Unsupervised Classification of Multilook Polarimetric SAR Data Using Spatially Variant Wishart Mixture Model with Double Constraints

Chi Liu, Wenzhi Liao, *Senior Member, IEEE*, Heng-Chao Li, *Senior Member, IEEE*, Kun Fu, and Wilfried Philips, *Senior Member, IEEE*

***Abstract*— This paper addresses the unsupervised classification problems for multilook Polarimetric synthetic aperture radar (PolSAR) images by proposing a patch-level spatially variant Wishart mixture model (SVWMM) with double constraints. We construct this model by jointly modeling the pixels in a patch (rather than an individual pixel) so as to effectively capture the local correlation in the PolSAR images. More importantly, a responsibility parameter is introduced to the proposed model, providing not only the possibility to represent the importance of different pixels within a patch but also the additional flexibility for incorporating the spatial information. As such, double constraints are further imposed by simultaneously utilizing the similarities of the neighboring pixels, respectively, defined on two different parameter spaces (i.e., the hyperparameter in the posterior distribution of mixing coefficients and the responsibility parameter). Furthermore, the variational inference algorithm is developed to achieve effective learning of the proposed SVWMM with the closed-form updates, facilitating the automatic determination of the cluster number. Experimental results on several PolSAR data sets from both airborne and spaceborne sensors demonstrate that the proposed method is effective and it enables better performances on unsupervised classification than the conventional methods.**

***Index Terms*— Polarimetric synthetic aperture radar (PolSAR), spatial constraint, unsupervised classification, variational inference (VI), Wishart mixture model (WMM).**

Manuscript received September 29, 2017; revised February 3, 2018; accepted March 14, 2018. This work was supported in part by the China Scholarship Council, in part by the FWO (Data Fusion for Image Analysis in Remote Sensing) under Grant G037115N, in part by the National Natural Science Foundation of China under Grant 61371165 and Grant 61331017, and in part by the Frontier Intersection Basic Research Project for the Central Universities under Grant A0920502051714-5. *(Corresponding author: Heng-Chao Li.)*

C. Liu is with the Sichuan Provincial Key Laboratory of Information Coding and Transmission, Southwest Jiaotong University, Chengdu 610031, China, and also with the Image Processing and Interpretation, imec Research Group at Ghent University, 9000 Ghent, Belgium (e-mail: chi.liu@ugent.be).

W. Liao and W. Philips are with the Image Processing and Interpretation, imec Research Group at Ghent University, 9000 Ghent, Belgium (e-mail: wenzhi.liao@ugent.be; wilfried.philips@ugent.be).

H.-C. Li is with the Sichuan Provincial Key Laboratory of Information Coding and Transmission, Southwest Jiaotong University, Chengdu 610031, China (e-mail: lihengchao_78@163.com).

K. Fu is with the Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China (e-mail: kunfuiecas@gmail.com).

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

Digital Object Identifier 10.1109/TGRS.2018.2819995

I. Introduction

INHERITING the advantages from synthetic aperture radar (SAR), Polarimetric SAR (PolSAR) can work all day and in all-weather condition due to its active imaging mechanism and the ability to penetrate to the earth surface [1]. Compared with SAR, PolSAR acquires data from different polarimetric channels instead of a single one and can provide more information of the scenes with different land-cover topologies. PolSAR images have been pervasively used in the applications, such as land cover classification, disaster management, environment monitoring, and urban planning [2]–[5]. One of the central tasks underlying these applications is the interpretation and analysis of PolSAR images, which has attracted extensive attentions from researchers.

Many techniques have been developed for the interpretation of PolSAR images ranging from supervised methods to unsupervised ones [5]–[15]. The supervised methods generally require a sufficiently large number of training samples to achieve good classification results. It is expensive and time-consuming to obtain these training samples. In contrast, the unsupervised methods eliminate the need for the label information, which could be more convenient for the automatic interpretation task.

As a significant analysis tool for the PolSAR data, polarimetric target decomposition (TD) theorems [1], [9] can provide insights into the scattering mechanism over the scenes with different land-cover topologies. Many unsupervised classification methods have been proposed based on polarimetric TD theorems [5], [12]–[15]. In the classification method based on the Cloude–Pottier decomposition [12], the physical scattering characteristics associated with PolSAR data were first represented by a pair of parameters (i.e., the entropy H and the α angle) and then assigned to one of the eight classes according to its location in the H/α plane. This idea was extended by introducing additional parameters such as the anisotropy parameter and SPAN (the total polarimetric power) to partition the feature space and to achieve the more detailed classification results [16], [17]. The significant improvement was observed by implementing the complex Wishart classifier on the resulting TD-based classification results, for which the statistical characteristics of PolSAR data were considered [5], [13], [14], [17].

Statistical characteristics of PolSAR data lay an alternative foundation for the development of the unsupervised classification approaches. It is shown in the literature that the complex Wishart distribution [18] is effective to model the multilook PolSAR data (e.g., the covariance matrix and the coherence matrix) [1], [13], [19]. Fuzzy clustering methods utilize the similarity/dissimilarity measures derived from this statistical model to perform the unsupervised classification of PolSAR data [20]–[22]. In contrast, the methods based on mixture models directly characterize the PolSAR data and achieve the classification results according to a decision rule. The Wishart mixture model, which is a weighted combination of the complex Wishart distribution, has been proposed to model the PolSAR data and perform the classification [19], [21]. The probability theory underlying the model-based method provides the foundation for constructing flexible models and allows for the possibility to incorporate the useful information.

Spatial information is favorable for the analysis and interpretation of PolSAR images. The framework of spatially variant mixture model (SVMM), which is obtained by associating each of the data with a unique set of mixing coefficients, provides the opportunity to incorporate the spatial information and has been widely used in an image classification task [23]–[25]. Within the framework, the spatial constraint can be taken into account by introducing the Markov random field (MRF) prior. In this MRF prior, the influence of each neighboring pixel on the centered pixel is evaluated according to the similarity in the neighborhood [26], [27]. As an alternative to the MRF, the mean template method is computationally efficient and is shown to be effective in facilitating the good classification [28]. The mean template method is utilized to capture another kind of similarity in the neighboring pixels, which declines with the increase of the spatial Euclidean distance from the centered pixel. However, these two kinds of methods have not fully exploited the similarities (e.g., the geometric similarity and the covariance matrix similarity) to characterize the local correlation. Meanwhile, the data are generally assumed to be Gaussian-distributed vectors or scalars, which cannot always deal with the multilook PolSAR data in the type of complex matrix and might not consider the statistical characteristics of the multilook PolSAR data.

To overcome the aforementioned limitations, a patch-level spatially variant Wishart mixture model (SVWMM) with double constraints is proposed for the unsupervised classification of PolSAR images. The proposed SVWMM describes the PolSAR image by jointly modeling the pixels in a patch (including the neighboring pixels and the centered pixel itself) instead of an individual pixel. With the conjecture that the pixels in a patch are not equally important, a responsibility parameter is introduced to the proposed SVWMM, such that the importance of different pixels in a patch can be taken into account. More importantly, the responsibility parameter provides additional flexibility for incorporating spatial information. Therefore, on two different parameter spaces (i.e., the hyperparameter in the posterior distribution of mixing coefficients and the responsibility parameter), the geometric similarity and the covariance matrix similarity are defined. As such, double constraints are imposed to incorporate the local correlation through the two defined similarities. Furthermore, a variational inference (VI) algorithm is developed for an effective implementation of the proposed method, where the update equations of the parameters are obtained in a closed form. Last but not least, the cluster number, as well as the other parameters in the proposed SVWMM, can be automatically determined from the PolSAR data. In this way, the proposed method provides a potential solution for some real applications, where the users may not have enough backgrounds on optimizing parameters (e.g., the cluster number). Several PolSAR data sets from both the airborne and the spaceborne sensors are exploited to validate the proposed method. The results of qualitative and quantitative analyses in the experiments demonstrate that the proposed method can provide smooth classification maps and high classification accuracy, implying that the proposed method is effective and can improve the performance of the unsupervised classification over the conventional methods.

The remainder of this paper is organized as follows. In Section II, the PolSAR data and the complex Wishart distribution are introduced. Moreover, we briefly present the SVMM and the framework of the VI. Section III is devoted to the methodology, including the proposed SVWMM and its VI algorithm. Section IV demonstrates the experimental results and discussion. Finally, the conclusions of this paper are drawn in Section V.

II. Preliminaries

A. PolSAR Data and Complex Wishart Distribution

PolSAR system measures the backscattering signals of the land covers in dual or quadratic polarizations to formulate the PolSAR data. A target is fully represented by a complex scattering vector (SV), which provides both the amplitude and the phase information [1]. For fully PolSAR data, the complex SV is given by

$$\mathbf{S} = [S_{\mathrm{HH}}, S_{\mathrm{HV}}, S_{\mathrm{VH}}, S_{\mathrm{VV}}]^T \tag{1}$$

where T is the transpose operation and the elements represent the measurements by the different combinations of transmitting and receiving polarizations (e.g., the horizontal polarization and the vertical polarization). The backscattering signal received from reciprocal media is assumed to be symmetric for S_{HV} and S_{VH}, i.e., $S_{\mathrm{HV}} = S_{\mathrm{VH}}$. In this case, the SV takes the form of $\mathbf{S} = [S_{\mathrm{HH}}, \sqrt{2}S_{\mathrm{HV}}, S_{\mathrm{VV}}]^T$.

PolSAR images are typically corrupted by the granular speckle noise, degrading the image quality. To facilitate the following processing, multilook sampling is commonly used to reduce speckle by averaging the complex covariance matrices or the complex coherence matrices of the neighboring pixels. Specifically, the multilook covariance matrix is obtained by

$$\mathbf{C} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{S}_i\mathbf{S}_i^H \tag{2}$$

where L is the number of looks and $(\cdot)^H$ is the Hermitian transpose. The statistical characteristics of the multilook covariance matrix in the PolSAR images can be described by

complex Wishart distribution [1], [18], which is given by

$$\mathcal{W}(\mathbf{C}; L, \Sigma) = \frac{L^{Ld}}{\Gamma_d(L)} \frac{|\mathbf{C}|^{L-d}}{|\Sigma|^L} \exp\{-L \cdot \mathrm{tr}(\Sigma^{-1}\mathbf{C})\} \quad (3)$$

where d is the number of elements in $\mathbf{S}_i$, $\Gamma_d(L) = \pi^{(d(d-1)/2)} \prod_{i=0}^{i=d-1} \Gamma(L-i)$, and $\Gamma(x) = \int_0^{+\infty} z^{x-1} \exp\{-z\} dz$. Σ is the expectation of $\mathbf{C}$. Here, $|\cdot|$ and $\mathrm{tr}(\cdot)$ are the operators of determinant and trace, respectively. This distribution is essential for modeling the multilook PolSAR data, and many similarity/dissimilarity metrics are derived based on this distribution.

B. Spatially Variant Mixture Model

A finite mixture model (FMM) is a superposition of multiple probability densities [29], [30], which provides a framework to build flexible models for various types of data by appropriately selecting the probability density function (pdf) of mixture components. An M-component FMM can be represented by

$$p(x; \Pi, \Lambda) = \sum_{i=1}^{M} \pi_i f(x; \lambda_i) \quad (4)$$

where $\Pi = \{\pi_1, \pi_2, \ldots, \pi_M\}$ is the set of the mixing coefficients and $f(x; \lambda_i)$ is the pdf of the ith component. $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_M\}$ includes the parameters in all the components. The mixing coefficients are subjected to $\sum_i \pi_i = 1$ and $\pi_i \geq 0$. With multiple components in (4), the FMM is able to accurately model various data, even those with more than one underlying modes.

To incorporate spatial context, the SVMM [23], [26] was proposed based on the FMM by associating pixels with their own mixing coefficients rather than the common mixing coefficients, which is given by

$$p(x_n; \Pi, \Lambda) = \sum_{i=1}^{M} \pi_{ni} f(x_n; \lambda_i) \quad (5)$$

where $\sum_{i=1}^{M} \pi_{ni} = 1$, $\pi_{ni} \geq 0$, and $\Pi = \{\pi_{ni}\}$. By introducing an MRF prior or implementing mean template, the context information in the neighboring pixels can be taken into consideration, which facilitates smoothing images and contributing to a classification map with better visual interpretation.

C. Variational Inference

VI is a Bayesian estimation method, which is able to alleviate the overfitting phenomenon for small sample size and automatically determine the number of components according to data [31]. The idea underlying VI is to minimize the Kullback–Leibler divergence between the true posterior distribution $p(\Xi|\mathbf{X})$ and the approximated posterior distribution $q(\Xi)$, i.e.,

$$\mathbf{KL}(q||p) = \int q(\Xi) \ln \frac{q(\Xi)}{p(\Xi|\mathbf{X})} \, d\Xi \quad (6)$$

where $\Xi = \{\Xi_i\}$ is the set of all the variables and $\mathbf{X} = \{x_1, x_2, \ldots, x_N\}$ is the set of N observations. Minimization of (6) is equivalent to maximizing the lower bound of $\ln p(\mathbf{X})$ [29], which is given by

$$\max_{q(\Xi)} \mathcal{L}(q) = \int q(\Xi) \ln \frac{p(\mathbf{X}, \Xi)}{q(\Xi)} d\Xi. \quad (7)$$

In practice, the solution to $q(\Xi)$ by maximizing (7) is rarely in a closed form for a complicated mixture model. Therefore, an approximation scheme is generally implemented to achieve analytically tractable solutions. In VI, a factorized approximation based on the mean field theory is effective and commonly used, where each variable is assumed to be independent [29], [31]–[33]. Thus, we have the approximated posterior distribution as

$$q(\Xi) \approx \prod_i q(\Xi_i). \quad (8)$$

With this approximation, the posterior distribution of Ξ_i can be obtained by

$$\ln q^*(\Xi_i) = \mathbf{E}_{j \neq i}[\ln p(\mathbf{X}, \Xi)] + \mathrm{Const} \quad (9)$$

where $\mathbf{E}_{j \neq i}[\ln p(\mathbf{X}, \Xi)] = \int \ln p(\mathbf{X}, \Xi) \prod_{j \neq i} q(\Xi_j) d\Xi_j$, and Const is the constant to enforce normalization. In conjunction with appropriate conjugate prior distributions for these variables, the analytically tractable solutions can be obtained by simply matching the form of the prior distributions with the corresponding posterior distributions.

The objective function in the maximization problem (7) can also be given by

$$\mathcal{L}(q) = \int q(\Xi) \ln p(\mathbf{x}|\Xi) d\Xi - \mathbf{KL}(q(\Xi)||p(\Xi)) \quad (10)$$

where the first term can be viewed as a log-likelihood term. The second term could penalize the overlapping components in mixture models, resulting in simple models with less components [31]. When VI is applied to the inference of SVMM, the cluster number could be automatically determined with the help of both the above-mentioned property of VI and the introduction of the Dirichlet prior distribution over the mixing coefficients [29].

III. Methodology

A. Proposed SVWMM

As elaborated earlier, the multilook PolSAR data are complex matrices, which can be appropriately described by the complex Wishart distribution. To formulate a model for the multilook PolSAR data, it is intuitive to select the complex Wishart distribution as the component density in the SVMM, yielding a standard SVWMM with M components as

$$p(C_n) = \sum_{i=1}^{M} \pi_{ni} \mathcal{W}\left(C_n; L, \Omega_i^{-1}\right) \quad (11)$$

where we use $\{\Omega_i^{-1}\}$ to facilitate assigning conjugate prior distributions. Under the assumption of independent and identically distributed (i.i.d.) multilook PolSAR data, the parameters in (11) can be estimated with the maximum likelihood estimation. However, there are some limitations in (11). $\{\pi_{ni}\}$ is the only appropriate parameter to incorporate spatial information, revealing that this model is not flexible enough to simultaneously incorporate the geometric similarity and the covariance matrix similarity. In addition, i.i.d. assumption implies that this model is still a pixel-based method and the local correlation is not explicitly indicated in (11).

The pixel-based SVWMM [see (11)] might not provide reliable results for the interpretation of PolSAR images due to the lack of the explicit indication of the local correlation. To effectively exploit the local correlation, we can jointly model the pixels in a patch (including the neighboring pixels and the centered pixel itself) obtained by implementing a spatial sliding window with the size of $k \times k$. Thus, the PolSAR image can be statistically described based on these patches (the grouped pixels) rather than individual pixels. In practice, each pixel in a patch is not generally of equal importance to its patch. Therefore, to effectively exploit the importance of different pixels in a patch, a responsibility parameter is introduced to the proposed SVWMM, which is given by

$$p(\mathcal{P}_n) = \sum_{i=1}^{M} \pi_{ni} \left\{ \prod_{m \in \mathcal{P}_n} \mathcal{W}(C_m; L, \Omega_i^{-1})^{\omega_{\mathrm{nm}}} \right\} \tag{12}$$

where $\mathcal{P}_n$ is the patch centered at pixel n. $\{\omega_{\mathrm{nm}}\}$ is the set of the responsibility parameters indicating the importance of the mth pixel in the nth patch, and $\sum_{m \in \mathcal{P}_n} \omega_{\mathrm{nm}} = 1$. In addition to the advantage of flexibly describing the patches, the responsibility parameter is more important in providing additional flexibility for incorporating the spatial information. Moreover, when all the PolSAR data in a patch are equal to each other, we have $C_m = C_s$ and $\omega_{\mathrm{nm}} = 1/k^2$ for $(m, s) \in \mathcal{P}_n$. Then, (12) reduces to the standard SVWMM [see (11)], which implies that the proposed SVWMM is more flexible in form and is an extension of the standard SVWMM. In Section III-C, the design of the responsibility parameter $\{\omega_{\mathrm{nm}}\}$ is elaborated.

B. Variational Inference for the Proposed SVWMM

Given a PolSAR image with N covariance matrices $\mathbf{C} = \{C_1, C_2, \ldots, C_N\}$, the log-likelihood function of the proposed SVWMM is given by

$$P(\mathbf{C}) = \sum_{n=1}^{N} \ln \left\{ \sum_{i=1}^{M} \pi_{ni} \{ \prod_{m \in \mathcal{P}_n} \mathcal{W}(C_m; L, \Omega_i^{-1})^{\omega_{\mathrm{nm}}} \} \right\}. \tag{13}$$

In view of the summation in the logarithm operation, it is challenging to estimate the parameters $\{\pi_{ni}\}$ and $\{\mathbf{\Omega}_i\}$ by directly maximizing (13). Therefore, hidden variables $\mathbf{Z} = \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n\}$ are introduced to mitigate this difficulty [29], [31]. Specifically, $\mathbf{z}_n = [z_{n1}, z_{n2}, \ldots, z_{nM}]$ is a binary vector for the nth observation, in which only one of the elements is 1 (i.e., $z_{ni} = 1$) and the others are 0 (i.e., $z_{nk} = 0$, $k \neq i$). By identifying the only element with the value of 1, this vector explicitly indicates which component an observation belongs to. Thus, the observations can be statistically described by a pdf conditioning on $\mathbf{Z}$. The multinomial distribution is assigned as the conditional distribution of $\mathbf{z}_n$ on $\boldsymbol{\pi}_n = [\pi_{n1}, \pi_{n2}, \ldots, \pi_{nM}]$. Furthermore, the Dirichlet distribution is selected as the prior distribution of the vector $\boldsymbol{\pi}_n$, which is conjugate with the multinomial distribution and enforces the summation of π_{ni} over i to be 1. More importantly, with this conjugate prior distribution, the closed-form update equations could be derived, contributing to an efficient inference algorithm. These distributions are given in the following forms as:

$$p(\mathbf{Z}|\Pi) = \prod_{n=1}^{N} \left\{ \prod_{i=1}^{M} \pi_{ni}^{z_{ni}} \right\} \tag{14}$$

$$p(\Pi; \{\alpha_{ni0}\}) = \prod_{n=1}^{N} \left\{ \frac{\Gamma(\sum_{i=1}^{M} \alpha_{ni0})}{\prod_{i=1}^{M} \Gamma(\alpha_{ni0})} \cdot \prod_{i=1}^{M} \pi_{ni}^{\alpha_{ni0}-1} \right\} \tag{15}$$

$$p(\mathbf{C}|\mathbf{Z}) = \prod_{n=1}^{N} \prod_{i=1}^{M} \left\{ \prod_{m \in \mathcal{P}_n} \mathcal{W}(C_m; L, \Omega_i^{-1})^{\omega_{\mathrm{nm}}} \right\}^{z_{ni}} \tag{16}$$

where every hyperparameter α_{ni0} is positive. By further assigning prior distributions to the parameters in the component densities, a fully Bayesian inference model is established. Consequently, we have the complete data log-likelihood function as

$$\begin{aligned} &\ln p(\mathbf{C}, \mathbf{Z}, \Pi, \mathbf{\Omega}) \\ &= \sum_{n=1}^{N} \sum_{i=1}^{M} z_{ni} \left\{ \sum_{m \in \mathcal{P}_n} \omega_{\mathrm{nm}} \cdot \ln \mathcal{W}(C_m; L, \Omega_i^{-1}) \right\} \\ &\quad + \sum_{i=1}^{M} \ln p(\Omega_i; \theta_{i0}) + \ln p(\mathbf{Z}|\Pi) + \ln p(\Pi; \{\alpha_{ni0}\}) \end{aligned} \tag{17}$$

where $p(\Omega_i; \theta_{i0})$ is the prior distribution of the parameter set in the density of ith component and θ_{i0} is the set of corresponding hyperparameters. This complete data log-likelihood function is significant both in learning model parameters and for the classification of observations.

To establish the Bayesian model for the inference of the proposed SVWMM, the prior distributions are required for the parameters in component densities. Specifically, L in the proposed SVWMM is the number of looks, which should be common for each component. Thus, L is not assigned prior distribution and is treated as a known parameter, which can be determined according to the data of the homogeneous area. As for $\{\mathbf{\Omega}_i\}$, the complex Wishart distribution $\mathcal{W}(\mathbf{\Omega}_i; \eta_{i0}, W_{i0}^{-1})$ is selected as its conjugate prior distribution facilitating the derivation of the closed-form update equations. By substituting the component density and the prior distributions into (17), we have the log-likelihood function of the proposed SVWMM as

$$\begin{aligned} &\ln p(\mathbf{C}, \Xi) \\ &= \sum_{n} \sum_{i} z_{ni} \left\{ \sum_{m \in \mathcal{P}_n} \omega_{\mathrm{nm}} [Ld \ln L - \ln \Gamma_d\ (L) \right. \\ &\qquad \left. + (L-d) \ln |C_m| + L \ln |\mathbf{\Omega}_i| - L \cdot \mathrm{tr}(\mathbf{\Omega}_i C_m)] \right\} \\ &\quad + \sum_{n} \sum_{i} z_{ni} \ln \pi_{ni} + \sum_{n} \sum_{i} (\alpha_{ni0} - 1) \ln \pi_{ni} \\ &\quad + \sum_{i} \{(\eta_{i0} - d) \ln |\mathbf{\Omega}_i| - \eta_{i0} \cdot \mathrm{tr}(W_{i0} \mathbf{\Omega}_i)\} + \mathrm{Const} \end{aligned} \tag{18}$$

where Const is the summation of the terms that are irrelevant to the variables $\Xi = \{\mathbf{Z}, \Pi, \mathbf{\Omega}\}$. The first term in (18) indicates

the contribution of the observations to the log-likelihood function. In particular, by observing this term, the contribution from a patch to the log-likelihood function is given by a weighted combination of the contribution of the pixels in the corresponding patch, which provides a flexible representation of a patch and can consider the spatial information in the neighboring pixels by designing the responsibility parameter $\{\omega_{\mathrm{nm}}\}$.

Following the framework of VI, we obtain the posterior distribution of $\mathbf{Z}$ in logarithm as:

$$\begin{aligned}
&\ln q^*(\mathbf{Z}) \\
&= \sum_n \sum_i z_{ni} \ln \rho_{ni} + \text{Const} \\
&= \sum_n \sum_i z_{ni} \left\{ \sum_{m \in \mathcal{P}_n} \omega_{\mathrm{nm}} (Ld \ln L - \ln \Gamma_d\ (L) \right. \\
&\quad + (L-d) \ln |C_m| + L\mathbf{E}[\ln |\mathbf{\Omega}_i|] - L \cdot \mathrm{tr}(\mathbf{E}[\mathbf{\Omega}_i] C_m)) \\
&\quad \left. + \mathbf{E}[\ln \pi_{ni}] \right\} + \text{Const} \qquad (19)
\end{aligned}$$

where $\mathbf{E}[\bullet]$ is the expectation. By further implementing normalization based on (19), the solution to $\mathbf{Z}$ can be obtained by

$$\ln q^*(\mathbf{Z}) = \sum_n \sum_i z_{ni} \ln r_{ni} \qquad (20)$$

where $r_{ni} = \rho_{ni} / \sum_i \rho_{ni}$ and $\mathbf{E}[z_{ni}] = r_{ni}$.

Similarly, the posterior distributions of Π and $\mathbf{\Omega}$ in logarithm are, respectively, given by

$$\ln q^*(\Pi) = \sum_n \sum_i (\alpha_{ni0} + \mathbf{E}[z_{ni}] - 1) \ln \pi_{ni} + \text{Const} \qquad (21)$$

and

$$\begin{aligned}
&\ln q^*(\mathbf{\Omega}) \\
&= \sum_i \left\{ (\eta_{i0} + \sum_n \sum_{m \in \mathcal{P}_n} L \cdot \mathbf{E}[z_{ni}] \cdot \omega_{\mathrm{nm}} - d) \ln |\mathbf{\Omega}_i| \right. \\
&\quad \left. - \mathrm{tr}\left(\left[\eta_{i0} W_{i0} + \sum_n \sum_{m \in \mathcal{P}_n} L\mathbf{E}[z_{ni}] \omega_{\mathrm{nm}} C_m \right] \cdot \mathbf{\Omega}_i \right) \right\} + \text{Const.}
\end{aligned} \qquad (22)$$

Since the conjugate prior distributions are assigned, these resulting posterior distributions keep the same function form as their prior distributions. With this knowledge, we can directly obtain the update equations of the hyperparameters in an elementwise form as

$$\alpha_{ni} = \alpha_{ni0} + \mathbf{E}[z_{ni}] \qquad (23a)$$

$$\eta_i = \eta_{i0} + \sum_n \sum_{m \in \mathcal{P}_n} L \cdot \mathbf{E}[z_{ni}] \cdot \omega_{\mathrm{nm}} \qquad (23b)$$

$$W_i = \frac{1}{\eta_i} \{ \eta_{i0} W_{i0} + \sum_n \sum_{m \in \mathcal{P}_n} L\mathbf{E}[z_{ni}] \omega_{\mathrm{nm}} \cdot C_m \}. \qquad (23c)$$

Therefore, the parameters in the proposed SVWMM can be estimated by alternatively evaluating (23a)–(23c) until the stop criterion is reached. The necessary expectations are calculated according to

$$\mathbf{E}[\ln \pi_{ni}] = \psi(\alpha_{ni}) - \psi\left(\sum_i \alpha_{ni} \right) \qquad (24)$$

$$\mathbf{E}[z_{ni}] = r_{ni} \qquad (25)$$

$$\mathbf{E}[\mathbf{\Omega}_i] = W_i^{-1} \qquad (26)$$

$$\mathbf{E}[\ln |\mathbf{\Omega}_i|] = \ln \left| W_i^{-1} \right| - d \ln \eta_i + \sum_{j=0}^{d-1} \psi(\eta_i - j) \qquad (27)$$

where $\psi(x) = d \ln \Gamma(x)/dx$ is the digamma function. With the resulting SVWMM, the label of nth pixel in the PolSAR image can be determined by identifying the largest element in $\mathbf{z}_n$ [34], [35], i.e.,

$$l_n = \arg\max_i \ \mathbf{E}[z_{ni}]. \qquad (28)$$

C. Double Constraints

The similarities in the neighborhood might not be fully exploited by the existing methods to perform the interpretation task of the multilook PolSAR images, for which only one of the geometric similarity and the covariance matrix similarity is considered. The geometric similarity increases as the spatial distance declines, and the covariance matrix similarity describes the variation of multilook PolSAR data in a local area. Simultaneous incorporation of these two similarities can benefit the analysis and interpretation of PolSAR images due to the sufficient consideration of the local correlation. To this end, the geometric similarity and the covariance matrix similarity are first defined from the spatial Euclidean distance and the distance measure of the covariance matrices. Then, double constraints are imposed by employing these two similarities to fully investigate the local correlation.

In the proposed method, two parameters (i.e., the hyperparameter in the posterior distribution of mixing coefficients and the introduced responsibility parameter) are available for incorporating the spatial information. The hyperparameter in the posterior distribution of mixing coefficients (i.e., $\{\alpha_{ni}\}$) is related to the count of pixels in each cluster, which is revealed by summing α_{ni} in (23a) over all the pixels, i.e.,

$$\begin{aligned}
\sum_n \alpha_{ni} &= \sum_n \alpha_{ni0} + \sum_n \mathbf{E}[z_{ni}] \\
&= \widetilde{\alpha}_{i0} + N_i, \qquad (29)
\end{aligned}$$

where N_i can be interpreted as the expected pixel number of cluster i in the PolSAR image. In contrast, according to (18), the responsibility parameter represents the portion of the contribution from each pixel to the patch in the log-likelihood function. In this regard, the responsibility parameter is the more appropriate choice for incorporating the spatial information.

To achieve the appropriate classification results of the PolSAR images, the top priority is given to the incorporation of the covariance matrix similarity, which can control the smoothness in the results. As such, this similarity is obtained by evaluating the responsibility parameter according to

$$\omega_{\mathrm{nm}} = \frac{\exp\{-\mathcal{D}_W(C_m, \mathbb{C}_n)\}}{\sum_{m \in \mathcal{P}_n} \exp\{-\mathcal{D}_W(C_m, \mathbb{C}_n)\}} \qquad (30)$$

Algorithm 1 Algorithm for the Inference of the Proposed SVWMM

Compute $\{\omega_{nm}\}$ by (30) and $\{\gamma_{nm}\}$ in (31)
$M \leftarrow$ initial number of components
$\{\alpha_{ni0}, \eta_{i0}, W_{i0}\} \leftarrow$ initial values of the hyperparameters in the prior distributions
$\mathbf{E}[\mathbf{Z}] \leftarrow$ initial expectation of the indicator variable
$t \leftarrow 0$
while Stop criterion is not satisfied, **do**
  Evaluate $\alpha_{ni}^{(t+1)}$ by (23a) and (31)
  Evaluate $\eta_i^{(t+1)}$ by (23b)
  Evaluate $W_i^{(t+1)}$ by (23c)
  Update $\mathbf{E}[z_{ni}]^{(t+1)}$ according to (19) and (20)
  $t \leftarrow t + 1$
end while
Output $\{\alpha_{ni}^*, \eta_i^*, W_i^*\} \leftarrow \{\alpha_{ni}^{(t)}, \eta_i^{(t)}, W_i^{(t)}\}$
Return $\mathbf{E}[z_{ni}]^{(t)}$

where $\mathbb{C}_n$ is the average of all the pixels in $\mathcal{P}_n$ and $\mathcal{D}_W(C_m, \mathbb{C}_n) = 1/2[\ln|C_m\mathbb{C}_n| + \mathrm{tr}(C_m^{-1}\mathbb{C}_n) + \mathrm{tr}(\mathbb{C}_n^{-1}C_m)]$ is a symmetric pairwise distance measure of similarity [36]. A lower value of the distance measure implies more similarity. Therefore, (30) entitles the pixels with more similarity to more importance in the corresponding patch.

The geometric similarity is defined from the spatial Euclidean distance and is incorporated by implementing the mean template on $\{\alpha_{ni}\}$ (i.e., the hyperparameter in the posterior distribution of mixing coefficients). Specifically, $\{\alpha_{ni}\}$ is updated in an elementwise form by

$$\alpha_{ni} = \sum_{m \in \mathcal{P}_n} \gamma_{\mathrm{nm}} \alpha_{ni} \tag{31}$$

where $\gamma_{\mathrm{nm}} = \sigma_{\mathrm{nm}} / \sum_{m \in \mathcal{P}_n} \sigma_{\mathrm{nm}}$ and $\sigma_{\mathrm{nm}} = 1/(1 + d_{\mathrm{nm}}^2)$. d_{nm} is the Euclidean distance between the centered pixel and the pixel m in patch n. With this implementation, the pixels nearer to the centered pixel n have more contribution to patch n. Furthermore, the geometric similarity is conveyed to $\mathbf{E}[z_{ni}]$ through the term $\mathbf{E}[\ln \pi_{ni}]$ as shown in (19), which finally affects the classification map resulting from the criteria such as that in (28).

By evaluating (30) and (31), double constraints are imposed in the proposed method by simultaneously considering the two similarities. The complete algorithm for the inference of the proposed SVWMM with double constraints is given in Algorithm 1. For the initialization of the hyperparameters in the prior distributions, $\alpha_{ni0} = 1 \times 10^{-1}$ and $\eta_{i0} = 5$ are the general choice. According to our experience, the appropriate value of W_{i0} could be selected in the range of $(1 \times 10^2 \times I_{\mathrm{avg}} \sim 1 \times 10^5 \times I_{\mathrm{avg}}) \cdot \mathbf{I}$, where I_{avg} is obtained by averaging the intensity of all the channels (e.g., the diagonal elements of the covariance matrix). The selection of the patch size depends on the data set. 3×3 and 5×5 windows are generally utilized. The use of larger patch size could provide smooth results in the homogeneous area, and the use of smaller patch size can preserve more details. Thus, W_{i0} can be appropriately selected, and the cluster number can be automatically determined according to the PolSAR data instead of being manually set.

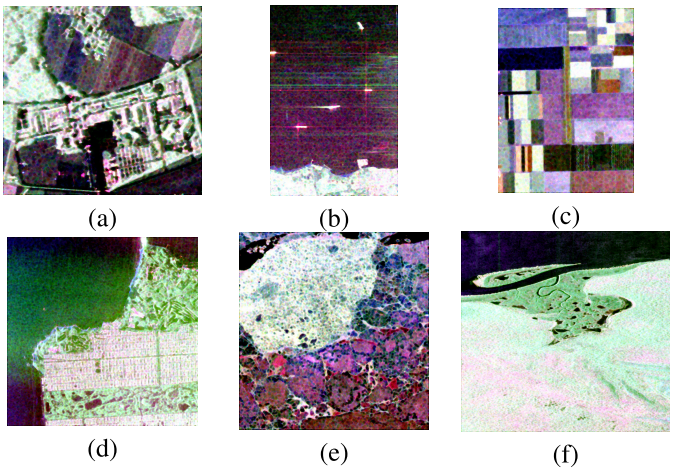

Fig. 1. Pauli RGB images over the test sites. (a) L-band PolSAR image with the size of 395×415 was acquired by Deutsches Zentrum für Luft- und Raumfahrt (DLR) ESAR over Oberpfaffenhofen, Germany. (b) Excerpt of the C-band PolSAR image with the size of 513×345 was acquired by RADARSAT-2 over Vancouver, BC, Canada. (c) This is a 383×281 excerpt of the PolSAR image over the agriculture area in Flevoland measured by NASA/JPL airborne synthetic aperture radar (AIRSAR) in the L-band. (d) 586×595 image over San Francisco Bay was obtained by NASA/JPL AIRSAR in the P-band. (e) 544×523 PolSAR image of sea ice was acquired by the Environment Canada SAR580-Convair in the Northumberland Strait off the Coast of New Brunswick, Canada. (f) NASA/JPL AIRSAR L-band Pauli RGB image over the savanna area in the north of Kakadu Northern Territory of Australia, the image size of which is 792×827.

IV. Experimental Results and Discussion

In the experiments, the effectiveness of the proposed method is verified based on six real PolSAR images, as shown in Fig. 1. Various types of terrains and objects are included ranging from ocean, forest, sea ice, agriculture area to ships, roads, streets, and buildings. Furthermore, these images are acquired in different frequencies (i.e., P-band, L-band, and C-band) and by either airborne or spaceborne sensor (i.e., NASA/Jet Propulsion Laboratory (JPL) AIRSAR, DLR E-SAR, SAR580-Convair, and RADARSAT-2), which facilitates an extensive validation of the proposed method in both the qualitative and the quantitative manner. The data sets in Fig. 1(a)–(d) are, respectively, named "Oberp" short for "Oberpfaffenhofen", "Vancouver", "Flevoland", and "San Francisco" according to their corresponding place names. The remaining data sets in Fig. 1(e) and (f) are referred to as "Sea Ice" and "Savanna" due to their landscapes. All the data sets are preprocessed by the refined Lee filter [37]. The labels of pixels in the classification maps resulting from both the variational WMM and the proposed SVWMM are obtained according to (28). The proposed method is implemented with MATLAB programming language, and the experiments are conducted with MATLAB R2015b on Ubuntu Linux operation system with 64-GB RAM and Intel Core i7-3930K CPU.

A. Smoothness and Cluster Number

We take the PolSAR image "Vancouver" as an example to see the influence of the parameter settings in the

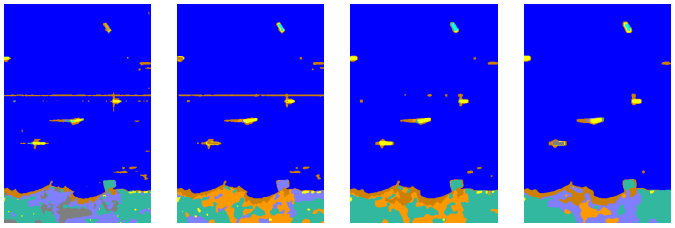

Fig. 2. Resulting classification maps with the patch sizes of (a) 3×3, (b) 5×5, (c) 7×7, and (d) 9×9. All the other parameters are fixed as $\{\alpha_{ni0} = 1 \times 10^{-5}, \eta_{i0} = 5, W_{i0} = 1 \times 10^{10} \cdot \mathbf{I}\}$ for all n and i values, where $\mathbf{I}$ is the identity matrix.

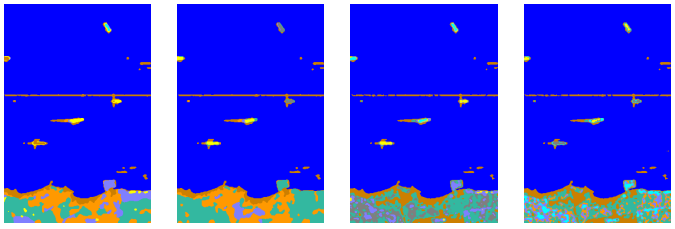

Fig. 3. Resulting classification maps by setting the parameter α_{ni0} with various values of (a) 1×10^{-5}, (b) 1×10^{-2}, (c) 1×10^{0}, and (d) 1×10^{2}. The patch size is fixed as 5×5, and the other parameters are fixed as $\{\eta_{i0} = 5, W_{i0} = 1 \times 10^{10} \cdot \mathbf{I}\}$ for all i values, where $\mathbf{I}$ is the identity matrix.

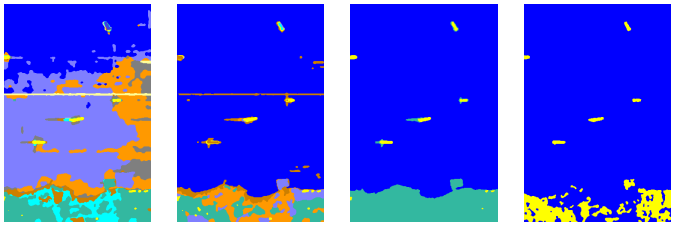

Fig. 4. Resulting classification maps using the parameter W_{i0} with various values of (a) $1 \times 10^{9} \cdot \mathbf{I}$, (b) $1 \times 10^{10} \cdot \mathbf{I}$, (c) $1 \times 10^{11} \cdot \mathbf{I}$, and (d) $1 \times 10^{12} \cdot \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. The patch size is fixed as 5×5, and the other parameters are fixed as $\{\alpha_{ni0} = 1 \times 10^{-5}, \eta_{i0} = 5\}$ for all n and i values.

proposed SVWMM. We vary only one of the parameters and keep the other parameters being fixed. In the experiments, the proposed SVWMM is initialized with 10 components, which could be much more than the underlying number of components. The parameters (i.e., $\{\alpha_{ni0}, \eta_{i0}, W_{i0}\}$) in the prior distributions are equally set for all the components so as to avoid the initial discrimination on any component. The resulting classification maps are illustrated in Figs. 2–4.

It can be observed from Fig. 2 that the land, the ocean, and the ships are recognized by the proposed method. For small patch sizes (e.g., 3×3 and 5×5), the pixels of the line in the middle of the classification map are identified to be different from the ocean, which is obviously the misclassification. With larger patch sizes, this misclassification can be corrected by incorporating more spatial context information from the surrounding pixels, and less miscellaneous pixels are observed [see Fig. 2(c) and (d)]. It is worth noting that the results with small patch sizes are very similar to the corresponding Pauli RGB image by observing Figs. 1(b) and 2(a) and (b), which could imply that a small patch size tends to identify more details although some of these details could not be necessary. However, the parameter $\{\alpha_{ni0}\}$ plays a different role in controlling the smoothness of the resulting classification map. Although the classification results in the ocean area (i.e., blue pixels) are similar for diversifying values of α_{ni0}, it is quite different in the land (the lower parts of the images), as shown in Fig. 3. Specifically, a relatively larger value of α_{ni0} (e.g., 1×10^{2}) leads to a complex structure in the land. By comparison, the simple structure is recognized for relatively small α_{ni0} according to Fig. 3(a) and (b), which are more convenient for the visual interpretation.

The classification maps for various values of W_{i0} are demonstrated in Fig. 4, from which the resulting cluster numbers can also be identified. The resulting cluster numbers in Fig. 4(a)–(d) are 10, 8, 3, and 2, respectively. As such, there is an evidence that the resulting cluster number decreases and the classification maps become smooth as the value of W_{i0} increases. Although large W_{i0} could result in oversmoothing results, the good classification map can be achieved by appropriately selecting the values of W_{i0}. In addition, it is found that the resulting number of clusters can be less than the initial 10 components. Therefore, the proposed method can automatically determine the cluster number according to the PolSAR data if the initial number of components is large enough.

B. Qualitative Analysis

The qualitative analysis of the proposed method is first performed by identifying the land use in data sets "Oberp", "San Francisco", and "Vancouver". Then, data sets "Sea Ice" and "Savanna" are employed to test the performance of the proposed method on the classification of natural areas. The proposed method is evaluated by comparing with: 1) the H/α-Wishart method proposed by Lee *et al.* [5]; 2) the Chernoff–Wishart method proposed by Dabboor *et al.* [14]; and 3) the variational Wishart mixture model (see the Appendix). For both the H/α-Wishart method and the Chernoff–Wishart method, the nearest two clusters are merged until the specified number of clusters is reached.

We first determine the specified cluster number for both the H/α-Wishart method and the Chernoff–Wishart method and then give the initial component number for the variational WMM and the proposed method. For data set "Oberp", we set the cluster number of both the H/α-Wishart method and the Chernoff–Wishart method to 8 to identify the rich details. According to the main objects and land use, the cluster numbers for data sets "San Francisco" and "Vancouver" are all specified as 3. Data set "Sea Ice" includes first year ice (FYI), rough FYI, ridged FYI, and leads. In data set "Savanna", water, mangroves, dense vegetation, and sparse vegetation are observed from the corresponding optical image in December 1996 provided by Google Earth. Therefore, their cluster number is identified as 4. For the variational WMM and the proposed method, the initial component numbers are set to be 10, which is a larger value than the

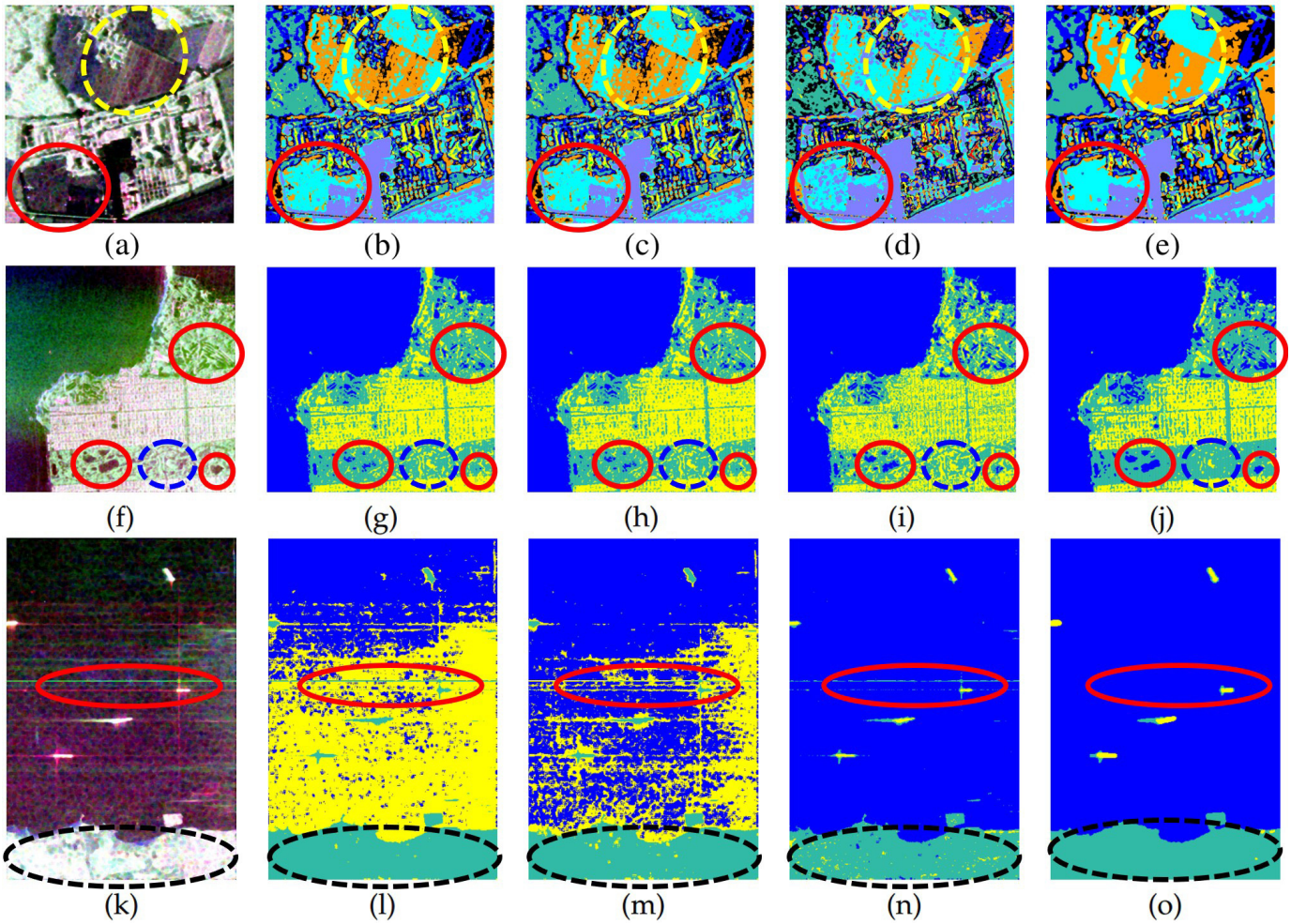

Fig. 5. Pauli RGB images of the test data sets and the classification maps resulting from the different methods. The first row corresponds to data set "Oberp", the second row to "San Francisco", and the third row to "Vancouver". (a), (f), and (k) Pauli RGB images of the test data sets. (b), (g), and (l) Classification maps using the H/α-Wishart method. (c), (h), and (m) Classification maps using the Chernoff–Wishart method. (d), (i), and (n) Classification maps using the variational WMM. (e), (j), and (o) Classification maps using the proposed method.

TABLE I

AVERAGE ELAPSED TIME FOR THE H/α-WISHART METHOD, THE CHERNOFF–WISHART METHOD, THE VARIATIONAL WMM, AND THE PROPOSED METHOD

Data Set	Average Elapsed Time (s)			
	H/α-Wishart method	Chernoff-Wishart method	Variational WMM	Proposed Method
San Francisco	13.10	339.36	29.60	240.78
Oberp	3.77	286.61	14.53	73.73
Vancouver	7.70	213.50	7.489	64.31

underlying cluster numbers of these data sets. The appropriate cluster number for these data sets can be automatically determined by the variational WMM and the proposed method.

1) Classification of Human Activity Areas: The classification maps resulting from all the four methods are presented in Fig. 5. According to Fig. 5(a)–(e), the buildings, the woodland, and the ground are identified and the rich details are demonstrated by all the methods. There are less miscellaneous pixels in the yellow circle and the red circle of Fig. 5(e), revealing that the proposed method can provide better classification results in the homogeneous area.

The Pauli RGB image and the corresponding classification maps of the P-band data set "San Francisco" are shown in Fig. 5(f)–(j). It is observed from Fig. 5(j) that the structure and the details in the red circles are explicitly exhibited, which is in accordance with those in the Pauli RGB image [see Fig. 5(f)]. In contrast, the H/α-Wishart method and the Chernoff–Wishart method cannot accurately capture these structures, leading to the loss of details according to Fig. 5(g) and (h). The variational WMM demonstrates the competitive performance in view of the good results in the red circles of Fig. 5(i). However, compared with the results of the proposed method, there are more miscellaneous and misclassified pixels for the variational WMM in the blue

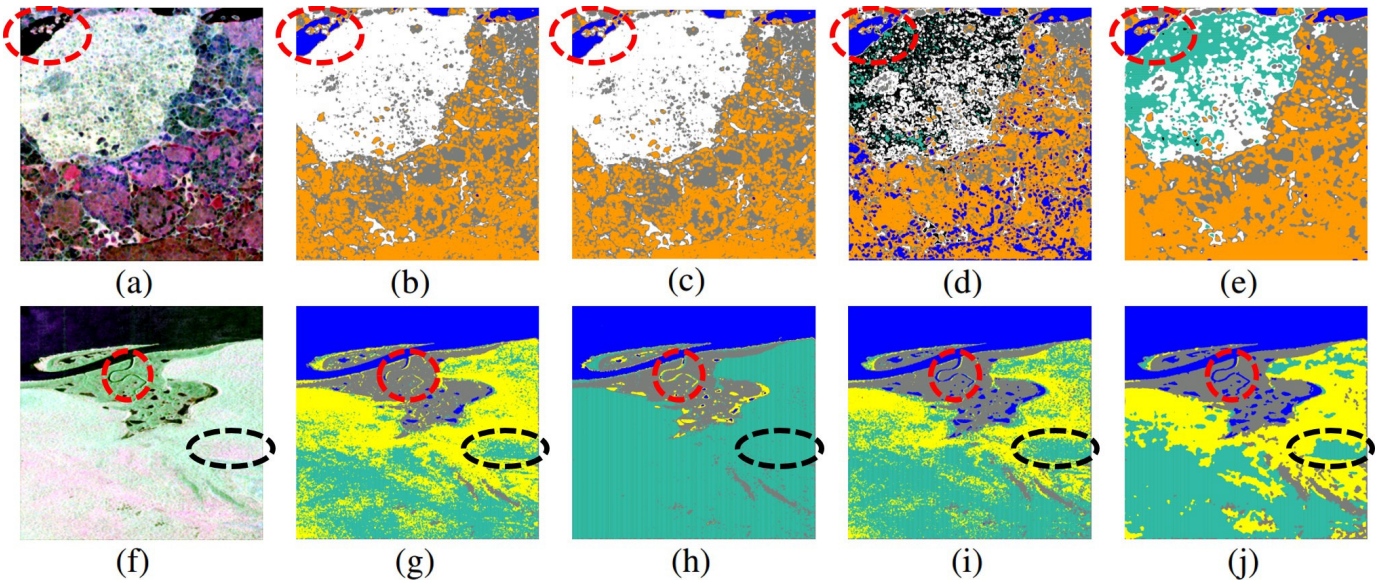

Fig. 6. Pauli RGB images of the test data sets and the classification maps resulting from the different methods. The first row corresponds to data set "Sea Ice" and the second row to "Savanna". (a) and (f) Pauli RGB images of the test data sets. (b) and (g) Classification maps using the H/α-Wishart method. (c) and (h) Classification maps using the Chernoff–Wishart method. (d) and (i) Classification maps using the variational WMM. (e) and (j) Classification maps using the proposed method.

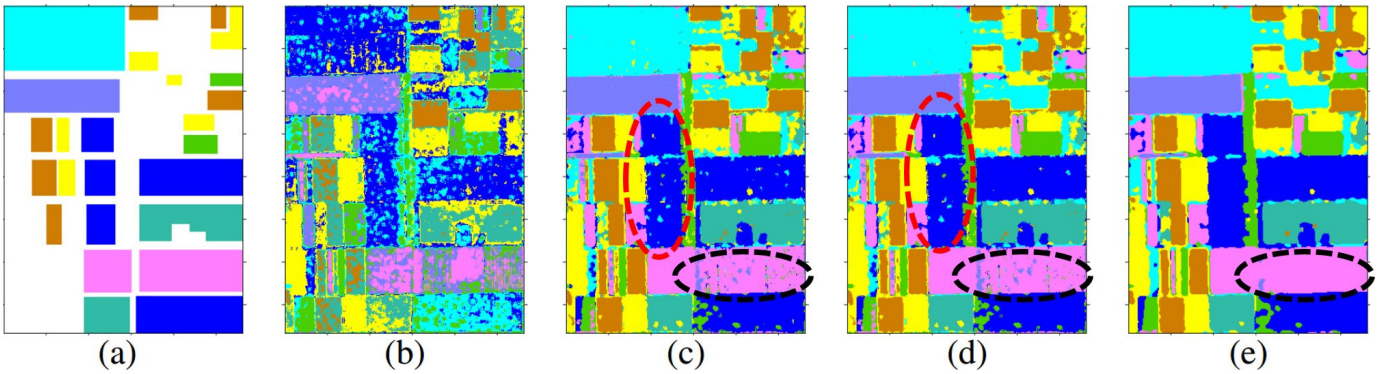

Fig. 7. (a) Ground truth for data set "Flevoland". The classification results of the proposed method (b) without considering both similarities, (c) considering only the geometric similarity, (d) considering only the covariance matrix similarity, and (e) considering both the similarities.

TABLE II

QUANTITATIVE EVALUATION OF THE EFFECT OF THE GEOMETRIC SIMILARITY AND THE COVARIANCE MATRIX SIMILARITY ON THE CLASSIFICATION OF DATA SET "FLEVOLAND"

	No Constraints	Only Geometric Similarity	Only Covariance Matrix Similarity	Double Constraints
OA	0.585	0.944	0.949	**0.965**
κ	0.515	0.933	0.939	**0.959**

dotted circle, implying that the incorporation of the spatial information in the proposed method is helpful in facilitating the visual interpretation of PolSAR images.

Data set "Vancouver" exhibits solecistic structures such as the horizontal lines in the red circle of Fig. 5(k), which brings about the challenge for its interpretation. For the H/α-Wishart method and the Chernoff–Wishart method [see Fig. 5(l) and (m)], the ships are correctly identified. However, the solecistic structures are preserved and the ocean area is classified as two categories, which are not desired. The WMM provides a better result as presented in Fig. 5(n), where the ocean area is identified as a whole part. Nevertheless, the pixels of the meaningless lines in the red circle are not completely assigned to the appropriate class. For this scenario, the spatial context is quite useful and can guide these pixels to be appropriately classified. As shown in Fig. 5(o), the proposed method appropriately identifies the pixels of the solecistic structures in the red circle as the ocean, contributing to a good classification map for visual interpretation. In addition, there are less miscellaneous pixels for the proposed method in the black dotted circle.

The average elapsed time of the four methods (i.e., the H/α-Wishart method, the Chernoff–Wishart method, the variational WMM, and the proposed method) over 30 independent

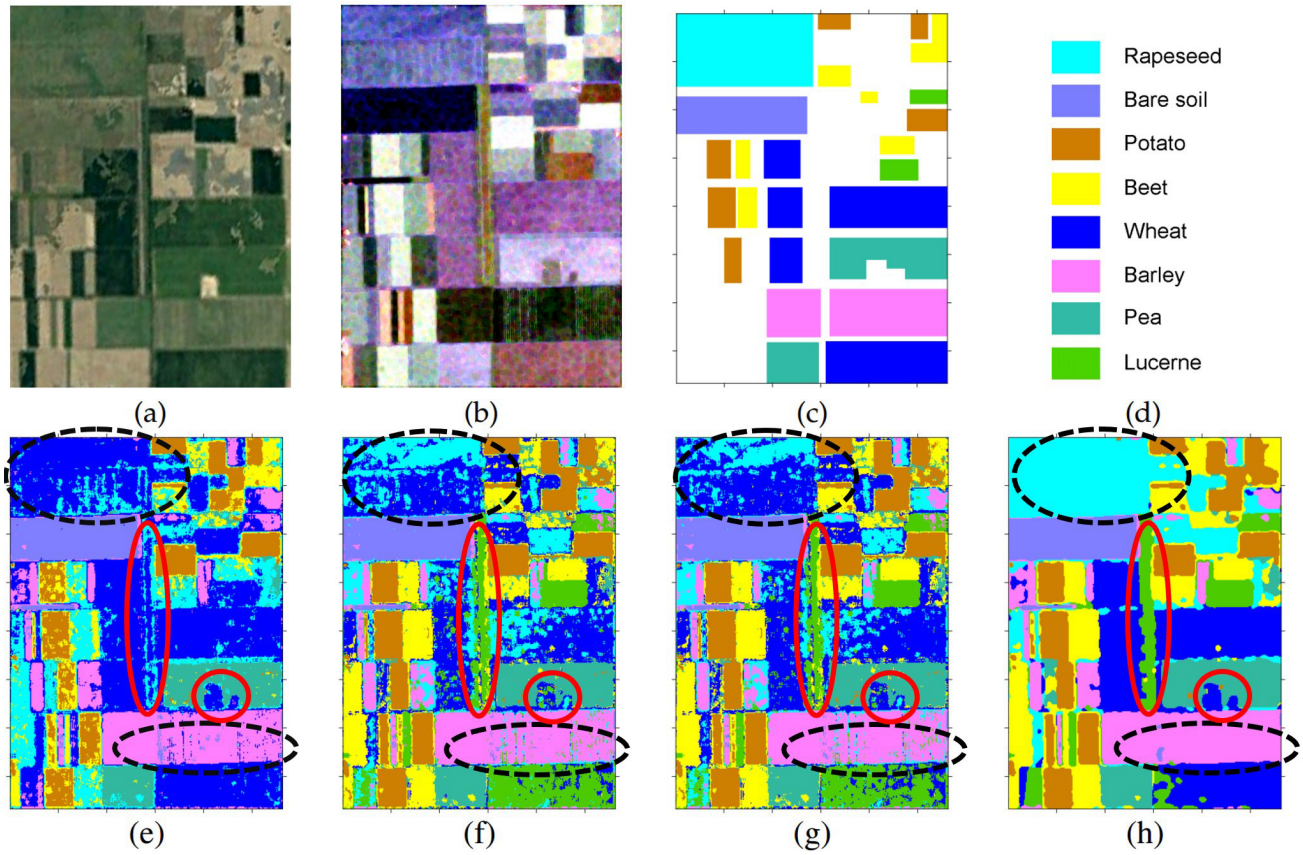

Fig. 8. Test data and the classification results. (a) Optical image over Flevoland, The Netherlands. (b) Pauli RGB image of the data set. (c) Ground truth. (d) Color code. The resulting classification maps by employing (e) H/α-Wishart method, (f) Chernoff–Wishart method, (g) Variational WMM, and (h) proposed method.

TABLE III

QUANTITATIVE EVALUATION OF THE H/α-WISHART METHOD, THE CHERNOFF–WISHART METHOD, THE VARIATIONAL WMM, AND THE PROPOSED METHOD BASED ON DATA SET "FLEVOLAND"

	Producer's Accuracy			
Class (Pixel Number)	H/α-Wishart	Chernoff-Wishart	WMM	Proposed Method
Rapeseed (11011)	0.136	0.480	0.373	**0.983**
Bare soil (5343)	0.967	0.962	0.961	**0.973**
Potato (5123)	0.800	0.972	0.972	**0.974**
Beet (4308)	0.425	0.811	0.828	**0.860**
Wheat (15208)	0.960	0.456	0.582	**0.962**
Barley (8729)	0.946	0.923	0.913	**0.985**
Pea (6842)	0.914	0.932	0.922	**0.973**
Lucerne (1450)	-	0.899	0.908	**0.964**
OA	0.719	0.717	0.728	**0.965**
κ	0.655	0.668	0.679	**0.959**

tests are presented in Table I. In view of the larger size of data set "San Francisco", all the four methods consume more time for this data set than the other two (i.e., "Oberp" and "Vancouver"), as shown in Table I. Although the H/α-Wishart method and the variational WMM are generally faster, the proposed method can incorporate local correlation and provide better classification results in the acceptable running time. Meanwhile, compared with the Chernoff–Wishart method, the proposed method is more computationally efficient according to Table I. The Chernoff–Wishart method always consumes much time. This is because the Chernoff–Wishart method merges two clusters from a large initial number (generally 48), which requires many iterations to reach the specified cluster number. In addition, the Wishart clustering procedure in the Chernoff–Wishart method, which is also an iterative procedure, is always performed after each merging operation, leading to more consumption of time.

2) Classification of Natural Areas: The Pauli RGB image of data set "Sea Ice" [see Fig. 6(a)] demonstrates complex structures in view of the many edges. Even the whole block of the ice in the upper part exhibits bright and dark parts. This structure is well identified by the proposed method, as shown in Fig. 6(e). In addition, the proposed method also correctly distinguishes the leads (such as that in the red circle), which are important for the navigation. For data set

TABLE IV

CONFUSION MATRICES RESULTING FROM THE H/α-WISHART METHOD, THE CHERNOFF–WISHART METHOD, THE VARIATIONAL WMM, AND THE PROPOSED METHOD BASED ON DATA SET "FLEVOLAND"

	Rapeseed	Bare soil	Potato	Beet	Wheat	Barley	Pea	Lucerne
	H/α-Wishart Method							
Rapeseed	1501	24	4	0	9211	227	44	0
Bare soil	2	5165	0	0	24	152	0	0
Potato	35	0	4098	983	1	0	6	0
Beet	2142	0	55	1829	14	0	268	0
Wheat	465	0	0	0	14593	141	9	0
Barley	59	147	0	1	263	8527	2	0
Pea	395	0	54	100	40	0	6253	0
Lucerne	93	0	0	1	1004	352	0	0
	Chernoff-Wishart Method							
Rapeseed	5288	19	3	24	5531	101	34	11
Bare soil	21	5141	0	0	4	166	0	11
Potato	0	0	4979	108	0	0	34	2
Beet	0	0	471	3495	21	0	296	25
Wheat	2765	0	0	46	6928	84	4	5381
Barley	270	63	0	39	25	8059	1	272
Pea	5	0	61	226	161	0	6377	12
Lucerne	0	0	0	7	2	138	0	1303
	Variational WMM							
Rapeseed	4102	17	3	30	6751	69	34	5
Bare soil	23	5132	0	0	7	171	0	10
Potato	0	0	4982	111	0	0	28	2
Beet	1	0	481	3565	20	0	218	23
Wheat	1727	0	0	51	8854	79	2	4495
Barley	417	50	0	40	31	7967	1	223
Pea	3	0	64	299	159	0	6308	9
Lucerne	0	0	0	7	2	124	0	1317
	Proposed Method							
Rapeseed	10825	0	20	0	0	161	5	0
Bare soil	9	5201	0	0	0	111	0	22
Potato	0	0	4989	134	0	0	0	0
Beet	2	0	584	3704	0	0	11	7
Wheat	20	0	0	30	14629	32	0	497
Barley	3	17	0	42	7	8601	0	59
Pea	24	0	39	103	18	0	6658	0
Lucerne	0	0	0	3	0	49	0	1398

"Savanna", the H/α-Wishart method does not identify the complete structure of the water according to the red circle of Fig. 6(g). Although the Chernoff–Wishart method can preserve this structure, the misclassification for the water is observed from Fig. 6(h) and the details are lost in the black circle. The proposed method provides the improved classification map in terms of image details' preservation [see Fig. 6(j)].

C. Quantitative Analysis

To quantitatively evaluate the proposed method, data set "Flevoland" is employed in view of its sufficient ground truth, as shown in Fig. 7(a). With this data set, two experiments are carried out to investigate the effect of two similarities and to evaluate the proposed method against the comparison methods, including the H/α-Wishart method, the Chernoff–Wishart method, and the variational WMM. The number of classes for the H/α-Wishart method and the Chernoff–Wishart method is specified as 8 [8]. For a fair comparison, the initial class number for both the variational WMM and the proposed method is also specified as 8. To calculate the metrics for the quantitative analysis, it is important to obtain an appropriate mapping of the generated labels by the unsupervised classification to the ground-truth labels. By exhaustively trying all the possible mappings from the generated labels to the ground-truth labels, the one with the highest overall accuracy (OA) is selected, and we can have the comparable classification map with the ground truth.

1) Effect of Double Constraints: By presenting the results with either the geometric similarity or the covariance matrix similarity, the effect of these two similarities can be investigated. The corresponding OA and the kappa coefficient (κ) are given in Table II. Under this circumstance, the proposed SVWMM reduces to (11) and the constraint can be imposed by implementing the mean template on $\{\alpha_{ni}\}$ (i.e., the hyperparameter in the posterior distribution of mixing coefficients). The result without considering both similarities could be obtained by further eliminating the implementation of the mean template. The ground truth of data set "Flevoland" and the classification maps are illustrated in Fig. 7.

As shown in Fig. 7(b), the result without considering both similarities exhibits many miscellaneous pixels, leading to the misclassification and the difficulty in the visual interpretation. Considering either the geometric similarity or the covariance matrix similarity will significantly improve the classification performance in view of the good classification maps in Fig. 7(c) and (d), which is also confirmed in Table II. The covariance matrix similarity can result in smoother results in the homogeneous area (e.g., the area in the red circle) than

the geometric similarity. Nevertheless, none of them provides the coinciding results with the ground truth in the black circles of Fig. 7(c) and (d). By observing Fig. 7(e), the smooth result is achieved in view of the less miscellaneous pixels in the pink blocks, implying that the proposed method can further improve the classification performance by simultaneously considering the two similarities.

2) Classification of Agriculture Area: In Fig. 8, the experimental results from the H/α-Wishart method, the Chernoff–Wishart method, the variational WMM, and the proposed method are illustrated. The metrics, including producer's accuracy, OA, and the κ coefficient, are summarized in Table III for the quantitative evaluation. Moreover, the confusion matrices of the H/α-Wishart method, the Chernoff–Wishart method, the variational WMM, and the proposed method are demonstrated in Table IV.

By observing Fig. 8(e)–(h), all of the four methods provide appropriate classification maps. However, according to Tables III and IV, the class "lucerne" is not identified by the H/α-Wishart method. The H/α-Wishart method can find at most eight classes according to the locations of pixels in the entropy and α plane [12]. This small number of classes inhibits its application in the classification of the complicated PolSAR images with more than eight classes. Moreover, the H/α-Wishart method cannot identify the details in the middle red circle by comparing Fig. 8(e) and (b). In contrast, the structures and details are explicitly recognized by the proposed method according to the results in the red circles of Fig. 8(h). In addition to this good preservation of structures, the proposed method allows for the smooth classification result in the homogeneous area. Specifically, less miscellaneous pixels and a smooth region are observed for the proposed method, as shown in the black dotted circles of Fig. 8(h). In addition, only the proposed method can correctly identify the whole block of rapeseed in the upper black circle by comparing the results of all the four methods with the ground truth in Fig. 8(c). According to Table IV, the diagonal elements in the confusion matrix of the proposed method achieve larger values and the values of other elements are small, implying that the proposed method achieves the good result. There is an evidence that the proposed method can provide appropriate classification results, which is confirmed by its large values of producer's accuracy, OA, and κ in Table III.

V. Conclusion

A patch-level SVWMM with double constraints has been proposed for the unsupervised classification of multilook PolSAR images. In the proposed method, the introduced responsibility parameter facilitates both imposing double constraints and utilizing the importance of different pixels in a patch. We consider not only the geometric similarity but also the covariance matrix similarity. The closed-form updates in the VI algorithm for the proposed SVWMM are achieved. The experiments were performed to verify the effectiveness and to evaluate the performance of the proposed method based on the P-band, L-band, and C-band PolSAR images acquired from either the airborne or spaceborne sensors. The patch size and the hyperparameter $\{\alpha_{ni0}\}$ in the prior distribution of the mixing coefficients $\{\pi_{ni}\}$ play different roles in controlling the smoothness of the classification results. The cluster number can be automatically determined according to the PolSAR data by the proposed method. The resulting cluster number is influenced by the hyperparameter $\{W_{i0}\}$ in the prior distribution of the density parameter $\{\mathbf{\Omega}_i\}$, for which a larger value of $\{W_{i0}\}$ can lead to the classification results with less clusters. The experimental results demonstrate that the proposed method can achieve appropriate classification results for the automatic interpretation of PolSAR images.

A future work will investigate the extension of the proposed method for the classification of high-resolution PolSAR images. In high-resolution PolSAR images, the scatterer number in one resolution cell is limited, under which the central limit theorem is not always valid. The zero-mean complex Gaussian distribution and the complex Wishart distribution could fail to capture the statistical characteristics of PolSAR data in high-resolution images. One potential solution is to develop models with more flexibility to characterize high-resolution PolSAR data. Another line of the future work will investigate the superpixel-level SVWMM to preserve clear boundaries and to develop fast learning algorithm for the PolSAR image classification. Finally, more PolSAR data sets from natural landscape areas will be exploited to test the performance of the proposed method in the scenarios with a complicated scattering behavior.

Appendix
Variational WMM

Within the framework of VI, we can conveniently have the update equations for the WMM. The formulation of the Bayesian model for the inference of the WMM is similar to the proposed SVWMM, where the latent variable $\mathbf{Z}$ is introduced and its prior distribution is endowed with the Dirichlet distribution. The number of looks is treated as a known parameter as that in the proposed SVWMM. The complex Wishart distribution is selected as the prior distribution for each random variable Σ_i^{-1}, leading to conjugacy between the prior distribution and the posterior distribution. On this basis, the WMM can be learned according to the following update equations:

$$\mathbf{E}[z_{ni}] = \rho_{ni} \Big/ \sum_i \rho_{ni} \tag{32a}$$

$$\alpha_i = \alpha_{i0} + \sum_n \mathbf{E}[z_{ni}] \tag{32b}$$

$$\eta_i = \eta_{i0} + \sum_n L \cdot \mathbf{E}[z_{ni}] \tag{32c}$$

$$W_i = \frac{1}{\eta_i} \left\{ \eta_{i0} W_{i0} + \sum_n L \cdot \mathbf{E}[z_{ni}] \cdot C_n \right\} \tag{32d}$$

where ρ_{ni} can be evaluated according to

$$\begin{aligned} \ln \rho_{ni} = \{ & Ld \ln L - \ln \Gamma_d\ (L) + (L-d) \ln |C_n| \\ & + L \cdot \mathbf{E}[\ln |\mathbf{\Omega}_i|] - L \cdot \mathrm{tr}(\mathbf{E}[\mathbf{\Omega}_i] C_n) \} \\ & + \mathbf{E}[\ln \pi_i]. \end{aligned} \tag{33}$$

The expectations $\mathbf{E}[\mathbf{\Omega}_i]$ and $\mathbf{E}[\ln |\mathbf{\Omega}_i|]$ can be calculated by (26) and (27), respectively. $\mathbf{E}[\ln \pi_i]$ can be obtained

according to the standard result of the Dirichlet distribution as $\mathbf{E}[\ln \pi_i] = \psi(\alpha_i) - \psi(\sum_i \alpha_i)$. As such, all the parameters can be estimated by alternatively evaluating (32a)–(32d) until the stop criterion is reached. With the resulting WMM, the classification map is obtained according to $\mathbf{E}[z_{ni}]$. One of the effective criteria is (28).

ACKNOWLEDGMENT

The authors would like to thank the European Space Agency, the Jet Propulsion Laboratory, and MacDonald, Dettwiler and Associates for providing the PolSAR images. They would also like to thank the editors and the anonymous reviewers for their careful reading and valuable comments, which are greatly helpful in improving the quality of this paper. W. Liao is a Post-Doctoral Fellow of the Research Foundation Flanders (FWO-Vlaanderen, Belgium) and acknowledges its support.

REFERENCES

[1] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*. Raleigh, NC, USA: SciTech, 2004.

[2] C. Hütt, W. Koppe, Y. Miao, and G. Bareth, "Best accuracy land use/land cover (LULC) classification to derive crop types using multitemporal, multisensor, and multi-polarization SAR satellite images," *Remote Sens.*, vol. 8, no. 8, p. 684, 2016.

[3] Y. Duguay, M. Bernier, E. Lévesque, and F. Domine, "Land cover classification in subarctic regions using fully polarimetric RADARSAT-2 data," *Remote Sens.*, vol. 8, no. 9, p. 697, 2016.

[4] C. Yonezawa, M. Watanabe, and G. Saito, "Polarimetric decomposition analysis of ALOS PALSAR observation data before and after a landslide event," *Remote Sens.*, vol. 4, no. 8, pp. 2314–2328, 2012.

[5] 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.

[6] H. Wang, Z. Zhou, J. Turnbull, Q. Song, and F. Qi, "Pol-SAR classification based on generalized polar decomposition of Mueller matrix," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 4, pp. 565–569, Apr. 2016.

[7] L. Zhang, L. Sun, B. Zou, and W. M. Moon, "Fully polarimetric SAR image classification via sparse representation and polarimetric features," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 8, pp. 3923–3932, Aug. 2015.

[8] 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.

[9] S. R. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 2, pp. 498–518, Mar. 1996.

[10] Q. Xu, Q. Chen, S. Yang, and X. Liu, "Superpixel-based classification using K distribution and spatial context for polarimetric SAR images," *Remote Sens.*, vol. 8, no. 8, p. 619, 2016.

[11] L. Shi, L. Zhang, J. Yang, L. Zhang, and P. Li, "Supervised graph embedding for polarimetric SAR image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 2, pp. 216–220, Mar. 2013.

[12] 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.

[13] 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.

[14] M. Dabboor, M. J. Collins, V. Karathanassi, and A. Braun, "An unsupervised classification approach for polarimetric SAR data based on the Chernoff distance for complex Wishart distribution," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4200–4213, Jul. 2013.

[15] S. Wang, K. Liu, J. Pei, M. Gong, and Y. Liu, "Unsupervised classification of fully polarimetric SAR images based on scattering power entropy and copolarized ratio," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 622–626, May 2013.

[16] E. Pottier, "The H/A/α polarimetric decomposition approach applied to POLSAR data processing," in *Proc. PIERS Workshop Adv. Radar Methods*, Baveno, Italy, Jul. 1998, pp. 120–122.

[17] F. Cao, W. Hong, Y. Wu, and E. Pottier, "An unsupervised segmentation with an adaptive number of clusters using the SPAN/H/α/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.

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

[19] X. Deng, C. López-Martínez, J. Chen, and P. Han, "Statistical modeling of polarimetric SAR data: A survey and challenges," *Remote Sens.*, vol. 9, no. 4, p. 348, 2017.

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

[21] 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.

[22] W. Wang, D. Xiang, J. Zhang, and J. Wan, "Integrating contextual information with $H/\bar{\alpha}$ decomposition for PolSAR data classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 2034–2038, Dec. 2016.

[23] S. Sanjay-Gopal and T. J. Hebert, "Bayesian pixel classification using spatially variant finite mixtures and the generalized EM algorithm," *IEEE Trans. Image Process.*, vol. 7, no. 7, pp. 1014–1028, Jul. 1998.

[24] K. Blekas, A. Likas, N. P. Galatsanos, and I. E. Lagaris, "A spatially constrained mixture model for image segmentation," *IEEE Trans. Neural Netw.*, vol. 16, no. 2, pp. 494–498, Mar. 2005.

[25] T. M. Nguyen and Q. M. J. Wu, "Gaussian-mixture-model-based spatial neighborhood relationships for pixel labeling problem," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 1, pp. 193–202, Feb. 2012.

[26] C. Nikou, N. P. Galatsanos, and A. C. Likas, "A class-adaptive spatially variant mixture model for image segmentation," *IEEE Trans. Image Process.*, vol. 16, no. 4, pp. 1121–1130, Apr. 2007.

[27] C. Nikou, A. C. Likas, and N. P. Galatsanos, "A Bayesian framework for image segmentation with spatially varying mixtures," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2278–2289, Sep. 2010.

[28] H. Zhang, Q. M. J. Wu, and T. M. Nguyen, "A robust fuzzy algorithm based on student's t-distribution and mean template for image segmentation application," *IEEE Signal Process. Lett.*, vol. 20, no. 2, pp. 117–120, Feb. 2013.

[29] C. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer-Verlag, 2006.

[30] G. McLachlan and D. Peel, *Finite Mixture Models*. New York, NY, USA: Wiley, 2000.

[31] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "The variational approximation for Bayesian inference," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 131–146, Jan. 2008.

[32] G. Parisi, *Statistical Field Theory*. New York, NY, USA: Addison-Wesley, 1988.

[33] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul, "An introduction to variational methods for graphical models," *Mach. Learn.*, vol. 37, no. 2, pp. 183–233, Nov. 1999.

[34] F. Forbes and N. Peyrard, "Hidden Markov random field model selection criteria based on mean field-like approximations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1089–1101, Sep. 2003.

[35] S. P. Chatzis and G. Tsechpenakis, "The infinite hidden Markov random field model," *IEEE Trans. Neural Netw.*, vol. 21, no. 6, pp. 1004–1014, Jun. 2010.

[36] S. N. Anfinsen, R. Jenssen, and T. Eltoft, "Spectral clustering of polarimetric SAR data with Wishart-derived distance measures," in *Proc. POLINSAR*, vol. 7. Frascati, Italy, Jan. 2007, pp. 1–9.

[37] 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.

Chi Liu received the B.Sc. and M.Sc. degrees from Southwest Jiaotong University, Chengdu, China, in 2009 and 2012, respectively. He is currently pursuing the Ph.D. degree in signal and information processing with the School of Information Science and Technology, Southwest Jiaotong University, and the Ph.D. degree in computer science engineering from Ghent University, Ghent, Belgium..

His research interests include machine learning, image classification, and remote sensing image analysis and processing.

Wenzhi Liao (S'10–M'14–SM'16) received the B.Sc. degree in mathematics from Hainan Normal University, Haikou, China, in 2006, the Ph.D. degree in engineering from the South China University of Technology, Guangzhou, China, in 2012, and the Ph.D. degree in computer science engineering from Ghent University, Ghent, Belgium, in 2012.

Since 2012, he has been holding a post-doctoral position at Ghent University and then a Post-Doctoral Research Fellow with the Research Foundation Flanders (FWO-Vlaanderen, Belgium). His research interests include pattern recognition, remote sensing, and image processing, and also mathematical morphology, multitask feature learning, multisensor data fusion, and hyperspectral image restoration.

Dr. Liao received the Best Paper Challenge Award from the 2013 IEEE GRSS Data Fusion Contest and the 2014 IEEE GRSS Data Fusion Contest. He is serving as an Associate Editor for the *IET Image Processing*.

Heng-Chao Li (S'06–M'08–SM'14) received the B.Sc. and M.Sc. degrees in information and communication engineering from Southwest Jiaotong University, Chengdu, China, in 2001 and 2004, respectively, and the Ph.D. degree in information and communication engineering from the Graduate University of Chinese Academy of Sciences, Beijing, China, in 2008.

From 2013 to 2014, he was a Visiting Scholar with Prof. W. J. Emery at the University of Colorado at Boulder, Boulder, CO, USA. He is currently a Professor with the Sichuan Provincial Key Laboratory of Information Coding and Transmission, Southwest Jiaotong University. His research interests include the statistical analysis of synthetic aperture radar images, remote sensing, image processing, and signal processing in communications.

Dr. Li received several scholarships or awards, especially including the Special Grade of the Financial Support from the China Post-Doctoral Science Foundation in 2009 and the New Century Excellent Talents in University from the Ministry of Education, China, in 2011. He has been a Reviewer for several international journals and conferences, such as the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, the IEEE TRANSACTIONS ON IMAGE PROCESSING, the *IET Radar, Sonar and Navigation*, and the *Canadian Journal of Remote Sensing*. He is currently serving as an Associate Editor for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.

Kun Fu received the B.Sc. and Ph.D. degrees from the National University of Defense Technology, Changsha, China, in 1995 and 2002, respectively.

He is currently the Head and a Professor with the Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, Institute of Electronics, Chinese Academy of Sciences, Beijing, China. He has authored two books and over 60 refereed publications. His research interests include geospatial data organization and visualization, computer vision, and remote sensing image interpretation.

Wilfried Philips (S'90–M'93–SM'10) was born in Aalst, Belgium, in 1966. He received the Diploma degree in electrical engineering and the Ph.D. degree in applied sciences from Ghent University, Ghent, Belgium, in 1989 and 1993, respectively.

Since 1997, he has been a Lecturer with the Department of Telecommunications and Information Processing, Ghent University. He is currently a Senior Full Professor with the Department of Telecommunications and Information Processing, Ghent University, where he heads the Image Processing and Interpretation Research Group. His research interests include the domains of image and video restoration and analysis.