Joint Polarimetric-Adjacent Features Based on LCSR for PolSAR Image Classification

Xiao Wang, *Student Member, IEEE*, Lamei Zhang, *Senior Member, IEEE*, Ning Wang, and Bin Zou, *Senior Member, IEEE*

***Abstract*—Image classification is a critical and important application in PolSAR image interpretation. Finding a feature extraction method, which can effectively describe the characteristics of the target, is an important basis for image classification. In addition to unique polarimetric features of PolSAR system, spatial adjacent features of image also need to be considered. So in this article, a joint polarimetric-adjacent features extraction method based on local convolution sparse representation is proposed for PolSAR image classification. Firstly, this article uses convolutional sparse representation to achieve the convolution of the image filters and the feature responses so as to achieve the effective combination of the polarimetric and adjacent information of the image. Meanwhile, construct and train the dictionary using local strategy in the original domain to avoid the high computational complexity and the confusion of different grounds caused by global dictionary. Finally, support vector machine (SVM) is used to combine the extracted features to achieve the classification. Three sets of full polarimetric data are used and the experiment results prove that the proposed method can effectively combine the polarimetric and adjacent information of data and have a good performance in PolSAR image classification.**

***Index Terms*—Local convolution sparse representation (LCSR), polarimetric-adjacent features, polarimetric SAR, target decomposition.**

I. Introduction

POLSAR system can obtain rich and comprehensive information of grounds through combinations of multiple polarimetric modes. How to effectively utilize the advantages of PolSAR image for classification has always been a practical and key branch in the application of PolSAR image. Currently, the classification methods of PolSAR image can be divided into three main categories: statistical model-based, polarimetric target decomposition (PTD)-based, and machine learning-based methods. Statistical model-based methods take advantage of the unique statistical distribution of PolSAR image for classification. This kind of methods focuses on the classification according to the mathematical statistical characteristics instead of the polarimetric scattering characteristics of the data itself. Representative methods are [1]–[3]. PTD, a widely used feature extraction method, is able to decompose PolSAR images into different scattering components that can intrinsically characterize the scattering properties of different grounds. PolSAR image classification can be achieved by directly or indirectly combining these different scattering features with classifiers, such as given in [4]–[6]. The methods based on machine learning introduce some mature classification methods in other fields for PolSAR image, and improve the methods to be applicable to PolSAR image, such as [7], [8].

PTD-based method is a widely used classification method, which is closely related to scattering mechanism. The classification results of PTD-based methods are highly dependent on the quality of the PTD. Currently, PTD methods can be divided into coherent target decomposition [9], [10] and incoherent target decomposition [11]–[19]. Among them, the latter is widely applied due to its significant advantages for identifying artificial targets, forest, and surface areas. In [11], a model-based target decomposition method was proposed, which treats the covariance matrix or coherence matrix as the superposition of volume scattering, surface scattering, and double bounce scattering model. These models are mathematical modeling of the basic scattering mechanism, without utilizing real ground measurements. After that, a large number of modified model-based decomposition methods were proposed, such as Yamaguchi *et al.* [12], [13], Zhang *et al.* [14], Van Zyl *et al.* [15], An *et al.* [16], Chen *et al.* [17], Bhattacharya *et al.* [18], and Singh *et al.* [19]. However, it is difficult to obtain better results based on polarimetric features alone. In addition to polarimetric information in PolSAR image, spatial structure information as an important feature should also be considered in classification. Commonly, numerous spatial feature extraction methods were proposed [20]–[25]. Generally, polarimetric features and spatial features are extracted independently and separately, which does not combine the polarimetric and spatial information of adjacent pixels in PolSAR image simultaneously.

Sparse representation (SR) [26] as an effective tool can be utilized to achieve the extraction of those features. It is a theory that uses a linear combination of a few basic signals as possible to express most or all of the original signals. These basic signals are called "atoms," which are from a redundant pool called "dictionary." SR can obtain a more concise representation of the signal, which makes it convenient to obtain the information

Manuscript received January 18, 2021; revised April 8, 2021 and April 29, 2021; accepted June 2, 2021. Date of publication June 8, 2021; date of current version June 30, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61871158, and in part by the Aeronautical Science Foundation of China under Grant 20182077008. *(Corresponding authors: Lamei Zhang; Bin Zou.)*

Xiao Wang, Lamei Zhang, and Bin Zou are with the Department of Information Engineering, Harbin Institute of Technology, Harbin 150001, China (e-mail: xiaowanghit@163.com; zzbei@hit.edu.cn; zoubin@hit.edu.cn).

Ning Wang is with the Beijing Institute of Radio Measurement, Beijing 100000, China (e-mail: wangningtop@163.com).

Digital Object Identifier 10.1109/JSTARS.2021.3087164

contained in the signal and facilitate further processing. So in the past years, SR models have brought a wealth of scientific research in image processing, such as image denoising [27], [28], inpainting [29], fusion [30], [31], and recognition [32].

However, when using SR in image, considering the complexity of calculation, most images are partitioned into blocks and coded separately [33], [34], which results in the potential loss of image structure information. In addition, due to the correlation of image blocks, redundant dictionaries learned through image blocks often contain repeated information. To improve this, convolution sparse representation (CSR) is proposed [35]. In this method, the image is represented globally as the approximation of the convolution sum of a set of filter dictionaries and sparse feature responses, which reduces the coding redundancy of image blocks and the loss of information in image segmentation. At present, the theory of CSR is widely used in image processing. Gu *et al.* [36] applied CSR to the superpixel processing, which preserved the texture and edge information of images well. Liu *et al.* [37] applied CSR to image fusion, proving that the CSR-based fusion method was obviously superior to SR-based method in the results of the evaluation and the visual effects. Chen *et al.* [38] applied CSR to the classification of handwritten fonts. The results demonstrated that dictionaries trained by convolution can obtain more representative image information and achieve better classification performance. Kavukcuoglu *et al.* [39] applied CSR to image detection to improve the detection effect.

Meanwhile, the solution procedure of CSR is complex, and a large number of algorithms have been proposed to solve CSR model [40], [41]. The most representative and widely used method is an efficient convolutional sparse coding method proposed by Wohlberg [42], [43], which transforms the image into the Fourier domain to reduce the calculation cost, and alternating direction method of multipliers (ADMM) [43], [44] algorithm is used to solve CSR. This method has great advantages, but introduces more variables and improves computational complexity and memory requirements.

In order to effectively combine the polarimetric and the spatial information to describe target features of the PolSAR system, considering the advantages of CSR, this article proposes a joint polarimetric-adjacent features extraction method based on local convolution sparse representation (LCSR) for PolSAR image classification. In this method, CSR is used to effectively combine the polarimetric information of PolSAR system and the spatial and adjacent information of the image. Considering different grounds, the PolSAR image has different scattering mechanisms, global convolution does not effectively utilize the local spatial information of the PolSAR image itself to achieve extraction of polarimetric-adjacent features, and local operations are performed on the image block, i.e., only local calculation rather than global calculation is used to train filters to extract features. Meanwhile, this approach avoids the problems of the ADMM method and the optimization process is in the original domain instead of mapping to frequency domain, which is more simple and easy to implement. Finally, SVM is used in this article for PolSAR image classification.

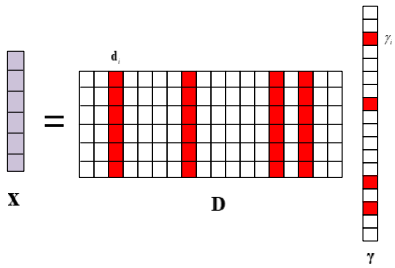

Fig. 1 Schematic diagram of SR.

The organization is as follows. In Section II, a brief introduction to CSR is introduced, Section III presents the detailed descriptions of polarimetric-adjacent features extraction based on LCSR. In Section IV, the procedure of PolSAR image classification via LCSR-based feature extraction and SVM classifier is given; the experiment results and discussion of the proposed algorithm are presented. Finally, Section V concludes the article.

II. Fundermental of Convolution Sparse Representation

SR is widely used in the field of signal and image processing. Subsequently, CSR appears on the basis of SR, which extends the one-dimensional product operation of image vector and dictionary atom to the convolution operation of the two-dimensional image and convolutional dictionary, overcoming the shortcoming of breaking the structural relationship by drawing the image into vectors.

A. Sparse Representation Theory

The signal is assumed that can be expressed by some basic signals, which is called "atoms" from a redundant pool called "dictionary." SR is a linear representation of the signal with as few atoms as possible in an overcomplete dictionary. Suppose the signal is $\mathbf{x} \in R^m$, $\mathbf{d}_i \in R^m$ is an m-dimensional feature vector, called atom, $\mathbf{D} \in R^{m \times K} (K \geq m)$ is called dictionary composed by K atoms, then $\mathbf{x}$ can be sparsely represented by

$$\mathbf{x} = \mathbf{D}\gamma = \sum_{i=1}^{K} \gamma_i \mathbf{d}_i \tag{1}$$

where $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \ldots, \gamma_K]^{\mathrm{T}}$ is sparse coefficient of $\mathbf{d}_i$. Due to the overcompleteness of the dictionary, the signals can be represented with countless solutions. The solution of SR is to find the most sparse solution, i.e., the solution with the least nonzero terms in $\boldsymbol{\gamma}$.

The schematic diagram of SR is shown in Fig. 1, where $\mathbf{x}$ represents the original signal, and γ_i represents the coefficient of each basis vector $\mathbf{d}_i$ in the expression of the original signal $\mathbf{x}$. The small red boxes represent basis vectors with non-zero coefficients, and others are zero.

However, due to the influence of noise, the signal is not in an ideal state in practical problems. Correspondingly, the SR model should be adjusted to fit the situation after adding noise. Then, in the case of the noise, using the Lagrange multiplier method, and under the premise that the λ value is properly selected, SR

model of the signal with noisy interference can be equivalently expressed by the unconstrained minimization as

$$\hat{\gamma} = \arg\min \|\mathbf{x} - \mathbf{D}\gamma\|_2^2 + \lambda\|\gamma\|_1 \tag{2}$$

where $\| * \|_1$ is l_1 norm. This is also a relatively simple SR model of the signal.

Solving SR mainly includes two parts: dictionary construction and sparse solving algorithm. Currently, there are two main methods for constructing the filter dictionary $\mathbf{D}$. One is the analysis dictionary, such as Fourier transform dictionary, discrete cosine transform (DCT) dictionary, and wavelet transform dictionary. These dictionaries are fixed. The other is the learning dictionary, such as the optimal direction method (MOD) [45], KSVD [46], and online dictionary learning algorithms [47]. These dictionaries use machine learning to construct the dictionary adaptively from a series of training samples. These methods can maximally match the dictionary according to the structural features of the original image. While the representative sparse solving algorithms are MP [48], OMP [49], GPSR [50], LASSO [51], LARS [52], etc.

B. Convolution Sparse Representation

In order to reduce the modeling and computational burden, the image blocks are individually encoded generally. Also, the traditional SR method only encodes one-dimensional signals separately, which ignores the spatial structure information of the two-dimensional image blocks. In remote sensing images (also natural optical images), neighboring pixels often belong to the same category. The individual pixels and neighborhoods generally have the same or similar physics attribute. In order to combine the adjacent information of the image effectively, convolution is introduced.

CSR is the convolutional form of SR. In other words, the convolution sum of the filter dictionaries and the feature responses is used to replace the product of the redundant dictionary and the sparse coefficient, so that the image can be sparsely encoded in the "whole" unit. The CSR model is defined as [42]

$$\arg\min_{\gamma_m} \left\| \mathbf{x} - \sum_{m=1}^{M} \mathbf{D}_m \otimes \gamma_m \right\|_2^2 + \lambda \sum_{m=1}^{M} \|\gamma_m\|_1 s.t. \|\mathbf{D}_m\|_2 = 1 \tag{3}$$

where "$\otimes$" is the convolution operation; $\mathbf{x}$ is the input image, $\{\mathbf{D}_m\}$ is a set of dictionary filters consisting of M basis functions, the size is $N \times N$; $\{\gamma_m\}$ is the corresponding sparse feature responses; λ is the regular term proportion; and $\| * \|_2^2$is the square of l_2 norm.

The image $\mathbf{x}$ is considered as the approximation of the convolution sum of a set of sparse feature responses $\{\gamma_m\}$ and filter dictionaries $\{\mathbf{D}_m\}$. Like the traditional SR, the sparse basis $\{\mathbf{D}_m\}$ obtained by the CSR also has a fixed space support. And unlike SR algorithm, the feature response $\{\gamma_m\}$ obtained by CSR is very different from the original image $\mathbf{x}$ and has higher dimension than the original image $\mathbf{x}$.

The solving of CSR is an iterative optimization process, in which the feature responses $\{\gamma_m\}$ and dictionary filters $\{\mathbf{D}_m\}$ need to be solved. The solution procedure is generally optimizing the two variables alternately until the objective function, as (3) described, is minimum through continuous iteration. Generally, ADMM is the most commonly used method to solve the optimization problem, and the optimization process is mapped to the frequency domain by fast fourier transform (FFT) to reduce the computation complexity.

III. Polarimetric-Adjacent Feature Extraction Based on LCSR

A. PolSAR Image Representation

PolSAR utilizes the reflected echoes of microwave to imaging. Due to the different structures and electromagnetic properties, the grounds exhibit different characteristics on PolSAR image, including grayscale, texture, polarimetric characteristics, and so on. The scattering matrix describing the scattering information of PolSAR image is expressed as

$$\mathbf{S} = \begin{pmatrix} S_{\mathrm{hh}} & S_{\mathrm{hv}} \\ S_{\mathrm{vh}} & S_{\mathrm{vv}} \end{pmatrix} \tag{4}$$

where h refers to the horizontal polarization, v refers to the vertical polarization. S_{hh} and S_{vv} are the copolarization components, S_{hv} and S_{vh} are the cross-polarization components. In the monostatic backscattering case, S is symmetric matrix, i.e., $S_{\mathrm{hv}} = S_{\mathrm{vh}}$. Pauli basis can be used to decompose S; the decomposed three-dimensional target vector $\mathbf{k}$ is obtained as

$$\mathbf{k} = \frac{1}{\sqrt{2}} \left[S_{\mathrm{hh}}, \sqrt{2} S_{\mathrm{hv}}, S_{\mathrm{vv}} \right]^{\mathrm{T}}. \tag{5}$$

Then, the covariance matrix of PolSAR is defined as

$$\mathbf{C}_3 = \left\langle \mathbf{k} \cdot \mathbf{k}^{\mathrm{H}} \right\rangle = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \tag{6}$$

where $\langle \cdot \rangle$ represents ergodic averaging, and the superscript H represents complex conjugate. The covariance matrix is a common data representation in PolSAR processing.

According to the covariance matrix, a variety of classical model-based PTD methods are proposed, which try to decompose the covariance matrix into the combination of different physical scattering mechanisms. Each scattering mechanism corresponds to a certain physical meaning. It can be expressed as

$$\mathbf{C} = \sum_{i=1}^{k} \mathbf{C}_i P_i. \tag{7}$$

In (7), $\mathbf{C}$ indicates the covariance matrix of PolSAR image, $\mathbf{C}_i$ is the different scattering model based on different scattering mechanisms, and P_i represents the power of scattering component $\mathbf{C}_i$. The model-based PTD methods have clear physical interpretation. The polarimetric information of the target echo can reflect the geometric and physical characteristics of the target. Using the scattering power to describe the target can effectively reduce the dimensionality of the data space, remove the redundant information, obtain the difference of different

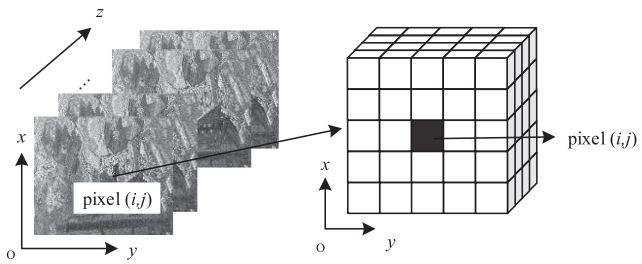

Fig. 2. Polarimetric-adjacent description of PolSAR image.

types of grounds in the image quickly and accurately, and reduce the complex background interference.

The traditional model-based PTD methods are dealt with pixel by pixel, which means that the information between each pixel in these methods is independent with each other. These methods do not take into account the spatial information of the image. As shown in Fig. 2, PolSAR image has three-dimensional information, where x and y are the rows and columns of the image, and z is the polarimetric information. So studying the joint polarimetric-adjacent PTD method provides a new idea for the feature extraction and the subsequent PolSAR image classification.

B. Polarimetic-Adjanct Feature Extraction Based on LCSR

Equation (7) is dealt with on each pixel of PolSAR image; in other words, P_i represents the decomposition result of a single pixel in the image. The decomposition result of each pixel P_i in the image constitutes the result of the whole PolSAR image. To distinguish the result between single pixel and two-dimensional image, the image form of decomposition result is written as $\mathbf{P}_i$ in this article. In order to realize the joint polarimetric-adjacent features extraction of PolSAR image, the adjacent information of the image also needs to be considered. In this article, the decomposition result image $\mathbf{P}_i$ can be represented by CSR model, i.e., (7) can be expressed as

$$\mathbf{C} = \sum_{i=1}^{k} \mathbf{C}_i \sum_{m=1}^{M} \mathbf{D}_{im} \otimes \boldsymbol{\gamma}_{im} \tag{8}$$

where $\{\mathbf{D}_{im}\}$ and $\{\boldsymbol{\gamma}_{im}\}$ are a set of filter dictionaries and feature responses corresponding to $\mathbf{P}_i$, respectively. Then, the objective function of (8) can be written as

$$\arg\min_{\mathbf{D}_{im},\gamma_{im}} \sum_{i=1}^{k} \mathbf{C}_i \left\| \mathbf{P}_i - \sum_{m=1}^{M} \mathbf{D}_{im} \otimes \gamma_{im} \right\|_2^2 + \sum_{i=1}^{k} \lambda_i \mathbf{C}_i \sum_{m=1}^{M} \|\gamma_{im}\|_1 \tag{9}$$

where λ_i is the regular term proportion of $\mathbf{P}_i$.

In (9), there are three unknowns need to be solved: $\{\mathbf{C}_i\}$, $\{\mathbf{D}_{im}\}$, and $\{\boldsymbol{\gamma}_{im}\}$. According to the theoretical basis of target decomposition of PolSAR image in the previous section, the classical target decomposition model can be adopted to $\{\mathbf{C}_i\}$. Then, $\{\mathbf{C}_i\}$ are fixed, and the feature responses $\{\boldsymbol{\gamma}_{im}\}$and the dictionary filters $\{\mathbf{D}_{im}\}$ can be solved by alternate optimization as mentioned in Section II.

The feature responses $\{\boldsymbol{\gamma}_{im}\}$ are high dimensional, and the solution in high dimension is difficult. Generally, the operations take place in the frequency domain to map the convolution to a product to reduce the dimension, and ADMM method is used to solve the optimization problem. ADMM is a traditional solution method for convex optimization problems, which has a fast convergence speed and low complexity for solving the feature responses in the above objective function. However, ADMM has some problems, as it requires the introduction of many intermediate and auxiliary variables, which increases its complexity and storage loss. In addition, these variables greatly affect the CSR results resulting in a complicated parameter selection process.

Furthermore, because different grounds in the PolSAR image have different scattering mechanisms, the global convolution does not portray different features to a sufficient extent, making it difficult to obtain satisfactory representation results and reduce the separability of different grounds. For these problems, considering that the method proposed by [53], [54] can solve CSR problem in the original domain instead of mapping to the frequency domain without introducing more variables, and leverage the localized strategy to operate on image locally, this article introduces the method of LCSR to solve the optimization problem.

C. Local Feature Responses $\{\boldsymbol{\gamma}_{im}\}$ Solution

As [53], [54] referred, CSR can simplify the convolution operation to the product operation to avoid mapping the solution procedure to the frequency domain. Then, $\mathbf{P}_i$ can be decomposed as matrix form

$$\mathbf{P}_i = \mathbf{D}_i \gamma_i \tag{10}$$

where $\mathbf{D}_i \in R^{N \times NM}$ is the concatenation of multiple banded matrices transformed by the dictionary filters $\{\mathbf{D}_{im}\}$; and the feature response $\boldsymbol{\gamma}_i \in R^{NM}$ is the alternant arrangement of all the feature responses $\{\boldsymbol{\gamma}_{im}\}$. In this way, the operations in high-dimensional can be transferred to the operations in two-dimensional to reduce the complexity.

In order to describe the adjacent information of different grounds accurately, the global feature response $\boldsymbol{\gamma}_i$ is divided into the sum of N local feature responses $\boldsymbol{\gamma}_{ij} \in R^M$ by local strategy. Then, the dictionary filter $\mathbf{D}_i$ can be decomposed into the sum of multiple subdictionary filters $\mathbf{D}_{ij} \in R^{N \times M}$ corresponding to the local feature responses $\boldsymbol{\gamma}_{ij}$. Equation (9) can be expressed as

$$\arg\min_{\gamma_{ij}} \sum_{i=1}^{k} \mathbf{C}_i \left\| \mathbf{P}_i - \sum_{j=1}^{N} \mathbf{D}_{ij} \gamma_{ij} \right\|_2^2 + \sum_{i=1}^{k} \lambda_i \mathbf{C}_i \sum_{j=1}^{N} \|\gamma_{ij}\|_1 \tag{11}$$

where $\sum_{m=1}^{M} \mathbf{D}_{im}$ and $\sum_{m=1}^{M} \boldsymbol{\gamma}_{im}$ are decomposed into N local dictionaries $\{\mathbf{D}_{ij}\}$and feature responses $\{\boldsymbol{\gamma}_{ij}\}$, respectively. $\mathbf{P}_{\mathbf{r}ij}$ is defined as the component without the contribution of

jth feature response $\boldsymbol{\gamma}_{ij}$.

$$\mathbf{P}_{\mathbf{r}ij} = \mathbf{P}_i - \sum_{\substack{l=1 \\ l\neq j}}^{N} \mathbf{D}_{il}\gamma_{il}. \tag{12}$$

Then, the solution for each slice can be expressed as

$$\arg\min_{\gamma_{ij}} \sum_{i=1}^{k} \mathbf{C}_i \left\| \left(\mathbf{P}_i - \sum_{\substack{l=1 \\ l\neq j}}^{N} \mathbf{D}_{il}\gamma_{il} \right) - \mathbf{D}_{ij}\gamma_{ij} \right\|_2^2 + \sum_{i=1}^{k} \lambda_i \mathbf{C}_i \sum_{j=1}^{N} \|\gamma_{ij}\|_1. \tag{13}$$

Finally, (11) can be written as

$$\arg\min_{\gamma_{ij}} \sum_{i=1}^{k} \mathbf{C}_i \|\mathbf{P}_{\mathbf{r}ij} - \mathbf{D}_{ij}\gamma_{ij}\|_2^2 + \sum_{i=1}^{k} \lambda_i \mathbf{C}_i \|\gamma_{ij}\|_1. \tag{14}$$

Therefore, each local feature response can be regarded as a slice of the global feature response and each slice is independent and can be optimized separately instead of optimizing the global feature response. The solution for the global feature response becomes the solution for each slice of local feature responses. When the dictionary filters are known, the solution of the feature responses is the same as the traditional SR methods. To solve this problem and improve the computational efficiency, least angle regression (LARS) [52] is used to solve the optimization problem, which is simple and can update efficiently in parallel.

D. Dictionary Filter $\{\mathbf{D}_{im}\}$ Update

When $\boldsymbol{\gamma}_{ij}$ is fixed, the objective function in (11) can be expressed as

$$\arg\min_{\mathbf{D}_{ij}} \sum_{i=1}^{k} \mathbf{C}_i \left\| \mathbf{P}_i - \sum_{j=1}^{N} \mathbf{D}_{ij}\gamma_{ij} \right\|_2^2, s.t. \|\mathbf{D}_{ij}\|_2 = 1. \tag{15}$$

Let $vec(\mathbf{D}_{ij} \otimes \boldsymbol{\gamma}_{ij}) = \boldsymbol{\gamma}_i \mathbf{d}_i$, where $\mathbf{d}_i$ is the vectorization of $\{\mathbf{D}_{ij}\}_{j=1,\ldots,N}$, $\boldsymbol{\gamma}_i = [\boldsymbol{\gamma}_{i1}, \ldots, \boldsymbol{\gamma}_{ij}, \ldots \boldsymbol{\gamma}_{iN}]$. Then, (15) is expressed as

$$\arg\min_{\mathbf{d}_i} \sum_{i=1}^{k} \mathbf{C}_i \left\| \mathbf{P}_i - \sum_{j=1}^{N} \gamma_i \mathbf{d}_i \right\|_2^2 \tag{16}$$

When updating the dictionary, a proximal gradient descent method is used to ensure that the total loss of the dictionary filters is minimized, where $\mathbf{d}_i$ can be updated as

$$\begin{aligned} \mathbf{d}_i^{t+0.5} &= \mathbf{d}_i^t - \mu \nabla \mathbf{d}_i^t \\ \mathbf{d}_i^{t+1} &= \Pr ox_{\mu\|\cdot\|=1}\left(\mathbf{d}_i^{t+0.5}\right) \end{aligned} \tag{17}$$

where $\Pr ox_{\mu\|\cdot\|=1}(\cdot)$is the unit ball projection operator, μ is called learning rate (or step size) of the proximal gradient descent, and t is the iterations. $\nabla \mathbf{d}_i$ in (17) is solved as

$$\begin{aligned} \nabla \mathbf{d}_i &= -\mathbf{C}_i \left(\sum_{j=1}^{N} \gamma_i \right)^T \left(\mathbf{P}_i - \sum_{j=1}^{N} \gamma_i \mathbf{d}_i \right) \\ &= -\sum_{j=1}^{N} \mathbf{C}_i \gamma_i^T \left(\mathbf{P}_i - \sum_{j=1}^{N} \gamma_i \mathbf{d}_i \right) \\ &= -\sum_{j=1}^{N} \mathbf{C}_i \gamma_i^T \left(\mathbf{P}_i - \hat{\mathbf{P}}_i \right) \end{aligned} \tag{18}$$

where $\hat{\mathbf{P}}_i$ is the estimate of $\mathbf{P}_i$.

By alternately constructing the dictionary filters and solving the feature responses, the features combining polarimetric and adjacent information of PolSAR image can be obtained. The schematic diagram of solution procedure is shown in Fig. 3.

IV. PolSAR Image Classification Based on Polarimetric-Adjacent Features

In this article, the classical SVM classifier is used to combine the extracted polarimetric-adjacent features as described in Section III for PolSAR image classification. The schematic diagram of classification procedure is shown in Fig. 4. The procedure of the proposed method is presented in Algorithm 1.

V. Experiments and Discussion

Three real PolSAR data are used to validate the proposed method qualitatively and quantitatively. Half Moon Bay and Salvador area datasets can be available from https://uavsar.jpl.nasa.gov. Oberpfaffenhofen area datasets can be available from http://envisat.esa.int/polsarpro/datasets.html. To illustrate the effectiveness of the proposed method further, six sets of comparative experiments are also conducted in this section.

A. Results of Half Moon Bay Area

The first data used in the experiment is the complex covariance format of L-band PolSAR image of Half Moon Bay (37.515°N–122.497°W), Haward area, America, which was acquired by UAVSAR system in November 12, 2014. The nominal one-look spatial resolution is $1.6\,\mathrm{m} \times 0.6\,\mathrm{m}$. The image is 3377 pixels in row and 5501 pixels in line. The amplitude image of HH channel is shown in Fig. 5(a). To present the abstract of different classes' distribution, the optical image (not the same time) from Google Earth is shown in Fig. 6, and the image with modified aspect ratio is shown in Fig. 7(a) to make easy to compare with PolSAR image. It is an area with about six classes, including forest, bare land, urban, farmland, road, and ocean. Due to the complexity of the grounds in the area, the groundtruth is difficult to obtain, so only part labels are marked, which is shown in Fig. 5(b). Each category selects 400 pixels as training samples. The classical surface scattering, double bounce scattering, volume scattering, and helix scattering model in Yamaguchi decomposition model [12] are used in the experiments as the scattering model $\{\mathbf{C}_i\}$ in this article. The experiments set $N = 4$, $M = 16$, $\lambda = 0.01$ and the iterations set 10 for polarimetric-adjacent features extraction, then SVM classifier is used with the proposed features for classification.

The classification result is shown in Fig. 7(h); the result shows that the proposed method can effectively realize PolSAR image

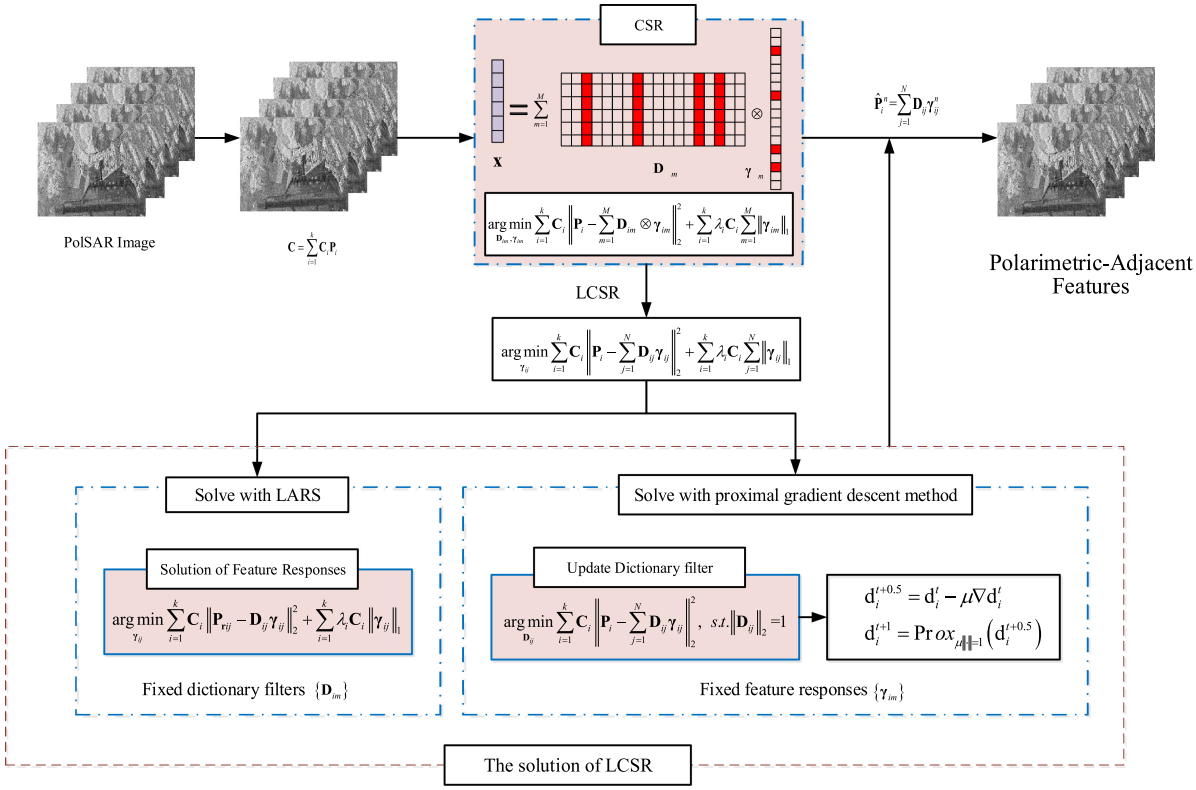

Fig. 3. Schematic diagram of the solution procedure for polarimetric-adjacent features extraction.

Algorithm 1 Procedure of PolSAR Image Classification via LCSR-Based Features Extraction and SVM Classifier

Input: PolSAR covariance matrix $\mathbf{C}$, the target decomposition model $\mathbf{C}_i$, the training samples of each category.

Initialization: initialize the parameters $\mathbf{D}_{im}$, $\{\gamma_{im}\}$, $n_{\max}$, λ, μ, $\mathbf{P}_i$ by $\mathbf{C}_i$.

for $t = 1 : n_{\max}$, **do**

1. Calculate the local feature response γ_{ij}^n with (14).
2. Update the dictionary filters with (17).
3. Calculate the reconstructed features with (19)

$$\hat{\mathbf{P}}_i^n = \sum_{j=1}^{N} \mathbf{D}_{ij}\gamma_{ij}^n \tag{19}$$

end

4. Obtain the joint polarimetric and adjacent features $\hat{\mathbf{P}}_i$.
5. Use SVM classifier for PolSAR image classification.

Output: the classification result of PolSAR image

classification, different grounds can be effectively distinguished, and the boundary is clear. The comparative experiments are also conducted in Fig. 7. The training samples used in the comparison experiments are same.

The result in Fig. 7(h) shows that some seawater are misclassified as roads. The main reason is that the seawater and roads are both relatively smooth in the PolSAR image. Because the radar echo is less and surface scattering is strong, the scattering properties of seawater and roads are similar. Also, some waves exist in the sea when acquiring the image, which results the scattering components of seawater changing. The double bounce scattering component increases and the scattering mechanism of seawater is closer to bare land's. The great change of spatial texture also leads to a certain degree of misclassification.

Fig. 7(b) and (c) shows the comparative experiments of classical classification methods, where Fig. 7(b) is the result based on Wishart classifier. The result is poor and forest, crop and bare land are difficult to distinguish. Fig. 7(c) is the result based on spatial constrained SR classifier [55], which takes advantage of the spatial domain correlation of the image. The result can classify the image effectively but partial misclassification still exists. Fig. 7(d)–(g) shows the comparison with different features. Fig. 7(d) is the result based on original covariance matrix and SVM classifier; the result shows that the seawater, roads as well as bare land are misclassified with each other. In Fig. 7(e) and (f), the classical polarimetric features are utilized for classification.

Fig. 7(e) is the result based on traditional Yamaguchi decomposition and SVM classifier, Fig. 7(f) is the result based on Zhang decomposition and SVM classifier. It can be seen that the classification based on the traditional target decomposition method can basically distinguish different grounds, and the misclassification of seawater, roads, and bare land can be improved, but the effect is not significant. Moreover, the result

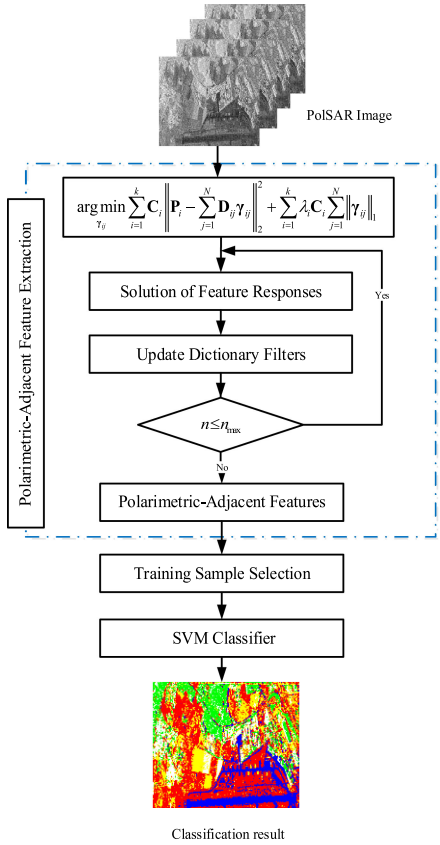

Fig. 4. Schematic diagram of PolSAR image classification via LCSR-based feature extraction and SVM classifier.

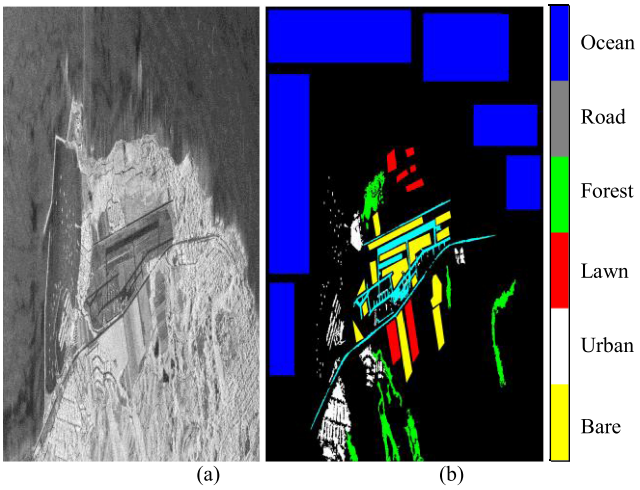

Fig. 5. UAVSAR PolSAR image of Half Moon Bay test area. (a) Amplitude image of HH channel. (b) Groundtruth.

of Fig. 7(f) is worse than Fig. 7(e), especially the confusion between forest and buildings is serious. Fig. 7(g) is the result based on SVM classifier with Yamaguchi and GLCM features [23]. Energy, entropy, and contrast features in GLCM are used in this article, which also utilize the spatial adjacent information to represent the texture of image. The result shows that the

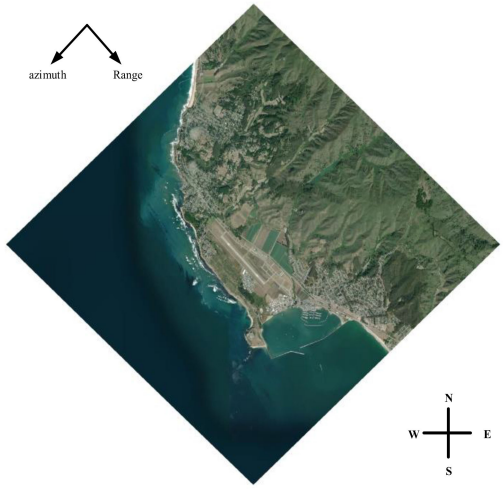

Fig. 6. Optical image of Half Moon Bay test area from Google Earth.

classification using the polarimetric and adjacent characteristics can improve the classification effect, and the misclassification of seawater, bare land near the airport, and roads have been effectively controlled. However, some forest, bare land, and lawn are incorrectly classified as urban. The texture between lawn and bare land is similar, which also leads to the misclassification of these two areas. The result based on the proposed method is shown in Fig. 7(h), which obtains the best classification result compared with the other experiments. Different grounds can be distinguished better; the confusion of forest, roads, and bare land has been effectively improved; and the scatterings are also controlled.

In order to illustrate the advantages of the proposed method, the quantitative evaluation is given in this article. The overall accuracy (OA), average accuracy (AA), and kappa coefficient, which are the most commonly used evaluation index of classification accuracy, are shown in Table I. Training samples are not included in the quantitative statistics. The conclusions are similar as the figures shown, the classification results based on Wishart and based on covariance matrix are similar and the worst, the OA, AA are only about 50% and the kappa coefficient are both low. The method based on Zhang is improved slightly, but the result is still poor. The methods based on spatial constrained SR and based on Yamaguchi can both be better, but the accuracies are not high. The classification result based on Yamaguchi and GLCM features is the best in the comparison methods, but the AA and OA are only 86.13% and 81.65%. The method proposed in this article can obviously and effectively improve AA and OA of the classification. The AA and OA of the proposed methods are 92.12% and 88.83%, respectively, and the classification accuracy of most categories is improved. The accuracy of urban is declined contrasting with the result based on Yamaguchi and GLCM features. It is because that other categories are misclassified as urban, resulting in the increase of urban accuracy. The result of the proposed method also leads to an improvement in the accuracy to the seawater and bare land. The misclassification is reduced. Meanwhile, the

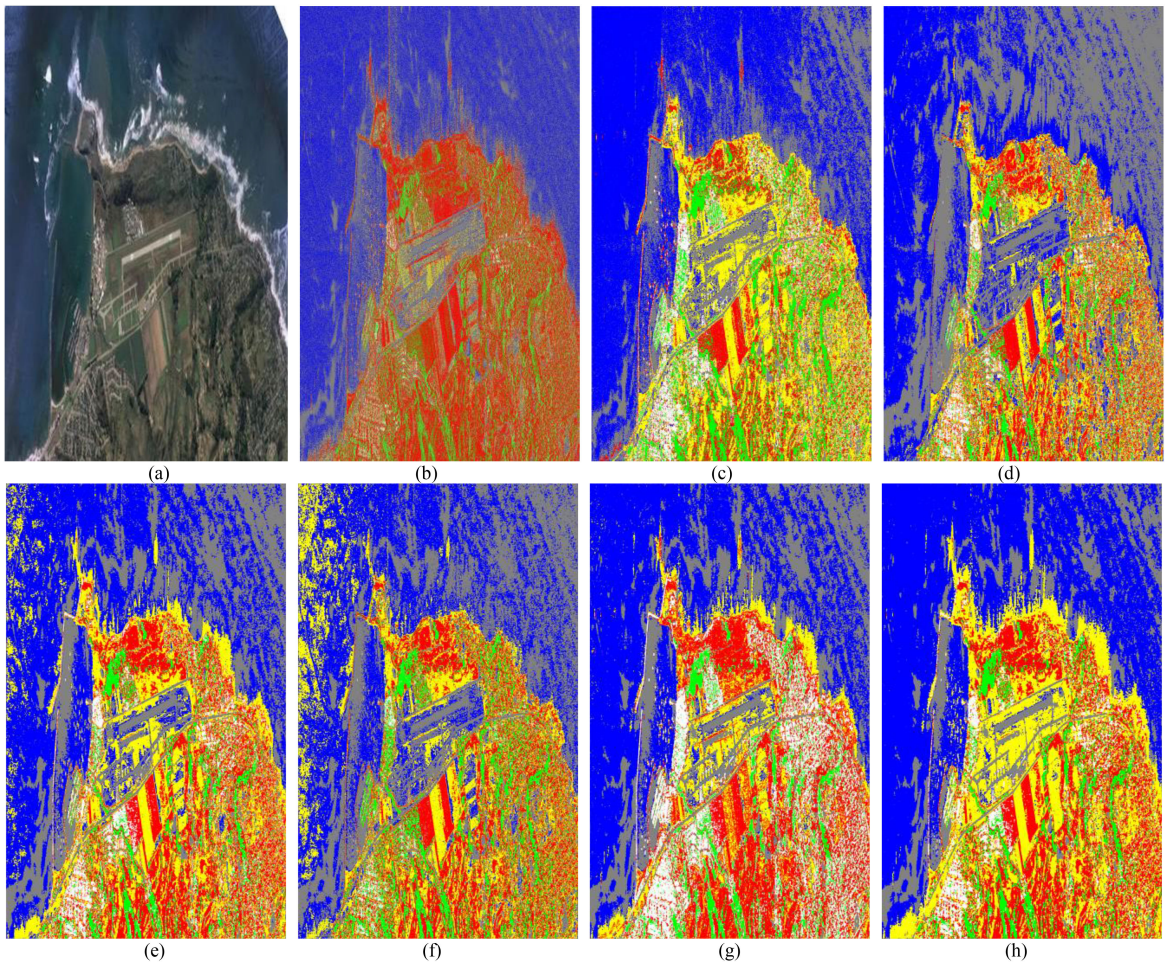

Fig. 7. Classification results of Half Moon Bay area based on different methods. (a) Optical image proportional to the SAR image resolution of test site. (b) Wishart classification. (c) Spatial constrained SR classification. (d) SVM classification with covariance matrix. (e) SVM classification with Yamaguchi decomposition. (f) SVM classification with Zhang decomposition. (g) SVM classification with Yamaguchi decomposition and GLCM. (h) Proposed method.

TABLE I
QUANTITATIVE COMPARISON OF HALF MOON BAY AREA CLASSIFICATION RESULTS BY DIFFERENT METHODS

Method	Ocean	Road	Forest	Lawn	Urban	Bare land	AA	OA	Kappa
Wishart	63.28%	63.26%	52.58%	78.70%	33.78%	43.38%	55.83%	59.96%	0.35
Spatial Constrained SR	80.55%	91.66%	89.24%	82.59%	72.95%	77.47%	82.41%	81.01%	0.68
SVM with Covariance matrix	63.36%	75.39%	42.98%	60.19%	47.98%	36.24%	54.36%	59.81%	0.35
SVM with Yamaguchi	68.35%	95.79%	83.96%	88.67%	84.03%	69.70%	81.75%	72.32%	0.54
SVM with Zhang	55.64%	94.86%	82.76%	90.35%	42.69%	50.18%	69.41%	59.15%	0.37
SVM with Yamaguchi and GLCM	79.86%	92.45%	86.43%	89.16%	**93.51%**	75.39%	86.13%	81.65%	0.67
proposed method	**87.28%**	**97.14%**	**93.33%**	**92.23%**	91.12%	**91.63%**	**92.12%**	**88.83%**	**0.79**

The bold entities means that this method has the best result of the comparison methods.

kappa coefficient of the proposed method is the highest, which proves that the proposed method has the highest consistency. The classification results also prove that the proposed method can effectively combine the polarimetric and adjacent characteristics of PolSAR image with LCSR, meanwhile, improving the ability in PolSAR image classification.

B. Results of Salvador Area

The second data used in the experiment is the complex covariance format of L-band PolSAR image of Salvador (13.742°N–89.393°W), which was acquired by UAVSAR system in April 10, 2015. The nominal one-look spatial resolution is

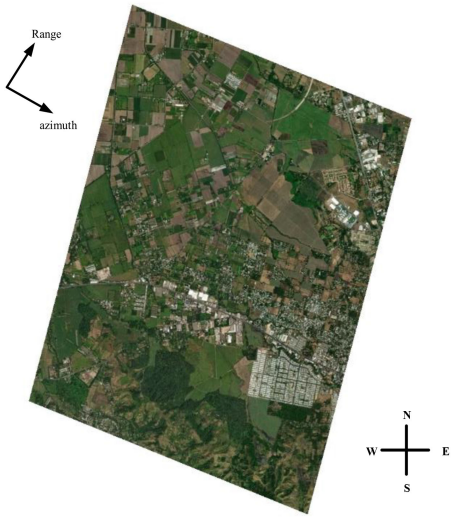

Fig. 8. Optical image of Salvador test area from Google Earth.

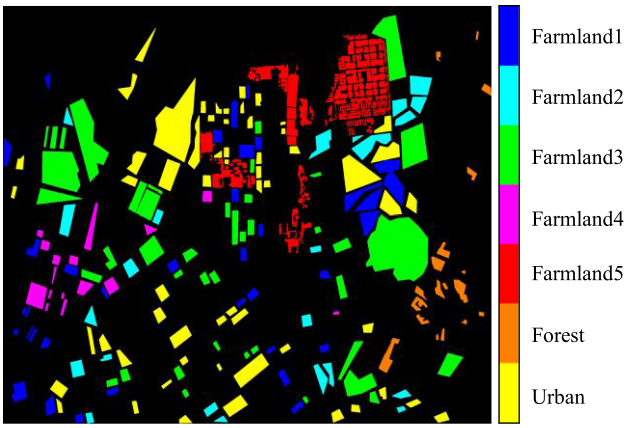

Fig. 9. Groundtruth of Salvador test area.

$1.6\,\mathrm{m} \times 0.6\,\mathrm{m}$. The aspect ratio of the image has been improved to better match the natural grounds. The image is 3745 pixels in row and 3200 pixels in line. The optical image (not the same time) from Google Earth is shown in Fig. 8. The amplitude image of HH channel is shown in Fig. 10(a). It is an area with about seven classes, including urban, forest, and five kinds of farmlands. The groundtruth, which is marked partly, is shown in Fig. 9. Each category selected 500 pixels as training samples and the classification result is shown in Fig. 10(h).

The comparative experiments are also conducted in Fig. 10. Similar with the first experiment results, Fig. 10(b) and (c) shows the comparison of classical classification methods. The result based on Wishart classifier is poor, with a large number of scattered points. The method based on spatial constrained SR is effective for classification and not has many scattered points. But some misclassifications are obvious. Fig. 10(d)–(g) shows the results of comparative experiments with different features. Fig. 10(d) is the result based on original covariance matrix and SVM classifier; the result shows that the classification is hardly achieved only by covariance matrix. The classification results are almost all scattered points and similar with Fig. 10(b). Fig. 10(e) is the result based on traditional Yamaguchi decomposition and SVM classifier, Fig. 10(f) is the result based on Zhang decomposition and SVM classifier. The results show that a large number of farmlands are misclassified with each other. The main reason is that the scattering characteristics of these farmlands are similar. It is easy to be mixed when only the polarimetric information of image is considered. Therefore, the spatial information of the PolSAR image is utilized in Fig. 10(g), which is the result based on SVM classifier with Yamaguchi and GLCM features. The result shows that the injection of adjacent information can improve the classification effect, and the misclassification can be partially controlled. The result of the proposed method is shown in Fig. 10(h), which has the best classification result and different grounds can be effectively distinguished. The results prove that the proposed method can effectively realize PolSAR image classification.

The quantitative evaluation is also given in this article. The OA, AA, and kappa coefficient of different methods are shown in Table II. Similar conclusions can also be obtained with the results shown in the figures. In the above experiment results, the proposed method leads to a great improvement in the accuracy of each category, OA, AA, and kappa coefficient. The OA and AA of the proposed methods are 93.60% and 94.31%, respectively, the kappa coefficient is 0.92, and the classification accuracy of each category is improved. The accuracy of urban is declined contrasting with the result based on Yamaguchi and GLCM features, and the reason is the same as the experiments of Half Moon Bay. The results prove the validity and high consistency of the proposed method as well as the ability in PolSAR image classification.

C. *Results of Oberpfaffenhofen area*

The third experimental data used in this article is the complex covariance format of L-band PolSAR image of Oberpfaffenhofen area (48.088°N–11.273°E), Germany, which was acquired by the German DLR SAR system ESAR. The nominal one-look spatial resolution is $3\,\mathrm{m} \times 3\,\mathrm{m}$. The image is 1540 pixels in row and 2816 pixels in line. The optical image (not the same time) from Google Earth is shown in Fig. 11. The amplitude image of HH channel is shown in Fig. 13(a). It is an area with about five classes including forest, bare land, urban, crop, and road. The groundtruth, which is marked partly, is shown in Fig. 12. Each category selected 200 pixels as training samples, and the classification result is shown in Fig. 13(h). The comparative experiments are also conducted in Fig. 13.

Fig. 13(b) and (c) shows the comparison of classical classification methods. The similar conclusion can also be drawn. Fig. 13(d)–(g) shows the results of comparative experiments

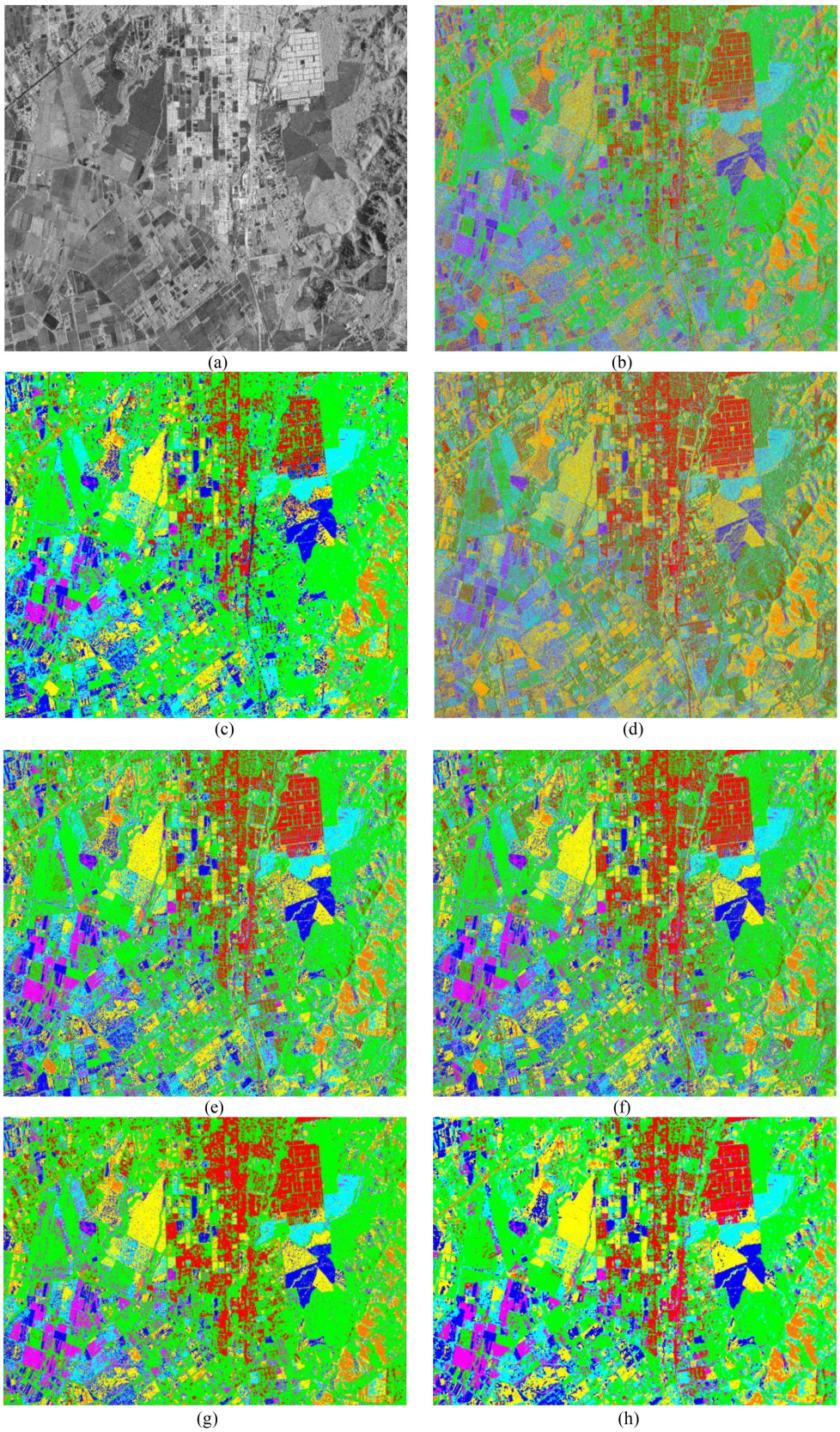

Fig. 10. Classification results of Salvador area based on different methods. (a) Amplitude image of HH channel. (b) Wishart classification. (c) Spatial constrained SR classification. (d) SVM classification with covariance matrix. (e) SVM classification with Yamaguchi decomposition. (f) SVM classification with Zhang decomposition. (g) SVM classification with Yamaguchi decomposition and GLCM. (h) Proposed method.

TABLE II
QUANTITATIVE COMPARISON OF SALVADOR AREA CLASSIFICATION RESULTS BY DIFFERENT METHODS

Method	Farm1	Farm2	Farm3	Farm4	Farm5	Forest	Urban	AA	OA	Kappa
Wishart	49.51%	38.77%	45.03%	83.37%	34.90%	56.27%	53.48%	51.62%	48.82%	0.39
Spatial Constrained SR	91.34%	91.69%	77.28%	74.94%	75.59%	93.72%	72.71%	82.47%	84.54%	0.81
SVM with Covariance matrix	44.52%	38.62%	38.15%	80.86%	48.43%	39.80%	65.73%	50.87%	47.41%	0.37
SVM with Yamaguchi	86.60%	81.69%	88.20%	88.47%	76.38%	87.42%	84.23%	84.71%	83.74%	0.80
SVM with Zhang	78.08%	73.65%	85.18%	83.87%	79.55%	82.09%	86.32%	81.25%	81.06%	0.76
SVM with Yamaguchi and GLCM	82.40%	80.38%	86.76%	88.77%	78.19%	87.40%	**95.05%**	85.56%	85.14%	0.81
proposed method	**95.25%**	**93.27%**	**95.53%**	**97.07%**	**90.89%**	**94.44%**	93.74%	**94.31%**	**93.60%**	**0.92**

The bold entities means that this method has the best result of the comparison methods.

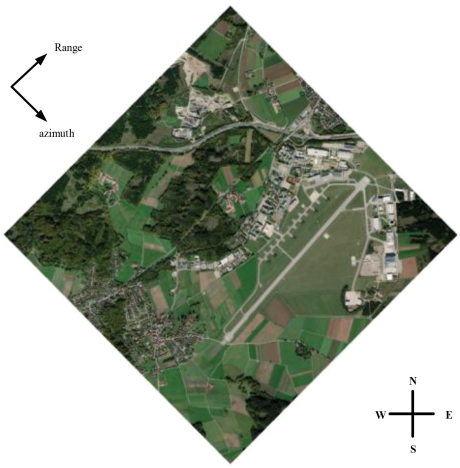

Fig. 11. Optical image of Oberpfaffenhofen test area from Google Earth.

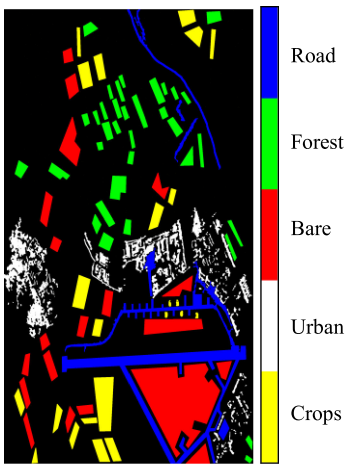

Fig. 12. Groundtruth of Oberpfaffenhofen area.

with different features, where Fig. 13(d) is the result based on original covariance matrix and SVM classifier, and the result is poor. In particular, the forest and the urban area are indistinguishable totally. Large numbers of scattered points and misclassification exist. Fig. 13(e) is the result based on traditional Yamaguchi decomposition and SVM classifier, and Fig. 13(f) is the result based on Zhang decomposition and SVM classifier. The forest and the urban area are basically distinguishable. But forest, some farmlands, and roads are still misclassified as urban. Fig. 13(e) is worse than Fig. 13(f), but Fig. 13(f) has more scattered points than Fig. 13(e). It is because that the scattering characteristics of forest, roads, and building are similar; the three kinds of grounds are prone to be misclassified and mixed in the classification. Fig. 13(g) is the result based on SVM classifier with Yamaguchi and GLCM features. The result shows that the method using the polarimetric and adjacent characteristics of image can improve the classification effect, but the improvement is not obvious. The method, which combines the polarimetric and adjacent information simultaneously in this article, is shown in Fig. 13(h). The method can get the best classification result compared with the above methods. Different grounds can be distinguished better and fewer scattered points exist. The results show that the proposed method can effectively realize the classification of PolSAR image; different grounds can be effectively distinguished and the boundary is clear.

The quantitative evaluation is also given in this article. The OA, AA, and kappa of different methods are shown in Table III. The AA and OA of the proposed methods are 91.60% and 91.84%, respectively, and most classification accuracies of each category are improved. The accuracy of urban is declined contrast with the result with Yamaguchi features. The reason is similar with the UAVSAR data, namely, the other categories are misclassified as urban, resulting in the increase of urban accuracy. Similar to the conclusions obtained in the figure, the proposed method also leads to a great improvement in the accuracy, OA, AA, as well as kappa coefficient.

Meantime, Y4O and Zhang's methods are decomposition methods based on pixel, which can get the results quickly. GLCM-based method takes longer time because of the addition of window operation. The sparse coefficients solution in the proposed method are conducted by LARS and proximal gradient descent method, and the proposed polarimetric-adjacent feature extraction method is an iteration procedure. Therefore, it will take more time inevitably. However, each local feature response can be regarded as a slice of the global feature responses, and each slice is independent and can be optimized separately, so the proposed method can process multiple sparse solutions in parallel by directly operating on image blocks, which also optimizes the speed to a certain extent. With the development of

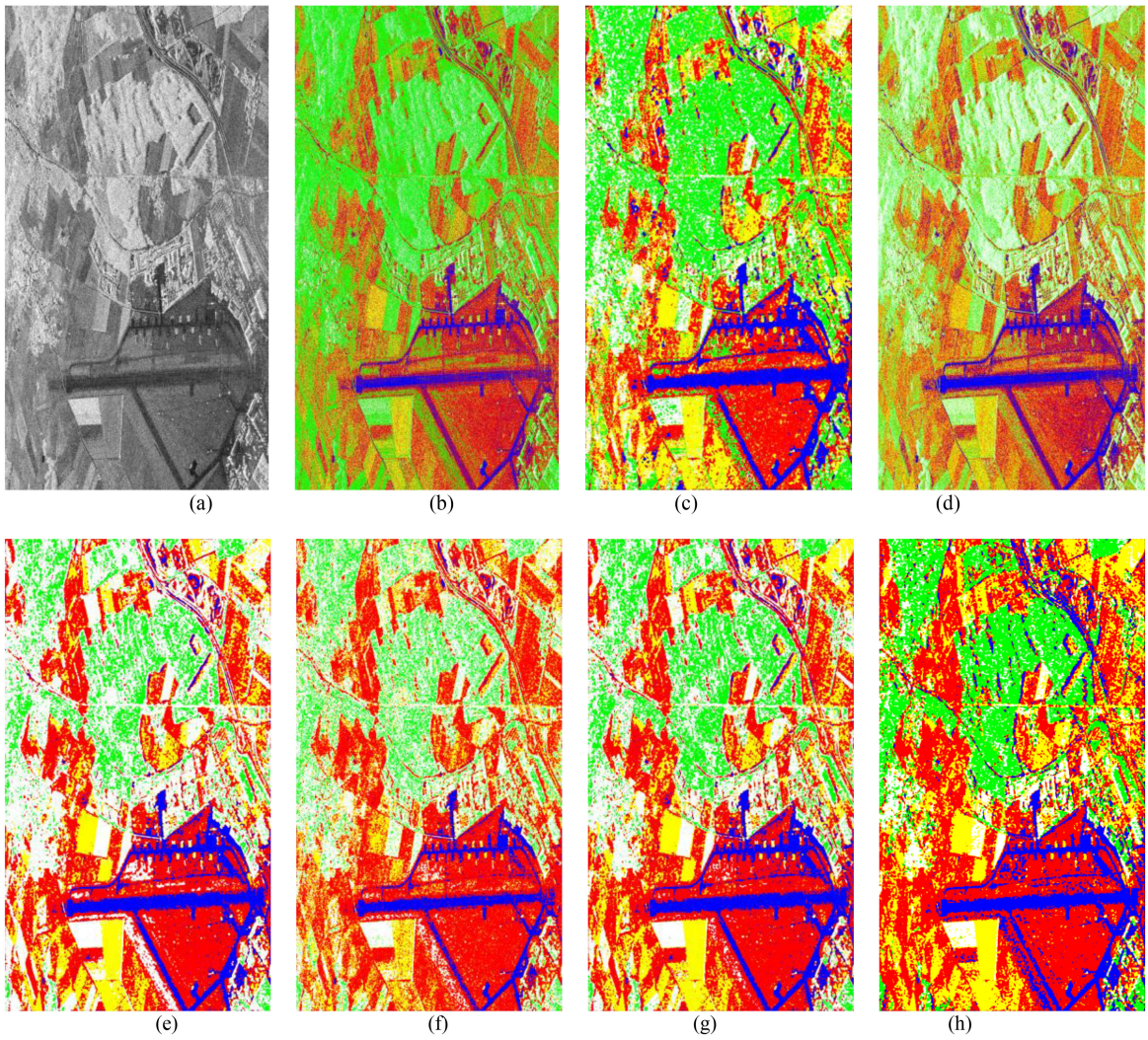

Fig. 13. Classification results of Oberpfaffenhofen area based on different methods. (a) Optical image of test site. (b) Wishart classification. (c) Spatial constrained SR classification. (d) SVM classification with covariance matrix. (e) SVM classification with Yamaguchi decomposition. (f) SVM classification with Zhang decomposition. (g) SVM classification with Yamaguchi decomposition and GLCM. (h) Proposed method.

TABLE III
QUANTITATIVE COMPARISON OF OBERPFAFFENHOFEN AREA CLASSIFICATION RESULTS BY DIFFERENT METHODS

Method	Road	Forest	Bare land	Urban	Crop	AA	OA	Kappa
Wishart	54.92%	70.42%	63.01%	40.29%	49.79%	55.69%	56.63%	0.45
Spatial Constrained SR	85.05%	85.13%	77.97%	71.10%	71.72%	78.19%	78.38%	0.74
SVM with Covariance matrix	70.29%	29.45%	56.43%	64.58%	52.91%	54.73%	56.34%	0.44
SVM with Yamaguchi	86.67%	59.08%	91.19%	**95.93%**	73.75%	81.32%	83.88%	0.79
SVM with Zhang	71.18%	46.89%	88.79%	82.46%	57.62%	69.39%	73.47%	0.65
SVM with Yamaguchi and GLCM	87.00%	71.96%	92.05%	93.68%	81.60%	85.26%	86.86%	0.83
proposed method	**91.23%**	**91.03%**	**93.08%**	90.40%	**92.25%**	**91.60%**	**91.84%**	**0.90**

The bold entities means that this method has the best result of the comparison methods.

hardware, the proposed method can be conducted on the better performance CPUs or GPUs to consume less time. Considering the consumption and the performance in PolSAR image classification, the tradeoff is reasonable. The three experiment data all prove that the proposed method can effectively combine the polarimetric and adjacent characteristics of PolSAR image with LCSR, meanwhile improving the classification ability for PolSAR image classification.

VI. Conclusion

In this article, a joint polarimetric-adjacent features extraction method based on LCSR is proposed for PolSAR image classification. The main contribution of the proposed method is using CSR to achieve the effective combination of the polarimetric and adjacent information of the image. Meanwhile, the convolution dictionaries and feature responses are constructed by local strategy to avoid the confusion of different grounds caused by global dictionary, which can also use the local spatial features of the image itself effectively. The extracted joint polarimetric-adjacent features are combined with SVM for PolSAR classification. Three sets of fully polarimetric SAR images of different system are used in the experiments. The qualitative and quantitative results validate the utility and potential of the proposed method in PolSAR image classification. The results indicate that the proposed approach can effectively combine the polarimetric and adjacent information of data and have potential in PolSAR remote sensing data classification.

References

[1] 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, Oct. 1994.

[2] 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.

[3] W. B. Silva, C. C. Freitas, S. J. S. San'Anna, and A. C. Frery, "Classification of segments in PolSAR imagery by minimum stochastic distances between Wishart distributions," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 3, pp. 1263–1273, Jun. 2013.

[4] 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.

[5] L. Zhang, B. Zou, J. Zhang, and Y. Zhang, "Classification of polarimetric SAR image based on support vector machine using multiple-component scattering model and texture features," *EURASIP J. Advances Signal Proc.*, vol. 9, no. 5, pp. 1–9, Jan. 2010.

[6] M. Pham, "Fusion of polarimetric features and structural gradient tensors for VHR PolSAR image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 10, pp. 3732–3742, Oct. 2018.

[7] Z. Zhang, H. Wang, F. Xu, and Y. Q. Jin, "Polarimetric SAR image classification using deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1935–1939, Dec. 2016.

[8] Z. Zhang, H. Wang, F. Xu, and Y. Q. Jin, "Complex-valued convolutional neural network and its application in polarimetric SAR image classification, " *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 7177–7188, Dec. 2017.

[9] R. Touzi, "Target scattering decomposition in terms of roll-invariant target parameters," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 73–84, Jan. 2007.

[10] B. Zou, D. Lu, L. Zhang, and W. M. Moon, "Independent and commutable target decomposition of PolSAR data using a mapping from SU(4) to SO(6)," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 6, pp. 3396–3407, Jun. 2017.

[11] A. Freeman, and S. L. Durden, "A three-component scattering model for polarimetric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 3, pp. 963–973, May 1998.

[12] Y. Yamaguchi, T. Moriyama, M. Ishido, and H. Yamada, "Four-component scattering model for polarimetric SAR image decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 8, pp. 1699–1706, Aug. 2005.

[13] Y. Yamaguchi, A. Sato, W. M. Boerner, R. Sato, and H. Yamada, "Four-component scattering power decomposition with rotation of coherency matrix," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2251–2258, Jun. 2011.

[14] L. Zhang, B. Zou, H. Cai, and Y. Zhang, "Multiple-component scattering model for polarimetric SAR image decomposition," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 603–607, Oct. 2008.

[15] J. J. van Zyl, M. Arii, and Y. Kim, "Model-based decomposition of polarimetric SAR covariance matrices constrained for nonnegative eigenvalues," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 9, pp. 3452–3459, Sep. 2011.

[16] W. An, Y. Cui, and J. Yang, "Three-component model-based decomposition for polarimetric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2732–2739, Jun. 2010.

[17] S. W. Chen, X. S. Wang, S. P. Xiao, and M. Sato, "General polarimetric model-based decomposition for coherency matrix," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1843–1855, Mar. 2014.

[18] A. Bhattacharya, A. Muhuri, S. De, S. Manickam, and A. C. Frery "Modifying the Yamaguchi four-component decomposition scattering powers using a stochastic distance," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 7, pp. 3497–3506, Jul. 2015.

[19] G. Singh *et al.*, "Seven-component scattering power decomposition of POLSAR coherency matrix," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 8371–8382, Nov. 2019.

[20] S. Suvorova, and J. Schroeder, "Automated target recognition using the Karhunen–Loeve transform with invariance," *Digit. Signal Process.*, vol. 12, no. 1, pp. 295–306, Apr. 2002.

[21] R. Zhang, J. Hong, and F. Ming, "An improved PCA based features for SAR ATR," in *Proc. IEEE IET Int. Radar Conf.*, 2009, pp. 1–3.

[22] G. Vasile, "Independent component analysis based incoherent target decompositions for polarimetric SAR data-practical aspects," in *Proc. IEEE Geoscience Remote Sens. Soc.*, 2018, pp. 5859–5862.

[23] A. Mahmoud, S. Elbialy, B. Pradhan, and M. Buchroithner, "Field-based landcover classification using TerraSAR-X texture analysis," *Adv. Space Res.*, vol. 48, no. 5, pp. 799–805, Sep. 2011.

[24] G. Liu, R. Wang, Y. Deng, R. Chen, Y. Shao, and Z. Yuan, "A new quality map for 2-D phase unwrapping based on gray level co-occurrence matrix," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 2, pp. 444–448, Feb. 2014.

[25] S. P. A. R. Putra, S. C. Keat, K. Abdullah, L. H. San, and M. N. M. Nordin, "Texture analysis of AIRSAR images for land cover classification, " in *Proc. IEEE IconSpace*, 2011, pp. 243–248.

[26] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing.*, Switzerland: Springer Science & Business Media, 2010.

[27] S. Liu, M. Liu, P. Li, J. Zhao, Z. Zhu, and X. Wang, "SAR image denoising via sparse representation in shearlet domain based on continuous cycle spinning," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2985–2992, May 2017.

[28] Y. Liu, S. Canu, P. Honeine, and S. Ruan, "Mixed integer programming for sparse coding: Application to image denoising," *IEEE Trans. Comput. Imag*, vol. 5, no. 3, pp. 354–365, Sep. 2019.

[29] L. Zhuang, and J. M. Bioucas-Dias, "Fast hyperspectral image denoising and inpainting based on low-rank and sparse representations," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 730–742, Mar. 2018.

[30] K. Zhang, M. Wang, S. Yang, and L. Jiao, "Convolution structure sparse coding for fusion of panchromatic and multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1117–1130, Feb. 2019.

[31] G. He, J. Ji, D. Dong, J. Wang, and J. Fan, "Infrared and visible image fusion method by using hybrid representation learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 11, pp. 1796–1800, Nov. 2019.

[32] M. Liu, S. Chen, J. Wu, F. Lu, X. Wang, and M. Xing, "SAR target configuration recognition via two-stage sparse structure representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2220–2232, Apr. 2018.

[33] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image superresolution via sparse representation," *IEEE Trans. Imag. Proc*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[34] J. Mairal, F. Bach, and J. Ponce, "Sparse modeling for image and vision processing," *Foundations Trends Comput. Graph. Vis.*, vol. 8, no. 2–3, pp. 85–283, Nov. 2014.

[35] A. Szlam, K. Kavukcuoglu, and Y. Lecun, "Convolutional matching pursuit and dictionary training," 2010, *arXiv:1010.0422*. [Online]. Available: https://arxiv.org/abs/1010.0422

[36] S. H. Gu, W. M. Zuo, Q. Xie, D. Y. Meng, X. C. Feng, and L. Zhang, "Convolutional sparse coding for image super-resolution," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1823–1831.

[37] Y. Liu, X. Chen, R. K. Ward, and Z. J. Wang, "Image fusion with convolutional sparse representation," *IEEE Signal Process. Lett.*, vol. 23, no. 12, pp. 1882–1886, Dec. 2016.

[38] B. H. Chen, J. Li, B. Y. Ma, and G. Wei, "Convolutional sparse coding classification model for image classification," in *Proc. IEEE Int. Conf. Image Process.*, 2016, pp. 1918–1922.

[39] K. Kavukcuoglu, P. Sermanet, Y. L. Boureau, K. Gregor, M. Mathieu, and Y. LeCun, "Learning convolutional feature hierarchies for visual recognition," in *Proc. Neural Inf. Process. Syst.*, 2010, pp. 1090–1098.
[40] H. Bristow, A. Eriksson, and S. Lucey, "Fast convolutional sparse coding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 391–398.
[41] F. Heide, W. Heidrich, and G. Wetzstein, "Fast and flexible convolutional sparse coding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 5135–5143.
[42] B. Wohlberg, "Efficient convolutional sparse coding," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, 2014, pp. 7173–7177.
[43] B. Wohlberg, "Efficient algorithms for convolutional sparse representations," *IEEE Trans. Imag. Process.*, vol. 25. no. 1, pp. 301–315, Jan. 2016.
[44] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
[45] K. Engan, S. O. Aase, and J. H. Husøy, "Frame based signal compression using method of optimal directions (MOD)," in *Proc. IEEE Int. Symp. Circuits Syst.*, 1999, pp. 1–4.
[46] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
[47] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *J. Mach. Learn. Res.*, vol. 11, no. 1, pp. 19–60, Feb. 2010.
[48] S. G. Mallat, and Z. F. Zhang, "Matching pursuit with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.
[49] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. IEEE 27th Asilomar Conf. Signals Syst. Comput.*, 2002, pp. 40–44.
[50] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 586–598, Dec. 2007.
[51] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *J. Roy. Statist. Soc.*, vol. 58, no. 1, pp. 267–288,1996.
[52] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, and R. Tibshirani, "Least angle regression," *Ann. Statist.*, vol. 32, no. 2, pp. 407–499, Apr. 2004.
[53] V. Papyan, Y. Romano, M. Elad, and J. Sulam, "Convolutional dictionary learning via local processing," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 5306–5314.
[54] E. Zisselman, J. Sulam, and M. Elad, "A local block coordinate descent algorithm for the CSC model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 8200–8209.
[55] L. Zhang, L. Sun, and W. M. Moon, "Polarimetric SAR image classification based on contextual sparse representation," in *Proc. IEEE Geosci. Remote Sens. Soc.*, 2015, pp. 1837–1840.

Xiao Wang (Student Member, IEEE) received the B.S. and M.Sc. degrees in electronics and communication engineering from the Harbin Institute of Technology, Harbin, China, in 2014 and 2016, respectively. She is currently working toward the Ph.D. degree with the School of Information and Communication Engineering, Harbin Institute of Technology.

Her research interests include feature extraction and classification of PolSAR images.

Lamei Zhang (Senior Member, IEEE) received the B.S., M.Sc., and Ph.D. degrees in information and communication engineering from the Harbin Institute of Technology, Harbin, China, in 2004, 2006, and 2010, respectively.

Currently, she is an Associate Professor with the Department of Information Engineering, Harbin Institute of Technology. She serves as the Secretary of the IEEE Harbin Geoscience and Remote Sensing (GRSS) Chapter. Her research interests include remote sensing images processing, information extraction, and intelligent interpretation of high-resolution SAR, polarimetric SAR, and polarimetric SAR interferometry.

Ning Wang received the B.S. and M.Sc. degrees in mathematics, and the Ph.D. degree in computer science and technology from the Harbin Institute of Technology, Harbin, China, in 2006, 2008, and 2013, respectively.

Currently, he is a Senior Engineer of Radar Electrical system with the Beijing Institute of Radio Measurement, Beijing, China. His research interests include the areas of space target radar signal processing, recognition, ISAR imaging, three-dimensional reconstruction, and pose estimation.

Bin Zou (Senior Member, IEEE) received the B.S. degree in electronic engineering from the Harbin Institute of Technology, Harbin, China, in 1990, the M.Sc. degree in space studies from International Space University, Strasbourg, France, in 1998, and the Ph.D. degree in information and communication engineering from the Harbin Institute of Technology, in 2001.

He is a Professor with the Department of Information Engineering, Harbin Institute of Technology. He serves as the Chairman of the IEEE Harbin Geoscience and Remote Sensing (GRSS) Chapter. His research interests include remote sensing images processing, electromagnetic simulation calculation, and characteristic analysis of radar target.