Superpixel-Based Classification With an Adaptive Number of Classes for Polarimetric SAR Images

Bin Liu, *Student Member, IEEE*, Hao Hu, Huanyu Wang, Kaizhi Wang, Xingzhao Liu, *Member, IEEE*, and Wenxian Yu

Abstract—Polarimetric synthetic aperture radar (PolSAR) image classification, an important technique in the remote sensing area, has been deeply studied for a couple of decades. In order to develop a robust automatic or semiautomatic classification system for PolSAR images, two important problems should be addressed: 1) incorporation of spatial relations between pixels; 2) estimation of the number of classes in the image. Therefore, in this paper, we present a novel *superpixel-based* classification framework *with an adaptive number of classes* for PolSAR images. The approach is mainly composed of three operations. First, the PolSAR image is partitioned into superpixels, which are local, coherent regions and preserve most of the characteristics necessary for image information extraction. Then, the number of classes and each class center within the data are estimated using the pairwise dissimilarity information between superpixels, followed by the final classification operation. The proposed framework takes the spatial relations between pixels into consideration and makes good use of the inherent statistical characteristics and contour information of PolSAR data. The framework is capable of improving the classification accuracy, making the results more understandable and easier for further analyses, and providing robust performance under various numbers of classes. The performance of the proposed classification framework on one synthetic and three real data sets is presented and analyzed; and the experimental results show that the framework provides a promising solution for unsupervised classification of PolSAR images.

Index Terms—Number-of-classes (NoC) estimation, polarimetric synthetic aperture radar (PolSAR) images, superpixel, unsupervised classification.

I. INTRODUCTION

THE SYNTHETIC aperture radar (SAR) system provides a day-or-night, all-weather means of remote sensing and produces high-resolution images of the land under illumination of radar beams. Polarimetric SAR (PolSAR) is the advanced form of SAR and focuses on emitting and receiving polarized radar waves to characterize observed land covers and targets.

In the last two decades, researches have shown that PolSAR is capable of providing more useful information than single polarization cases in target detection [1], [2], terrain classification [3], parameter inversion [4], topography extraction [5], and so on. Nowadays, several spaceborne platforms, such as TerraSAR-X, RADARSAT-2, and ALOS-PALSAR, continuously provide enormous amount of PolSAR data. It is ineffective to interpret these large volumes of highly complicated images manually. Therefore, developing automatic or semiautomatic systems for PolSAR image interpretation and information mining is urgently required and widely studied.

In this paper, we focus on terrain classification for PolSAR images, which is arguably the most important application of PolSAR [6] and fundamental to exploiting the enormous amount of PolSAR data. It is a key requirement in both military and civil sectors, a highly desired goal for developing intelligent databases and also a necessary process for target detection and recognition. Many algorithms have been developed for supervised and unsupervised classification for the last two decades. Techniques, such as image processing, physical scattering analysis, and statistical analysis, are applied for PolSAR image classification. A detailed overview can be found in [6], [7], and Section II of this paper.

In these developed classification methods, although pixel-based classification methods have achieved great results, due to speckle noise, the traditional pixel-based classification still has some drawbacks [8]–[10]. In contrast, region-based classification is a promising scheme. Images are first segmented into many homogeneous regions. The following classification is based on regions instead of pixels. Though region-based approach can achieve improved and more understandable classification results, it relies on the accuracy of segmentation to a great extent. Imprecise segmentation can cause information loss, even lead to incorrect classification results. Until now, automatic segmentation for PolSAR images is still a challenging task. Many studies have been carried out on this topic recently. Related segmentation methods for PolSAR images are overviewed in Section II. In this paper, based on the idea of superpixel, we add inherent statistical characteristics of PolSAR data into the contour information, use the normalized cuts, and propose a superpixel-based classification method for PolSAR images.

In addition to incorporation of spatial relations between neighboring pixels, another problem in unsupervised classification for PolSAR images is that the number of classes is generally unknown, and it will affect the performance of classification algorithms greatly [11], [12]. In this paper,

B. Liu, H. Hu, K. Wang, X. Liu, and W. Yu are with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: liubin22810@gmail.com; waszxq@sjtu.edu.cn; kz_wang@sjtu.edu.cn; xzhliu@sjtu.edu.cn; wxyu@sjtu.edu.cn).

H. Wang is with Shanghai Laboratory, Fujitsu Research and Development Center Co., Ltd., Shanghai 201204, China (e-mail: bothwinfx@gmail.com).

based on the generated superpixels, we propose a preclustering number-of-classes (NoC) estimation scheme for PolSAR images.

Therefore, in this paper, we present a novel *superpixel-based* classification framework *with an adaptive number of classes* for PolSAR images. We first calculate the edge maps, based on which the superpixels are generated. Then, the NoC estimation operation estimates the number of classes within the image and extracts each class center, followed by the final classification operation. The proposed framework has several advantages as follows.

— The framework is based on superpixels instead of individual pixels. Hence, it takes the spatial relations between pixels into account, which makes the classification process more effective and the results more understandable.
— The framework incorporates the merits of graph partitioning schemes, makes good use of the statistical characteristics and contour information of PolSAR data, and improves the classification accuracy.
— The framework extracts the number of classes from the data, thus it estimates each class center before the final classification more accurately. It provides robust performance under various numbers of classes, which is very important for unsupervised classification of PolSAR images.

In this paper, the performance of the proposed framework is presented and analyzed on one synthetic and three real experimental PolSAR data sets.

The remainder of this paper is organized as follows. Section II presents the related work of this paper. Section III introduces the PolSAR data for completeness. Sections IV–VI discuss in detail the proposed framework. The implementation procedures are provided in Section VII. Experiments and discussions are given in Section VIII. Conclusion appears in Section IX.

II. RELATED WORK

PolSAR image classification follows three major approaches. One type of research focuses on analyzing the polarimetric scattering mechanisms, which has the merits that some information about class type is provided. Typical studies include the classification schemes proposed by van Zyl [13], Cloude and Pottier [14], and Ferro-Famil *et al.* [15]. The second type of algorithm is mainly based on the statistical characteristics of PolSAR images. For single-look PolSAR data, Kong *et al.* [16] derived a maximum likelihood (ML) classifier based on the complex Gaussian distribution. For multilook PolSAR data represented in coherence or covariance matrices, Lee *et al.* [17] have derived a Wishart distance to classify PolSAR data. The variants of this method, such as using expectation maximization schemes [18] and fuzzy decision rules [19], [20], are found in the literature. In the third category, both the analyses of scattering mechanisms and statistical information are combined. Lee *et al.* [21] have used the Cloude-Pottier decomposition to initialize the iterated Wishart clustering. The related approaches can be found in [15] and [22].

All these aforementioned algorithms can be categorized as pixel-based classification, in which each pixel is treated independently. As pointed out in Section I, pixel-based classification methods still have some limitations. The inclusion of contextual information in the statistical decision concerning class membership is reasonable. It can improve classification accuracy and result understandability.

A common way to include spatial relations between pixels is to model the labeling process as a Markov random field (MRF) [23], [24]. Dong *et al.* [25] proposed a segment-based classification method for PolSAR images based on the Gaussian MRF model; and Wu *et al.* [9] introduced a region-based classification scheme for PolSAR images based the Wishart MRF (WMRF) model.

Other approaches for integrating contextual information can be found in the literature. Reigber *et al.* [8] presented an approach for incorporating spatial context called probabilistic label relaxation. Hoekman *et al.* [26] and Li *et al.* [10] used region growing technique as an important preprocessing operation to improve classification accuracy and increase computational efficiency of following steps. Benz and Pottier [27] developed an object-based analysis method for PolSAR data using the eCognition software. It is worth noting that Ersahin *et al.* [28], [29] did pioneering research of PolSAR image segmentation based on spectral graph partitioning and contour information to realize an object-based classification. The differences between their work and this paper are discussed detailedly at the end of Section IV.

In this paper, we oversegment PolSAR images into superpixels to integrate contextual information of neighborhood and improve classification accuracy. The idea of superpixel is originally proposed by Ren and Malik [30]. It represents a local, coherent region, which preserves most of the characteristics necessary for image information mining, and they are roughly homogeneous in size and shape just like pixels, particularly the number of the superpixels is very high and in homogeneous areas. Hence, they are called as superpixels. It is rather a vivid description. Superpixels are generated using the normalized cuts algorithm [31]. Based on the idea of superpixel, Ren and Malik [30] learned a classification model for segmentation; and Mori [32] demonstrated using superpixels to improve the efficiency and accuracy of model search in an image.

As introduced in Section I, the number of classes is a crucial parameter in the unsupervised classification of PolSAR images [11], [12], and it can affect the quality of the final classification map greatly. This topic, however, is not widely studied in the literature.

Tran *et al.* [11] and Cao *et al.* [12] have proposed unsupervised classification with an adaptive number of clusters for PolSAR images. Their methods are based on agglomerative hierarchical clustering (AHC) and postclustering validation. The AHC is applied to homogeneous regions or clusters in the PolSAR image, generating a dendrogram, from the predefined maximum number of clusters to the minimum number of clusters. After visiting every level of the dendrogram, the number of clusters is obtained using a validation function, such as the pseudo-likelihood information criterion [11] and the data log-likelihood function [12]. Different from these methods, in this

paper, we directly estimate the number of classes existing in the scene and each class center. The advantages and implementation details of this method are discussed in Section V.

III. POLARIMETRIC SAR DATA

A. Wishart Distribution

For a reciprocal medium illuminated by a monostatic SAR, a complex scattering vector in the linear basis is

$$\mathbf{u} = [S_{\mathrm{HH}} \quad \sqrt{2}S_{\mathrm{HV}} \quad S_{\mathrm{VV}}]^{\mathrm{T}} \tag{1}$$

where the superscript "T" denotes the matrix transpose.

Most SAR data are multilook-processed for speckle reduction. The multilook PolSAR data can be represented by a polarimetric covariance matrix $\mathbf{C}$, and

$$\mathbf{C} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{u}_i\mathbf{u}_i^{*\mathrm{T}} \tag{2}$$

where $\mathbf{u}_i$ denotes the ith sample of vector $\mathbf{u}$, the superscript "$*$" denotes the complex conjugate, and n is the number of looks.

In homogeneous areas with fully developed speckle and no texture, the polarimetric covariance matrix $\mathbf{C}$ has a complex Wishart distribution [33]. Let $\boldsymbol{\Sigma} = E\{\mathbf{u}\mathbf{u}^{*\mathrm{T}}\}$. The probability density function (PDF) for the covariance matrix $\mathbf{C}$ is

$$p_{\mathbf{C}}(\mathbf{C}|n, \boldsymbol{\Sigma}) = \frac{n^{qn}|\mathbf{C}|^{n-q}\exp\left[-n\cdot\mathrm{Tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C})\right]}{K(n,q)|\boldsymbol{\Sigma}|^n}$$

$$K(n,q) = \pi^{\left(\frac{1}{2}\right)q(q-1)}\Gamma(n)\dots\Gamma(n-q+1) \tag{3}$$

where $q = 3$ for the monostatic SAR on a reciprocal medium and $q = 4$ for the bistatic SAR, $\mathrm{Tr}(\cdot)$ is the trace of a matrix, $K(n,q)$ is a normalization factor, and $\Gamma(\cdot)$ is the gamma function [17]. Therefore, each point feature in the PolSAR image is modeled as a complex covariance matrix obeying the Wishart PDF.

B. Dissimilarity Measures Between Two Regions

Measuring pairwise dissimilarities between two regions in the PolSAR image is essential to the proposed framework, and it is used in the edge map calculation and NoC estimation. Therefore, this subsection discusses the dissimilarity measures between two regions in the PolSAR image for completeness.

Inspired by [34], the Wishart test statistic is utilized to measure the pairwise dissimilarities. The problem is formulated as a likelihood-ratio function to test the equality of center covariance matrices of two regions. The hypothesis test is

$$H_0 : \boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}_j \text{ versus } H_1 : \boldsymbol{\Sigma}_i \neq \boldsymbol{\Sigma}_j \tag{4}$$

where $\boldsymbol{\Sigma}_i$ and $\boldsymbol{\Sigma}_j$ are the center covariance matrices of the ith and jth regions, respectively.

Let Θ_i and Θ_j be the sample covariance matrix data sets of the ith and jth regions, respectively. It is assumed that sample

covariance matrices are spatially independent, therefore, the conditional PDFs of Θ_i and Θ_j are

$$p(\Theta_i|\boldsymbol{\Sigma}_i) = \prod_{l=1}^{N_i} p_{\mathbf{C}_l}(\mathbf{C}_l|n, \boldsymbol{\Sigma}_i) \tag{5}$$

$$p(\Theta_j|\boldsymbol{\Sigma}_j) = \prod_{l=1}^{N_j} p_{\mathbf{C}_l}(\mathbf{C}_l|n, \boldsymbol{\Sigma}_j) \tag{6}$$

where N_i and N_j are the numbers of samples in the ith and jth regions, respectively. The likelihood function under H_0 is

$$L_{H_0}(\boldsymbol{\Sigma}|\Theta_i, \Theta_j) = \prod_{l=1}^{N_i} p_{\mathbf{C}_l}(\mathbf{C}_l|n, \boldsymbol{\Sigma}) \prod_{l=1}^{N_j} p_{\mathbf{C}_l}(\mathbf{C}_l|n, \boldsymbol{\Sigma}) \tag{7}$$

where $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}_j$, and the likelihood function under H_1 is

$$L_{H_1}(\boldsymbol{\Sigma}_i, \boldsymbol{\Sigma}_j|\Theta_i, \Theta_j) = \prod_{l=1}^{N_i} p_{\mathbf{C}_l}(\mathbf{C}_l|n, \boldsymbol{\Sigma}_i) \prod_{l=1}^{N_j} p_{\mathbf{C}_l}(\mathbf{C}_l|n, \boldsymbol{\Sigma}_j). \tag{8}$$

The ML estimators of $\boldsymbol{\Sigma}_i$, $\boldsymbol{\Sigma}_j$, and $\boldsymbol{\Sigma}$ are $(1/N_{\mathrm{SCM}})\sum_{l=1}^{N_{\mathrm{SCM}}}\mathbf{C}_l$, where N_{SCM} is N_i, N_j, and $N_i + N_j$ for $\boldsymbol{\Sigma}_i$, $\boldsymbol{\Sigma}_j$, and $\boldsymbol{\Sigma}$, respectively. They are derived by taking the logarithmic derivatives of (7) and (8), and setting them to zeros.

According to [34] and [12], the dissimilarity measure between the ith and jth regions can be derived from the likelihood ratio test. The test statistic is

$$Q_1 = \frac{L_{H_0}(\hat{\boldsymbol{\Sigma}}|\Theta_i, \Theta_j)}{L_{H_1}(\hat{\boldsymbol{\Sigma}}_i, \hat{\boldsymbol{\Sigma}}_j|\Theta_i, \Theta_j)}$$

$$= \frac{\prod_{l=1}^{N_i} p_{\mathbf{C}_l}(\mathbf{C}_l|n, \hat{\boldsymbol{\Sigma}}) \prod_{l=1}^{N_j} p_{\mathbf{C}_l}(\mathbf{C}_l|n, \hat{\boldsymbol{\Sigma}})}{\prod_{l=1}^{N_i} p_{\mathbf{C}_l}(\mathbf{C}_l|n, \hat{\boldsymbol{\Sigma}}_i) \prod_{l=1}^{N_j} p_{\mathbf{C}_l}(\mathbf{C}_l|n, \hat{\boldsymbol{\Sigma}}_j)}$$

$$= \frac{|\hat{\boldsymbol{\Sigma}}_i|^{nN_i}|\hat{\boldsymbol{\Sigma}}_j|^{nN_j}}{|\hat{\boldsymbol{\Sigma}}|^{n(N_i+N_j)}} \tag{9}$$

where $\hat{\boldsymbol{\Sigma}}_i$, $\hat{\boldsymbol{\Sigma}}_j$, and $\hat{\boldsymbol{\Sigma}}$ are ML estimators of $\boldsymbol{\Sigma}_i$, $\boldsymbol{\Sigma}_j$, and $\boldsymbol{\Sigma}$, respectively. If the value of Q_1 is too low, the null hypothesis $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}_j$ is rejected. Thus, the dissimilarity measure between the ith and jth regions can be defined as

$$D_{\mathrm{S}}(S_i, S_j) = -\frac{1}{n}\ln Q_1$$

$$= (N_i + N_j)\ln|\hat{\boldsymbol{\Sigma}}| - N_i\ln|\hat{\boldsymbol{\Sigma}}_i| - N_j\ln|\hat{\boldsymbol{\Sigma}}_j| \tag{10}$$

where S_i and S_j represent the ith and jth regions, respectively. Details of the aforementioned derivation can be found in [12]. The dissimilarity measure D_{S} is symmetric. If $i = j$, $D_{\mathrm{S}}(S_i, S_j)$ has a minimum value, zero. If the ith and jth regions are more dissimilar, the value of $D_{\mathrm{S}}(S_i, S_j)$ is higher.

Assume that $\boldsymbol{\Sigma}_j$ is known for hypotheses H_0 and H_1, the hypothesis test becomes a general binary hypothesis test for a

given class j. The test statistic can be given as [18], [35]

$$Q_2 = \frac{L_{H_0}(\hat{\Sigma}_j|\Theta_i)}{L_{H_1}(\hat{\Sigma}_i|\Theta_i)} = \frac{\prod\limits_{l=1}^{N_i} p_{\mathbf{C}_l}(\mathbf{C}_l|n, \hat{\Sigma}_j)}{\prod\limits_{l=1}^{N_i} p_{\mathbf{C}_l}(\mathbf{C}_l|n, \hat{\Sigma}_i)}$$

$$= \frac{|\hat{\Sigma}_i|^{nN_i}}{|\hat{\Sigma}_j|^{nN_i}} \exp\left\{-nN_i \cdot \left(\text{Tr}\left((\hat{\Sigma}_j)^{-1}\hat{\Sigma}_i\right) - q\right)\right\}. \quad (11)$$

If the value of Q_2 is too low, the null hypothesis $\Sigma_i = \Sigma_j$ is rejected. Correspondingly, the dissimilarity measure between the ith and jth regions is defined as

$$D_{\text{RW}}(S_i, S_j) = -\frac{1}{nN_i} \ln Q_2$$

$$= \ln\left(\frac{|\hat{\Sigma}_j|}{|\hat{\Sigma}_i|}\right) + \text{Tr}\left((\hat{\Sigma}_j)^{-1}\hat{\Sigma}_i\right) - q. \quad (12)$$

Since the revised Wishart (RW) distance $D_{\text{RW}}(S_i, S_j)$ is not symmetric, it is modified to [36]

$$D_{\text{SRW}}(S_i, S_j) = \frac{1}{2}\left(D_{\text{RW}}(S_i, S_j) + D_{\text{RW}}(S_j, S_i)\right)$$

$$= \frac{1}{2}\text{Tr}\left((\hat{\Sigma}_j)^{-1}\hat{\Sigma}_i + (\hat{\Sigma}_i)^{-1}\hat{\Sigma}_j\right) - q. \quad (13)$$

If $i = j$, the symmetric RW (SRW) distance $D_{\text{SRW}}(S_i, S_j)$ has a minimum value, zero. If the ith and jth regions are more dissimilar, the value of $D_{\text{SRW}}(S_i, S_j)$ is higher.

The test on equality of two covariance matrices here is the Box's M test. We found that, in the literature, Formont *et al.* [37]–[39] have done some excellent work using the Box's M test for classifying high-resolution PolSAR images. It is believed that it is promising to use the Box's M test and the corresponding distribution distances in high-resolution PolSAR image information extraction.

It is worth noting that the dissimilarity measure D_{RW} has the same form as the Kullback-Leibler (KL) divergence measure between PolSAR images proposed in [40], and D_{SRW} has the same form as the symmetrized version of the KL-divergence measure, since these measures discussed in this subsection share the same original information, called generalized variance ratio, with the KL-divergence-based measures introduced in [40].

In this subsection, dissimilarity measures between two regions in the PolSAR image, D_{S} and D_{SRW}, are discussed. They are both derived from the Wishart distribution, which incorporates the inherent statistical characteristics of PolSAR data into the dissimilarity measures. However, for multilook PolSAR data, the polarimetric covariance matrix $\mathbf{C}$ has a complex Wishart distribution, which is based on the assumption that the real and imaginary parts of complex scattering vector $\mathbf{u}$ have jointly circular Gaussian distribution. This is only satisfied for homogeneous areas with fully developed speckle and no texture [6], [41]. Therefore, the dissimilarity measures derived from the Wishart PDF might not be well adapted to some scenes with lots of details, such as urban areas in high-resolution PolSAR images. In that case, dissimilarity measures derived from other
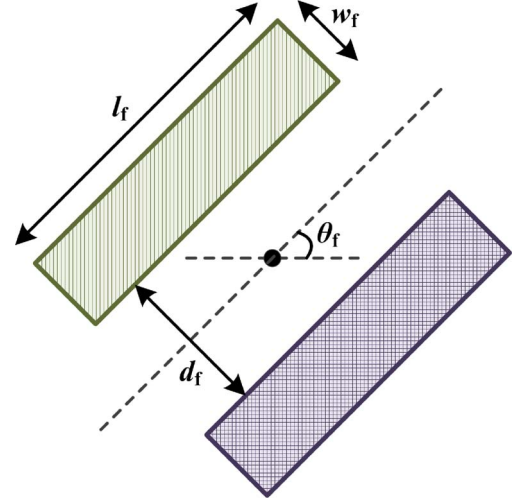

Fig. 1. Edge map calculation filter is characterized by the following parameters. l_{f} is the length of the filter, w_{f} is the width of each subpart of the filter, d_{f} is the spacing between the two subparts, and $\Delta\theta$ is the angular increment between two filter orientations. There are $N_{\text{f}} = \pi/\Delta\theta$ filters with the same parameters l_{f}, w_{f}, and d_{f}, but different orientations. The orientation of the filter is $\theta_{\text{f}} = i \cdot \Delta\theta$, $i = 0, 1, \ldots, N_{\text{f}} - 1$. Adopted from [45].

information may be more suitable or complementary, such as more complicated statistical models [35], [42], scattering mechanisms [43], texture, and linear features. This is application and problem oriented.

IV. SUPERPIXEL GENERATION

A. Calculation of Edge Maps

Inspired by [44] and [45], in the proposed framework, the edge map of a PolSAR image is calculated by visiting each pixel of the image sequentially, and applying a set of filters with different orientations. These filters extract the dissimilarity between each side of the center pixel, as a measure of probability of edge and an indicator of edge information in this pixel.

The filter to calculate edge maps is shown in Fig. 1, and it is controlled by four parameters, the length l_{f}, the width w_{f}, the spacing d_{f} between each side of the center pixel, and the angular increment $\Delta\theta$ between two filter orientations. Thus, there are $N_{\text{f}} = \pi/\Delta\theta$ filters with the same parameters l_{f}, w_{f}, and d_{f}, but different orientations.

The procedures of edge map calculation are described as follows.

1) Set the parameters l_{f}, w_{f}, and d_{f} for the filter.
2) For N_{f} filters with different orientations, calculate the dissimilarity $D_{\text{S}(\theta)}$ between these two regions using (10).
3) Find the maximum $D_{\text{S}(\theta)}$, $D_{\max}$, of the N_{f} values.
4) Save the orientation θ^* and the strength $D_{\max}$, move to the next pixel, and continue from Step 2).

Postprocessing of Edge Maps: The aim of edge map postprocessing is to locate the edge elements. We use the oriented nonmaximal suppression [46], [47], an algorithm frequently used in image processing and computer vision. For an arbitrary pixel p in the image, $D_{\max}(p)$ denotes the edge strength calculated using the aforementioned procedures and $\theta^*(p)$ is the corresponding orientation. Find two neighboring values of

$D_{\max}$ on either side of p along the line through p perpendicular to the orientation $\theta^*(p)$. If the edge strength of p is greater than or equal to each of the neighbors', it can be kept. Otherwise, it is set to zero. For example, suppose $\theta^*(p)$ is equal to $90°$, the edge strength in the pixel p can be kept only if its value is greater than or equal to each of the right and left neighbors'. After the postprocessing, the edge strength in the pixel p is modified to $D^*(p)$, which is ready to be used in the superpixel generation operation.

B. Superpixel Generation Based on Edge Maps

In this paper, superpixels are generated to improve the accuracy, efficiency, and understandability of PolSAR image NoC estimation and classification. The proposed operation of superpixel generation uses the normalized cuts algorithm [31] based on edge maps of PolSAR data.

Normalized Cuts: In the normalized cuts algorithm, Shi and Malik [31] formulated visual grouping as a graph partitioning problem. The basic principle of graph-based partitioning schemes is to represent a set of points in an arbitrary feature space using an undirected graph $G = \{V, E\}$, where V is for the vertices and E is for the edges between the vertices. Each vertex corresponds to a point in the feature space, and the edge between two vertices, e.g., x and y, is associated with a weight $W(x, y)$, that indicates the affinity of the pair. Image segmentation can be formulated as the best partitioning of the feature space into two regions, A and B, based on the minimum cut criterion. The cost function cut as follows is minimized:

$$cut(A, B) = \sum_{x \in A, y \in B} W(x, y) \tag{14}$$

where $W(x, y)$ is the pairwise affinity between x and y.

Since minimizing the cost function in (14) favors cutting out small and isolated partitions, Shi and Malik [31] proposed a new measure of partitioning cost. Instead of using value of total edge weight connecting the two regions, the new measure defines the partitioning cost as a fraction of the total edge connections to all the nodes in the graph. It is referred to as the normalized cut, $Ncut$, as follows:

$$
\begin{aligned}
Ncut(A, B) &= \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(B, A)}{assoc(B, V)} \\
&= \frac{cut(A, B)}{\sum\limits_{x \in A, v \in V} W(x, v)} + \frac{cut(B, A)}{\sum\limits_{y \in B, v \in V} W(y, v)}
\end{aligned} \tag{15}
$$

where $V = A \cup B$, and $assoc(A, V)$ is the total connection from nodes in A to all nodes in the graph and $assoc(B, V)$ is similarly defined. The principle of the normalized cuts algorithm is to minimize the partitioning cost measure $Ncut$. How to solve the minimization problem and more details related to the normalized cuts algorithm can be found in [31].

Calculation of the Weights: The grouping quality of normalized cuts depends heavily on the definition of the weight W between arbitrary pixels in the image. This procedure is rather flexible, since different information can be merged into the weights. Selection and fusion of features are problem oriented.
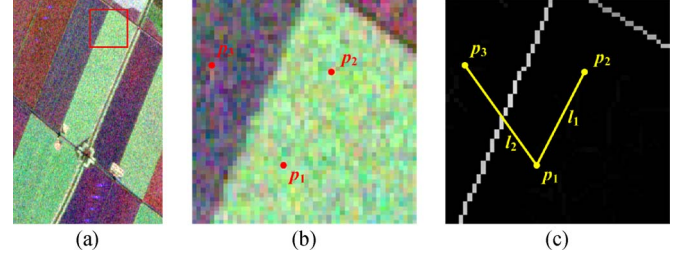

Fig. 2. Illustration of extracting the dissimilarity information from edge maps. (a) A subset of an L-band PolSAR image (PauliRGB). (b) Part of the original image (a) marked by the box; p_1 and p_2 are in the same class, and p_1 and p_3 are in different classes. (c) Edge map of image (b) after the oriented nonmaximal suppression; somewhere along l_2, the value of edge map strength is high, which suggests that p_1 and p_3 are in different classes; along l_1, the values of edge map strength are all low, proposing that p_1 and p_2 are probably in the same class. Adopted from [48].

In this paper, the inherent statistical characteristics of PolSAR data are added into the weights. The statistical characteristics of PolSAR data are obtained locally, while the dissimilarity matrix is global and represents pairwise distance between any two pixels of the image. The *intervening contour method* introduced in [48] gives us an important inspiration that the edge information can be used as a *bridge* between local features and global dissimilarity matrix. Moreover, edge information is considered critical and essential in the image segmentation literature. The basic principle of this method is illustrated in Fig. 2. The example is obtained using a subset of an L-band PolSAR image, acquired by the AIRSAR airborne platform in 1991 [49]. The subset image is shown in Fig. 2(a). Fig. 2(b) shows a magnified part of the subset image in the red box. We can observe from Fig. 2(b) that pixel p_1 is probably in the same class with p_2, and in different classes with p_3. Comparing Fig. 2(c) with (b), the basic principle of defining dissimilarity measures between arbitrary pixels in the PolSAR image is as follows: if the value of maximum edge map strength is high, the dissimilarity measure between these two pixels, e.g., x and y, should be high, thus the weight for affinity $W(x, y)$ should be low.

Therefore, the dissimilarity information is computed through edge maps generated using inherent statistical characteristics of PolSAR data. The dissimilarity of two pixels based on the contour information, $D_{\mathrm{C}}(x, y)$, is defined as [28], [29], [48]

$$
\begin{aligned}
D_{\mathrm{C}}(x, y) &= D^*(z^*), \\
z^* &= \arg\max_{z \in l} D^*(z)
\end{aligned} \tag{16}
$$

where $D^*(\cdot)$ denotes the strength of edge maps after the oriented nonmaximal suppression, l is the line joining x and y, and z^* is the location where the strength of edge maps after the oriented nonmaximal suppression is maximum along l. Then, the pairwise affinity is defined using a Gaussian kernel [28], [29], [31] as follows:

$$W(x, y) = \exp\left\{ \frac{-D_{\mathrm{C}}^2(x, y)}{2\sigma_{\mathrm{C}}^2} \right\} \tag{17}$$

where σ_{C} is the scaling parameter for the kernel.

After the pairwise affinity matrix is calculated, we apply the normalized cuts algorithm to produce the superpixel map. The number of superpixels is N_S.

As mentioned in Section II, Ersahin *et al.* [28], [29] have applied spectral graph partitioning to PolSAR image segmentation. The differences between their work and this paper are discussed as follows.

1) *The way of extracting contour information*—In their work, the strength of contour is obtained through the orientation energy [48], which was originally developed for gray-level natural images. In our method, the edge maps are calculated using inherent statistical characteristics. It is particularly developed and suitable for PolSAR images, especially in homogeneous areas.

2) *The way of calculating pairwise dissimilarities*—They select several channels of PolSAR data, such as power channels and the magnitude of the copolarized correlation coefficients. They treat each channel as a gray image separately and then combine pairwise dissimilarities calculated in each channel. We calculate pairwise dissimilarities in a different way. As we know, PolSAR data are completely represented by the polarimetric covariance matrix for homogeneous areas [50]. In order to incorporate the full polarimetric information into pairwise dissimilarities, we use the edge map as a *bridge* connecting local statistical information and global dissimilarity information: first, calculate edge maps locally by testing equality of covariance matrices, and then use the intervening contour method to consider the dissimilarity information from the entire image.

3) *The way of cutting image*—They use the multiclass spectral clustering proposed by Yu and Shi [51] to cut the image. In their approach, the number of partitions plays an important role in cutting the image. They considered the value of the parameter as known or provided by the user; this may limit practical applications. In our method, the aim of superpixel generation is to *oversegment* the image into many local, coherent regions. The experiments show that the number of superpixels can be chosen from a large range; this is more practical.

It is worth mentioning that, in this paper, superpixels are generated using the normalized cuts algorithm, and they are roughly homogeneous in size and shape. It is particularly suitable for image information extraction in homogeneous areas. As for more complicated scenes in the future research, such as urban areas in high-resolution PolSAR images, multi-scale edge maps, hierarchical segmentation, and spatial adaptivity should be considered in the oversegmentation process. The size of segments and the strength of relationships between a pixel and its neighborhood need to depend on the complexity of different parts of the whole scene.

V. NUMBER-OF-CLASSES ESTIMATION

The selection of the number of classes is an important and challenging issue in PolSAR data analysis, and it is crucial in building a robust automatic/semiautomatic classification system for PolSAR images. This topic, however, is not widely studied in the literature. As mentioned in Section II, Tran *et al.* [11] and Cao *et al.* [12] have proposed PolSAR image classification with an adaptive number of clusters. Their methods are based on AHC and postclustering validation. They attempt to choose the best partition from a set of alternative partitions based on some predefined criteria.

In contrast, motivated by [52] and [53], in this paper, we propose a NoC-adaptive classification method based on *preclustering estimation*. Since the proposed method attempts to estimate the number of classes before clustering occurs, we use the word *preclustering* to emphasize the characteristics of the method and to differentiate the method from postclustering-validation-based methods in [11] and [12]. The proposed preclustering estimation scheme has some advantages as follows.

- The preclustering estimation scheme does not require the data to be clustered first. The postclustering validation methods in [11] and [12] require the data to be clustered multiple times to generate a dendrogram and then find the most suitable number of clusters according to a predefined criterion. Repeated clustering may be time-consuming, whereas the preclustering estimation scheme is supposed to be more direct and more efficient in comparison.

- The preclustering estimation scheme is based on the dissimilarity matrix, thus, it makes the estimation operation flexible. Suppose, we need not only the statistical information, but also other features, such as the edge information, to be merged into the estimation, the pairwise dissimilarity between arbitrary superpixels can be represented by

$$D_{\text{fusion}}(S_i, S_j) = \eta_{\text{SRW}} \cdot D_{\text{SRW}}(S_i, S_j) + \eta_{\text{EM}} \cdot D_{\text{EM}}(S_i, S_j)$$
$$= \eta_{\text{SRW}} \cdot D_{\text{SRW}}(S_i, S_j) + \eta_{\text{EM}}$$
$$\cdot D_{\text{cor}}\left(F_{\text{EM}}(S_i), F_{\text{EM}}(S_j)\right) \quad (18)$$

where D_{SRW} and D_{EM} are dissimilarity measures derived from the statistical and edge information, respectively. η_{SRW} and η_{EM} are the respective impact factors of these dissimilarity measures. D_{cor} is the correlation distance measure between the feature vectors $F_{\text{EM}}(S_i)$ and $F_{\text{EM}}(S_j)$, which are histograms of edge maps, calculated in Section IV-A, of superpixels i and j, respectively.

Fig. 3 shows the flow chart of the proposed automatic preclustering NoC estimation scheme for PolSAR data, with a synthetic PolSAR image generated using the Monte Carlo method introduced in [6, ch. 4.5.2]. The simulation parameters are taken from a real PolSAR image. The scheme comprises four steps, which are elaborated as follows.

1) *Representation*—In order to represent the data structures within a PolSAR image, pairwise dissimilarity information is depicted as an $N_S \times N_S$ image, where N_S is the total number of superpixels as described in Section IV-B. The pairwise dissimilarities between arbitrary superpixels are calculated using D_{SRW} in (13). As shown in Fig. 3, a superpixel-based PolSAR image, the size of which is $N \times M$, is represented by an $N_S \times N_S$ dissimilarity image.
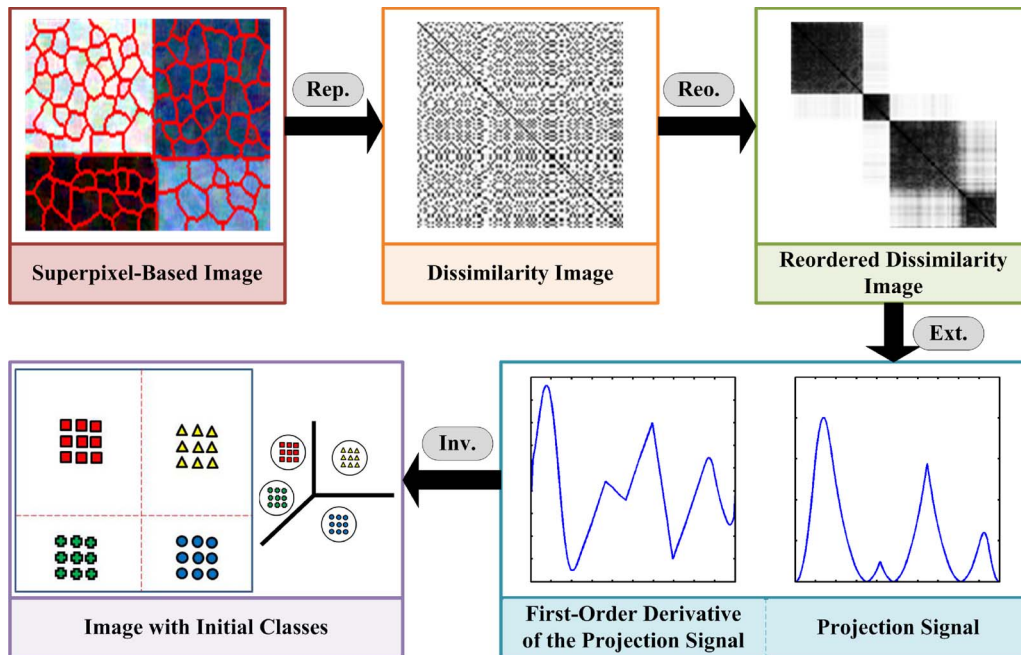
Fig. 3. Illustration of the preclustering estimation operation. A superpixel-based image, the size of which is $N \cdot M$, is represented (Rep.) by an $N_S \cdot N_S$ dissimilarity image. After the reordering (Reo.) step, the dissimilarity image is transformed into the reordered dissimilarity image (RDI), which reveals the class structures and information. The extraction (Ext.) step transforms the RDI into a projection signal, and the number of classes in the data corresponds to the number of *major* peaks in the projection signal. The inversion (Inv.) step uses the class information represented in the projection signal to achieve initial class parameters, and then some superpixels in the image have initial class labels.

2) *Reordering*—As shown in Fig. 3, unfortunately, the dissimilarity image cannot clearly illustrate the class information within the PolSAR data. In [54], Cattell first used an image to represent pairwise dissimilarities between arbitrary objects in a data set and *reordered* these objects suitably so that the resulting image is better able to highlight the potential cluster structure in the data. Different methods of implementing visual representation of pairwise dissimilarity information can be found in the literature. The fundamental principle of these methods is the reordered dissimilarity image (RDI). Based on the idea of RDI, by *reordering* the N_S superpixels suitably, it results in an image, which can stress the potential class structures in the PolSAR data, as shown in Fig. 3. The intensity of each point in the RDI corresponds to the dissimilarity between the pair of superpixels addressed by the coordinates of this point. A useful RDI can highlight the potential classes as a set of *dark blocks* along the diagonal of the image, indicating the set of superpixels with low dissimilarity. In the proposed framework, we use the visual assessment of cluster tendency (VAT) algorithm [52] to transform the dissimilarity image into the RDI. The detailed information about the VAT algorithm can be found in [52].

3) *Extraction*—As shown in Fig. 3, the RDI can highlight the potential classes as a set of *dark blocks* along the diagonal of the image. Human can simply read the number of classes in the RDI. When it comes to computer, the dark block extraction (DBE) method proposed in [53] is selected and specially tailored to automatically extract *dark blocks* along the diagonal of the RDI. The basic principle of the DBE method is to extract the number of

classes using several common image and signal processing techniques. The procedures of DBE are described as follows.

a) Perform image segmentation on the RDI to obtain a binary image, and then apply the directional morphological filters to the binary image.

b) Apply a distance transform to the filtered binary image, and then project the pixel values along the main diagonal axis of the image to form a projection signal.

c) We use a simple average filter to smooth the projection signal, and then use the first-order derivative of the smoothed projection signal to detect the *major* peaks and valleys of the projection signal. We use a size filter to remove relatively small peaks. The peak within the two neighboring valleys will be kept as a meaningful peak, only if the width between these two neighboring valleys is larger than $\gamma \cdot L_s$, where L_s is the length of the projection signal, and γ is set to 0.03 experimentally in this paper.

Therefore, the DBE method transforms the RDI into a 1-D projection signal, and the number of classes within the data corresponds to the number of *major* peaks in the projection signal as shown in Fig. 3. For more information about the DBE algorithm, one can refer to [53].

4) *Inversion*—In the projection signal, a peak between two neighboring valleys realistically represents a class in the data. After detecting the peaks and valleys in the projection signal, the inversion step uses the class information represented in the projection signal to achieve initial class parameters. Suppose that there are N_x superpixels between the xth valley and the $(x + 1)$th valley. Due to noise, the inversion step cannot simply determine that all
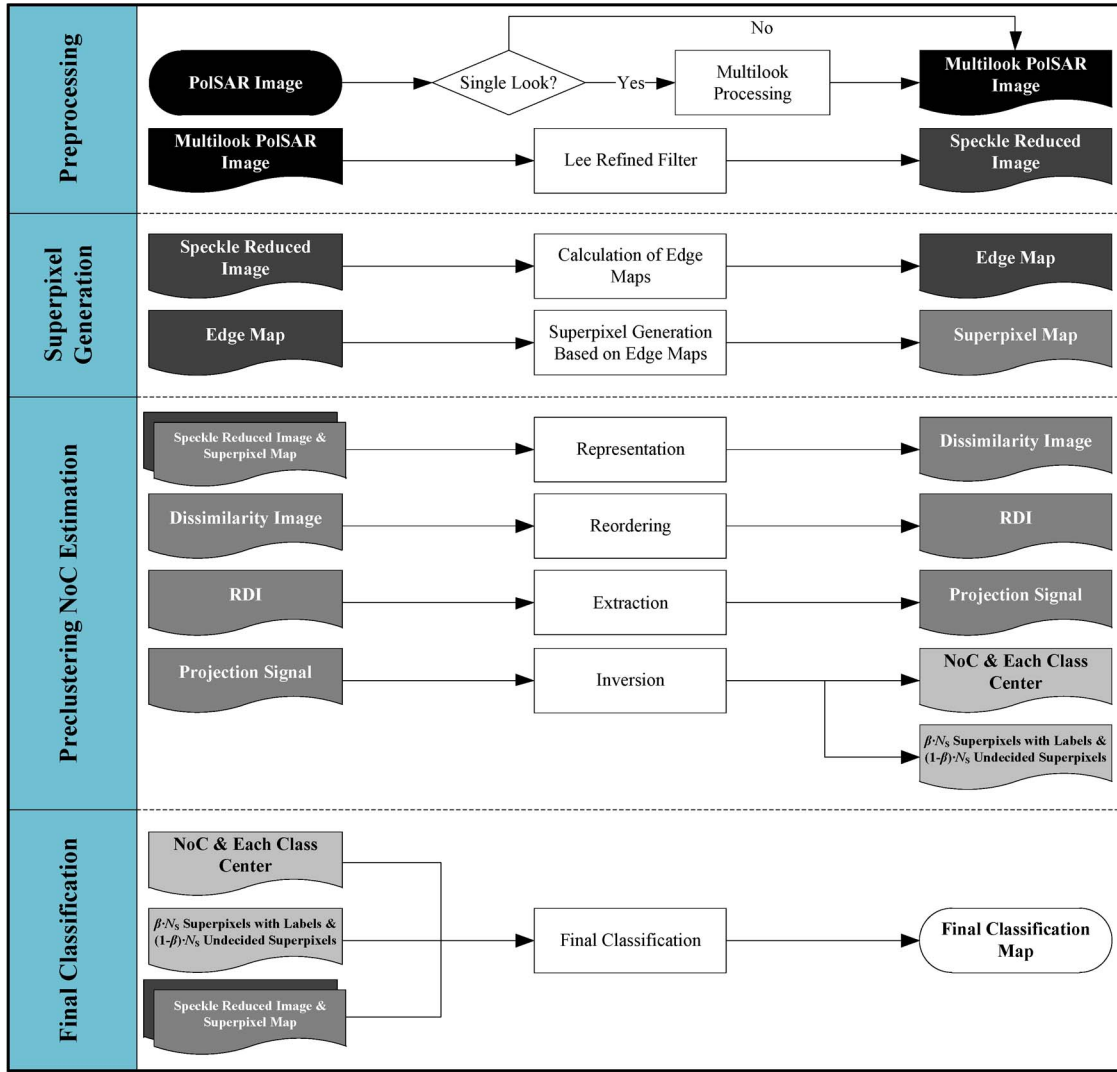
Fig. 4.　Schematic concept of the proposed superpixel-based classification with an adaptive number of classes for PolSAR images.

the N_x superpixels belong to the xth class. In our method, the inversion step labels the $\beta \cdot N_x$ superpixels nearest to the xth peak as elements of the xth class, where β is from 0 to 1. The other $(1 - \beta) \cdot N_x$ superpixels between the xth valley and the $(x + 1)$th valley can be seemed as a *buffer zone*. The superpixels in this area are categorized as *undecided*, and their labels will be determined in the final classification.

VI. FINAL CLASSIFICATION

As discussed in Section V, after the operations of superpixel generation and superpixel-based preclustering NoC estimation, there are still $(1 - \beta) \cdot N_S$ superpixels categorized as *undecided*, and their class labels are determined by the final classification operation. After the superpixel-based preclustering estimation operation, we get several important initial class parameters,

- The number of classes in the PolSAR image X;
- The initial classes $C_1, C_2, \ldots, C_X$;
- The ML estimators of the 1st, 2nd, $\ldots$, Xth class center covariance matrices.

The determination of class labels is inspired by the iterated Wishart clustering proposed by Lee *et al.* [17], [21]. For an arbitrary superpixel S in the PolSAR image, the estimation of class label of this superpixel is derived as follows:

$$\hat{x}_S = \arg \min_{x \in \{1,2,\ldots,X\}} \{D_{\mathrm{W}}(S, C_x)\}$$

$$D_{\mathrm{W}}(S, C_x) = \ln \left| \hat{\mathbf{\Sigma}}_{C_x} \right| + \mathrm{Tr}\left(\hat{\mathbf{\Sigma}}_{C_x}^{-1} \hat{\mathbf{\Sigma}}_S \right) \qquad (19)$$

where $D_{\mathrm{W}}(S, C_x)$ is the Wishart distance between superpixel S and class C_x derived from (3), and it is calculated using the ML estimators of the center covariance matrices of superpixel S and class C_x.

According to [17] and [21], iteration can improve the quality of classification map. We also apply iteration to the final classification operation to achieve more accurate class label estimation. The termination criterion is a combination of

1) percentage of number of superpixels switching classes per iteration P_{TC};
2) predefined number of iterations N_{TC}.

VII. IMPLEMENTATION

The presented research aims at providing a robust superpixel-based NoC-adaptive classification framework for PolSAR images. The whole processing chain in Fig. 4 consists of *four operations* as follows.

1) *Preprocessing*
   a) The single look PolSAR data are multilook-processed.
   b) The Lee refined filter [55] is selected and performed, since it is able to preserve polarimetric properties and statistical correlation between channels; while this step is optional.

2) *Superpixel generation*
   a) Calculation of edge maps—The edge information strength of each pixel in the PolSAR image is measured by the dissimilarity between both sides of this pixel with scanning window configuration.
   b) Superpixel generation based on edge maps—The PolSAR image is partitioned into homogeneous segments based on the edge map using the normalized cuts algorithm.

3) *Preclustering NoC estimation*—Using the pairwise dissimilarities between arbitrary superpixels, the PolSAR image is represented as a dissimilarity image, which is reordered into a RDI. The DBE step transforms the RDI into a projection signal. The class information represented in the projection signal is extracted to estimate the number of classes, each class center, and initial class labels of part of superpixels.

4) *Final classification*—Class labels of the *undecided* superpixels from the preclustering NoC estimation operation are determined and refined to get the final classification map.

After the operations of *preprocessing*, *superpixel generation*, *preclustering NoC estimation*, and *final classification*, the PolSAR image is classified unsupervisedly.

VIII. EXPERIMENTS AND DISCUSSIONS

A. Description of Experimental Data Sets

In this section, the performance of the proposed classification framework for PolSAR images is presented and analyzed on one synthetic and three real experimental data sets.

Synthetic Experimental Data Set: The first data set is a synthetic experimental data set. The data set is generated using the Monte Carlo method motivated by [6, ch. 4.5.2]. The detailed simulation design will be described in Section VIII-C. The simulation parameters are taken from a real PolSAR image, and the center covariance matrices of the nine classes are listed as follows. The sizes of all synthetic images are 500×700. The corresponding true classification map of the synthetic data set is shown in Fig. 5(a); and one example simulated image is shown in Fig. 5(b) (see equation at the bottom of the next page).

Real Experimental Data Sets in Agricultural Areas: The second data set is obtained from a subset of an L-band, multilook PolSAR image, acquired by the AIRSAR airborne platform in 1991 [49]. The scene covers Flevoland, Netherlands. The experimental image is shown in Fig. 6(a), the size of which is
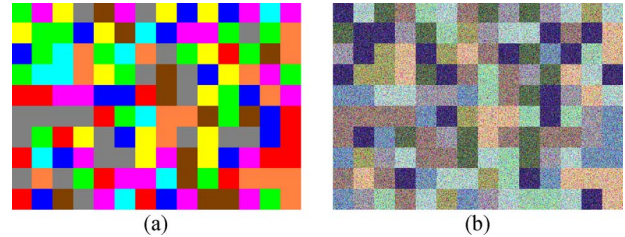

Fig. 5. Illustration of the synthetic data set. (a) True classification map of the synthetic data set. (b) One example simulated image (PauliRGB).
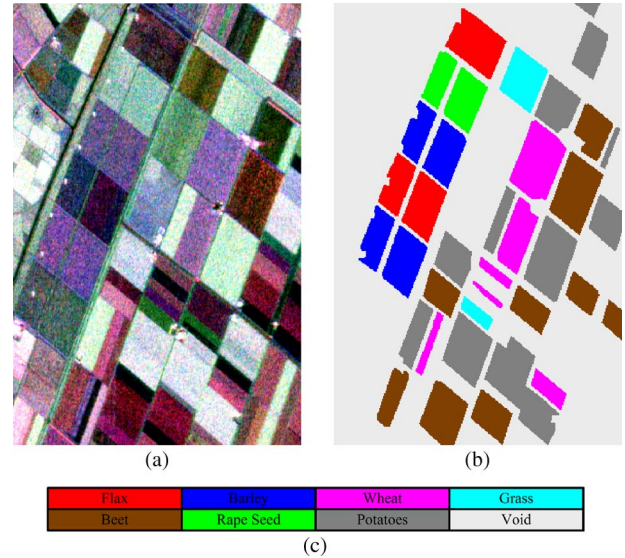

Fig. 6. L-band, multilook PolSAR image in Flevoland, Netherlands in 1991 of AIRSAR platform as the second data set used in the experiment. (a) PolSAR image (PauliRGB). (b) The ground truth map of (a). (c) Color code.
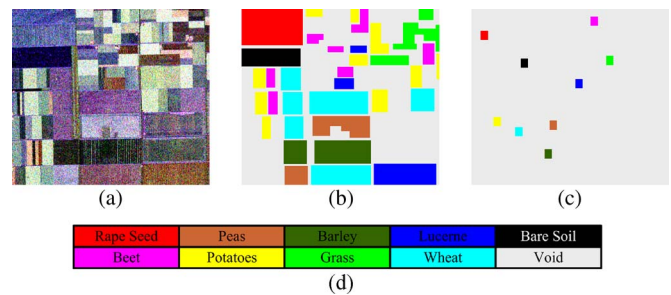

Fig. 7. L-band, multilook PolSAR image in Flevoland, Netherlands in 1989 of AIRSAR platform as the third data set used in the experiment. (a) PolSAR image (PauliRGB). (b) The ground truth map of (a). (c) Training data set. (d) Color code.

430×280. The ground truth map is gleaned from [56], and it is shown in Fig. 6(b). According to the ground truth map, there are seven classes in the second data set.

The third data set is obtained from a subset of an L-band, multilook PolSAR image, acquired by the AIRSAR airborne platform in 1989 [57]. The scene covers Flevoland, Netherlands, too. The experimental image is shown in Fig. 7(a), the size of which is 380×430. The ground truth map is gleaned from [3] and [9], and it is shown in Fig. 7(b). According to the ground truth map, there are seven classes of crops, one class of grass, and one class of bare soil in the third data set.
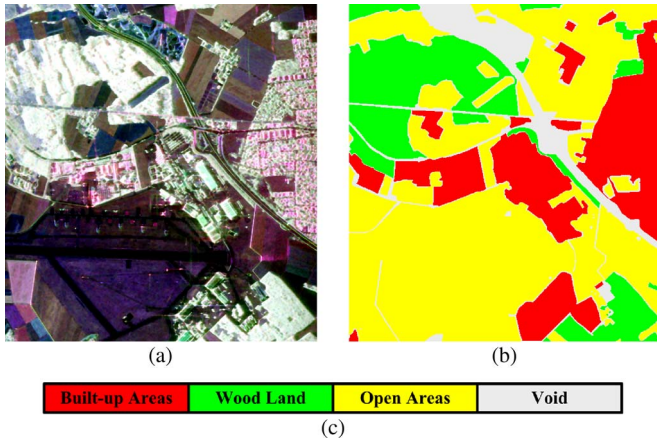
Fig. 8. L-band, multilook PolSAR image in Oberpfaffenhofen, Germany of ESAR platform as the fourth data set used in the experiment. (a) PolSAR image (PauliRGB). (b) The ground truth map of (a). (c) Color code.

Real Experimental Data Set in General Areas: The fourth data set is obtained from an L-band, multilook PolSAR image, acquired by the ESAR airborne platform [57]. The scene covers Oberpfaffenhofen, Germany. The fourth experimental image is shown in Fig. 8(a), the size of which is 1300×1200.

For our present purpose, manual classification according to the corresponding optical remote sensing imagery in Google Earth is used as the ground truth map, as shown in Fig. 8(b). According to the ground truth map, the scene is categorized into three basic classes: built-up areas, wood land, and open areas.

In the second, third, and fourth data sets, pixels without ground truth or lie near boundaries between classes are categorized as *void*. The void pixels are ignored in both training and testing.

In Section VIII-B, the third data set, which includes nine classes of land covers and is widely used in the literature, is selected to test the operation of superpixel generation as a tool for PolSAR image classification, analyze its performance under various input parameter configurations, and compare the superpixel-based classification method with other widely used classification methods.

In Section VIII-C, the synthetic data set is used to test the performance of the proposed superpixel-based NoC estimation method for PolSAR images.

In Section VIII-D, all the real experimental data sets are utilized to examine and present the performance and reliability of the whole superpixel-based NoC-adaptive classification framework for PolSAR images.

$$\mathbf{C}_1 = \begin{pmatrix} 0.0045 & 0.0001j & 0.0049 - 0.0005j \\ -0.0001j & 0.0004 & -0.0001 - 0.0001j \\ 0.0049 + 0.0005j & -0.0001 + 0.0001j & 0.0070 \end{pmatrix}$$

$$\mathbf{C}_2 = \begin{pmatrix} 0.0011 & 0 & 0.0014 - 0.0003j \\ 0 & 0.0001 & 0 \\ 0.0014 + 0.0003j & 0 & 0.0024 \end{pmatrix}$$

$$\mathbf{C}_3 = \begin{pmatrix} 0.0037 & -0.0002 - 0.0001j & 0.0030 - 0.0014j \\ -0.0002 + 0.0001j & 0.0005 & -0.0002 + 0.0002j \\ 0.0030 + 0.0014j & -0.0002 - 0.0002j & 0.0062 \end{pmatrix}$$

$$\mathbf{C}_4 = \begin{pmatrix} 0.0009 & 0 & 0.0006 - 0.0001j \\ 0 & 0.0002 & 0 \\ 0.0006 + 0.0001j & 0 & 0.0013 \end{pmatrix}$$

$$\mathbf{C}_5 = \begin{pmatrix} 0.0080 & 0.0001 + 0.0003j & 0.0068 - 0.0017j \\ 0.0001 - 0.0003j & 0.0017 & -0.0002 \\ 0.0068 + 0.0017j & -0.0002 & 0.0114 \end{pmatrix}$$

$$\mathbf{C}_6 = \begin{pmatrix} 0.0030 & -0.0001 + 0.0001j & 0.0002 - 0.0002j \\ -0.0001 - 0.0001j & 0.0006 & -0.0001j \\ 0.0002 + 0.0002j & 0.0001j & 0.0018 \end{pmatrix}$$

$$\mathbf{C}_7 = \begin{pmatrix} 0.0056 & -0.0002 + 0.0002j & 0.0036 \\ -0.0002 - 0.0002j & 0.0018 & -0.0001 - 0.0002j \\ 0.0036 & -0.0001 + 0.0002j & 0.0055 \end{pmatrix}$$

$$\mathbf{C}_8 = \begin{pmatrix} 0.0018 & 0 & 0.0008 - 0.0008j \\ 0 & 0.0003 & 0 \\ 0.0008 + 0.0008j & 0 & 0.0034 \end{pmatrix}$$

$$\mathbf{C}_9 = \begin{pmatrix} 0.0071 & 0.0002 & -0.0013 + 0.0026j \\ 0.0002 & 0.0010 & -0.0001 - 0.0001j \\ -0.0013 - 0.0026j & -0.0001 + 0.0001j & 0.0073 \end{pmatrix}$$
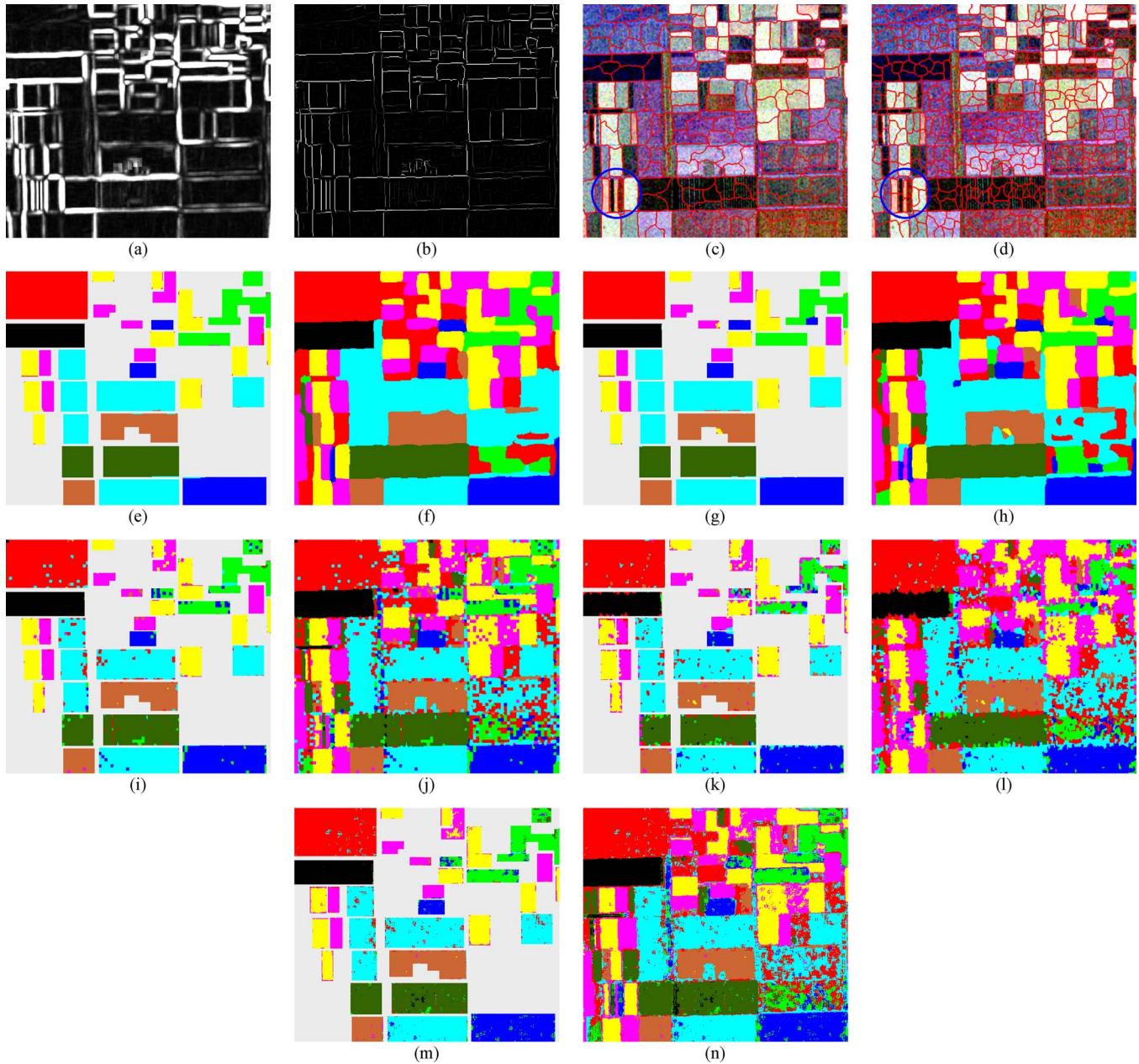
Fig. 9. Superpixel-based classification results on the third data set, and comparison of classification results using different methods. (a) and (b) are edge maps before and after the oriented nonmaximal suppression, respectively. (c) and (d) are superpixel generation results with 200 and 500 superpixels, respectively. (e) and (g) are superpixel-based classification method results with 200 and 500 superpixels, respectively. (i) WMRF-based method result. (k) Watershed-based method result. (m) Pixel-based method result. (f), (h), (j), (l), and (n) are the results without void mask of (e), (g), (i), (k), and (m), respectively.

B. Performance of Superpixel Generation and Superpixel-Based Classification

In this subsection, we use the third data set, including nine classes of land covers and widely used in the literature, to present the results of the superpixel-based classification method and analyze the performance of superpixel generation under various choices of input parameters.

In order to test the performance of superpixel generation as a tool for PolSAR image classification and compare the superpixel-based classification with other methods, we use supervised classification in this experiment.

In the preprocessing operation, the PolSAR image is processed with the Lee refined filter. For region-based methods,

speckle filters with large window may not be suitable, because they might blur images and smear edges, which are very important in region generation. Hence, in the preprocessing operation, we use relatively small window, 3×3, Lee refined filter for region-based methods. However, for the pixel-based method, in case the 3×3 Lee refined filter is not enough, we increase the size of the filter to 7×7.

After the operation of superpixel generation, the number of classes and initial class centers are determined by the training data set as shown in Fig. 7(c), and the final classification operation takes the superpixels as elements and uses the Wishart distance to classify the PolSAR data, as discussed in Section VI.

In this experiment, edge maps before and after the oriented nonmaximal suppression are illustrated in Fig. 9(a) and (b),

TABLE I
CLASSIFICATION ACCURACY (%) AS A FUNCTION OF l_f AND w_f

l_f ╲ w_f	3	5	7	9	11
7	77.11	94.99	99.34	98.84	97.81
9	75.47	95.85	99.06	99.01	98.38
11	76.64	95.05	99.16	98.53	99.17
13	76.12	91.56	99.43	98.51	98.89
15	78.10	95.69	99.47	98.50	98.92
17	77.39	94.74	99.40	99.16	98.96
19	83.22	94.83	99.15	98.99	99.06
21	85.71	95.32	99.28	98.87	98.74

TABLE II
COMPARISON OF CLASSIFICATION ACCURACY (%)

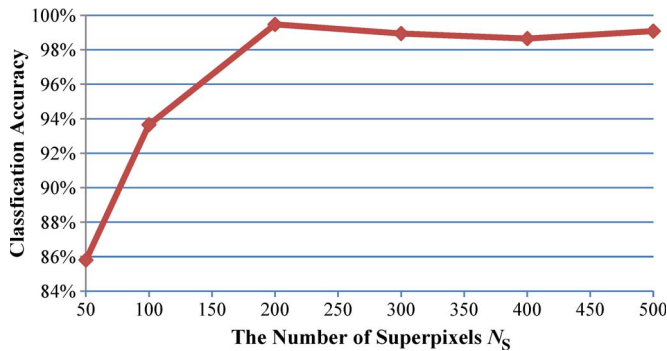| Class ╲ Method	Proposed Method	WMRF Based	Watershed Based	Pixel Based
Rape Seed	100.00	96.22	97.65	97.65
Bare Soil	100.00	99.74	94.05	99.32
Potatoes	98.90	96.64	87.37	92.91
Beet	98.21	93.99	87.08	92.56
Lucerne	99.15	91.37	87.59	91.34
Grass	99.33	88.49	60.98	89.03
Wheat	99.87	94.16	93.30	94.66
Peas	99.57	98.17	95.52	97.04
Barley	99.50	94.74	86.80	95.83
Total	99.47	94.80	89.31	94.54

Fig. 10. Classification accuracy (%) as a function of N_S.

respectively. Fig. 9(c) and (d) show the results of superpixel generation with 200 and 500 superpixels, respectively. The respective final classification maps are shown in Fig. 9(e) and (g). The parameters of edge map calculation, d_f and $\Delta\theta$, are set to 1 and $\pi/2$, respectively. We mainly test the effects of parameters l_f and w_f of edge map calculation, and the number of superpixels N_S here.

Table I lists the classification accuracy under various choices of edge map calculation parameters l_f and w_f. The accuracy is the percentage of image pixels belonging to each class that were categorized to the correct class. The higher are the values of l_f and w_f, the more accurate are the important and obvious edges. However, if the values of l_f and w_f are too high, some small edges might be erased, and it may cause undersegmentation, which should be avoided.

From the experimental results in Table I, the suitable choices of values of parameter l_f are between 13 and 17, and the suitable choices of values of parameter w_f are between 7 and 9. It is worth mentioning that l_f and w_f are set to 15 and 7, respectively, in Figs. 9 and 10, and Table II.

Fig. 10 shows the classification accuracy under various choices of the number of superpixels N_S, which is determined by the size and intrinsic complexity of the PolSAR image. If the value of N_S is too low, it might cause undersegmentation in some area, as comparison of Fig. 9(c) and (d) in blue circles. However, if the value of N_S is too high, large volume of memory and massive computations are required, and the characteristics of small superpixels are more easily affected by speckle.

Since in this paper, we aim at *oversegmenting* the PolSAR image into many superpixels, the number of superpixels can be

chosen from a large range as illustrated in Fig. 10. It is worth noting that N_S is set to 200 in Tables I and II.

For comparison, classification maps of the proposed superpixel-based, WMRF-based [9], watershed-based [58], and pixel-based [17] methods are shown in Fig. 9(e), (i), (k), and (m), respectively. The WMRF-based method is selected because it is a novel and effective MRF-based method. The watershed-based method is chosen since it is frequently used in image oversegmentation. In these four methods, the PolSAR data are all preprocessed using the Lee refined filter. In the WMRF-based and watershed-based methods, the final classification operation takes the small regions as elements and uses the Wishart distance, as discussed in Section VI.

Table II lists the classification accuracy obtained by the aforementioned four classification methods. From Table II, we conclude that the superpixel-based method outperforms the other three widely used methods for PolSAR image classification. The total accuracy of the proposed method is beyond 99%, which is about 4.5%, 10%, and 5% higher than other three methods, respectively, since spatial relations between pixels are taken into consideration, and the statistical characteristics and contour information of PolSAR data are made good use of.

From the experiment, the pixel-based method has rather good results. However, as we know, increasing the window size of filters is not always a suitable solution for terrain classification, since filters with large window size may blur the image and smear edges. It is believed that the one advantage of region-based classification methods is that they can provide satisfactory results in some areas where large window size of filters is not suitable.

C. Performance of Number-of-Classes Estimation

As discussed in Section V, the NoC estimation is based on pairwise dissimilarities. Thus, the performance of the NoC estimation depends on the distances between the covariance matrices of different class types. In this subsection, we use a Monte-Carlo-based design to synthesize a scenario with closer covariance matrices, and then to test the performance and robustness of the NoC estimation.

Based on the nine center covariance matrices $C_1, C_2, \ldots, C_9$ of respective nine class types extracted from the real PolSAR data, as described in Section VIII-A, we

use the mixture of these covariance matrices to simulate closer center covariance matrices, as summarized by the following *Experimental Design*.

Experimental Design. Simulation of PolSAR data with closer covariance matrices

---

INPUT
- Nine center covariance matrices $\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_9$ of respective nine class types as described in Section VIII-A.
- True classification map as shown in Fig. 5(a).
- Random mixing coefficient boundary R.

PROCEDURE
1. Compute the overall distance before mixing using

$$Dis_{\text{before}} = \sum_{1 \le i < j \le 9} D_{\text{SRW}}(C_i, C_j). \tag{20}$$

2. Compute the center covariance matrices of the mixed classes using

$$\mathbf{C}_x^{\text{mix}} = \mathbf{C}_x + (r_1 \ \ldots \ r_9)\begin{pmatrix} \mathbf{C}_1 \\ \vdots \\ \mathbf{C}_9 \end{pmatrix} \tag{21}$$

where $x = 1, 2, \ldots, 9$, and $r_1, r_2, \ldots, r_9$ are nine uniformly distributed random numbers from 0 to R, for each x.

3. Compute the overall distance after mixing using

$$Dis_{\text{after}} = \sum_{1 \le i < j \le 9} D_{\text{SRW}}\left(C_i^{\text{mix}}, C_j^{\text{mix}}\right). \tag{22}$$

4. For given center covariance matrix of each mixed class, compute $(\mathbf{C}_x^{\text{mix}})^{1/2}$, where

$$\left(\mathbf{C}_x^{\text{mix}}\right)^{\frac{1}{2}}\left[\left(\mathbf{C}_x^{\text{mix}}\right)^{\frac{1}{2}}\right]^{*\mathrm{T}} = \mathbf{C}_x^{\text{mix}}. \tag{23}$$

5. Simulate a complex random vector $\mathbf{v}$; independently generate the real and imaginary parts of each component of $\mathbf{v}$ that are statistically independent from a normal distribution with zero mean and 0.5 variance.

6. The complex single-look vector is obtained by

$$\mathbf{u} = \left(\mathbf{C}_x^{\text{mix}}\right)^{\frac{1}{2}}\mathbf{v}. \tag{24}$$

7. Compute the n-look covariance matrix using $\mathbf{C} = (1/n)\sum_{i=1}^{n} \mathbf{u}_i \mathbf{u}_i^{*\mathrm{T}}$.

OUTPUT
- Simulated PolSAR data according to the true classification map.
- Overall distances before and after the mixing.

Verification—From [6], polarimetric covariance matrices $\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_9$ are Hermitian positive semidefinite matrices. Therefore, we can make sure that the mixtures $\mathbf{C}_1^{\text{mix}}, \mathbf{C}_2^{\text{mix}}, \ldots, \mathbf{C}_9^{\text{mix}}$ are also Hermitian positive semidefinite matrices. The matrix $(\mathbf{C}_x^{\text{mix}})^{1/2}$ is obtained by using a unitary transform to diagonalize $\mathbf{C}_x^{\text{mix}}$.
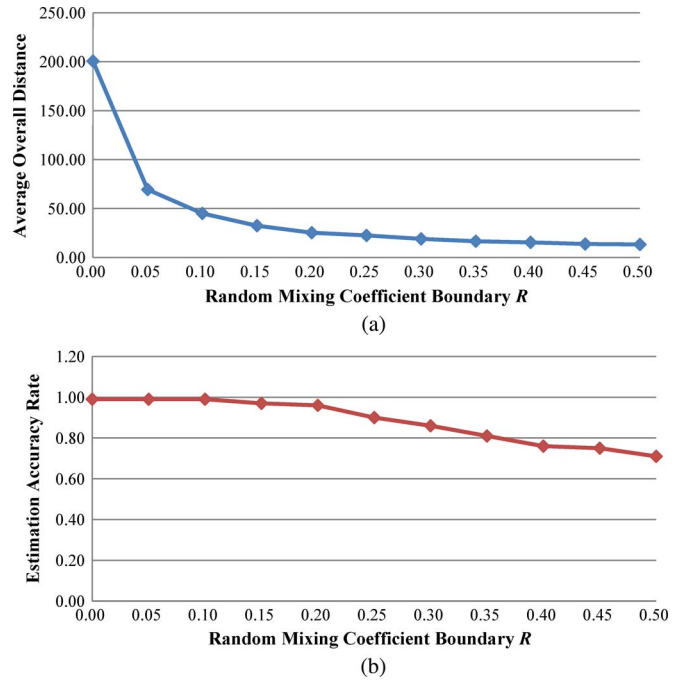

Fig. 11. Illustration of performance of the NoC estimation on simulated PolSAR data. (a) Average overall distance after the mixing under various random mixing coefficient boundaries R. (b) Estimation accuracy rate under various R.

Since the purpose here is to test the performance of NoC estimation, for all the experiments in this subsection, we use the same true classification map as shown in Fig. 5(a), and also the same superpixel map. For each random mixing coefficient boundary R, we run the simulation process for 100 times and perform the NoC estimation on these 100 synthetic PolSAR data. The average overall distances after the mixing and the estimation accuracy rate are shown in Fig. 11(a) and (b), respectively.

According to the experimental results illustrated in Fig. 11, the performance of the NoC estimation is discussed as follows.

- In all the experiments in this subsection, γ is set to 0.03 as discussed in Section V; and the whole estimation is done automatically.
- It shows the proposed NoC estimation method can provide rather good and promising results. The accuracy of NoC estimation is under 80% only if the covariance matrices are severely mixed and very close to each other. It also shows that the performance of NoC estimation relies on the pairwise dissimilarities greatly. Thus, the performance of NoC estimation can be improved by incorporating more suitable or complementary information into the dissimilarity measures, and we will discuss this further in the next subsection.

D. Performance of the Whole Framework

In this subsection, we use the all the real data sets to test the performance of the whole processing chain. In the preprocessing operation, the second and third experimental PolSAR images are processed with 3×3 Lee refined filter. For the fourth
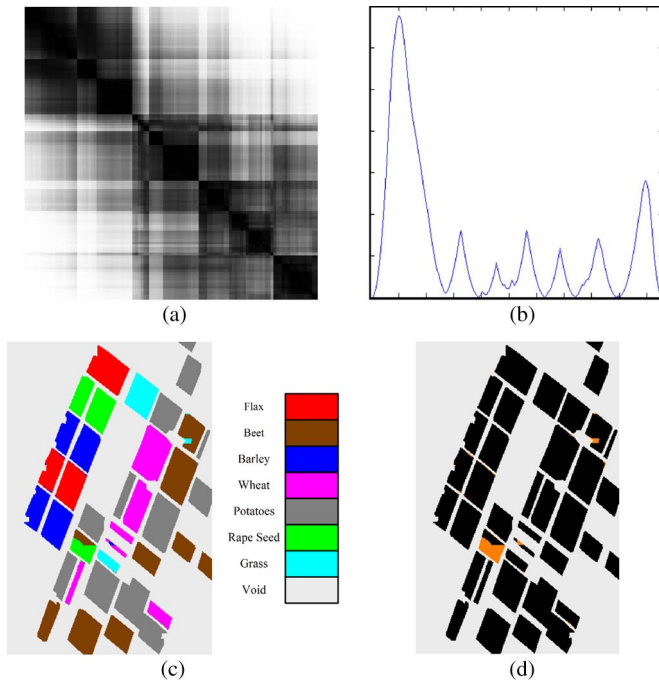
Fig. 12. Results of the whole classification framework on the second data set. (a) Reordered dissimilarity image. (b) One-dimensional projection signal. (c) Final classification map. (d) Error map, void in gray and errors in orange.
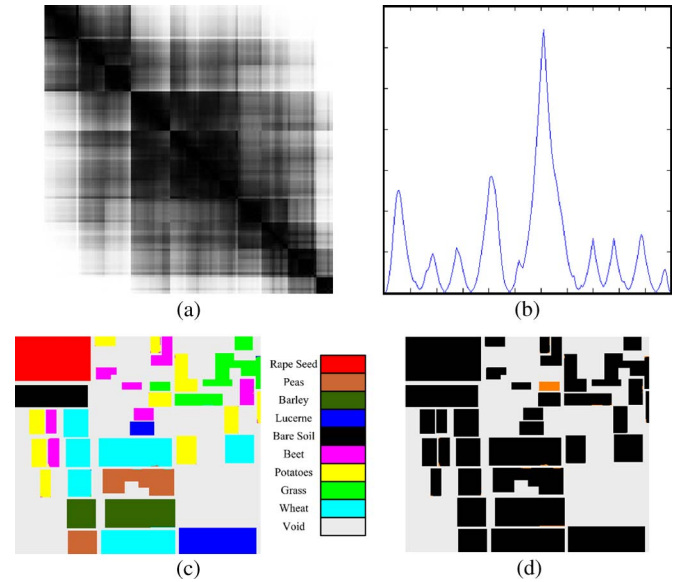

Fig. 13. Results of the whole classification framework on the third data set. (a) Reordered dissimilarity image. (b) One-dimensional projection signal. (c) Final classification map. (d) Error map, void in gray and errors in orange.

data set, the speckle noise has been well suppressed, thus, no more filter is needed. It is worth noting that, in the operation of superpixel generation, the parameters of edge map calculation l_f, w_f, d_f, and $\Delta\theta$ are set to 15, 7, 1, and $\pi/2$, respectively. After superpixel generation, the number of classes and initial class centers are estimated by the operation of preclustering NoC estimation, and then class labels of *undecided* superpixels are determined and refined by the final classification operation, just like the whole implementation procedures discussed in Section VII and illustrated in Fig. 4.

Real Experimental Data Sets in Agricultural Areas: The performance of the proposed framework on the real experimental data sets in agricultural areas, the second and the third, is listed in Figs. 12 and 13 and Tables III and IV.

The reordered pairwise dissimilarity images between arbitrary superpixels in the experimental images are shown in Figs. 12(a) and 13(a). As discussed in Section V, the RDI highlights the potential classes existing in the PolSAR data as a set of *dark blocks* along the diagonal of the image, indicating the set of superpixels with low dissimilarity. From the RDIs in Figs. 12(a) and 13(a), the class structures in the experimental images are depicted.

The operation of preclustering NoC estimation transforms the RDI to a 1-D projection signal, and the number of classes in the PolSAR image corresponds to the number of *major* peaks in the signal, as shown in Figs. 12(b) and 13(b). Although there are several very small peaks in the signals in Figs. 12(b) and 13(b), they can easily be eliminated by the smoothing and size filters as discussed in Section V. The operation of preclustering NoC estimation determines that there are seven and nine classes existing in the second and third data sets, respectively, and the estimations are consistent with the ground truth maps in

Figs. 6(b) and 7(b). As discussed in Section V, the operation of preclustering estimation labels the $\beta \cdot N_x$ superpixels nearest to the xth peak as elements of the xth class, where β is set to 0.7 experimentally.

After the preclustering estimation operation, there are still $(1 - \beta) \cdot N_S$ superpixels categorized as *undecided*, and their class labels are determined by the final classification operation. The termination criterion of the final classification is a combination of the percentage of number of superpixels switching classes per iteration P_{TC} and a predefined number of iterations N_{TC}. P_{TC} and N_{TC} are set to 0 and 20, respectively. The final classification maps are shown in Figs. 12(c) and 13(c).

According to the ground truth maps in Figs. 6(b) and 7(b), the total accuracy of the whole proposed framework on the second and third data sets is 98.06% and 99.04%, respectively, and the respective error maps are shown in Figs. 12(d) and 13(d), confusion matrices listed in Tables III and IV.

From Fig. 12 and Table III, in the second data set, part of Beet mistakenly classified as Rape Seed is due to the large within-class variance of the class Beet; from Fig. 13 and Table IV, in the third data set, part of Lucerne mistaken as Grass is because of very small inter-class difference of the classes Lucerne and Grass. These can be observed from the respective PauliRGB images in Figs. 6(a) and 7(a).

Real Experimental Data Set in General Areas: The processing chain on the fourth data set is much like the processing on the second and third data sets as aforementioned. The edge map after the oriented nonmaximal suppression and superpixel map are shown in Fig. 14(a) and (b), respectively.

In this experiment, it is worth stressing that information fusion in the preclustering estimation can achieve better result.

In the NoC estimation, if only the statistical information, the SRW distance D_{SRW} in (13), is used, the RDI is shown in Fig. 14(c); the NoC is estimated as three; and the final classification result is in Fig. 14(e).

TABLE III
CONFUSION MATRIX (%) OF THE WHOLE CLASSIFICATION FRAMEWORK ON THE SECOND DATA SET

	Potatoes	Rape Seed	Wheat	Barley	Grass	Beet	Flax
Potatoes	**100.00**	0.00	0.37	0.00	0.32	0.26	0.00
Rape Seed	0.00	**100.00**	0.12	0.00	0.00	6.17	0.06
Wheat	0.00	0.00	**98.95**	0.00	0.00	0.00	0.00
Barley	0.00	0.00	0.56	**100.00**	0.00	0.00	1.22
Grass	0.00	0.00	0.00	0.00	**99.68**	0.76	0.00
Beet	0.00	0.00	0.00	0.00	0.00	**92.81**	0.00
Flax	0.00	0.00	0.00	0.00	0.00	0.00	**98.72**

TABLE IV
CONFUSION MATRIX (%) OF THE WHOLE CLASSIFICATION FRAMEWORK ON THE THIRD DATA SET

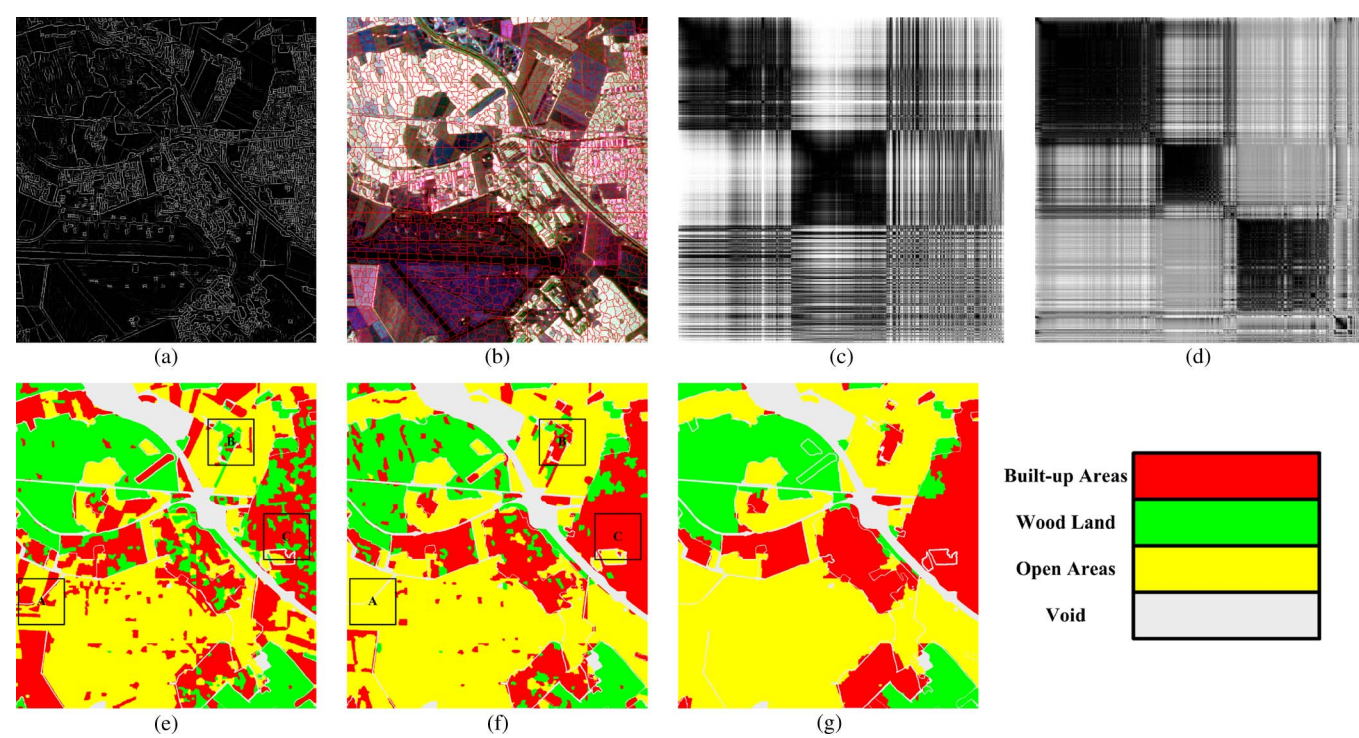|	Potatoes	Beet	Peas	Rape Seed	Wheat	Lucerne	Grass	Barley	Bare Soil
Potatoes	**99.53**	0.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Beet	0.41	**99.58**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Peas	0.01	0.00	**99.57**	0.00	0.09	0.00	0.00	0.00	0.00
Rape Seed	0.01	0.05	0.00	**100.00**	0.00	0.00	0.00	0.00	0.00
Wheat	0.03	0.00	0.43	0.00	**99.87**	0.00	0.38	0.50	0.00
Lucerne	0.00	0.00	0.00	0.00	0.01	**92.68**	0.20	0.00	0.00
Grass	0.01	0.10	0.00	0.00	0.03	7.32	**99.42**	0.00	0.00
Barley	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**99.50**	0.00
Bare Soil	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**100.00**

Fig. 14. Results of the whole classification framework on the fourth data set. (a) The edge map after the oriented nonmaximal suppression. (b) Superpixel map. (c) The RDI from statistical information only. (d) The RDI from fusion of statistical and edge information. (e) The final classification map from the NoC estimation using statistical information only. (f) The final classification map from the NoC estimation using fusion of statistical and edge information. (g) The final result after using majority vote on (f).

TABLE V
Confusion Matrix (%) of the Whole Classification
Framework on the Fourth Data Set

	Wood Land	Open Areas	Built-up Areas
Wood Land	94.68	0.61	2.51
Open Areas	1.76	95.10	4.19
Built-up Areas	3.56	4.29	93.30

With the distance fusing the statistical information and edge information D_{fusion} as described in (18), where η_{SRW} and η_{EM} are set to 1 and 1, respectively, the RDI is shown in Fig. 14(d); the NoC is estimated as three; and the final classification result is in Fig. 14(f).

Compare the final classification results in Fig. 14(e) and (f); though the estimated NoCs are both consistent with the ground truth map, the final classification result in (f) is better than result in (e), particularly in some strongly scattered farmland and built-up areas, such as in areas A, B, and C in Fig. 14(e) and (f). The experimental results show that, with inclusion of edge information, the class centers can be estimated more accurately. Since the edge information is an effective indicator of built-up areas, the merging of edge information in the NoC estimation can extend the distances between built-up areas and open areas, and between built-up areas and wood land.

Since the fourth data set is more complicated than the second and third data sets, a majority vote is applied to the final classification map in Fig. 14(f) as postprocessing to generate the final result, as shown in Fig. 14(g). According to the ground truth map in Fig. 8(b), the total accuracy of the whole proposed framework on the fourth data set is 94.42%. Corresponding confusion matrix is listed in Table V.

For the fourth data set, the image is automatically classified according to the basic scattering characteristics in general areas. This automatic operation gives a solid foundation for further analyses. For each basic land cover type, it can be further divided, such as what we did in agricultural areas on the second and third data sets, which is problem and application oriented.

According to the experimental results on the second, third, and fouth data sets, the performance of the whole processing chain is discussed as follows.

1) The proposed framework correctly estimates the number of classes within the PolSAR data.
2) The framework provides an effective and robust solution for PolSAR image unsupervised classification under various numbers of classes.
3) Since the framework takes spatial relations between pixels into account and makes good use of statistical characteristics and contour information, the classification results are easy to understand and for further analyses.

IX. Conclusion

In this paper, we present a novel superpixel-based classification framework with an adaptive number of classes for PolSAR images. First, an edge map is calculated using inherent statistical characteristics of the PolSAR image, which is then partitioned into superpixels based on the edge map. Next, the number of classes is extracted from the data, and each class center is estimated. Finally, the class labels are determined and refined using iterated Wishart clustering.

Since the whole framework is based on superpixels, it takes spatial relationship between pixels into consideration, and it is robust to speckle noise. The framework makes good use of the statistical characteristics and contour information of the PolSAR data. The experimental results show that the framework improves classification performance and result understandability.

Estimating the number of classes existing in the PolSAR image plays a major part in the whole processing chain. The framework determines the number of classes and each class center before the final classification and provides a robust unsupervised classification under various numbers of classes. The estimation operation is based on superpixel, which makes it more effective and efficient.

The whole framework is based on pairwise dissimilarities, which makes the framework rather flexible. It is independent of the representation of features, as long as a pairwise dissimilarity matrix is available to represent the relations between arbitrary feature points in the PolSAR image. Therefore, multiple features can be selected and fused into the framework easily.

In the future, the framework needs to be applied on an enormous PolSAR image database to develop a fully operational procedure and allow users to benefit from the image contents. Several problems should be considered and addressed: 1) the number of classes within the image is very large; 2) the relations among different classes are unbalanced, i.e., some class might be mistaken as noise for a very small area.

In order to overcome these problems, we consider dividing the whole scene into several subscenes and then combining and validating the results of subscenes using some criteria, and adding spatial adaptivity into the oversegmentation operation. This will be reflected in our following research.

References

[1] J. C. Souyris, C. Henry, and F. Adragna, "On the use of complex SAR image spectral analysis for target detection: Assessment of polarimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 12, pp. 2725–2734, Dec. 2003.

[2] C. Liu and C. H. Gierull, "A new application for PolSAR imagery in the field of moving target indication/ship detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3426–3436, Nov. 2007.

[3] J.-S. Lee, M. R. Grunes, and E. Pottier, "Quantitative comparison of classification capability: Fully-polarimetric versus partially polarimetric SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2343–2351, Nov. 2001.

[4] P. C. Dubois, J. J. van Zyl, and T. Engman, "Measuring soil moisture with imaging radars," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 4, pp. 915–926, Jul. 1995.

[5] D. L. Schuler, J.-S. Lee, and G. De Grandi, "Measurement of topography using polarimetric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 5, pp. 1210–1221, Sep. 1996.

[6] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton, FL: CRC Press, 2009.

[7] J.-S. Lee and T. Ainsworth, "An overview of recent advances in polarimetric SAR information extraction: Algorithms and applications," in *Proc. IEEE IGARSS*, Honolulu, HI, Jul. 2010, pp. 851–854.

[8] A. Reigber, M. Jäger, M. Neumann, and L. Ferro-Famil, "Classifying polarimetric SAR data by combining expectation methods with spatial context," *Int. J. Remote Sens.*, vol. 31, no. 3, pp. 727–744, Apr. 2010.

[9] Y. Wu, K. Ji, W. Yu, and Y. Su, "Region-based classification of polarimetric SAR images using Wishart MRF," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 668–672, Oct. 2008.

[10] H. T. Li, H. Y. Gu, Y. S. Han, and J. H. Yang, "Object-oriented classification of polarimetric SAR imagery based on statistical region merging and support vector machine," in *Proc. Int. Workshop Earth Observ. Remote Sens. Appl.*, Beijing, China, Jul. 2008, pp. 147–152.

[11] T. N. Tran, R. Wehrens, D. H. Hoekman, and L. M. C. Buydens, "Initialization of Markov random field clustering of large remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 8, pp. 1912–1919, Aug. 2005.

[12] F. Cao, W. Hong, Y. Wu, and E. Pottier, "An unsupervised segmentation with an adaptive number of clusters using the $SPAN/H/\alpha/A$ space and the complex Wishart clustering for fully polarimetric SAR data analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3454–3467, Nov. 2007.

[13] J. J. van Zyl, "Unsupervised classification of scattering behavior using radar polarimetry data," *IEEE Trans. Geosci. Remote Sens.*, vol. 27, no. 1, pp. 36–45, Jan. 1989.

[14] S. R. Cloude and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 68–78, Jan. 1997.

[15] L. Ferro-Famil, E. Pottier, and J.-S. Lee, "Unsupervised classification of multifrequency and fully polarimetric SAR images based on the H/A/Alpha Wishart classifier," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2332–2342, Nov. 2001.

[16] J. A. Kong, A. S. Swartz, H. A. Yueh, L. M. Novak, and R. T. Shin, "Identification of terrain cover using the optimal terrain classifier," *J. Electromagn. Waves Appl.*, vol. 2, no. 2, pp. 171–194, Jan. 1988.

[17] J.-S. Lee, M. R. Grunes, and R. Kwok, "Classification of multi-look polarimetric SAR imagery based on the complex Wishart distribution," *Int. J. Remote Sens.*, vol. 15, no. 11, pp. 2299–2311, Jul. 1994.

[18] 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.

[19] L.-J. Du and J.-S. Lee, "Fuzzy classification of earth terrain covers using multilook polarimetric SAR image data," *Int. J. Remote Sens.*, vol. 17, no. 4, pp. 809–826, Mar. 1996.

[20] C.-T. Chen, K.-S. Chen, and J.-S. Lee, "The use of fully polarimetric information for the fuzzy neural classification of SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 2089–2100, Sep. 2003.

[21] J.-S. Lee, M. R. Grunes, T. L. Ainsworth, L.-J. Du, D. L. Schuler, and S. R. Cloude, "Unsupervised classification using polarimetric decomposition and the complex Wishart classifier," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2249–2258, Sep. 1999.

[22] J.-S. Lee, M. R. Grunes, E. Pottier, and L. Ferro-Famil, "Unsupervised terrain classification preserving polarimetric scattering characteristics," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 4, pp. 722–731, Apr. 2004.

[23] S.-Z. Li, "Markov random field models in computer vision," in *Proc. IEEE ECCV*, Stockholm, Sweden, May 1994, vol. 2, pp. 361–370.

[24] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.

[25] Y. Dong, A. K. Milne, and B. C. Forster, "Segmentation and classification of vegetated areas using polarimetric SAR image data," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 321–329, Feb. 2001.

[26] D. H. Hoekman, M. Vissers, and T. Tran, "Unsupervised full-polarimetric SAR data segmentation as a tool for classification of agricultural areas," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 2, pp. 402–411, Jun. 2011.

[27] U. Benz and E. Pottier, "Object based analysis of polarimetric SAR data in alpha-entropy-anisotropy decomposition using fuzzy classification by eCognition," in *Proc. IEEE IGARSS*, Sydney, Australia, Jul. 2001, vol. 3, pp. 1427–1429.

[28] K. Ersahin, I. G. Cumming, and R. K. Ward, "Segmentation of polarimetric SAR data using contour information via spectral graph partitioning," in *Proc. IEEE IGARSS*, Barcelona, Spain, Jul. 2007, pp. 2240–2243.

[29] K. Ersahin, I. G. Cumming, and R. K. Ward, "Segmentation and classification of polarimetric SAR data using spectral graph partitioning," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 164–174, Jan. 2010.

[30] X. Ren and J. Malik, "Learning a classification model for segmentation," in *Proc. 9th ICCV*, Nice, France, Oct. 2003, vol. 1, pp. 10–17.

[31] 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.

[32] G. Mori, "Guiding model search using segmentation," in *Proc. 10th ICCV*, Beijing, China, Oct. 2005, vol. 2, pp. 1417–1423.

[33] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)," *Ann. Math. Stat.*, vol. 34, no. 1, pp. 152–177, Mar. 1963.

[34] K. Conradsen, A. A. Nielsen, J. Schou, and H. Skriver, "A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 1, pp. 4–19, Jan. 2003.

[35] G. Vasile, J.-P. Ovarlez, F. Pascal, and C. Tison, "Coherency matrix estimation of heterogeneous clutter in high-resolution polarimetric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 1809–1826, Apr. 2010.

[36] S. N. Anfinsen, R. Jenssen, and T. Eltoft, "Spectral clustering of polarimetric SAR data with Wishart-derived distance measures," in *Proc. 3th Int. Workshop PolInSAR*, Frascati, Italy, Jan. 2007, pp. 1–8.

[37] P. Formont, J.-P. Ovarlez, F. Pascal, G. Vasile, and L. Ferro-Famil, "A new method for high resolution polarimetric SAR data classification based on the M-Box test," in *Proc. EUSAR*, Aachen, Germany, Jun. 2010, pp. 1–4.

[38] P. Formont, J.-P. Ovarlez, F. Pascal, G. Vasile, and L. Ferro-Famil, "A test statistic for high resolution polarimetric SAR data classification," in *Proc. IEEE IGARSS*, Honolulu, HI, Jul. 2010, pp. 1871–1874.

[39] P. Formont, F. Pascal, G. Vasile, J.-P. Ovarlez, and L. Ferro-Famil, "Statistical classification for heterogeneous polarimetric SAR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 3, pp. 567–576, Jun. 2011.

[40] E. Erten, O. Chesnokova, C. Rossi, and I. Hajnsek, "A polarimetric temporal scene parameter and its application to change detection," in *Proc. IEEE IGARSS*, Vancouver, BC, Canada, Jul. 2011, pp. 1091–1094.

[41] S. N. Anfinsen, T. Eltoft, and A. Doulgeris, "A relaxed Wishart model for polarimetric SAR data," in *Proc. 4th Int. Workshop PolInSAR*, Frascati, Italy, Apr. 2009, pp. 1–7.

[42] L. Bombrun and J.-M. Beaulieu, "Fisher distribution for texture modeling of polarimetric SAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 3, pp. 512–516, Jul. 2008.

[43] B. Liu, H. Wang, K. Wang, X. Liu, and W. Yu, "A foreground/background separation framework for interpreting polarimetric SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 2, pp. 288–292, Mar. 2011.

[44] C. J. Oliver, D. Blacknell, and R. G. White, "Optimum edge detection in SAR," *Proc. Inst. Elect. Eng.—Radar, Sonar, Navig.*, vol. 143, no. 1, pp. 31–40, Feb. 1996.

[45] J. Schou, H. Skriver, A. A. Nielsen, and K. Conradsen, "CFAR edge detector for polarimetric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 1, pp. 20–32, Jan. 2003.

[46] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.

[47] J. Malik, S. Belongie, T. Leung, and J. Shi, "Contour and texture analysis for image segmentation," *Int. J. Comput. Vis.*, vol. 43, no. 1, pp. 7–27, Jun. 2001.

[48] T. Leung and J. Malik, "Contour continuity in region-based image segmentation," in *Proc. ECCV*, Freiburg, Germany, Jun. 1998, vol. 1, pp. 544–559.

[49] Alaska Satellite Facility User Remote Sensing Access. [Online]. Available: https://ursa.asfdaac.alaska.edu/cgi-bin/airsar_mission_list/guest/

[50] J. J. van Zyl and F. T. Ulaby, "Scattering matrix representation for simple targets," in *Radar Polarimetry for Geoscience Applications*, F. T. Ulaby and C. Elachi, Eds. Norwood, MA: Artech House, 1990.

[51] S. X. Yu and J. Shi, "Multiclass spectral clustering," in *Proc. 9th ICCV*, Nice, France, Oct. 2003, pp. 313–319.

[52] J. C. Bezdek and R. Hathaway, "VAT: A tool for visual assessment of (cluster) tendency," in *Proc. IJCNN*, Honolulu, HI, May 2002, pp. 2225–2230.

[53] L. Wang, C. Leckie, K. Ramamohanarao, and J. Bezdek, "Automatically determining the number of clusters in unlabeled data sets," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 3, pp. 335–350, Mar. 2009.

[54] R. B. Cattell, "A note on correlation clusters and cluster search methods," *Psychometrika*, vol. 9, no. 3, pp. 169–184, Sep. 1944.

[55] J.-S. Lee, M. R. Grunes, and G. de Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2363–2373, Sep. 1999.

[56] D. H. Hoekman and M. A. M. Vissers, "A new polarimetric classification approach evaluated for agricultural crops," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 12, pp. 2881–2889, Dec. 2003.

[57] Spaceborne and Airborne PolSAR Images. [Online]. Available: http://envisat.esa.int/polsarpro/datasets.html

[58] S. Beucher and F. Meyer, "The morphological approach to segmentation: The watershed transformation," in *Mathematical Morphology in Image Processing*, E. Dougherty, Ed. New York: Marcel Dekker, 1992.

Kaizhi Wang received the B.Sc. degree in electronic engineering from Xidian University, Xi'an, China, in 1999, the M.Sc. degree in biomedical engineering from Southeast University, Nanjing, China, in 2002, and the Ph.D. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2007.

Currently, he is an Assistant Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University. His research interests include synthetic aperture radar signal processing and time frequency analysis.

Bin Liu (S'11) was born in Hunan, China, in 1985. He received the B.Sc. and M.Sc. degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2007 and 2009, respectively, where he is currently working toward the Ph.D. degree in signal and information processing, Department of Electronic Engineering.

His research interests include synthetic aperture radar (SAR) image information mining and understanding, particularly SAR and polarimetric S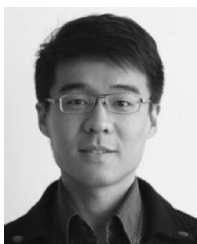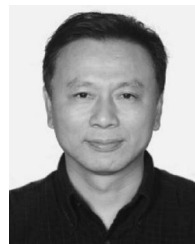AR image segmentation and classification, target detection and recognition, remote sensing software development, and high-performance computing with graphics processing unit.

Mr. Liu is a Reviewer of IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He is a Student Member of the IEEE Geoscience and Remote Sensing Society.

Xingzhao Liu (M'04) received the B.Sc. and M.Sc. degrees in electronic engineering from the Harbin Institute of Technology, Harbin, China, in 1984 and 1992, respectively, and the Ph.D. degree in electronic engineering from the University of Tokushima, Tokushima, Japan, in 1995.

Currently, he is with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China, where he is a Professor and serves as the Head of the department. His research interests mainly include HF and synthetic aperture radar signal processing.

Hao Hu received the B.Sc. degree in communication engineering from Sun Yat-Sen University, Guangzhou, China, in 2007 and the M.Sc. degree in signal and information processing from Shanghai Jiao Tong University, Shanghai, China, in 2012, where he is currently working toward the Ph.D. degree in signal and information processing, Department of Electronic Engineering.

His research interests are in the domain of synthetic aperture radar image interpretation (segmentation, road extraction, and classification).

Huanyu Wang received the B.Sc. degree in information engineering from Xi'an Jiaotong University, Xi'an, China, in 2009, and the M.Sc. degree in signal and information processing from Shanghai Jiao Tong University, Shanghai, China, in 2012.

Currently, he is working as a Microelectronics Research Engineer at Shanghai Laboratory, Fujitsu Research and Development Center Co., Ltd., Shanghai. His research interests are in the domain of synthetic aperture radar image interpretation (segmentation, clustering, and semantic annotation) and video processing (inpainting, super-resolution, and stereo camera 3-D reconstruction).

Wenxian Yu was born in Shanghai, China, in 1964. He received the B.Sc., M.Sc., and Ph.D. degrees from the National University of Defense Technology, Changsha, China, in 1985, 1988, and 1993, respectively.

From 1996 to 2008, he was a Professor with the College of Electronic Science and Engineering, National University of Defense Technology. Currently, he is with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, where he is a Professor and Head of the research part of the school. His current research interests include radar target recognition, remote sensing information processing, multi-sensor data fusion, etc. In these areas, he has published more than 200 research papers.

Dr. Yu currently serves in the expert group of major project of the second-generation satellite navigation system, and also in the expert group of major project of high-resolution Earth observation system. He serves as the Director of Shanghai Key Laboratory of Navigation and Location Based Services. He serves in the editorial board of the *Chinese Journal of Electronics* and *Journal of Infrared and Millimeter Waves*. He was the Deputy Director of the expert group for Earth Observation and Navigation Technology, National High-Tech Research and Development Program (863 Program) in China, from 2006 to 2011. He was also the Chief Scientist of Information Acquisition and Processing Technology of the 863 Program in China, from 2001 to 2005.