provide diversity to ensemble. There are four parameters that must be indicated to generate a GLCM, i.e., the interpixel orientation, distance, gray-level quantization, and window size [31]. Here, we set the interpixel orientation to $0°$ for convenient calculation. Short interpixel distances typically generate the preferred texture features in SAR image analysis [32], so we set interpixel distance to one. The role of different values for gray levels and windows size with respect to statistics from GLCM has been investigated in many literatures [30], [33]. According to their analysis and fine-tune experiments, in this paper, we set the image quantization to 16 and the window size to 9×9.

2) Wavelet Energy Features: Wavelet transform has the ability to examine a signal at different scales [29]. In this paper, the undecimated wavelet-based feature vector composed by the energies of the subband coefficients is used. Here, we implement three-level wavelet decomposition on each square local area with size of 16×16. The features of each pixel can be represented as a ten-dimension vector $(e_{\mathrm{LL}-1}, e_{\mathrm{LH}-1}, e_{\mathrm{HL}-1}, e_{\mathrm{HH}-1}, e_{\mathrm{LH}-2}, e_{\mathrm{HL}-2}, e_{\mathrm{HH}-2}, e_{\mathrm{LH}-3}, e_{\mathrm{HL}-3}, e_{\mathrm{HH}-3})$, in which, for example $e_{\mathrm{LL}-1}$ denotes the energy of the LL subimage in the first level. Wavelet transform generates localized spatial and spectral information simultaneously. The energy of the low-pass subimage $e_{\mathrm{LL}-1}$ describes the spectral information. The other features except $e_{\mathrm{LL}-1}$ characterize the textural properties [34]. Incorporating all the features will enhance the performance.

TABLE I
SC ALGORITHM

Procedure:
Input: data set $X = \{x_1, x_2, \cdots, x_n\} \in R^d$, number of clusters k
Step1. Form the affinity matrix $W \in R^{n \times n}$ where w_{ij} is defined by equation (1) if $i \neq j$, and $W_{ii} = 0$.
Step2. Define D to be the diagonal matrix where $D(i,i)$ is the sum of W's ith row, and construct the matrix $L = D^{-1/2} W D^{-1/2}$.
Step3. Compute the first k eigenvectors $s_1, s_2, \cdots, s_k$ of L, and form the matrix $S = [s_1, s_2, \cdots, s_k] \in R^{n \times k}$ as columns.
Step4. Construct the matrix $Y \in R^{n \times k}$ from S by normalizing each row in S to have unit length, i.e. $y_{ij} = s_{ij} / \left(\sum_j s_{ij}^2 \right)^{1/2}$.
Step5. Treating each row $y_i \in R^k$ of Y as a point, cluster them into k clusters via k-means or other algorithm.
Output: k clusters of X

III. SC

In general, the SC first embeds data points into a new space by the eigenvectors of an affinity matrix. This affinity matrix defines the similarities in the data. Moreover, a Gaussian function was often used as the similarity function with form

$$w_{ij} = \exp\left(-\|x_i - x_j\|^2 / 2\sigma^2\right) \qquad (1)$$

where σ is the scaling parameter, and W_{ij} the similarity of samples x_i and x_j. The top k eigenvectors are used as k-dimension indicator vectors for samples. A simple clustering algorithm such as k-means clustering is used to get k clusters. The algorithm is summarized in Table I.

The main trick of the SC is to map data points into a new space with k-dimension by means of eigenvector decomposition, wherein the data points form tight clusters. Therefore, we can get the natural clusters of data points with k-means clustering in this space, as performed in step 5.

However, given a data set with n samples, the size of the affinity matrix will be $n \times n$, and the complexity of eigenvector calculations in high dimensionality is $O(n^3)$. In order to reduce the computational cost, the SC using Nyström approximation was proposed in literature [12]. It works by first solving the grouping problem for a small random subset of pixels and then extrapolating this solution to the full sample set. Its computational complexity is reduced to $O(m^2 n)$, where m is the number of samples randomly chosen from n samples. It was shown that 100 randomly chosen pixels are sufficient to capture the salient groups in typical natural images [12]. Therefore, $m \ll n$ in applications. In this paper, the SC using Nyström method is used as the base-level clustering of the ensemble for SAR image segmentation.

IV. SCE LEARNING

A. Clustering Ensemble Problem

Ensemble learning has been successfully applied in the supervised learning by using various strategies [35], [36]. Inspired by the success of classifier ensemble, research on the

unsupervised ensemble or clustering ensemble is receiving more attention gradually. It is deemed that clustering ensembles can go beyond what is typically achieved by a single clustering algorithm in several respects, such as robustness, stability and confidence estimation, and parallelization and scalability [37]. Recently, many promising results have been reported in [24], [38]–[42].

The clustering ensemble problem can be described as follows. Let r be the number of component clusterings in the considered ensemble, and Π denote the set of the component clustering results $\Pi = \{\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(r)}\}$, in which ith clustering solution $\lambda^{(i)} = \{C_1^{(i)}, C_2^{(i)}, \dots, C_{k_i}^{(i)}\}$ partitions the data set $\boldsymbol{X}$ into k_i disjoint clusters. The final partition of $\boldsymbol{X}$ will be gotten from the available r component clustering results in Π.

Because there is no corresponding relationship between two clustering results, combining multiple label vectors in clustering ensemble is more difficult than in classifier ensemble. For instance, there are two clustering results of the same data set $\boldsymbol{X}$, whose label vectors are $\lambda^{(1)} = [1, 1, 1, 2, 2, 3, 3]^{\mathrm{T}}$ and $\lambda^{(2)} = [2, 2, 2, 3, 3, 1, 1]^{\mathrm{T}}$, respectively. Although the two label vectors are different in appearance, they are logically identical, namely, they denote the same partition of $\boldsymbol{X}$. For this challenging problem, several approaches are available, such as the evidence accumulation approach [41], [42], hypergraph operations-based method [24], mixture model-based aggregation [37], voting methods [39], [43], etc.

In this paper, we can get two benefits for SAR image segmentation by constructing an SCE. On one hand, better segmentation result is obtained. On the other hand, some problems of the SC can be solved to some degree, for instance, the selection of the scaling parameter σ in (1) for individual clustering and the instability due to the Nyström approximation.

The construction of the SCE consists of two phases, the generation of component clusterings and the combination of multiple clustering results.

B. Generation of Diverse Component Clusterings

Combining several identical predictors produces no gain [35]; therefore, the diversity of components is very important for improving the performance of an ensemble. To construct a clustering ensemble, the diversity of component learners can be obtained through two ways: the input perturbation and the clustering algorithm perturbation. To the input perturbation, the component clusterings can be gotten from different feature subsets of the same samples, which were generated by feature selection technique, random subspace method [24], the random projection method [25], etc. In addition, resampling methods such as bagging [39] or boosting [40] technique can be used to get different input as well. To the clustering algorithm perturbation, the different clustering algorithms [38] and the same clustering algorithm with different initialization or parameters are both useful for improving the diversity of the components. In this paper, multimodal perturbations for component clusterings are adopted, which are convenient to perform for the SC.

1) Random Scaling Parameter: Parameter selection is a popular problem in machine learning. It directly impacts the performance of the learning machine.

Let us take a supervised learning, such as a support vector machine (SVM) [44] with Gaussian kernel, serve as an illustration. In SVM, the magnitude of penalty factor C and the scaling parameter σ need to be selected carefully, and high generalization ability can be obtained with appropriate parameter settings. Generally, the proper learning parameters could be selected by the q-fold cross-validation on the training set. The process takes q times of computation to evaluate a parameter set. It is exhaustive. However, the intractable problem is becoming resolvable.

Comparing with the supervised learning, it is harder to select the appropriate parameters in the unsupervised learning, because there is no prior information about the data set. Generally, the parameter is manually determined. The SC is very sensitive to the scaling parameter. Obviously, such a fact is not desirable for applications but useful for constructing an ensemble. Therefore, we randomly pick the parameter σ_i of ith component SC from the preestablished interval $[\sigma_{\min}, \sigma_{\max}]$ with uniform distribution in our method. Using random scaling parameter in component clustering not only avoids the cost of accurate parameter selection for each individual SC, but also provides the required diversity for the clustering ensemble.

2) Instability From Nyström Method: The SC belongs to pairwise grouping methods, which requires comparing all possible data points for composing the affinity matrix. For instance, a given image of 256×256 pixels has $65\,536$ points, and the size of the generated affinity matrix will be $65\,536 \times 65\,536$. This requires great cost of computation and storage. The SC using Nyström method [12] is a better solution for the application of the SC in the image segmentation. However, the random sampling in the Nyström approximation will lead to different results in different runs, which will be demonstrated in Section V-A. Fortunately, this fact is also helpful for constructing an efficient ensemble because the difference between components is essential for an ensemble. In consideration of computation and diversity, we apply the SC using Nyström method as the base-level learner of the ensemble.

3) Random Initialization: SC can get the global-optima in the relaxed continuous domain by eigendecomposing. In order to get a discrete solution from eigenvectors, it is required to solve another clustering problem in a lower dimension space. Generally, k-means clustering is employed here. It is well known that k-means clustering is sensitive to the initialization. Therefore, we will do nothing to interfere with the initialization of k-means in this paper because the random initial values could be helpful in improving diversity as well.

4) Random Subspace: Using various features can improve the performance of SAR image segmentation. In Section II, we reviewed two types of features: the GLCM-based statistic features and the wavelet energy features. Not all the features are useful for different types of land cover in different SAR images. Also, there are usually some features irrelevant to the learning task. Feature selection is a good solution for finding the optimal feature subset to a given task. It is also effectively used in the classifier ensemble, in which the diversity and accuracy of the component learners with selected feature subsets can be evaluated on the training set by the holdout technique. Here, we cannot evaluate the feature subsets for unsupervised ensemble.
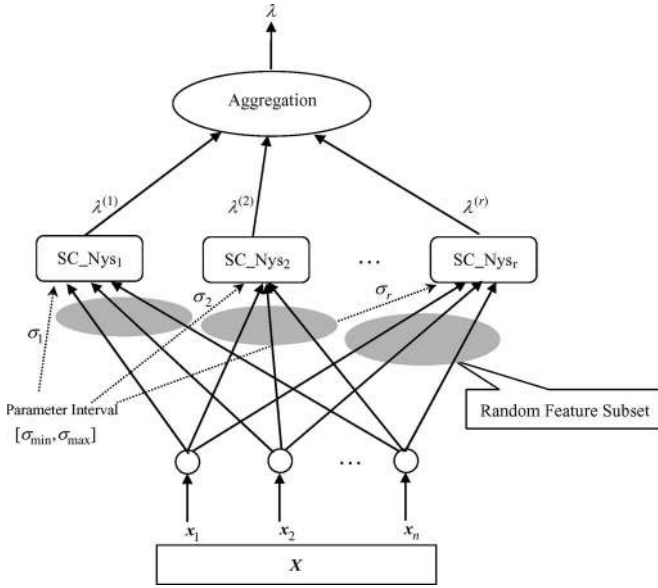
Fig. 1.   Spectral clustering ensemble construction.

It is doubtless that different feature subspaces will give the different observations of objects, and the component learners with different feature subspaces might be quite diverse. In this paper, the random subspace approach [26] is used in the construction of the clustering ensemble for its convenience and effectiveness. In our experiment, a feature subset is randomly chosen from all features for each component SC, and the number of features is randomly determined by d_i.

From the above four points, we can describe the construction of SCE as follows. For each pixel of the SAR image, we extract features with a moving window first. Also, all pixels represented by randomly selected features from all features are the input of each individual clustering. The SC using Nyström method is applied as the base-level clustering whose scaling parameter is randomly chosen from $[\sigma_{\min}, \sigma_{\max}]$. This procedure can be seen in Fig. 1.

C. Combination of Multiple Clustering Results

The combination of multiple results is a difficult task because the class label is only a symbolic in clustering.

The majority voting is one of the most widely used aggregation methods in the classifier ensemble. However, it cannot be used in the clustering ensemble directly. We should find the correspondence of all labels first. It is carried out under the supposition that the most similar clusters should contain maximum number of the similar objects. All component clustering results are relabeled according to a given reference partition. Also, the performance of the reference partition ties up with the ensemble result. In our algorithm, we get a reference partition by clustering the samples with all extracted features, and all the components are relabeled according to their best agreement with the reference partition. Suppose there are two clustering results whose label vectors are $\lambda^{(a)}$ and $\lambda^{(b)}$, respectively. The data set $\boldsymbol{X}$ is partitioned into k disjoined clusters which is denoted by $\{C_1^{(a)}, C_2^{(a)}, \ldots, C_k^{(a)}\}$ and $\{C_1^{(b)}, C_2^{(b)}, \ldots, C_k^{(b)}\}$. We then get a symmetric matrix S with size $k \times k$ whose

value s_{ij} is the number of overlapped data points in the pair of clusters $C_i^{(a)}$ and $C_j^{(b)}$. The pair of clusters, whose number of the overlapped data points is the largest, is corresponding in the way that they are designated to the same label. Such a process is repeated until all the clusters are corresponding. After the alignment, the majority voting strategy is carried out to get the final prediction similar to the supervised ensemble.

However, it is impossible that one combination method is effective for all data sets. The majority voting method is effective for problems with small number of classes. When the number of classes is large, the corresponding strategy in the majority voting may bring a risk for the final result. Therefore, other kind of combination methods is needed for remediation.

As mentioned before, the segmentation of SAR image is a large-scale clustering problem. Aggregation algorithms based on the similarity matrix, such as the coassociation matrix-based consensus function [42] with complexity $O(n^2)$ and cluster-based similarity partitioning algorithm (CSPA) [24] with complexity $O(kn^2 r)$, cannot be used because of their large demand of computation and storage. Among the hypergraph-based consensus functions, metaclustering algorithm (MCLA) with complexity $O(k^2 nr^2)$ has been proved to be an efficient one [24]. Therefore, it is used in this paper to combine r clustering results.

In the MCLA, the label vector set is first transformed into a hypergraph representation. MCLA is to group and collapse related hyperedges and assign each object to the collapsed hyperedge in which it participates most strongly. The hyperedges that are considered to be related for the purpose of collapsing are determined by a graph-based clustering of hyperedges. Each cluster of hyperedges is referred to as a metacluster $C^{(M)}$. Collapsing reduces the number of hyperedges from rk to k.

For various data sets, it is hard to conclude that which type of consensus function is more efficient. Therefore, two different types of consensus function, MCLA and the majority voting method, are both used to combine the component clustering results. The better solution is the one that shares the most information with the original clusterings. Mutual information provides a sound indication of the shared information between a pair of clusterings [24]. The normalized mutual information between two clustering label vectors $\lambda^{(a)}$ and $\lambda^{(b)}$ can be estimated by

$$
\phi^{(\mathrm{NMI})}(\lambda^{(a)}, \lambda^{(b)})
$$
$$
= \frac{\sum\limits_{i=1}^{k} \sum\limits_{j=1}^{k} n_{ij} \log\left(\frac{n_{ij} n}{n_i^{(a)} n_j^{(b)}}\right)}{\sqrt{\left(\sum\limits_{i=1}^{k} n_i^{(a)} \log \frac{n_i^{(a)}}{n}\right)\left(\sum\limits_{j=1}^{k} n_j^{(b)} \log \frac{n_j^{(b)}}{n}\right)}} \quad (2)
$$

where $n_i^{(a)}$ is the number of objects in cluster i according to $\lambda^{(a)}$, $n_j^{(b)}$ is the number of objects in cluster j according to $\lambda^{(b)}$, n_{ij} is the number of objects that are in cluster i according to $\lambda^{(a)}$ as well as in cluster j according to $\lambda^{(b)}$, and $\phi^{(\mathrm{NMI})}$ is a $r \times r$ symmetrical matrix and the diagonal elements are ones because $\phi^{(\mathrm{NMI})}(\lambda^{(a)}, \lambda^{(a)}) = 1$. Then, the average mutual

TABLE II
PSEUDOCODE OF THE SCE ALGORITHM

Procedure:

Input: data set X, number of clusters k, number of components r, the spectral clustering using Nyström method C

Step1. generation of component clusterings for ensemble

for $i=1$: r

X_i: random subspace for all samples

σ_i: randomly selected from $[\sigma_{\min}, \sigma_{\max}]$

$\lambda^{(i)} = C(X_i, \sigma_i)$

end for

Step2. combining the r label vectors $\{\lambda^{(1)}, \lambda^{(2)}, \cdots, \lambda^{(r)}\}$ by MCLA, we get $\lambda^{(MCLA)}(X)$

Step3. combining the r label vectors by majority voting

get the reference clustering: $\lambda^{(ref)} = C(X, \sigma)$ where σ is randomly selected in $[\sigma_{\min}, \sigma_{\max}]$.

for $i=1$ to r

$\lambda^{(i)}$ is re-labeled according to their best agreement with $\lambda^{(ref)}$

end for

$\lambda^{(vote)}(X)$ is got by using the majority voting

Step4. choose the final clustering result $\lambda^*(X)$ from $\lambda^{(MCLA)}(X)$ and $\lambda^{(vote)}(X)$ according to the normalized mutual information

Output: k clusters of X

---

information between the final clustering result λ^* of the ensemble and the component clustering results set Λ is computed by

$$\phi^{(ANMI)}(\Lambda, \lambda^*) = \frac{1}{r} \sum_{j=1}^{r} \phi^{(NMI)}(\lambda^{(j)}, \lambda^*). \tag{3}$$

The result with larger average mutual information is selected from the results of MCLA and the voting approach as the final decision.

The SCE algorithm is summarized in Table II.

V. EXPERIMENTAL STUDY

A. Performance Analysis on UCI Data Sets

To elucidate relative advantages of the SCE with respect to single SC, the results of different algorithms on six data sets from UCI machine learning repository [45] are presented first.

Performance of the SC depends on the selection of the scaling parameter σ. On the following data sets, the high impact σ on the SC is evaluated. It is also performed to compare the SCE with the k-means clustering, the SC, and the SC using Nyström method.

Iris data set is extensively used in classifier comparisons. It consists of three types of Iris plants (Setosa, Versicolor, and Virginica), with 50 instances per class, represented by four features. It has one class (Setosa) linearly separable from the remaining two classes, whereas the other two classes partially overlap in the feature space.

Sonar data set consists of 208 instances, in which 111 instances are obtained by bouncing sonar signal off a metal cylinder, and 97 instances were obtained from rocks. Each instance is represented by 60 feature values.

Breast-cancer-W is the breast cancer data provided by the University of Wisconsin hospitals. It consists of 699 instances where 16 instances with missing feature values are removed. Each instance is described by nine features and belongs to one of two possible classes: benign or malignant.

Chess (kr-vs-kp) data set consists of 3196 instances which could be classified into two categories (win or no win). Each instance describes the board for this chess endgame using 36 features.

Segmentation data set consists of 2310 instances, drawn randomly from a data set of seven outdoor images, brickface, sky, foliage, cement, window, path, and grass. There are 330 instances per class, and each instance is represented by 19 features.

Pen digits data set consists of 10 992 instances. Each instance with 12 features describes one of the handwritten numbers 0–9.

On the first five data sets, four algorithms are used for clustering analysis: 1) k-means clustering; 2) the SC without approximation; 3) the SC using Nyström method; 4) the proposed SCE. On the pen digits data set, the SC without approximation is not performed because of large computational cost. The error rates by matching the clustering results with the available label vectors are computed to evaluate the final partitioning.

The k-means clustering is known to be sensitive to initialization. For comparison, we performed 50 runs of it with random initial values of cluster centers, and retained only the result with minimum error rates attained over these 50 experiments, although a large variance on the error rates was observed. The results are recorded in Table III.

For each data set, the SC is carried out with each value of σ in the interval with the step length listed in Table III. Results presented in Table III are the minimum error rates obtained.

The SC with Nyström method (SC_Nys) requires sampling a part of data points in the approximation procedures. For each data set, 50 data points are randomly sampled from the original data set in the approximation method. The scale parameter interval and the step length used are similar to those used in the SC without approximation. 50 runs of this method are performed. Note that the 50 data subsets sampled are fixed for different parameter to investigate the influence of the scaling parameter on the clustering performance. The results of SC_Nys demonstrated in Figs. 2–5 are the average values of 50 runs of this method. Also, the minimum value of these averages for each data set is listed in Table III.

The SCE combines 30 component SC_Nys. Each component is carried out with a scaling parameter randomly selected in the given interval listed in Table III. In SC_Nys, 50 data points are randomly sampled in the approximation method. We perform ten runs of SCE independently and plot the average values in Figs. 2–5. Table III records the average error rate of the ten runs for each data set also.

From the results of SC and SC_Nys on these data sets shown in Figs. 2–5, it is found that the error rates of both methods change with the scaling parameter greatly. Therefore,

TABLE III
ERROR RATES OBTAINED BY THE *k*-MEANS: SC, SC_Nys, AND SCE

Data set	*k*-means clustering (%)	Spectral Clustering			SCE	
		Interval and step length of scaling parameter	SC (%)	SC_Nys (%)	MCLA (%)	Voting (%)
Iris	10.67	0.2:0.02:1	9.33	10.15	8.70	8.70
Sonar	44.23	0.5:0.02:2	43.75	44.26	43.41	43.03
Breast-cancer-W	4.25	0.1:0.02:2	3.07	2.97	3.12	3.07
Chess(kr-vs-kp)	46.25	0.1:0.02:2	45.71	44.56	44.31	44.11
Segmentation	45.45	1:0.1:10	39.00	40.64	37.00	42.72
Pen digits	22.17	0.1:0.1:5	-	25.94	23.40	24.26
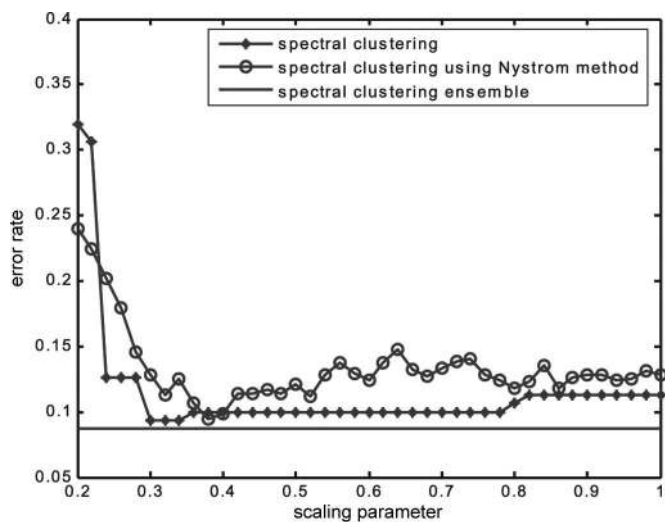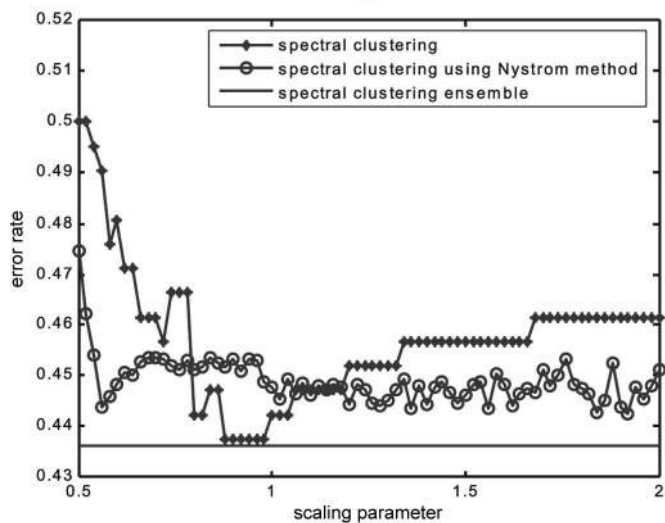
Fig. 2. Error rates of SC and SC_Nys with different scaling parameters, and the error rate of SCE in the interval. (a) Iris data set. (b) Sonar data set.

Fig. 3. Average values and standard deviation of error rates by SC_Nys with the changing of the scaling parameter on the Breast-cancer-W data set.

the performance of SC and SC_Nys depends on the setting of scaling parameter in practice.

In Fig. 2, the error rates of three methods, SC, SC_Nys, and SCE on Iris and Sonar data sets, are demonstrated. From the co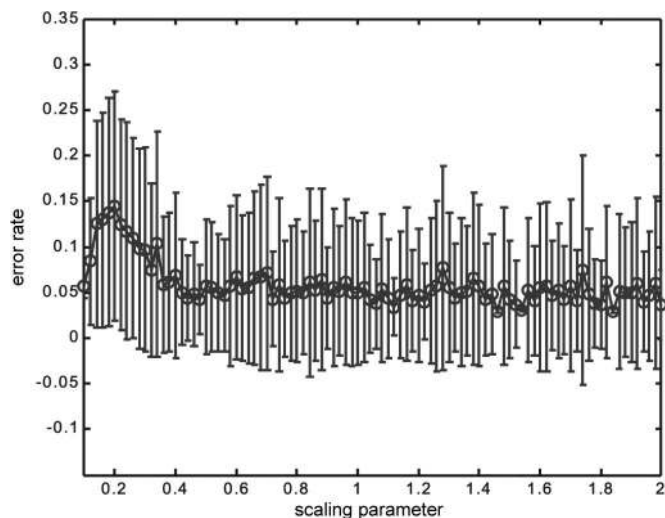mparison, it is clear that SC is better than SC_Nys at some special points, and SCE obtained the best results over the given parameter interval.

Fig. 3 shows the error rate curve of SC_Nys with the scaling parameter in the interval [0.1, 2] with step length 0.02 on the Breast-cancer-W data set. Each plot shows the mean and the standard deviation over 50 runs of this method. The curve of the average error rates shows that the performance of SC_Nys changes with the scaling parameter. Large standard deviation demonstrates the instability of SC_Nys. Certainly, the number of randomly sampled data points also influences the error rates and its standard deviation.

From the comparison results in Fig. 4, it is found that SC outperforms SC_Nys with most of the scaling parameters in the interval, and the performance of SCE is better than that of SC and SC_Nys except at some special points. However, these special points are difficult to be selected in practice. Therefore, we can conclude that SCE is more efficient than SC and SC_Nys in term of the stability to the scaling parameter. The results in Fig. 5 illustrate this point further.

In addition, in order to investigate the influence of different parameter intervals on the ensemble performance, we perform SCE with different parameter intervals on Segmentation and Pen digits data sets. The results are given in Fig. 5(a) and (b). On the Segmentation data set, SCE is performed with the intervals [1, 5], [5, 10], and [1, 10], respectively. Also, on the
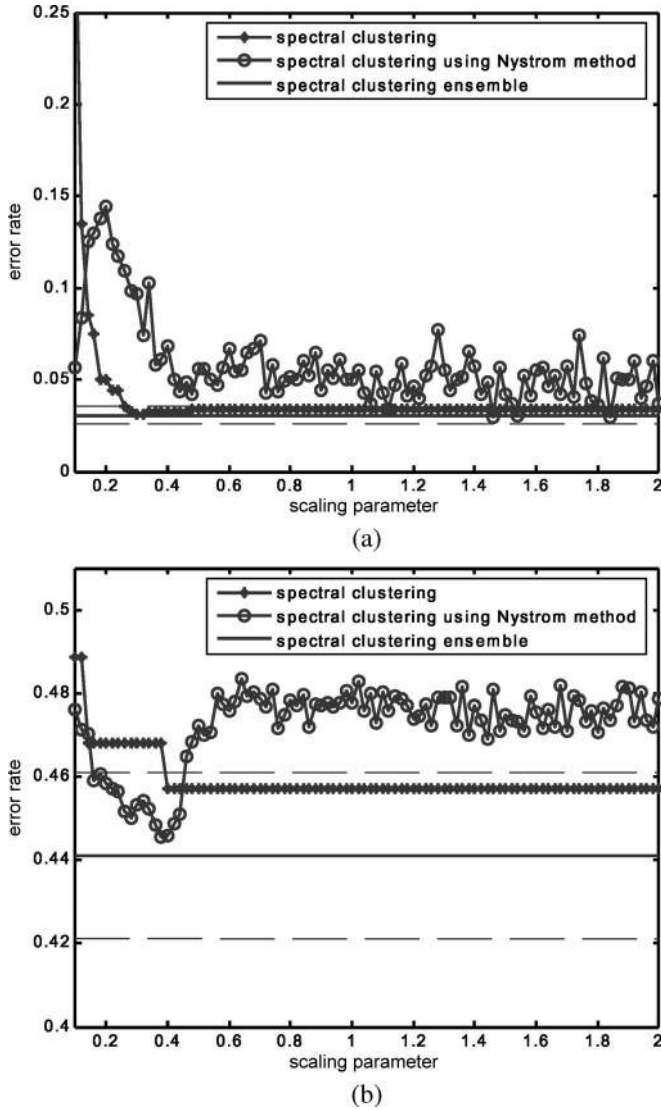
Fig. 4. Error rates of SC and SC_Nys with different scaling parameters, and the error rate of SCE in the interval. The dashed denotes the standard deviation of error rate by SCE. (a) Breast-cancer-W data set. (b) Chess data set.
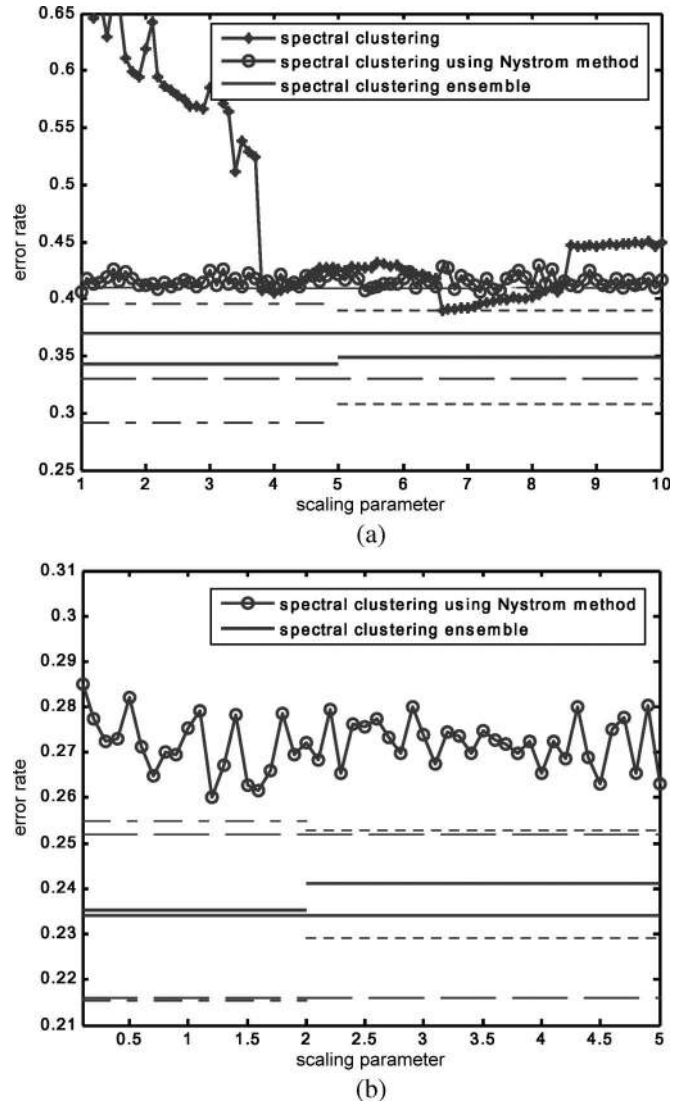

Fig. 5. Comparison of clustering results of SC_Nys and SCE with the different intervals of scaling parameter. The dashed denotes the standard deviation of error rate by SCE. (a) Segmentation data set. (b) Pen digits data set.

Pen digits data set, the parameter intervals are [1, 2], [2, 5], and [1, 5], respectively. From the results, we can find that the clustering results of SCE with different parameter intervals are comparatively stable, and the interval of the scaling parameter slightly influences the performance.

Table III records the minimum error rates obtained by the k-means, SC, SC_Nys, and the average error rates of ten runs of SCE. It can be found that SCE using either MCLA or voting algorithm outperforms the k-means clustering substantially except for the Pen digits data set. In comparison to SC with the accurately selected scaling parameter in the given interval, SCE achieves the minimum error rates except for the Chess (kr-vs-kp) data set. Furthermore, from the results of these two combination methods, we found that the voting algorithm works slightly better than MCLA for the data sets with fewer classes including Sonar, Breast-cancer-W, and Chess (kr-vs-kp). Also, for the data sets with more classes including Segmentation and Pen digits data sets, MCLA performs better than the voting algorithm. Therefore, for various data sets, using

the normalized mutual information to get the final result from the two combination results is more reliable.

The computational cost of the algorithm was assessed via running time on a personal computer with dual-core 1.86 GHz Pentium IV processors, 1 GB memory, and Windows XP operating system, using programs written by Matlab 7.01. Table IV shows the running time of different methods. The listed running

TABLE IV
COMPUTATION TIME OF SINGLE RUN OF SC, SC_Nys, AND SCE ON THE
DATA SETS. TIME DENOTES THE CPU TIME IN SECONDS

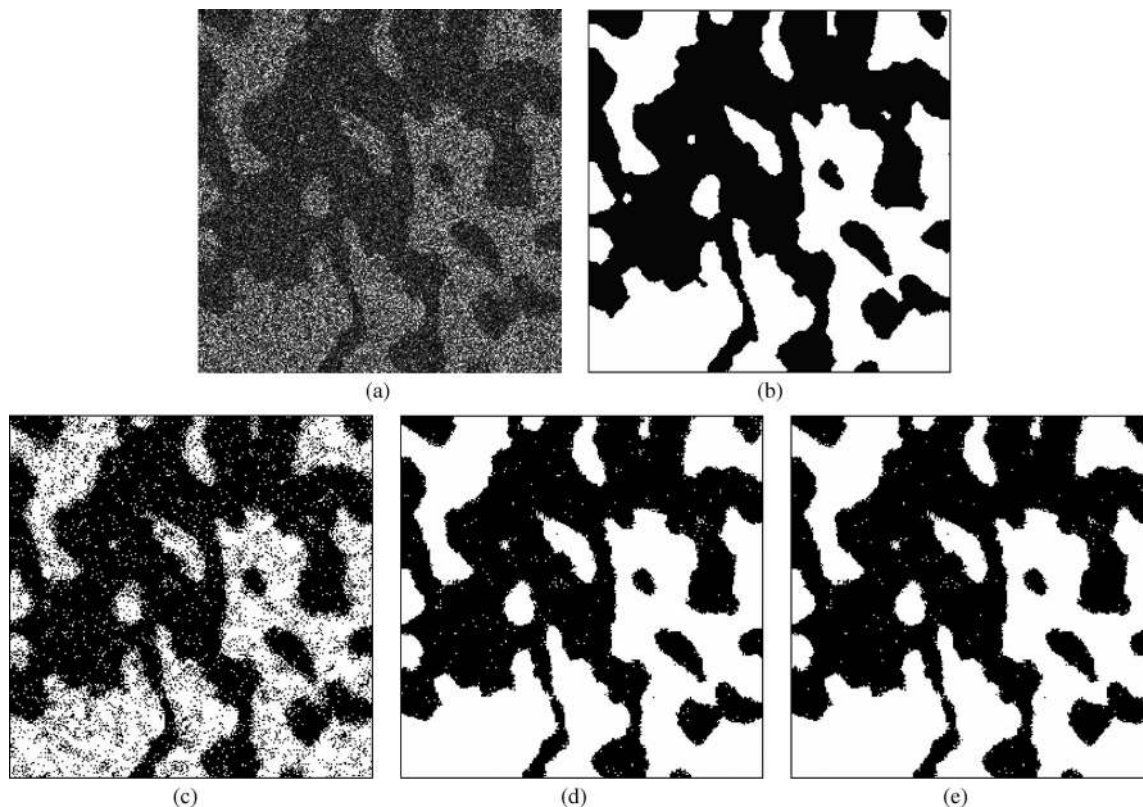| Data set	Spectral Clustering		SCE
	SC	SC_Nys	
Iris	0.13	0.172	5.09
Sonar	0.21	0.213	5.67
Breast-cancer-W	1.12	0.218	6.33
Chess(kr-vs-kp)	33.69	0.594	13.92
Segmentation	23.32	0.582	15.09
Pen digits	-	3.25	82.41

Fig. 6. (a) Three-look simulated SAR image (256 × 256 pixels). (b) Ground truth. (c) Segmentation obtained by the k-means (error rate: 12.73%; the number of misclassified pixels: 8342). (d) Segmentation obtained by SC_Nys (error rate: 4.02%; the number of misclassified pixels: 2635). (e) Segmentation obtained by SCE (error rate: 3.84%; the number of misclassified pixels: 2516).

time of SC and SC_Nys is the average value to perform the algorithm once. The time of SCE is for one trial. Obviously, the number of component clusterings and the quantity of sampled data points in each component SC_Nys would influence the computation time of SCE. From the statistical results in Table IV, we found that the running time of our algorithm is constant multiple of that of SC_Nys. On the medium-scale data sets including Chess (kr-vs-kp), Segmentation and Pen digits data sets, our algorithm is faster than SC.

B. Segmentation of Simulated SAR Image

When we deal with real SAR images segmentation, the ground truth corresponding to the SAR images being segmented are absent generally. In this case, the evaluation of the segmentation result is based on visual inspection of the segmented images. In order to objectively evaluate the performance of the segmenting method, an experiment on a simulated three-look SAR image is carried out. The error rate is computed by matching the segmentation result with the ground truth to evaluate the method.

Following the work of Carincotte *et al.* [46], the generation procedure of the simulated SAR image was inspired by radar image formation phenomena. The ground truth image comes from a two-class Gibbs field, as shown in Fig. 6(b), which is used to calculate the error rates of the segmentations obtained by different algorithms. The corresponding three-look noisy image, as shown in Fig. 6(a), is generated by averaging three independent realizations of speckle.

The texture features are not used in the segmentation of this image for its generation previously mentioned. For each pixel, a two-dimension feature vector is obtained by combining the gray value and the spectral information estimated by $e_{\text{LL}-1}$ of wavelet transform. Also, the features are scaled to the interval [0, 1].

Three algorithms are used for segmentation, respectively: 1) the k-means clustering; 2) SC_Nys; 3) SCE. We performed ten runs of k-means. Fig. 6(c) shows the segmentation result with the minimum error rate of ten runs. Fig. 6(d) shows the segmentation result of SC_Nys, which is the best one of all results obtained by repeatedly performing the algorithm with each parameter from the interval [1, 10] with step length 0.5. One hundred pixels are sampled in the approximation method of SC_Nys. In the application of SCE to image segmentation, 30 component SC_Nys are combined, and 100 pixels are sampled in each component SC_Nys. The scaling parameter σ for each component SC_Nys is randomly selected from the interval [1, 10]. Fig. 6(e) shows the best segmentation result among ten runs of SCE according to the error rate. However, we found that the standard deviation of error rates over ten runs of SCE for image segmentation is small, 0.21, which demonstrates that our algorithm is stable for SAR image segmentation.

Visually, the segmentation of the k-means, which is shown in Fig. 6(c), is seriously spotty in consistent regions. Many pixels in two classes are confused. SC_Nys performs better than the k-means, as shown in Fig. 6(d). Therefore, SC_Nys is more robust to the noise than the k-means. In the result, the
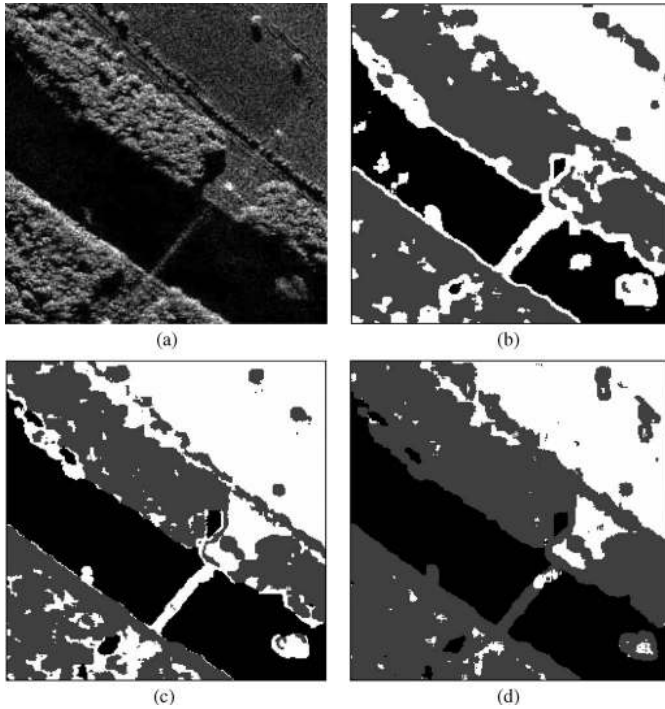
Fig. 7. (a) Ku-SAR image in the area of Rio Grande River near Albuquerque, NM (256 × 256). (b) Segmentation obtained by the k-means. (c) Segmentation obtained by SC_Nys. (d) Segmentation obtained by SCE.

Fig. 8. (a) X-SAR subimage of Switzerland (256 × 256). (b) Segmentation obtained by k-means. (c) Segmentation obtained by SC_Nys. (d) Segmentation obtained by SCE.

misclassified pixels mainly locate in the black regions of Fig. 6(a), and the gray regions of Fig. 6(a) are well segmented. Compared with SC_Nys, SCE reduces the number of the misclassified pixels, as shown in Fig. 6(e).

The overall error rates are reduced from 12.73% to 4.02% by using SC_Nys and to 3.84% by using SCE. Therefore, SCE is better than SC_Nys; SC_Nys and SCE outperform the k-means clustering.

C. Segmentation of Real SAR Images

To demonstrate the applicability of the SC and the SCE to real SAR image segmentation, experiments on three SAR images are performed.

In the experiments, the segmentation result of k-means is selected according to visual performance from ten runs of the algorithm. SC_Nys is performed with each parameter in the interval [1, 10] with step length 0.5. The final segmentation is selected from multiple results visually. Since SCE is stable for SAR image segmentation, we get the segmentation result by performing the algorithm once.

The image, as shown in Fig. 7(a), is a part of a Ku-band SAR image with 1-m spatial resolution in the area of Rio Grande River near Albuquerque, NM. This image consists of three types of land cover: water, vegetation, and crop.

The segmentation obtained by the k-means is shown in Fig. 7(b). The boundary between the water and the vegetation is not correctly defined, and the vegetation near the water is classified as the crop. Furthermore, some local regions in the water area are also misclassified. The segmentation obtained by SC_Nys, as shown in Fig. 7(c), improves the uniformity in the water region. However, there is serious misclassification in
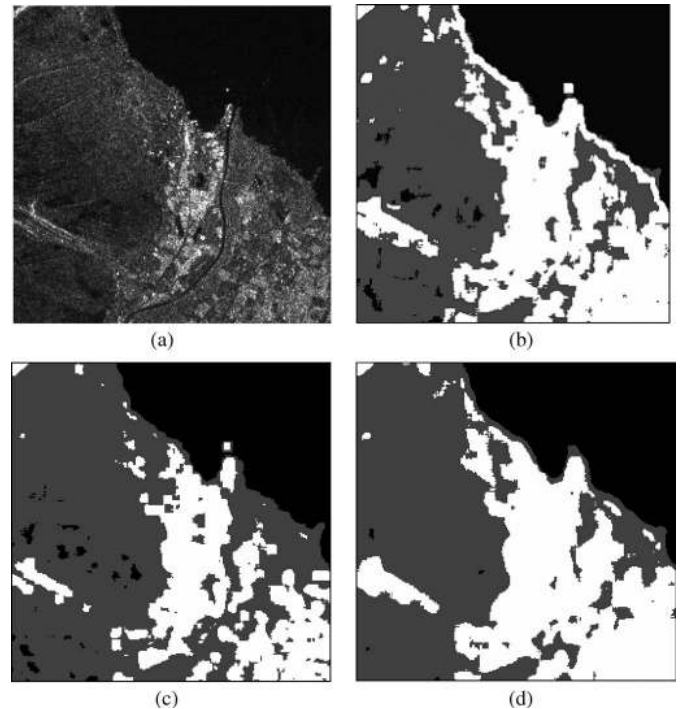
the vegetation region. SCE gets the best segmentation result, as shown in Fig. 7(d). The boundaries of unique regions are well defined, and three types of land cover are consistently identified as corresponding regions. In particular, the only algorithm that correctly identifies the small part of vegetation in the water region is SCE.

Another experiment is carried out on a three-look X-band SAR subimage of Switzerland obtained by Space Radar Laboratory Missions in 1994, as shown in Fig. 8(a). The image consists of three types of land cover: water, urban area, and mountain. The left part of this image is the mountain area, the water is on the top right, and the bottom right is the urban area. Visually, it would be difficult to segment the urban area and the mountain area in this image. The segmentation obtained by the k-means clustering is shown in Fig. 8(b). The water area is well separated. However, the boundary between the water and the mountain is not well defined. Besides this, a part of local regions in the mountain area are classified as the water, and a small part of the mountain areas are classified as the urban areas. Therefore, the k-means clustering is not effective for this image. SC_Nys improves the segmentation result to some degree, as shown in Fig. 8(c). The water and the mountain areas are well distinguished. In particular, the misclassified regions in the mountain area using the k-means are segmented correctly. However, the urban area is segmented badly by SC_Nys, because some regions are misclassified as the mountain area. The segmentation of SCE shows an effective segmentation result in comparison with those of the k-means and SC_Nys, as shown in Fig. 8(d). The uniformity in the urban area and the mountain area is improved, and the mountain area and the water area are identified as well.
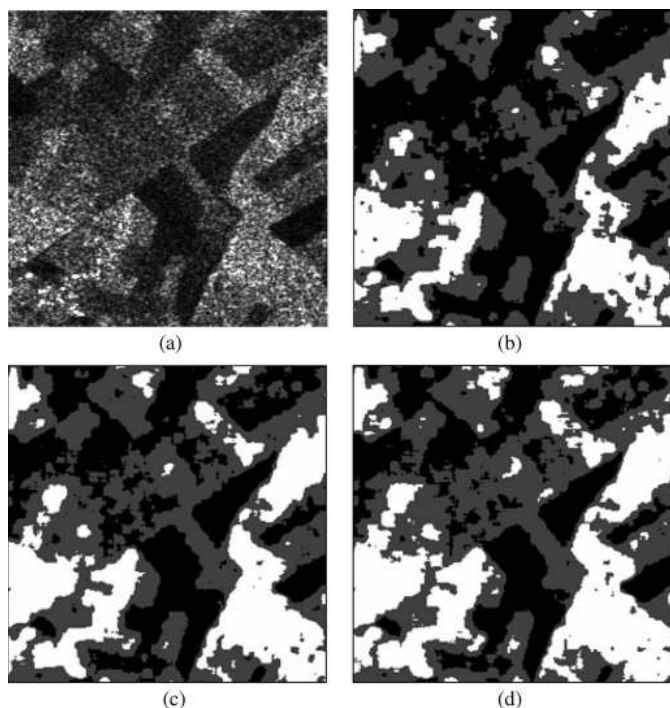
Fig. 9. (a) ERS-2 SAR image (256 × 254). (b) Segmentation obtained by *k*-means. (c) Segmentation obtained by SC_Nys. (d) Segmentation obtained by SCE.

The last experiment is carried out on a four-look second European Remote Sensing Satellite (ERS-2) SAR image with much speckle noise, as shown in Fig. 9(a). It consists of three types of crops, represented by white, gray, and black. It is not easy to visually identify some regions in this image. Fig. 9(b) shows the segmentation obtained by the *k*-means. In Fig. 9(b), the crop regions represented by white in Fig. 9(a) are seriously misclassified as gray crop regions, particularly in the top of Fig. 9(b). In the central area of Fig. 9(b), the gray crop regions in Fig. 9(a) are misclassified as black crop regions. Also, in the right of Fig. 9(b), the white crop regions are not well segmented either. Some small white crop regions are misclassified as gray crop regions. Compared with the *k*-means, SC_Nys produces a better segmentation for the same areas in Fig. 9(c). However, the white crop regions in the bottom right of Fig. 9(a) are still not well defined in Fig. 9(c). In Fig. 9(d), SCE improved the segmentation. The result of SCE is better than the result of SC_Nys and greatly improved than the result of the *k*-means in upper mentioned areas.

In the above experiments, to segment a SAR image of 256 × 256 pixels, the average running time of single SC_Nys is around 23.03 s, whereas SCE needs about 561.79 s. It seems that SCE requires large running cost because the ensemble combines lots of components. However, in the application of SC_Nys, we need to perform many times of SC_Nys to select the final result. The procedure requires multiple running time of single SC_Nys. Therefore, the computation time cost of SCE is comparable to that of the SC.

VI. CONCLUSION

We have developed an ensemble algorithm based on the SC for the segmentation of SAR image. In comparison with the *k*-means clustering, the SCE achieves better performance on the cases we have studied. It also outperforms the SC with the accurately selected scaling parameter in some cases. Furthermore, the SCE avoids the selection of the scaling parameter for single SC.

The diversity is crucial for ensemble learning. From the experiments, we found that diversity of the components can be achieved by the strategies, which are mentioned in this paper. However, how to evaluate the diversity of individual clustering is a challenging problem. It is expected that using a diversity evaluation to guide the ensemble construction will improve the performance further.

REFERENCES

[1] R. J. Dekker, "Texture analysis and classification of ERS SAR images for map updating of urban areas in The Netherlands," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1950–1958, Sep. 2003.

[2] Q. Zhao and J. C. Principe, "Support vector machines for SAR automatic target recognition," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, no. 2, pp. 643–654, Apr. 2001.

[3] J. S. Lee and I. Jurkevich, "Segmentation of SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 27, no. 6, pp. 674–680, Nov. 1989.

[4] A. E. Zaart, D. Ziou, S. Wang, and Q. Jiang, "Segmentation of SAR images using mixture of gamma distribution," *Pattern Recognit.*, vol. 35, no. 3, pp. 713–724, 2002.

[5] P. R. Kersten, J.-S. Lee, and T. L. Ainsworth, "Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 519–527, Mar. 2005.

[6] W. Chumsamrong, P. Thitimajshima, and Y. Rangsanseri, "Synthetic aperture radar (SAR) image segmentation using a new modified fuzzy c-means algorithm," in *Proc. IEEE Symp. Geosci. Remote Sens.*, Honolulu, HI, 2000, pp. 624–626.

[7] R. Samadani, "A finite mixtures algorithm for finding proportions in SAR images," *IEEE Trans. Image Process.*, vol. 4, no. 8, pp. 1182–1185, Aug. 1995.

[8] Y. Dong, B. C. Forster, and A. K. Milne, "Comparison of radar image segmentation by Gaussian- and Gamma-Markov random field models," *Int. J. Remote Sens.*, vol. 24, no. 4, pp. 711–722, Feb. 2003.

[9] H. Deng and D. A. Clausi, "Unsupervised segmentation of synthetic aperture radar sea ice imagery using a novel Markov random field model," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 528–538, Mar. 2005.

[10] C. Lemaréchal, R. Fjørtoft, P. Marthon, E. Cubero-Castan, and A. Lopes, "SAR image segmentation by morphological methods," *Proc. SPIE*, vol. 3497, pp. 111–121, 1998.

[11] B. Ogor, V. Haese-coat, and J. Ronsin, "SAR image segmentation by mathematical morphology and texture analysis," in *Proc. IGARSS*, 1996, pp. 717–719.

[12] C. Fowlkes, S. Belongie, F. Chung, and J. Malik, "Spectral grouping using the Nyström method," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 214–225, Feb. 2004.

[13] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[14] M. Fiedler, "Algebraic connectivity of graphs," *Czechoslovak Math. J.*, vol. 23, pp. 298–305, 1973.

[15] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Advances in Neural Information Processing Systems (NIPS)*, vol. 14, T. Dietterich, S. Becker, and Z. Ghahramani, Eds. Cambridge, MA: MIT Press, 2002.

[16] C. Ding, X. He, H. Zha, M. Gu, and H. Simon, "A min-max cut algorithm for graph partitioning and data clustering," in *Proc. IEEE Int. Conf. Data Mining*, 2001, pp. 107–114.

[17] S. X. Yu and J. Shi, "Multiclass spectral clustering," in *Proc. Int. Conf. Comput. Vis.*, 2003, vol. 1, pp. 313–319.

[18] F. R. Bach and M. I. Jordan, "Blind one-microphone speech separation: A spectral learning approach," in *Advances in Neural Information Processing Systems*, vol. 17. Cambridge, MA: MIT Press, 2005.

[19] P. K. Chan, M. D. F. Schlag, and J. Y. Zien, "Spectral k-way ratio-cut partitioning and clustering," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 13, no. 9, pp. 1088–1096, Sep. 1994.

[20] L. Zelnik-Manor and P. Perona, "Self-tuning spectral clustering," in *Advances in Neural Information Processing Systems*, vol. 17, L. K. Saul, Y. Weiss, and L. Bottou, Eds.   Cambridge, MA: MIT Press, 2005.

[21] M. Breitenbach and G. Z. Grudic, "Clustering through ranking on manifolds," in *Proc. 22nd Int. Conf. Mach. Learn.*, Bonn, Germany, 2005, pp. 73–80.

[22] A. Krogh and J. Vedelsby, "Neural network ensembles, cross validation, and active learning," in *Advances in Neural Information Processing Systems*.   Cambridge, MA: MIT Press, 1995, pp. 231–238.

[23] L. Hansen and P. Salamon, "Neural network ensembles," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 10, pp. 993–1001, Oct. 1990.

[24] A. Strehl and J. Ghosh, "Cluster ensembles—A knowledge reuse framework for combining multiple partitions," *J. Mach. Learn. Res.*, vol. 3, no. 3, pp. 583–617, Mar. 2002.

[25] X. Z. Fern and C. E. Brodley, "Random projection for high dimensional data clustering: A cluster ensemble approach," in *Proc. 20th Int. Conf. Mach. Learn.*, 2003, pp. 186–193.

[26] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832–844, Aug. 1998.

[27] A. H. S. Solberg and A. K. Jain, "Texture fusion and feature selection applied to SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 2, pp. 475–479, Mar. 1997.

[28] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. 3, no. 6, pp. 610–621, Nov. 1973.

[29] S. Fukuda and H. Hirosawa, "A wavelet-based texture feature set applied to classification of multifrequency polarimetric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 8, pp. 2282–2286, Sep. 1999.

[30] D. A. Clausi and B. Yue, "Comparing cooccurrence probabilities and Markov random fields for texture analysis of SAR sea ice imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 1, pp. 215–228, Jan. 2004.

[31] D. A. Clausi and Y. Zhao, "Rapid extraction of image texture by co-occurrence using a hybrid data structure," *Comput. Geosci.*, vol. 28, no. 6, pp. 763–774, Jul. 2002.

[32] D. G. Barber and E. F. LeDrew, "SAR sea ice discrimination using texture statistics: A multivariate approach," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 4, pp. 385–395, 1991.

[33] D. A. Clausi, "An analysis of co-occurrence texture statistics as a function of grey level quantization," *Can. J. Remote Sens.*, vol. 28, no. 1, pp. 45–62, 2002.

[34] L. J. Du, J. S. Lee, K. Hoppel, and S. A. Mango, "Segmentation of SAR images using the wavelet transform," *Int. J. Imaging Syst. Technol.*, vol. 4, no. 4, pp. 319–329, 1992.

[35] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996.

[36] Y. Freund, "Boosting a weak learning algorithm by majority," *Inf. Comput.*, vol. 121, no. 2, pp. 256–285, Sep. 1995.

[37] A. Topchy, A. K. Jain, and W. Punch, "A mixture model for clustering ensembles," in *Proc. SIAM Int. Conf. Data Mining*, 2004, pp. 379–390.

[38] M. H. C. Law, A. P. Topchy, and A. K. Jain, "Multiobjective data clustering," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2004, vol. 2, pp. 424–430.

[39] S. Dudoit and J. Fridlyand, "Bagging to improve the accuracy of a clustering procedure," *Bioinformatics*, vol. 19, no. 9, pp. 1090–1099, Jun. 2003.

[40] A. Topchy, B. Minaei-Bidgoli, A. K. Jain, and W. Punch, "Adaptive clustering ensembles," in *Proc. Int. Conf. Pattern Recog.*, Cambridge, U.K., 2004, pp. 272–275.

[41] A. L. N. Fred, "Finding consistent clusters in data partitions," in *Proc. 3rd Int. Workshop Multiple Classifier Syst.,* F. Roli and J. Kittler, Eds., 2001, vol. 2364, pp. 309–318.

[42] A. L. N. Fred and A. K. Jain, "Combining multiple clusterings using evidence accumulation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 6, pp. 835–850, Jun. 2005.

[43] Y. Jiang and Z.-H. Zhou, "SOM ensemble-based image segmentation," *Neural Process. Lett.*, vol. 20, no. 3, pp. 171–178, Nov. 2004.

[44] V. Vapnik, *The Nature of Statistical Learning Theory*.   Berlin, Germany: Springer-Verlag, 1999.

[45] C. Blake, E. Keogh, and C. J. Merz, *UCI Repository of Machine Learning Databases*.   Irvine, CA: Dept. Inf. Comput. Sci., Univ. California, 1998. [Online]. Available: http://www.ics.uci.edu/~mlearn/MLRepository.html

[46] C. Carincotte, S. Derrode, and S. Bourennance, "Unsupervised change detection on SAR images using fuzzy hidden Markov chains," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 432–441, Feb. 2006.

Xiangrong Zhang (M'07) received the B.S. and M.S. degrees from the School of Computer Science, Xidian University, Xi'an, China, in 1999 and 2003, respectively, and the Ph.D. degree from the School of Electronic Engineering, Xidian University, in 2006.

She currently works as a Lecturer with the Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education of China, Institute of Intelligent Information Processing, Xidian University. Her current research interests include synthetic aperture radar image analysis and understanding, pattern recognition, and machine learning.

Licheng Jiao (SM'89) received the B.S. degree from Shanghai Jiaotong University, Shanghai, China, in 1982, and the M.S. and Ph.D. degrees from Xi'an Jiaotong University, Xi'an, China, in 1984 and 1990, respectively.

From 1990 to 1991, he was a Postdoctoral Fellow with the National Key Laboratory for Radar Signal Processing, Xidian University, Xi'an. From June 1992 to March 2003, he was a Professor with the National Key Laboratory of Radar Signal Processing. Since 2004, he has been a Distinguished Professor with Xidian University, where he is currently the Dean of the Electronic Engineering School and the Director of the Institute of Intelligent Information Processing. He is the author or coauthor of over ten books and more than 200 papers in journals and conferences. His current research interests include signal and image processing, machine learning, natural computation, and intelligent information processing.

Fang Liu (M'07–SM'07) received the B.S. degree in computer science and technology from Xi'an Jiaotong University, Xi'an, China, in 1984, and the M.S. degree in computer science and technology from Xidian University, Xi'an, in 1995.

Currently, she is a Professor with the School of Computer Science, Xidian University. Her research interests include synthetic aperture radar image processing, multiscale geometry analysis, optimization problems, and data mining.

Liefeng Bo received the B.S. degree from the School of Science, Xidian University, Xi'an, China, in 2002, and the Ph.D. degree in circuits and systems from the Institute of Intelligent Information Processing, Xidian University, in 2007.

He is currently a Postdoctoral Scholar with the Toyota Technological Institute at Chicago (TTI-C), Chicago, IL. His current research interests include kernel-based learning, manifold learning, neural networks, and computer vision. He is the author of several papers in published some leading journals such as *Neural Computation* and IEEE TRANSACTIONS ON NEURAL NETWORKS.

Maoguo Gong (M'07) received the B.S. degree in 2003 from Xidian University, Xi'an, China, where he is currently working toward the Ph.D. degree at the Institute of Intelligent Information Processing.

He currently works as a Lecturer under the Innovative Research Team of the Ministry of Education of China, Xidian University, Xi'an, China. His research interests are broadly in the area of computational intelligence and hybrid intelligent systems. The areas of special interests include artificial immune systems, evolutionary computation, data mining, optimization, and some other related areas. He has published more than 30 papers in journals and conferences.