IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 24, 2017, accepted October 22, 2017, date of publication November 10, 2017,
date of current version November 28, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2768575

Context-Based Max-Margin for PolSAR Image Classification

SHUYIN ZHANG, BIAO HOU, (Member, IEEE), LICHENG JIAO, (Senior Member, IEEE), QIAN WU, CHEN SUN, AND WEN XIE

[1]Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center of Intelligent Perception and Computation, Xidian University, Xi'an 710071, China
[2]International Collaboration Joint Lab in Intelligent Perception and Computation, Xidian University, Xi'an 710071, China

Corresponding author: Biao Hou (avcodec@163.com)

This work was supported in part by the National Basic Research Program (973 Program) of China under Grant 2013CB329402, in part by the National Natural Science Foundation of China under Grant 61671350, Grant 61573267, and Grant 61572383.

ABSTRACT Context-based method for classification has been successfully applied in image. However, most of these classifiers work in stages. This paper presents a novel discriminative model named context-based max-margin to perform the task of classification for polarimetric synthetic aperture radar (PolSAR) images. Based on the max-margin frame, support vector machine (SVM), and conditional random fields (CRF) are used to describe the spectral and spatial information of polarimetric synthetic aperture radar (PolSAR) image, respectively. First, the probabilistic result which is obtained from SVM can be applied as the spectral term of the discriminative classifier. Second, CRF is used to describe the spatial information of PolSAR image. The contextual information of both label and observation field are built as the spatial term, by which the smoother region is obtained and the spatial information is preserved. Finally, a discriminative classifier can be learned by means of integrating the spectral and spatial terms. Compared with other state-of-the-art classification methods, our method exhibits higher accuracy, which indicating the effectiveness of our scheme. Here, the total classification accuracy of the proposed model increases by about 10% and 3% compared with the other methods for two data sets.

INDEX TERMS CRF, max-margin, PolSAR image, Wishart distance.

I. INTRODUCTION

The fully polarimetric synthetic aperture radar (PolSAR) provides useful information on describing the observed targets than traditional synthetic aperture radar (SAR) and the application of PolSAR has intensively attracted much attention in the last two decades [1]–[3]. Therefore, PolSAR plays an important role in many fields such as agriculture, geology and military. The land cover classification of a PolSAR image is one of the most crucial tasks in remote sensing image field [4]–[6].

In recent years, in order to improve the classification accuracy of PolSAR image, many researchers have forward the application of spatial information into PolSAR image classification [7]–[9]. Generally speaking, classification methods for PolSAR image which include the spatial information are divided into three main categories.

The first one is to make the spatial information used for pre-processing. Liu et al. propose a novel superpixel-based classification method with an adaptive number of classes for PolSAR images [10]. The PolSAR image is first partitioned into superpixels. Then, the number of classes and each class center within the data are estimated. A classification method based on multilayer auto-encoders and superpixels is proposed by Hou *et al.* [11]. Firstly, the RGB image formed with Pauli decomposition is used to produce superpixels. Secondly, multilayer auto-encoders network is used to learn the features. Song *et al.* proposed a segmentation and classification method which is implemented by fusing the Dirichlet process mixture model (DPMM) model and a similarity measure scheme into the MRF framework [12].

The second approach is to use spatial information for post-processing. A support vector machine (SVM) and markov random field (MRF) Based Method for PolSAR classification is proposed by Wu *et al.* [13]. An initial classification result is obtained by SVM. Then, MRF method is used to revise the initial classification. Liu *et al.* propose a method which combines the Wishart classifier and triplet Markov field for PolSAR image classification [14]. Shen *et al.* present a classification method that combines the MRF and v-SVM for sea ice PolSAR image [15].

The third method uses spatial information directly for classification. A new classifier is proposed by Wu *et al.* for PolSAR image [16]. In this method, Wishart distribution is applied into MRF to classify the PolSAR image. Masjedi *et al.* presented a method adds the texture features into the contextual classification [17]. A classification method which adds local spatial information into stacked sparse auto-encoder is proposed by Zhang *et al.* [18].

The spatial information used as either pre-processing or post-processing, they belong to the method which is two-step. The results of these methods significantly depend on the first step extremely. Therefore, the results cannot be revised completely by the second step. In addition, only the label information or the local information is used by the third method mentioned.

In order to mitigate the problem and retain the benefits of the aforementioned methods, a discriminative model, namely Context-based Max-Margin (CMM) is proposed. The proposed new classifier relates the context information to the Max-Margin framework. Both spatial and spectral information are combined in the classifier. Therefore, the proposed method can learn the spatial and spectral characteristics from the training samples directly. At the same time, the character of PolSAR data is considered.

In the proposed method, the characteristic function is designed by combining the spectral and spatial characteristics of PolSAR image under the Max-Margin framework. There are two essential parts designed for PolSAR image classification. The first part is a spectral term with the classical discriminative classifier SVM which models the relationship between the observed data and the corresponding label. The second part is spatial term applied here which is used to describe the spatial information and obtain the smoother regions. Especially, the spatial information is applied during the learning of the classifier, which is quite different from the two-stage classifier since the spatial information is used to revise the misclassification. In that case, the result of CMM is not dependent on the spectral term too much. Numerous descriptive and generative models have been widely utilized in the domain of image classification, such as the MRF [19]–[21]. In MRF, the contextual information is limited to the labeling field because of the conditional independent hypothesis on the observed data. Therefore, the variation of conditional random fields (CRF) is applied to extend the direct contextual information in both labeling field and observation field [22]. In the meantime, Wishart distance is applied as the similarity measurement.

There are several advantages of the proposed model. The main novel contribution of this work is:

Firstly, it is a discriminative linear classification framework designed from the Max-Margin framework for easy understanding.

Secondly, different characteristic function can be designed for different kinds of data based on the Max-Margin framework. In CMM, there are two different terms designed for characteristic function according to the PolSAR data. They are spatial term and spatial term. Therefore, selecting the appropriate spectral term under the limited number of labeled training samples and selecting the appropriate spatial term in which both PolSAR data character and context information can be taken into account are crucial.

The remainder of this paper is organized as follows. In Section II, a brief presentation of back ground is given and the proposed model is discussed in detail. Experiments and analysis are given in Section III. The conclusion is will be drawn Section IV.

II. CONTEXT-BASED MAX-MARGIN METHOD

A. MAX-MARGIN STRUCTURE IN POLSAR CLASSIFICATION PROBLEMS

Linearly separable classification task is designed to learn an optimal hyperplane that maximally separates the positive examples from the negative ones, as measured by the geometric margin [23]. The optimal hyperplane is learned by the training samples and the corresponding labels. Then the optimal hyperplane is used to classify the test samples. The task of hyperplane learning is to maximize the margin ρ ($\mathbf{W}$):

$$\max_{\mathbf{W}} \rho\ (\mathbf{W}) = \max_{\mathbf{W}} (\min_{i} \frac{\boldsymbol{y}_i \mathbf{W} \boldsymbol{x}_i}{\|\mathbf{W}\|}) \tag{1}$$

where x_i is the *ith* training sample, and y_i is its label. This task is transformed into two steps. Firstly, hyperplanes can be find to found to separate the training samples with different classes. This step is implemented by $\min_i \boldsymbol{y}_i \mathbf{W} \boldsymbol{x}_i$. Secondly, there are many hyperplanes qualified. Thus, the max margin $\mathbf{W}$ should be found by optimizing Eq.(1).Then, the following invariance is used: for every solution, an existing a solution achieves the same target function value, but with a margin of Eq. 1. Then the classification framework can be written as a constrained optimization problem:

$$\max_{\mathbf{W}} \|\mathbf{W}\|^{-1} \quad s.t.\ \boldsymbol{y}_i\ (\mathbf{W} \boldsymbol{x}_i) \geq 1 \tag{2}$$

For the PolSAR image classification problem, samples usually are divided into several different classes. The main task is to learn a function $h : \mathcal{X} \mapsto \mathcal{Y}$. For example: $T = (x_i, y_i)_{i=1}^{N}$, N is the number of the training samples. The classification function h is trained from the parametric family $\mathcal{H}$. A linear family is usually chosen: $f_j : \mathcal{X} \times \mathcal{Y} \mapsto \mathrm{IR}$, a hypothesis is defined as:

$$\begin{aligned} h_{\mathrm{W}}(x) &= \arg\max_{y} \sum_{j=1}^{t} w_j f_j\,(\mathrm{x}, \mathrm{y}) \\ &= \arg\max_{y} \mathbf{w}^T f\,(\mathrm{x}, \mathrm{y}) \end{aligned} \tag{3}$$

where $h_{\mathrm{W}} \in \mathcal{H}$, w_j is the coefficients and t is the number of w_j. f(x, y) is the characteristic function for classification issue. Max-Margin is a linear classification framework. On the base of this framework, the different characteristic function can be designed by combining different kinds of data. By combining data and spatial characters of PolSAR image, spectral and spatial term are designed in this article.

B. CRF MODEL

Conditional random fields (CRF) is a classical discriminative model based on the conditional probability and corresponds to an undirected graphical model. CRF is a popular tool for remote sensing images [22]. It can offer the ability to propagate contextual information among the labeling field and the observation field.

Our focuses are on the problems of the PolSAR image classification. Let $X = \{x_1, x_2, \ldots x_i, \ldots x_N\}_{i \in S}$ be the input image as the observation field and $\boldsymbol{x}_i = (x_{i1}, x_{i2}, \ldots, x_{i9})$ is the vector of the C matrix, where $S = \{1, 2, \ldots, N\}$ is the set of image pixel indices and N is the number of training samples. Suppose $Y = \{y_1, y_2, \ldots y_i, \ldots y_N\}_{i \in S}$ is the labeling field and $\boldsymbol{y}_i = (y_1, y_2, \ldots, y_m)$ where m is the number of classes. CRF is one of the undirected graphical models. $G = (V, E)$ is defined as an undirected graphical model, where V represents the set of vertexes and E represents the set of the edges. Assuming that each node in Y corresponds the vertex in V and if the observation field X is regarded as the global condition as well as the labeling field Y obeys the Markov property, (X, Y) forms a CRF. The frame of CRF is the posterior probability $P(Y|X)$ which satisfies the Gibbs distribution and it can be written as:

$$P(Y|X) = \frac{1}{Z(X, \theta)} \prod_{c \in C} \psi_c(y_c, X, \theta) \tag{4}$$

where $Z(X, \theta) = \sum_y \prod_{c \in C} \psi_c(y_c, X, \theta)$ is a partition function, C is a set of clique types and θ is the parameter of the model. $\psi_c(y_c, X, \theta)$ is a potential function with the parameter θ and it is based on the observed data and their corresponding labels for different types of cliques C. If we only consider the clique potential with unary and pairwise, the CRF model can be rewritten as Eq. (5)

$$P(Y|X) = \frac{1}{Z(X, \theta)} \exp\{[\sum_{i \in V} \phi_i(y_i, x_i)] + [\sum_{(i,j) \in E} \varepsilon_{i,j}(X, y_i, y_j)]\} \tag{5}$$

where $\phi_i(y_i, x_i)$ is the unary potential function which establishes the relationship of the vertexes between the observed data and its label. $\varepsilon_{i,j}(X, y_i, y_j)$ is the pairwise potential function. To take of the observation and the labeled fields into account, it represents the relationship between the current vertex and its neighborhood. i is the location of the vertex, y_i is the label of the location i. In the neighborhood system, j and i can be connected to be an edge.

C. THE PROPOSED METHOD FOR POLSAR CLASSIFICATION

Considering the characteristic of PolSAR image, a new classifier is proposed, namely context-based Max-Margin (CMM). This is a discriminative model in which the pixel-based and region-based methods are combined on the basis of Max-Margin frame to perform the task of classification of PolSAR image. In this article, the PolSAR image data is represented in coherency matrix and each pixel is a 3×3 complex matrix. The complex matrix is vectorized into 9 dimensional vector which is used as the input $\mathbf{x} \cdot \mathbf{y}$ is the label of $\mathbf{x}$. In the new frame, the CRF model can be rewritten as Maximum posterior probability (MAP):

$$\max \ P(Y|X) \propto \exp[\sum_i \phi_i(y_i, x_i) + \sum_{i,j} \varepsilon_{i,j}(\boldsymbol{x}, y_i, y_j)] \tag{6}$$

where $P(Y|X)$ is the maximum posterior probability, X is the observation vector of the image which is vectorized from the complex matrix and Y is the label of the observation vector. The label can be obtained from the MAP function.

In the max-margin frame, the problem can be re-described as Eq. (7):

$$\max_y \mathbf{w}^T [\sum_i \phi_i(y_i, x_i) + \sum_{i,j} \varepsilon_{i,j}(\mathbf{x}, y_i, y_j)] \tag{7}$$

where $\mathbf{f}(\mathbf{x}, \mathbf{y}) = [\sum_i \phi_i(y_i, x_i) + \sum_{i,j} \varepsilon_{i,j}(\mathbf{x}, y_i, y_j)]$ is the characteristic function which is redesigned, $\phi_i(y_i, x_i)$ represents the spectral term which describe the relationship between the training samples and its label and $\varepsilon_{i,j}(\mathbf{x}, y_i, y_j)$ represents the spatial term which measure the pair-site label and sample interaction for PolSAR image classification.

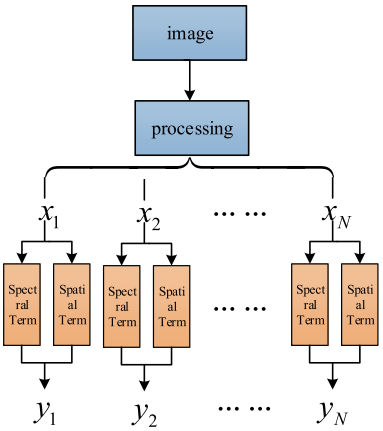

FIGURE 1. The classification framework.

The classification framework is shown in Fig.1. The input is the PolSAR image with whitening. $\{x_1, x_2, \ldots x_i, \ldots x_N\}$ are the pixels. $\{y_1, y_2, \ldots y_i, \ldots y_N\}$ are the labels of these pixels that are from the CMM method.

The proposed CMM is illustrated in Fig.2, exhibiting the combination of the spectral and spatial terms. As is shown in Fig.2, the red dot P5 is the pixel which will be classified. The brown dot is the posterior probability of the pixel. PV5 is the posterior probability obtained from the spectral term, which is related to the single spectral feature. PE5 is the posterior probability obtained from the spatial term which is decided by the 4-neighborhood. The spectral and spatial term will be introduced in the following section in details.

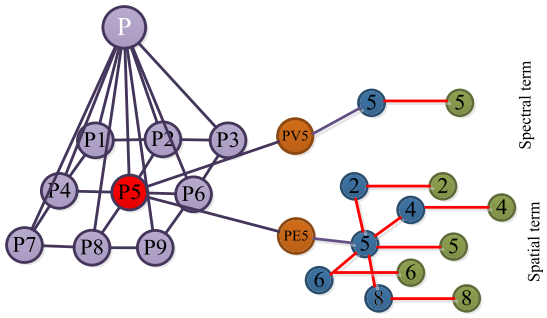

FIGURE 2. Illustration of CMM.

1) SPECTRAL TERM

The spectral term models the relationship between the pixels and their labels. In the previous model, the minimum distance from the pixel to its class center is defined as the spectral term. However, it requires abundant training samples to achieve ideal results. For PolSAR image, the limited number of labeled training samples is a critical problem. Therefore, a discriminative classifier support vector machine (SVM) is applied here, which provides an effective method to learn a max-margin decision boundary. The SVM benefits from its generality and requires a few training samples. More importantly, SVM can exploit a very high-dimensional by adopting the kernel function. The expression of a spectral term is:

$$\phi_i(y_i, x_i) = \sum_{l=1}^{L} P_{SVM}(y_i = l|x_i) \tag{8}$$

where $\phi_i(y_i, x_i)$ is the probability in which the pixel x_i belongs to a certain class y_i, $P_{SVM}(y_i = l|x_i)$ is the probability SVM model. This model is suited to the binary classification problem. For the multi-classification problem, the one-against-one model is used. L is the number of classes.

2) SPATIAL TERM

Because of the speckle in the PolSAR image, it is difficult to obtain the continuous classification results if the spectral term is only taken into account. Generally, PolSAR image has obvious spatial information, geometric structure and rich object details. Hence, a spatial term is designed to smooth the region. In most of cases, the classification of a pixel depends on its neighborhood. In a certain kind of terrain, the spectral variation of a pixel may follow some underlying patterns in the neighboring region, instead of being random. Furthermore, the underlying patterns of different classes may also be different. In this paper, Wishart distance and a smooth term are used to model the spatial term. The coherency matrix of PolSAR is satisfied with the complex Wishart distribution. In the traditional PolSAR data processing, Wishart distribution can describe the statistical property of the PolSAR data very well. We use Wishart distance as the similarity measurement that can find the underlying relationship between two observed vectors. The Wishart distance between two pixels of the PolSAR can be written as Eq. (9)

$$w_D(x_i, x_j) = \frac{1}{2} Tr((\widehat{\sum}_j)^{-1} \widehat{\sum}_i + (\widehat{\sum}_i)^{-1} \widehat{\sum}_j) - q \tag{9}$$

The spatial term is written as:

$$\begin{aligned} \varepsilon_{i,j}(\mathbf{x}, y_i, y_j) &= P_{i,j}(y_i, y_j|x_i, x_j) \\ &= \mu^l (1 - \exp(\frac{w_D(x_i, x_j)}{2\sigma^2}) \delta(y_i, y_j)) \end{aligned} \tag{10}$$

where $\varepsilon_{i,j}(\mathbf{x}, y_i, y_j)$ is the posterior probability which constructs the interaction between the single site and its neighborhoods, which can obtain the underlying pattern of the observation data and the corresponding label. μ^l is a weight of the spatial term which is large in a smooth region and small in a non-smooth region. $w_D(x_i, x_j)$ represents the Wishart distance between the pixels x_i and x_j. The closer the distance between pixel x_i and x_j is, the more similar they are. $\sigma^2 = \frac{1}{4N} \sum_{i=1}^{N} \sum_{j \in \partial_i} (x_i - x_j)^T (x_i - x_j)$ is the mean value of $(x_i - x_j)^T (x_i - x_j)$. There is a strong possibility that the two pixels have the same label. If y_i and y_j have the same label, $\delta(y_i, y_j)$ is equal to 0 and the region can be regarded to be sufficiently smooth enough. If y_i and y_j have the different labels, $\delta(y_i, y_j)$ is equal to 1. Then, the contextual information in the observation field should be considered. If the pixels have similar feature vectors, the value of exp $(\frac{w_D(x_i, x_j)}{2\sigma^2})$ is close to 0. Accordingly, the spatial term is dominant. If the pixels have dissimilar feature vectors, exp$(\frac{w_D(x_i, x_j)}{2\sigma^2})$ is close to 1. The value of $1 - \exp(\frac{w_D(x_i, x_j)}{2\sigma^2}) \delta(y_i, y_j)$ is close to 0. Thus, the spectral term is dominant.

The detailed steps of this algorithm are shown in Table 1.

TABLE 1. Detailed steps of this algorithm.

Algorithm: Context-based Max-Margin
Input: The training samples of PolSAR image: x. The label of the training samples: y
Step1: initialize the Wishart distance $w_D(x_i, x_j)$ by Eq. (9), the iteration times K and μ
Step2: for $k = 1$ to K
Step3: perform the probability of SVM by Eq. (8)
Step4: perform the spatial term by Eq. (10)
Step5: sum the result of step3 and step4
Step6: update the label by Eq. (7)
if the number of the labels changes less than 50. Stop.
else go to step3
end
end
Output: the final label for PolSAR

III. EXPERIMENTS AND DISCUSSIONS

There are two set of data used in our experience. The first set is Flevoland with the dimension of 750 × 1024, acquired by the AIRSAR platform on 495 August 16, 1989, as shown in

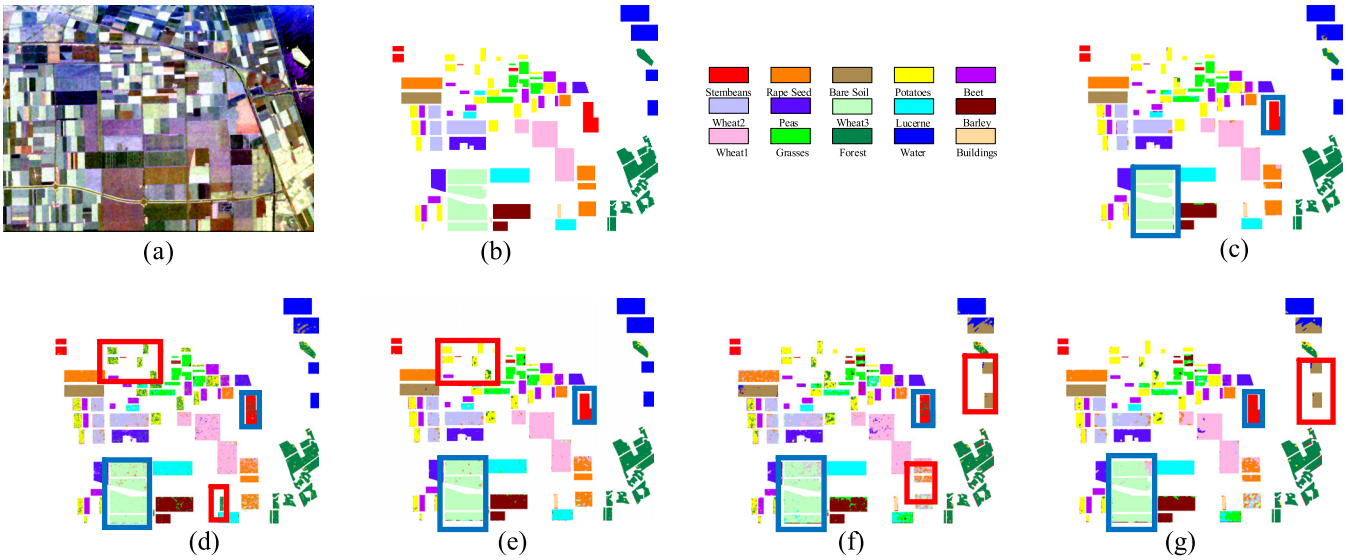

FIGURE 3. (a) the PolSAR image (PauliRGB) of Flevoland. (b) The ground truth Map. (c) Our method. (d) SVM method. (e) SVM+MRF method. (f) Wishart ML method. (f) Wishart-based MRF method.

Fig.3.(a). It is a four-look polarimetric L-band scene. The second set is Jingkun Highway in western-Xi'an-Area that is provided by our laboratory. It is acquired by RADARSAT-2, and obtained from a subset of a C-band single-look PolSAR image which size is of 512 × 512. A refined Lee filter with 5 × 5 window size is chosen to reduce the noise for both Flevoland and western-Xi'an-Area sets of data.

To show the advantages of our frame, the traditional SVM models [24], Wishart-based maximum likelihood (Wishart ML) [25] and Wishart-based Markov random fields (Wishart MRF) [16] are used for comparison.

These two experiments have the same parameters and the parameter settings are as follows:

In the experiments, SVM parameters are adapted by five-fold cross validation. The regularization parameter in the SVM is tuned in the range of $\{10^{-2}, 10^{-1}, \ldots, 10^{3}\}$. The Wishart ML method is a maximum likelihood classifier based on Wishart distribution. The Wishart MRF method is a region-based classifier. The window size is chosen as 5×5 for segmenting in Wishart MRF and SVM+MRF. In the proposed method, only one parameter μ^l needs to be set and μ^l is selected adaptively. For a reasonably fair comparison, the window size is set as 5 × 5. If the number of pixels in the same block belongs to the same class more than 20, the region of this block can be regarded as a smooth region. Otherwise, the region is considered as a non-smooth region. In smooth region, the value of μ^l is set to 5. Spatial term is the major factor for classification. On the contrary, μ^l is set to the reciprocal of the value in a smooth region, and the value of μ^l is set to 1/5 and spectral term will be dominated.

A. EXPERIMENT ON FLEVOLAND IMAGE

Fig.3. (a) and (b) show the corresponding Pauli RGB images of Flevoland and the ground truth respectively. The ground truth is manually labeled. There are 15 classes in this scene including Stem Beans, Rape Seed, Bare Soil, Potatoes, Beet, Wheat, Wheat2, Wheat3, Peas, Barely, Lucerne, Grass, Forest Water and Buildings. Three classical algorithms are used to make a comparison with our method. In our experiments, we choose 10% samples randomly for each class for training and remaining 90% for testing.

Results of these methods for comparison and the proposed method are shown Fig.3. (c)∼(g). It can be observed that, our method shows better results than the other ones. Fig.3. (d) shows the classification result with traditional SVM. As shown in the top red rectangle in Fig.3. (d), Potatoes, Forest, and Stem beans cannot be distinguished from each other very well for SVM method. In addition, Buildings cannot be recognized at all in the same figure in the bottom red rectangle. Furthermore, as a pixel-based classification method, many pixels are misclassified and the region is not smooth enough. Fig.3. (e) is the classification result of SVM+MRF. It is clear that, compared with SVM, the effect of classification is improved in many categories, such as Stem Beans, Potatoes, Peas and Water. However, the result of SVM+MRF depends on the classification effects of SVM. Therefore, many pixels that are misclassified cannot be revised. Fig.3. (f) is the classification result of Wishart ML. In the red rectangle in Fig.3. (f), the Rape Seed is wrongly classified into Wheat1 and Wheat3. Because of lacking training data in these area, the estimation of parameters is unreliable. The classification result with Wishart MRF is shown in Fig.3. (g). Compared with SVM and Wishart ML methods, the region is smoother, benefiting from the region-based classification method. However, in the top right corner, Water is recognized as Bare Soil. The classification result of Wishart MRF depends on the center label. If the center label is incorrect, the whole block will be misclassified. Fig.3. (c) exhibits the

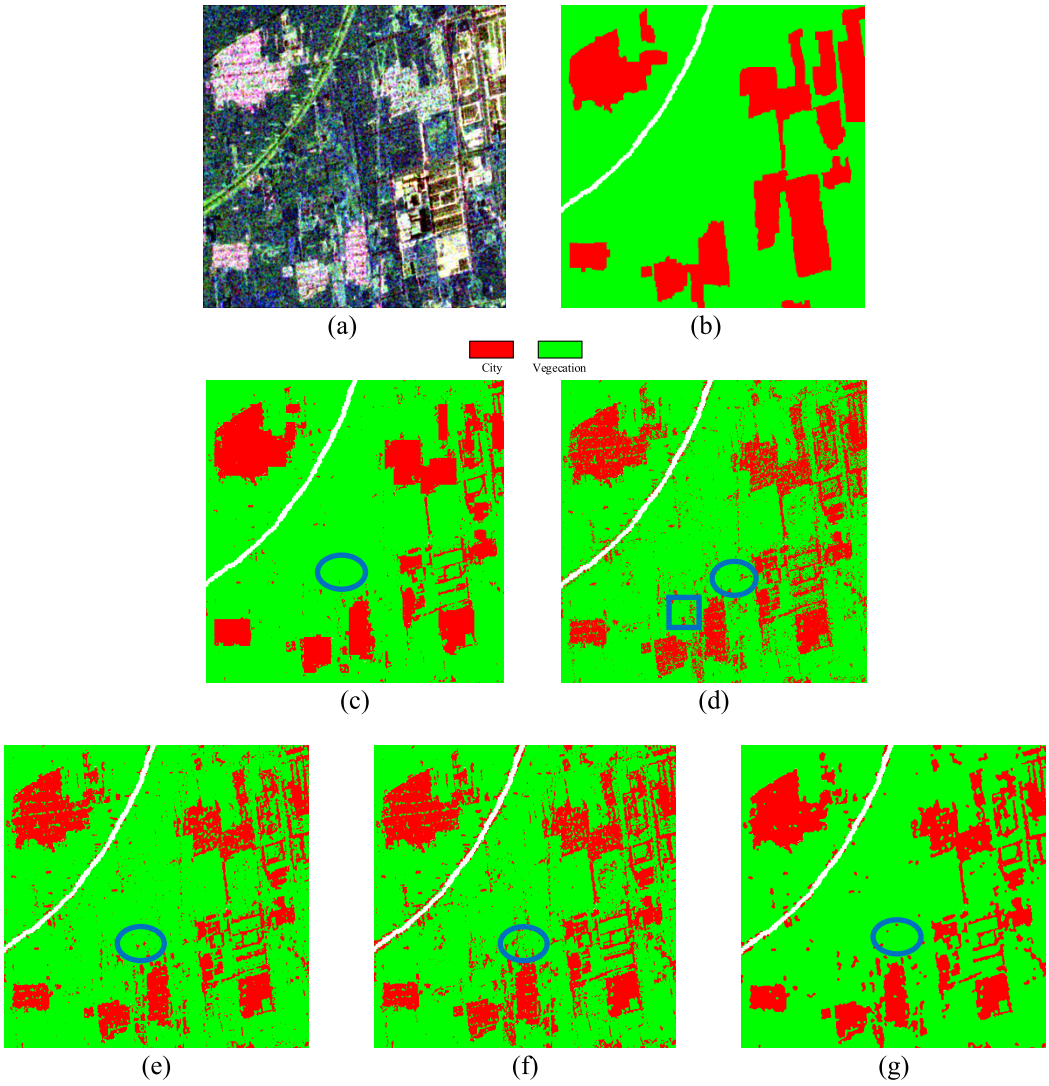

FIGURE 4. (a) the PolSAR image (PauliRGB) of Xi'an-area. (b) The ground truth Map. (c) Our method. (d) SVM method. (e) SVM+MRF method. (f) Wishart ML method. (f) Wishart-based MRF method.

result of proposed method, which is better than the other three methods. In particular, we got the smoother region and higher classification accuracy. As shown in small blue rectangle in Fig.3. (c)∼(g), the Steam Beans is classified almost by CMM and Forest is classified to Steam Beans by the other methods. In bigger blue rectangle in Fig.3. (d)∼(g), Wheat1 is recognized as Wheat1 more or less. CMM shows an exciting result.

We record the classification accuracy in TABLE 2. The overall accuracy (OA) is used to evaluate the performance of different methods. Accuracy is calculated by Eq. (11):

$$\mathrm{A} = \frac{N_t}{N_g} \tag{11}$$

where A is the accuracy, N_t is the number of pixels that are classified correctly for a certain class, N_g is the total number of pixels in this class in ground truth. For the total accuracy, N_t is the number of correct classifications for all the training samples and N_g is the total number of pixels in ground truth.

TABLE 2. Classification accuracy of different methods for Flevoland.

Method / Region	CMM	SVM	SVM+ MRF	Wishart ML	Wishart MRF
Stem Beans	**0.9921**	0.8722	0.9804	0.8867	0.9713
Rape Seed	**0.9626**	0.8521	0.9122	0.5347	0.7885
Bare Soil	0.9931	0.9922	0.9836	0.9286	0.9980
Potatoes	**0.9720**	0.6661	0.9253	0.8304	0.9311
Beet	0.9594	0.9224	0.9753	0.8515	0.9363
Wheat2	**0.9336**	0.8323	0.8852	0.8030	0.8746
Peas	**0.9918**	0.8997	0.9752	0.9409	0.9690
Wheat3	**0.9975**	0.9505	0.9707	0.8795	0.9459
Lucerne	**0.9604**	0.9277	0.9387	0.8183	0.7894
Barley	0.9716	0.9483	0.9820	0.8808	0.9623
Wheat1	**0.9879**	0.9152	0.9703	0.9411	0.9348
Grasses	**0.9538**	0.8614	0.9124	0.5968	0.7210
Forest	**0.9799**	0.9347	0.9566	0.8924	0.9190
Water	0.9837	0.9478	0.9957	0.5117	0.5360
Buildings	**0.8776**	0.1646	0.8476	0.8204	0.8095
Total	**0.9752**	0.8868	0.9530	0.8074	0.8746

We can see that, for the experiment of Flevoland, the classification accuracy of CMM is the highest compared with the other three methods for most classes and the best

TABLE 3. Classification accuracy of different methods for Xi'an-area.

Method / Region	CMM	SVM	SVM+ MRF	Wishart ML	Wishart MRF
City	0.7100	0.7153	0.6670	0.7087	0.7102
Vegetation	**0.9777**	0.9417	0.9514	0.9153	0.9584
Total	**0.9104**	0.8848	0.8824	0.8633	0.8960

results are in bold. CMM is a discriminative classification framework with spectral and spatial information for classifier learning. Therefore, the better results can be obtained compared with the other methods. Especially for the categories that CMM recognizes very well, including Stem Beans, Peas and Wheat3, with the accuracy values of 0.9921, 0.9918 and 0.9975. For the total classification accuracy, the proposed method achieved 0.9752. The total classification accuracy is enhanced by about 10%. Nevertheless, not all of the pixels have the true labels for the whole image, so the accurate calculation is based on the pixels with the true label as shown in Fig.3. (b).

B. EXPERIMENT ON XI'AN-AREA

There are two classes of Jingkun Highway in western-Xi'an-Area in the Fig.4. (a). They are City and Vegetation. The ground truth is shown in Fig.4. (b) . The ground truth is manually labeled. We also choose 10% samples randomly for training and remaining 90% samples for testing.

The results of classification are shown in Fig.4. (c)-(g). Fig.4. (d) shows the classification result of SVM., exhibiting that such a pixel-based method causes, many speckle noise in the figure. In the blue rectangle in Fig.4. (d), Vegetation is regarded as City. Fig.4. (e) shows the result of SVM+MRF, exhibiting that the region is not smooth enough compared with the result of CMM. Fig.4. (f) is the result of of Wishart ML, in which more isolated regions exist and the region is not sufficiently smooth. Fig.4. (g) shows a smooth region from the result of WMRF. This method is liable to misclassify the pixels, because it exploits only spatial information and however the spectral information is not taken into account. The result of our proposed method exhibits much better classification result of Vegetation that the other methods on the visual effects, as shown in Fig.4. (c). Especially in the middle of the figure, as shown in the ellipse region, Vegetation cannot distinguish from City well by contrast. Our proposed method achieves obvious improvement in Vegetation.

The classification accuracy of these four methods is listed in Table 3. The accuracy is calculated by Eq. (11). For the classification of Vegetation, the accuracy of CMM can reach 0.9777 and the total accuracy 0.9104 is highest. The experimental results show that the CMM is able to achieve excellent classification accuracy and outstanding visual effect.

IV. CONCLUSION

This work proposed a new classification method applied to the PolSAR data. The method is a novel discriminative model which is used to enhance the classification accuracy of PolSAR image. On the basis of Max-Margin frame, the characteristic function is redesigned and conditional random fields is employed to propagate the contextual information in both labeling field and observation field. Spectral term and spatial term are two important parts of the model. SVM is used as spectral term which provides an effective method of learning a maximum-margin decision boundary and cover the speckle-like classification under the limited training samples. CRF is applied as the spatial term which is used to describe the spatial information of PolSAR and smoothen the region during the classification. In the spatial term, Wishart distance is used to describe the statistical property of PolSAR image. At the same time, the pixel and label information is combined to smoothen the region. The experimental results show that our method offers potential application to the PolSAR classification and we can get clearer classification maps compared with the other three classification methods.

REFERENCES

[1] F. T. Ulaby and C. Elachi, *Radar Polarimetry for Geoscience Applications*, vol. 5. Norwood, MA, USA: Artech House, 1990, p. 38.
[2] E. Attema *et al.*, "ENVISAT ASAR science and applications," *ESA Publ. SP*, vol. 1225, no. 1, p. 59, Nov. 1998.
[3] P. Zhang, M. Li, Y. Wu, L. An, and L. Jia, "Unsupervised SAR image segmentation using high-order conditional random fields model based on product-of-experts," *Pattern Recognit. Lett.*, vol. 78, pp. 48–55, Jul. 2016.
[4] C. Sun, L. Jiao, H. Liu, and S. Yang, "New classifier based on compressed dictionary and LS-SVM," *Neurocomputing*, vol. 216, pp. 617–626, Dec. 2016.
[5] B. Hou, C. Chen, X. Liu, and L. Jiao, "Multilevel distribution coding model-based dictionary learning for PolSAR image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 11, pp. 5262–5280, Nov. 2015.
[6] W. Xie *et al.*, "POLSAR image classification via Wishart-AE model or Wishart-CAE model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3604–3615, Aug. 2017.
[7] H. Bi, J. Sun, and Z. Xu, "Unsupervised PolSAR image classification using discriminative clustering," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 6, pp. 3531–3544, Jun. 2017.
[8] F. Liu, L. Jiao, B. Hou, and S. Yang, "POL-SAR image classification based on Wishart DBN and local spatial information," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3292–3308, Jun. 2016.
[9] X. Niu and Y. Ban, "A novel contextual classification algorithm for multitemporal polarimetric SAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 3, pp. 681–685, Mar. 2014.
[10] B. Liu, H. Hu, H. Wang, K. Wang, X. Liu, and W. Yu, "Superpixel-based classification with an adaptive number of classes for polarimetric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 907–924, Feb. 2013.
[11] B. Hou, H. Kou, and L. Jiao, "Classification of polarimetric SAR images using multilayer autoencoders and superpixels," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 7, pp. 3072–3081, Jul. 2017.
[12] W. Song, M. Li, P. Zhang, Y. Wu, L. Jia, and L. An, "Unsupervised PolSAR image classification and segmentation using Dirichlet process mixture model and Markov random fields with similarity measure," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3556–3568, Aug. 2017.
[13] Z. Wu and Q. Ouyang, "SVM- and MRF-based method for contextual classification of polarimetric SAR images," in *Proc. Int. Conf. Remote Sens., Environ. Transp. Eng.*, Jun. 2011, pp. 818–821.
[14] G. Liu, M. Li, Y. Wu, P. Zhang, L. Jia, and H. Liu, "PolSAR image classification based on wishart TMF with specific auxiliary field," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 7, pp. 1230–1234, Jul. 2014.
[15] S. Yang, W. Lang, J. Wu, and X. Yang, "Combining MRF and *v*-SVM for SAR sea ice image classification," *J. Remote Sens.*, vol. 19, no. 5, pp. 844–855, 2015.

[16] 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.
[17] A. Masjedi, M. J. V. Zoej, and Y. Maghsoudi, "Classification of polarimetric SAR images based on modeling contextual information and using texture features," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 932–943, Feb. 2016.
[18] L. Zhang, W. Ma, and D. Zhang, "Stacked sparse autoencoder in PolSAR data classification using local spatial information," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 9, pp. 1359–1363, Sep. 2016.
[19] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, "SVM- and MRF-based method for accurate classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 736–740, Oct. 2010.
[20] B. Tso and R. C. Olsen, "A contextual classification scheme based on MRF model with improved parameter estimation and multiscale fuzzy line process," *Remote Sens. Environ.*, vol. 97, no. 1, pp. 127–136, 2005.
[21] A. Sarkar, M. K. Biswas, B. Kartikeyan, V. Kumar, K. L. Majumder, and D. K. Pal, "A MRF model-based segmentation approach to classification for multispectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 5, pp. 1102–1113, May 2002.
[22] Y. Zhong, X. Lin, and L. Zhang, "A support vector conditional random fields classifier with a mahalanobis distance boundary constraint for high spatial resolution remote sensing imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1314–1330, Apr. 2014.
[23] B. Taskar, C. Guestrin, and D. Koller, "Max-margin Markov networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2003, pp. 25–32.
[24] C. Lardeux *et al.*, "Support vector machine for multifrequency SAR polarimetric data classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 12, pp. 4143–4152, Dec. 2009.
[25] J. S. Lee, M. R. Grunes, and R. Kwok, "Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution," *Int. J. Remote Sens.*, vol. 15, no. 11, pp. 2299–2311, 1994.

SHUYIN ZHANG was born in Weinan, China, in 1982. She received the B.S. degree in measurement and control technology and instrumentation program and the M.S. degree in microelectronics and solid state electronics from Liaoning University, Shenyang, China, in 2004 and 2008, respectively. Since then, she has been taking doctoral programs with the Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University, Xi'an, China.

Her research interests include machine learning and polarimetric synthetic aperture radar image classification.

BIAO HOU (M'07) was born in Baoji, China, in 1974. He received the B.S. and M.S. degrees in mathematics from Northwest University, Xi'an, China, in 1996 and 1999, respectively, and the Ph.D. degree in circuits and systems from Xidian University, Xi'an, in 2003.

Since 2003, he has been with the Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University, where he is currently a Professor. His research interests include dictionary learning, deep learning, compressive sensing, and synthetic aperture radar image processing.

LICHENG JIAO (SM'89) received the B.S. degree in electric engineering from Shanghai Jiaotong University, Shanghai, China, in 1982, and the M.S. and Ph.D. degrees in signal, circuit and system from Xi'an Jiaotong University, Xi'an, China, in 1984 and 1990, respectively.

Since 1992, he has been a Professor with the School of Electronic Engineering, Xidian University, Xi'an, where he is currently the Director of the Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education of China with the International Research Center of Intelligent Perception and Computation. His current research interests include intelligent information processing, image processing, machine learning, and pattern recognition.

Dr. Jiao is a member of the IEEE Xi'an Section Execution Committee, the President of the Computational Intelligence Chapter, the IEEE Xi'an Section, and IET Xi'an Network, the Chairman of the Awards and Recognition Committee, the Vice Board Chairperson of the Chinese Association of Artificial Intelligence, a Councilor of the Chinese Institute of Electronics, a Committee Member of the Chinese Committee of Neural Networks, and an Expert of the Academic Degrees Committee of the State Council.

QIAN WU received the B.S. degree from Xidian University, Xi'an, China, in 2013, where she is currently pursuing the Ph.D. degree with the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China. Her current research interests include Bayesian learning, polarimetric synthetic aperture radar image classification, and deep learning.

CHEN SUN received the B.S. degree from Hebei University, Hebei, China, in 2011. She is currently pursuing the M.S. and Ph.D. degrees in circuit and system from the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, Xidian University, Xi'an, China.

Her current research interests include pattern recognition, machine learning, and compressed sensing.

WEN XIE was born in Weinan, China, in 1989. She received the B.S. degree in automation from Xidian University, Xi'an, China, in 2011. Since then, she has been taking successive post-graduate and doctoral programs with the Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University.

Her research interests include sparse representation, deep learning, and polarimetric synthetic aperture radar image classification.