



Optimal Combination of Polarimetric Features for Vegetation Classification in PolSAR Image

Q. Yin, W. Hong, F. Zhang, E. Pottier

► To cite this version:

Q. Yin, W. Hong, F. Zhang, E. Pottier. Optimal Combination of Polarimetric Features for Vegetation Classification in PolSAR Image. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12 (10), pp.3919-3931. 10.1109/JSTARS.2019.2940973 . hal-02442763

HAL Id: hal-02442763

<https://univ-rennes.hal.science/hal-02442763>

Submitted on 30 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimal Combination of Polarimetric Features for Vegetation Classification in PolSAR Image

Qiang Yin , Member, IEEE, Wen Hong, Senior Member, IEEE, Fan Zhang , Senior Member, IEEE, and Eric Pottier, Fellow, IEEE

Abstract—Polarimetric features of PolSAR images include inherent scattering mechanisms of terrain types, which are important for classification and other Earth observation applications. By using target decomposition methods, many polarimetric scattering components can be obtained. Besides, the elements of a coherency/covariance matrix, as well as polarimetric descriptors, such as SPAN, single-bounce eigenvalue relative difference/double-bounce eigenvalue relative difference, etc., can also provide characteristic information. In fact, more and more polarimetric decomposition components and descriptors have been proposed; the computation cost increases if all of them are employed as the input of the classification process. Although all these features obtained from the coherency/covariance matrix are not independent, still, finding out which ones are significant for the classification of different terrain types will improve the understanding of scattering mechanisms. In this article, the effective polarimetric feature combination is studied based on the vegetation classification performance of support vector machine (SVM) and nearest-regularized subspace (NRS) machine learning approaches, as well as their combinations with a Markov random field (MRF). A framework on the basis of similarity and the orthogonal subspace projection (OSP) method in a hyperspectral area is used to select the polarimetric features. For the airborne PolSAR data in Flevoland, The Netherlands, 107 polarimetric features are extracted, including matrix elements, target decomposition components, and polarimetric descriptors. A subset is selected by using the proposed and OSP methods. They have a good classification accuracy evaluated by SVM+MRF and NRS+MRF classifiers. However, when the SVM and the NRS are used without combining spatial information of the MRF, the features selected by the proposed framework with correlation coefficient criteria have much better classification performance than those of OSP and principal component analysis.

Index Terms—Classification, feature combination, nearest-regularized subspace (NRS), PolSAR, support vector machine (SVM).

Manuscript received September 29, 2018; revised March 11, 2019 and July 3, 2019; accepted August 29, 2019. This work was supported in part by the National Natural Science Foundation of China under Grants 61801015, 61871413, and 61431018, and in part by the Beijing Natural Science Foundation under Grant 4184094. (Corresponding author: Fan Zhang.)

Q. Yin and F. Zhang are with the College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China (e-mail: yinq@mail.buct.edu.cn; zhangf@mail.buct.edu.cn).

W. Hong is with the Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China (e-mail: whong@mail.ie.ac.cn).

E. Pottier is with the Institut d'Electronique et de Telecommunications de Rennes, UMR CNRS 6164, Université de Rennes 1, 35042 Rennes, France (e-mail: eric.pottier@univ-rennes1.fr).

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

Digital Object Identifier 10.1109/JSTARS.2019.2940973

I. INTRODUCTION

POLARIMETRIC features provide scattering mechanisms of land cover, which make unsupervised classification possible. Many studies focusing on polarimetric features can be found in the literature; however, the main role of different features has not been revealed. More and more polarimetric decomposition components and descriptors have been proposed; the computation cost increases if all of them are employed as the input of the classification process. Meanwhile, there is no direct answer about their importance and specific roles in different applications yet. It is meaningful to study which features are important for certain applications. Some of the features may be irrelevant to the classification task, and others may be redundant. Thus, it is useful to exploit the discriminative power offered by the selection and combination of these features [1]. The idea of this article is to find out whether a small subset of features could achieve similarly good classification results as the whole feature sets. Actually, it depends on the features of specific data as well as the classifier, but there might also exist some regular patterns underlying.

The scheme of normal supervised classification includes two parts: feature extraction/selection and machine learning classifiers. The focus of this article is on the polarimetric features of PolSAR images, rather than the classifiers. We employ some classifiers for evaluating the performance of feature combinations.

Several groups work on the optimal feature selection for PolSAR image classification. Feature reconstruction methods, such as independent component analysis [2] and principal component analysis (PCA) [3], have been found applicable, but they will change the physical meaning of the original data because new channels do not correspond to an original polarimetric feature, but their linear combinations. In this way, we cannot figure out which polarimetric features play a significant role in applications. Chen *et al.* [4] applied a convex framework for joint learning of the optimal feature weights and support vector machine (SVM) parameters. Haddadi *et al.* [5] used a combination of a genetic algorithm and an artificial neural network for extracting optimized features of PolSAR images that are required for classification. Some methods need to rearrange the PolSAR image to vectors, which may lose the spatial information of neighboring pixels. For this consideration, Tao *et al.* proposed a tensor-based dimension reduction technique, using 48 polarimetric decomposition features of synthesized PolSAR data [6]. Based on the Touzi decomposition parameters,

Banerjee *et al.* [7] proposed a third-order class-dependent mutual information (MI) method and another eigenspace decomposition of the conditional MI matrix method to select a feature subset. Besides, a feature selection separability index criterion with the naive Bayes classifier was utilized in [8], where ten polarimetric features and 27 other features were examined.

The analysis on polarimetric features is necessary for classification and other related applications, since machine learning as well as deep learning classifiers have been used a lot; however, their requirement on input features is not clear. Several classifiers such as nearest neighbor, naive Bayes, and neural network (NN) have been successfully used in the literature for characterizing SAR data [9], [10]. The SVM has outperformed the traditional classifiers in almost all the domains in terms of the generalization accuracy as well as time and space complexities. The SVM has been used in this article, as it has a convex formulation, requires lesser numbers of training samples, and has a simple but mathematically sound statistical learning-based background theory. In some cases, the recently proposed representation-based classifications provide better performance than the classic NN [11] and SVM [9] in that they require few labeled data instead of the traditional training–testing mode. Representation-based classification is essentially based on the concept that a pixel can be represented as a linear combination of labeled samples via the sparse regularization methods, such as the l_0 -norm, l_1 -norm, and l_2 -norm regularization. In this manner, an approximation of the pixel is generated from labeled samples of each class independently, and the class label is then derived according to the class of the minimum representation error. In this article, the supervised classifiers of SVM, nearest-regularized subspace (NRS), as well as SVM+MRF [12] and NRS+MRF [13] are used. By the way of involving the MRF model, the spatial information between pixels is reserved. The Markov random field (MRF) approach is a popular model for incorporating spatial information into image classification [12], [14], [15].

In this article, we focused on the polarimetric feature analysis and its effective combinations for the classification of PolSAR data. The classifiers are employed only for the performance evaluation of the features' capability on terrain classification. They are independent of the feature selection process. A set of 107 polarimetric features is extracted, and the polarimetric feature vector is constructed, which cover almost all the representative and recent ones [16]. A polarimetric feature selection method based on the similarity between each two features is proposed, which takes into consideration of both representatives of highly related features and the unrelated ones. As summarized in the review of [17], the current feature selection methods based on the similarity all use a criterion or a parameter, such as Laplacian Score, SPEC, Fisher Score, Trace ratio, Relief F, etc., to compute the weight of each feature. Afterward, with the obtained weights, all the features are ranked, so the important ones can be selected out. However, the key point of our suggested method is to compute the similarity between each pair of features, for example, the correlation coefficient (CR). Hence, what we got is the correlation between each pair of two features, rather than the weight of each feature. In the obtained feature subset, the features with different kinds of scattering mechanisms are

preserved. In addition, another similarity-based band selection method proposed for a hyperspectral image is introduced to the PolSAR image for the first time. Experiments are implemented to assess these two methods, as well as to compare them with the principal components constructed by the PCA approach. It is shown that the feature subset obtained by the proposed CR criteria keeps better polarimetric scattering information at the pixel level for the application of vegetation classification.

The rest of this article is organized as follows. In Section II, the basic of polarimetric features of SAR data is briefly introduced. Afterward, in Section III, two feature selection methods are presented. The first proposed criterion is based on CR and the second one is introduced from hyperspectral image processing called similarity-based band selection. Supervised classifiers used for classification are briefly introduced in Section IV. Then, the experiment data, extracted polarimetric features, as well as the corresponding classification results with different selection methods are shown in Section V. For the aim of comparison, the PCA method is also employed to form the same number of new features and then to be used for the classification application. Finally, the conclusion is given in Section VI.

II. POLARIMETRIC FEATURE EXTRACTION

The information of the PolSAR data is contained in four channels, respectively, HH , HV , VH , and VV , which indicate four linear orthogonal polarization combinations. Assuming that the data are processed with spatial average under a monostatic mode, the complex backscattering matrix $[S]$ for each image pixel at a specific incidence angle can be expressed as

$$[S] = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} \quad (1)$$

where the subscript h indicates the horizontal polarization channel, and the subscript v represents the vertical polarization channel. Under the monostatic backscattering case, there exists a reciprocity theorem, i.e., $S_{hv} = S_{vh}$. We discuss this backscattering case in the following part of this article.

From the scattering matrix $[S]$, polarimetric information is usually interpreted by the second-order polarimetric descriptor coherent and covariance matrices. The multilook coherent matrix $[T]$ and covariance matrix $[C]$ can be expressed as follows:

$$[T] = \langle \vec{k} \cdot \vec{k}^H \rangle \quad (2)$$

$$[C] = \langle \vec{\Omega} \cdot \vec{\Omega}^H \rangle \quad (3)$$

where $\langle \cdot \rangle$ denotes the ensemble average, the superscript H indicates the complex conjugation and transpose of vector and matrix, $\vec{k}$ is the Pauli-based scattering vector, and $\vec{\Omega}$ is the Lexicographic scattering vector. These two vectors can be specified by vectorizing the scattering matrix $[S]$ and are defined in the backscattering case as follows:

$$\vec{k} = \frac{1}{\sqrt{2}} [S_{hh} + S_{vv}, S_{hh} - S_{vv}, 2S_{hv}]^T \quad (4)$$

$$\vec{\Omega} = [S_{hh}, \sqrt{2}S_{hv}, S_{vv}]^T \quad (5)$$

TABLE I
EXTRACTED POLARIMETRIC FEATURES

Category	Feature	No
Matrix elements	Coherency Matrix, Covariance Matrix	10
Decomposition components	Barnes,Cloude,Freeman, Freeman-Durden,H/a/A, Holm,Huynen,Krogager, Vanzyl,Yang,Yamaguchi, Mcsm5,Neuman,Tsvm	79
Polarimetric descriptors	Span,SERD/DERD, Polarization Fraction, Polarization Asymmetry, Shannon Entropy, Radar Vegetation Index, Pedestal Height, Alpha Approximation, Entropy Approximation, Scattering Mech. Entropy, Kozlov Anisotropy, Lueneburg Anisotropy, Conformity Coefficient, Scattering Predominance	18
Total		107

where the superscript T indicates the transpose of the vector. Accordingly, the coherency matrix $[T]$ and the covariance matrix $[C]$ are 3×3 in the monostatic backscattering case.

Three categories of polarimetric features are used in this article: matrix elements, decomposition components, and polarimetric descriptors, as shown in Table I. The first category includes ten elements of coherency and covariance matrices. The second category includes 79 decomposition components of different target decomposition methods. In addition, the third category has 18 polarimetric descriptors, such as SPAN, polarimetric fraction, polarimetric asymmetry (PA), scattering predominance, etc.

In the first category of polarimetric features, seven elements of the coherency matrix are extracted, which are T_{11} , T_{22} , T_{33} , and the amplitude and phase of T_{13} and T_{23} . Other three elements are obtained from the covariance matrix, which are the amplitude of C_{12} , C_{13} , and C_{23} .

Within the second category, there have been many proposed target decomposition theorems that can be classified into four main types, which are those based on the Kennaugh matrix, those using an eigenvector or eigenvalues analysis of the covariance matrix or coherency matrix, those based on a model-based decomposition of the covariance matrix or the coherency matrix, and those employing coherent decomposition of the scattering

matrix [16]. Among these four types of target decomposition algorithms, the former three belong to the incoherent decomposition, which is usually used to process the coherency matrix, the covariance matrix, and the Kennaugh matrix for representing them as a linear combination. The coherent target decomposition algorithm is used to process the scattering matrix for representing the polarimetric information of the point target and the distributed target [18]–[31].

In the third category of polarimetric descriptors, most are developed based on the eigenvalues and eigenvectors of the coherency matrix, such as SPAN, single-bounce eigenvalue relative difference (SERD), double-bounce eigenvalue relative difference (DERD), Shannon entropy, polarization fraction, PA, target randomness, radar vegetation index, pedestal height, etc. These new parameters have been proved to be useful when analyzing scattering characteristics of PolSAR data, together with other polarimetric features [32]–[36].

All these 107 polarimetric features are in fact extracted originally from the coherency matrix $[T]$, or the equivalent covariance matrix $[C]$, which means they are not independent. However, different features can reflect scattering mechanism details from different perspectives, so we can study about which features are significant for land classifications.

III. FEATURE SELECTION METHOD

The proposed method for polarimetric feature selection is based on the similarity measured by the CR; then, the orthogonal subspace projection (OSP) similarity-based approach in the hyperspectral area is introduced into the polarimetric feature. Besides, PCA is applied to form the same number of principal components for comparison. The data processing flowchart is shown in Fig. 1.

A. Proposed Framework Based on Similarity

Our suggested method is based on the similarity between each pair of features. For instance, the CR can be employed to measure the similarity between every two features. Hence, what we got is the correlation between each pair of two features, rather than the weight of each feature. This is the main difference compared to other existing methods, which usually use a criterion or a parameter, such as Laplacian Score, SPEC, Fisher Score, Trace ratio, Relief F, etc., to compute the weight of each feature. The difference leads to the following steps of the feature selection process. It is because we cannot simply rank the correlation of each pair. Our strategy in the following steps is: on one hand, with the similarity between each pair of features, above a certain threshold, those pairs containing the same feature are put into one group. By this step, several groups are formed, within which the highly related features are put together. Then, in each group, the appearance frequency of each feature is calculated, and the one with the maximum frequency in each feature group is selected out. It means that the representative features are kept from each group for classification. On the other hand, the features under the threshold are also kept, since the unrelated ones may contain specific information, which may be useful for the discrimination of classes. The whole framework is shown in

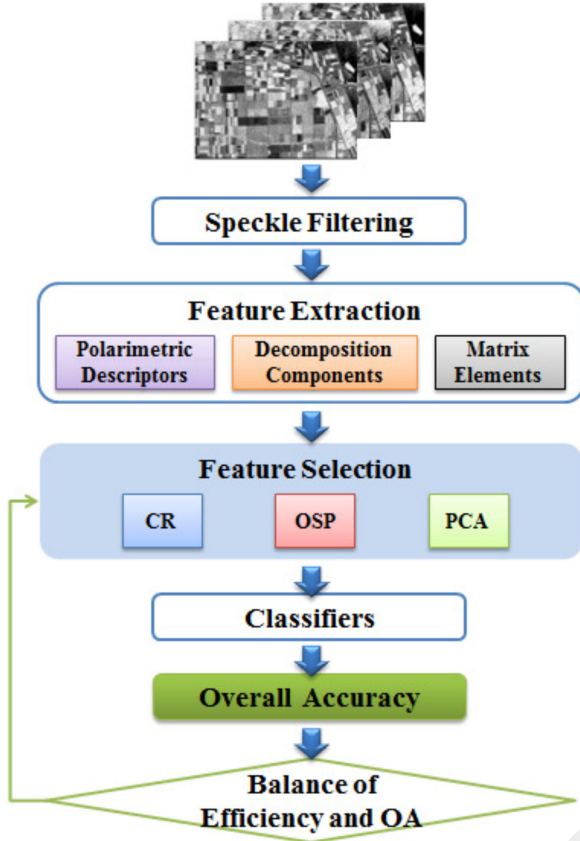


Fig. 1. Flowchart of data processing.

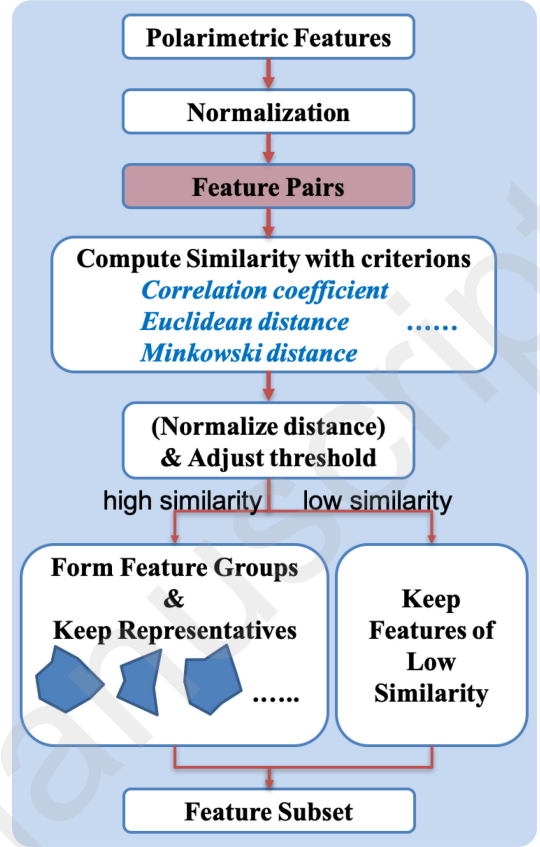


Fig. 2. Framework of the proposed method.

Algorithm 1: Proposed Method for Feature Selection.

Require: Extracted polarimetric features of the PolSAR image

Ensure: Selected features in the subset

- 1: Feature vector construction ← stacking all the categories of polarimetric features;
- 2: The similarity measurement between any two features is calculated;
- 3: The feature pairs with the CR above a certain threshold are examined;
- 4: Those pairs containing the same feature are put into one group, and the one with the maximum appearance frequency in each group is chosen into the feature subset;
- 5: The features under the threshold, which has not appeared above the threshold, are kept into the subset.

Fig. 2. And the general algorithm for feature selection is listed in the following.

One of the basic criteria of measuring the similarity between a pair of features is the CR, when it is employed to evaluate the similarity among different features. First, the CR between any pair of two features is calculated by the following equation:

$$\text{Coef}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \quad (6)$$

where $\text{Cov}(X, Y)$ is the covariance between the single feature vectors X and Y , and Var is the variance of the feature vector. Second, a proper threshold should be set. Thereafter, the feature pairs with the CR above this threshold are examined. Those pairs containing the same feature are put into one group, and the one with the maximum appearance frequency in each group is chosen into the feature subset. Meanwhile, the features under the threshold, which has not appeared above the threshold, are kept into the subset as well.

In our suggested framework, besides the CR, any criterion measuring the similarity could be used, such as Euclidean distance, Minkowski distance, Manhattan distance, and so on. Therefore, it is possible to change the criterion for similarity under the general framework. It provides a scheme to select feature subsets based on the similarity between two features. Some classic distances are listed as follows:

$$\text{EuDist}(X, Y) = \sqrt{(X - Y)^2} \quad (7)$$

$$\text{MinkDist}(X, Y) = ((X - Y)^p)^{1/p} \quad (8)$$

$$\text{ManhDist}(X, Y) = |(X - Y)|. \quad (9)$$

In this article, the Euclidean distance is also used for data experiments, which is shown and analyzed in Section V. When the distance is used to obtain the similarity among features, the computed values are not the same as with the correlation, which lies between 0 and 1. So, it is necessary to do normalization

Algorithm 2: OSP Similarity Band Selection.**Require:** Extracted polarimetric features of the SAR image**Ensure:** Selected features in the subset

- 1: Feature vector construction ← stacking all the categories of polarimetric features;
- 2: Choose an initial pair of features;
- 3: Find a third one feature that is the most dissimilar to all others in the current subset by using the OSP criterion;
- 4: Continue on Step 3 until the number of features is large enough.

before setting the threshold. In addition, the indication of distance is opposite to correlation. The larger the distance, the lower the similarity. Usually, the negative sign can be added to the normalized distances for next selection steps. In this way, the pairs above the threshold have higher similarity, while the pairs under have lower correlation.

B. Similarity-Based Band Selection (OSP)

In this article, we introduce the similarity-based band selection method, which is widely used in hyperspectral image processing into polarimetric SAR feature selections. Generally, there are two well-known algorithms, which are unsupervised fully constrained least-squares linear unmixing in [37] and OSP in [38]. Due to a large number of original bands, the exhaustive search for optimal band combinations is computationally prohibitive. For this group of similarity-based approaches, the band similarity is evaluated jointly instead of pairwise. The sequential forward search can save significant computation time. It begins with the best two-band combination, and then, this two-band combination is subsequently augmented to three, four, and so on until the desired number of bands is selected [39].

Here, the OSP band selection is used to compute the similarity between a single feature and multiple features in step 3. Assume that there are two features B_1 and B_2 in Φ . To find a feature that is the most dissimilar to B_1 and B_2 , an orthogonal subspace of B_1 and B_2 is constructed as

$$\mathbf{P} = \mathbf{I} - \mathbf{Z}(\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \quad (10)$$

where $\mathbf{I}$ is an $N \times N$ identity matrix, and $\mathbf{Z}$ is an $N \times 2$ matrix, whose first column includes all the pixels in B_1 and the second column includes all the pixels in B_2 . Then, the projection $\mathbf{y}_o = \mathbf{P}^T \mathbf{y}$ is computed, where $\mathbf{y}$ includes all the pixels in B and $\mathbf{y}_o$ is the component of B in the orthogonal subspace of B_1 and B_2 . The feature that yields the maximum orthogonal component $\|\mathbf{y}_o\|$ is considered as the most dissimilar band to B_1 and B_2 and will be selected as B_3 for Φ .

The initial two features of M feature dataset whose dissimilarity is the largest can be found with the following steps.

- 1) Randomly select a band A_1 , and project all the other $M - 1$ bands to its orthogonal subspace $\langle A_1 \rangle^\perp$.
- 2) Find the band A_2 with the maximum projection in $\langle A_1 \rangle^\perp$, which is considered as the most dissimilar to A_1 .

- 3) Project all the other $M - 1$ bands to the orthogonal subspace $\langle A_2 \rangle^\perp$, and find the band A_3 with the maximum projection.
- 4) If $A_3 = A_1$, A_1 and A_2 are confirmed to be the pair with the most significant dissimilarity, and the algorithm is terminated; if $A_3 \neq A_1$, go to the next step.
- 5) Continue the algorithm until $A_{i+1} = A_{i-1}$; then, either A_{i-1} or A_i can be used as the band selection initial B_1 .

IV. SUPERVISED CLASSIFIER

The classification methods employed in this article are SVM, NRS, as well as their corresponding combinations with the MRF method, i.e., SVM+MRF and NRS+MRF, under the consideration of reserving spatial information of pixels in images. The performance of these latter two classifiers have been studied in [13] with different datasets. The principles of NRS and MRF are stated in the following.

Assume a given dataset with training samples $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^n$ in $\mathbb{R}^d$ and the class label $w_i \in \{1, 2, \dots, C\}$, where d is the number of spectral variables (bands), C is the number of classes, and n is the total number of the training samples.

An approximation of the test sample $\mathbf{y}$ is represented via a linear combination of available training samples per class. For each class, we can calculate the approximation $\mathbf{y}_l$ as

$$\mathbf{y}_l = \mathbf{X}_l \cdot \boldsymbol{\alpha}_l \quad (11)$$

where $\mathbf{X}_l$ is a matrix with size of $d \times n_l$, n_l is the number of available training samples for class l , $\sum_{l=1}^C n_l = n$, and $\boldsymbol{\alpha}_l$ represents the weight vector coefficients with size of $n_l \times 1$ for the linear combination. Suppose we have obtained the weight vector; the label of the test sample is determined by the residual between $\mathbf{y}$ and $\mathbf{y}_l$, which is represented as

$$r_l(\mathbf{y}) = \|\mathbf{y} - \mathbf{y}_l\|_2 = \|\mathbf{y} - \mathbf{X}_l \cdot \boldsymbol{\alpha}_l\|_2. \quad (12)$$

Then, the class label is derived according to the class of the most accurate representation (i.e., the minimum value of the residuals for all the classes) as

$$\text{class}(\mathbf{y}) = \arg \min_{l=1,2,\dots,C} r_l(\mathbf{y}). \quad (13)$$

In NRS [40], the weight vector for the linear combination is solved as follows:

$$\boldsymbol{\alpha}_l = \arg \min \|\mathbf{y} - \mathbf{X}_l \cdot \boldsymbol{\alpha}_l\|_2^2 + \lambda \|\boldsymbol{\Gamma}_{l,\mathbf{y}} \cdot \boldsymbol{\alpha}_l\|_2^2 \quad (14)$$

where $\boldsymbol{\Gamma}_{l,\mathbf{y}}$ is a biasing Tikhonov matrix specific to each class l and test sample $\mathbf{y}$, and λ is a global regularization parameter, which is used to balance the minimization between the residual and regularization terms. For the NRS classifier, we design the diagonal elements of matrix $\boldsymbol{\Gamma}_{l,\mathbf{y}}$ in the form of

$$\boldsymbol{\Gamma}_{l,\mathbf{y}} = \begin{bmatrix} \|\mathbf{y} - \mathbf{x}_{l,1}\|_2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \|\mathbf{y} - \mathbf{x}_{l,n_l}\|_2 \end{bmatrix} \quad (15)$$

where $\mathbf{x}_{l,1}, \mathbf{x}_{l,2}, \dots, \mathbf{x}_{l,n_l}$ are the columns of $\mathbf{X}_l$ for the class l . Commonly, the Euclidean distance will be used for measuring the similarity between the testing sample and the

training sample and then to vote for the test sample belong to the proper label. According to formulas defined in (14) and (15), the approximation $\mathbf{y}_l$ of each test sample $\mathbf{y}$ can be expressed as

$$\mathbf{y}_l = \mathbf{X}_l (\mathbf{X}_l^T \mathbf{X}_l + \lambda \mathbf{\Gamma}_{l,y}^T \mathbf{\Gamma}_{l,y})^{-1} \mathbf{X}_l^T \mathbf{y}. \quad (16)$$

The pixelwise NRS classifier has been demonstrated to be effective [40]; however, for the PolSAR classification, the image representation of the same category object is always the block-wise manner, under which the spatial information may have the potential to further improve the classification accuracy.

In the MRF model, we have the probability of the test pixel $\mathbf{y}_p$ (note that $\mathbf{y}_p$ has the same meaning as aforementioned $\mathbf{y}$) for each class, and the energy function $E(\cdot)$ of $\mathbf{y}_p$ for class l can be represented as

$$E(\mathbf{y}_p | w_p = l) = \mu E_{\text{data}}(\mathbf{y}_p | w_p = l) + E_{\text{smooth}}(\mathbf{y}_p | w_p = l) \quad (17)$$

$$E_{\text{data}}(\mathbf{y}_p | w_p = l) = \sum_{p \in P} D_p(\mathbf{y}_p | w_p = l) \quad (18)$$

where D_p is the spectral energy function for $\mathbf{y}_p$, and μ is the balance parameter. $E_{\text{smooth}}(\mathbf{y}_p | w_p = l)$ means the extent on piecewise smoothness:

$$E_{\text{smooth}}(\mathbf{y}_p | w_p = l) = \sum_{q \in N_p} B_{\langle p,q \rangle} (1 - \delta(w_p, w_q)) \quad (19)$$

where w_p and w_q are the labels, δ is the Kronecker delta function ($\delta(w_p, w_q) = 1$ if $w_p = w_q$, and $\delta(w_p, w_q) = 0$ otherwise), $B_{\langle p,q \rangle}$ is viewed as the penalty between $\mathbf{y}_p$ and $\mathbf{y}_q$ [41], $\mathbf{y}_q$ represents the adjacent pixel to $\mathbf{y}_p$, and N_p is the set of neighbors for the given pixel $\mathbf{y}_p$. Commonly, the more similar these neighboring pixels are, the larger the value it will be. If they are totally different, $B_{\langle p,q \rangle}$ will be definitely zero.

The performance of these four supervised classifiers has been demonstrated in [13]. In general, the MRF can greatly decrease the misclassifications based on the good homogeneousness assumption. Hence, SVM+MRF and NRS+MRF can achieve high classification accuracy than single SVM and NRS. Usually, SVM is better than NRS. But with a combination of MRF, NRS+MRF performs even better than SVM+MRF.

V. EXPERIMENT

A. Datasets

The PolSAR dataset is the widely used L-band data acquired by the NASA/JPL AIRSAR system in Flevoland, The Netherlands, in August 1989. There are 11 different terrain types marked in ground truth, most of which are agricultural classes, such as stem beans, forest, wheat, bare land, rapeseed, pea, and so on. The size of this dataset is 750×1024 pixels, in which the ground truth data have around 68 188 pixels. The Pauli image and the ground truth image are shown in Fig. 3. The ground truth region mask is generated manually. To avoid the edge effect between different agricultural areas caused by the speckle filtering process, the mask of each region is set a bit smaller than the natural borders. Accordingly, the pixels in the mask region

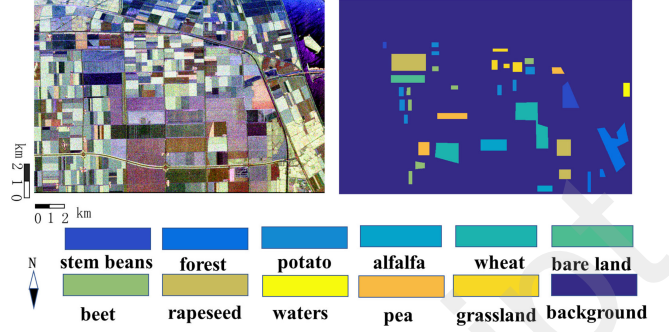


Fig. 3. Flevoland dataset: (left) Pauli image and (right) ground truth image.

are used to calculate the classification overall accuracy in the following experimentations.

B. Polarimetric Features

After the speckle filtering process with the refined Lee method of a 7×7 window size, all of these features are extracted through the recent version of open-source software PolSARpro v5.0 funded by the European Space Agency [42]. This set of 107 features includes almost all the polarimetric characteristics for classification, as listed in Table I. The polarimetric properties and statistical correlation between channels filtered by the refined Lee filter are well preserved, through the experimentation of many airborne and spaceborne polarimetric SAR data experiments in the literature of last 20 years. The filtering is applied to each term of the covariance matrix, including all the matrix elements. According to Lee *et al.*, to preserve the polarimetric signatures, all elements are filtered in the same way by the same amount. The span image is used to compute the weight in a selected edge aligned window [43].

Again, these 107 polarimetric features are not independent, since they are extracted originally from the coherency matrix $[T]$ or the equivalent covariance matrix $[C]$. However, we can still investigate on which ones play important role for land classifications.

C. Feature Combinations

1) *Feature Combination by the Proposed Method:* With the proposed selection strategy based on the CR, ten features are chosen out of the total 107 features under the correlation threshold of 50%, as shown in Table II. In fact, for the Flevoland data, different threshold has been studied between 0.5 and 0.8. Although the obtained feature subset changes a bit, the whole classification performances basically keep the same level. In our proposed method, the idea to set this threshold is to find out the representative features in the highly related feature groups and, meanwhile, keep the unrelated features. As shown in Fig. 4, the number of feature groups containing more than 20 feature pairs above the threshold changes from 6 to 4 as the threshold changes from 0.5 to 0.8. Considering both the representative related and unrelated features inside the whole feature set, the threshold is set to 0.5 in this article, in order to keep six features above it and four features below it.

TABLE II
TEN FEATURES SELECTED BY THE CR

Number	Feature
1	Yang3 Odd
2	Cloude T33
3	Combination HA
4	Entropy
5	Scattering mech. Entropy
6	Shannon Entropy I
7	Freeman Odd
8	VanZyl3 Odd
9	Yamaguchi3 Dbl
10	Yamaguchi3 Vol

TABLE III
TEN FEATURES SELECTED BY THE EUCLIDEAN DISTANCE

Number	Feature
1	C13 real
2	VanZyl3 Dbl
3	Yamaguchi Y4O Dbl
4	Holm2 T22
5	Barnes1 T22
6	T22
7	SPAN
8	Combination HA
9	Krogager Kd
10	MCSM Vol

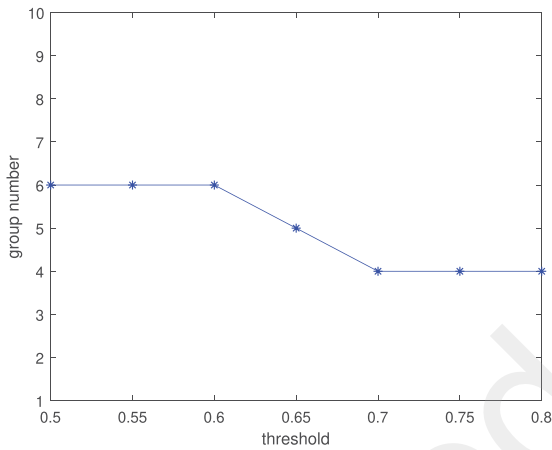


Fig. 4. Feature group number changes as the threshold increases.

It can be seen that the subset contain eight decomposition components (Yang3 Odd, Cloude T33, Combination HA, Entropy, Freeman Odd, VanZyl3 Odd, Yamaguchi3 Dbl, and Yamaguchi3 Vol) and two polarimetric descriptors (Scattering mech. Entropy and Shannon Entropy I). Within these eight target decomposition components, different kinds of scattering mechanisms are preserved. Yang Odd, Freeman Odd, and VanZyl Odd refer to single bounce or surface scattering, Yamaguchi Dbl corresponds to double bounce scattering, and Cloude T33 and Yamaguchi Vol are volume scattering. Besides, Entropy and Combination HA are the important parameters of eigenvalue decomposition. For this feature selection method, if the threshold is lower, the fewer features are preserved.

Yang decomposition is based on the Huynen decomposition, which is proposed as phenomenological theory. From this theory, it is possible for a single stationary target to define the “target structure diagram” and the nine “Huynen parameters,” which are all tied to a physical property of the target. The basic idea of the Huynen target decomposition theorem is to separate from the incoming data stream apart, which would be identified with a single average target and a residue

component called “N-target.” But if only the Odd component is observed, actually it is the T11 element itself, while VanZyl is one of the eigenvector-based target decomposition models. Since the eigenvalue problem is automatically basis invariant, such decompositions have been suggested as alternatives to the Huynen approach. This decomposition shows that the first two eigenvectors represent equivalent scattering matrices that can be interpreted in terms of odd and even numbers of reflections, so the odd component corresponds to the multiscattering with odd times. As to Freeman decomposition, which is one of the scattering-model-based decompositions, the three components correspond to three scattering mechanism models. The first component is described by the first-order Bragg surface scattering model of a slightly rough surface, in which the cross-polarized component is negligible. Because the above three decomposition methods belong to three types of methods, their odd components have some differences. Hence, it is possible to be kept in our selected feature subsets.

Besides the CR, the criterion of the Euclidean distance is also employed in our proposed framework to choose a polarimetric feature subset. When we use distance as the similarity measurement, it should be noted that normalization is very important in two aspects. On one hand, since the previous CR is a normalized parameter, which lies between 0 and 1, there is no need to do normalization. But for various distances, the values sometimes are very large. It is necessary to do normalization before setting the threshold of distances. In addition, the indication of distance is opposite to correlation. The larger the distance, the lower the similarity. Usually, the negative sign can be added to the normalized distances for next selection steps. On the other hand, while the distance is used, the features input to the classifiers should be normalized before the computation of distances. This is due to the fact that the distance is obtained between different pairs of features. Finally, ten features are chosen out of total 107 features under the threshold of 40%, as shown in Table III.

In this feature subset chosen by our proposed framework with the Euclidean distance, there are two matrix elements (C13 real and T22), seven decomposition components (VanZyl3 Dbl,

TABLE IV
TEN FEATURES SELECTED BY OSP

Number	Feature
1	MCSM Vol
2	Gamma
3	Polarisation Fraction
4	Huynen T11
5	Holm2 T22
6	Pedestal Height
7	MCSM Hlx
8	Neumann delta mod
9	MCSM Wire
10	Krogager Ks

Yamaguchi Y4O Dbl, Holm2 T22, Barnes1 T22, Krogager Kd, and MCSM Vol), and two polarimetric descriptors (SPAN and Combination HA). Different scattering mechanisms could be observed in this feature combination.

2) *Feature Combination by the OSP Method*: In order to compare different feature selection approaches, the same number of features is kept in the subset in the following experiments. With the introduced selection approach based on OSP, ten features are chosen out of the total 107 features, as shown in Table IV. Here, the subset contains eight decomposition components (MCSM Vol, Gamma, Huynen T11, Holm2 T22, MCSM Hlx, Neumann delta mod, MCSM Wire, and Krogager Ks) and two polarimetric descriptors (Polarization fraction and Pedestal). Again, different scattering mechanisms are preserved. Huynen T11 and Kragager Ks are single-bounce scattering; Holm T22 refers to double-bounce scattering. MCSM Vol, Hlx, and Wire corresponds to volume, helix, and wire scattering. There are also the components of eigenvalue decomposition and Neumann decomposition kept in the subset. Since there is a sequence of the selected features in this band selection method, once the number of required features is set, the selection process stops until the number reaches to the required number.

In the step of feature selection, the CR method mainly computes the correlation between each pair of two features, and it costs 184 s for the Flevoland data. The time for choosing features under a threshold is very short—less than 1 s, so as to be neglected. Compared to it, the OSP method costs 1190 s to select ten features for the same data.

D. Classification Results

First, the polarimetric SAR feature space is constructed by 107 features for each pixel. In the SVM classifier, sigma is equal to 0.1. In the NRS, the regularization parameter is 2.5, and the tuning parameter is -0.5 . For all the classification experiments in this article, the number of training samples is set to 300 for each class. And the number of test samples is shown in Table V.

1) *Classification of the Whole Feature Set*: In Fig. 5, the classification results of all 107 features using four classifiers

TABLE V
NUMBER OF TEST SAMPLES FOR EACH CLASS

Class	Test
stem beans	41211
forest	10109
potato	4848
alfalfa	5132
wheat	14587
bare land	3451
beet	3977
rapeseed	12469
water	5337
pea	2938
grassland	1219

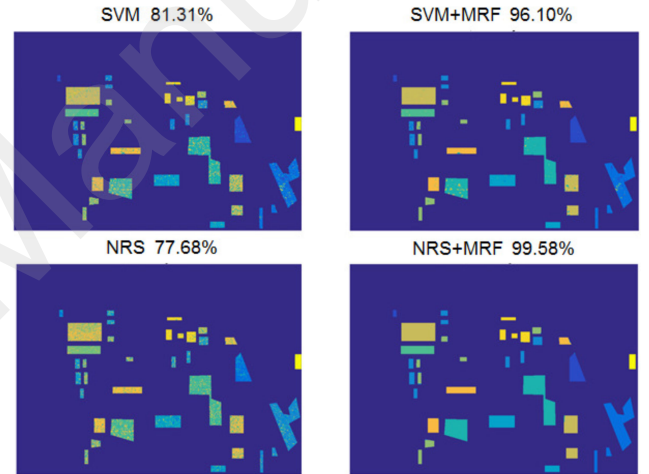


Fig. 5. Classification results using 107 features: (top left) SVM, (top right) SVM+MRF, (bottom left) NRS, and (bottom right) NRS+MRF.

are given. Using the SVM, the classification results is 81.31%, while the NRS gets lower accuracy of 77.68%. Both SVM+MRF and NRS+MRF classifiers have very fine overall classification accuracies of 96.1% and 99.58%, respectively, on the condition of training ratio lower than 4.5%.

2) *Classification of Feature Combination by the Proposed Method*: The classification results of the feature subset selected by CR criteria are shown in Fig. 6. It can be seen that for all these four classifiers, this feature combination can achieve similar classification performances. Using the SVM, the classification result is 81.20%, while the NRS gets the accuracy of 75.56%. Both SVM+MRF and NRS+MRF classifiers have very fine overall classification accuracies of 89.78% and 95.76%, respectively. Especially, for NRS+MRF, except that only a part of an area covered by wheat is wrongly classified into grassland, almost all pixels in other areas are classified correctly.

It should be noted that without MRF improvement, only SVM or NRS can still get good results from this feature subset.

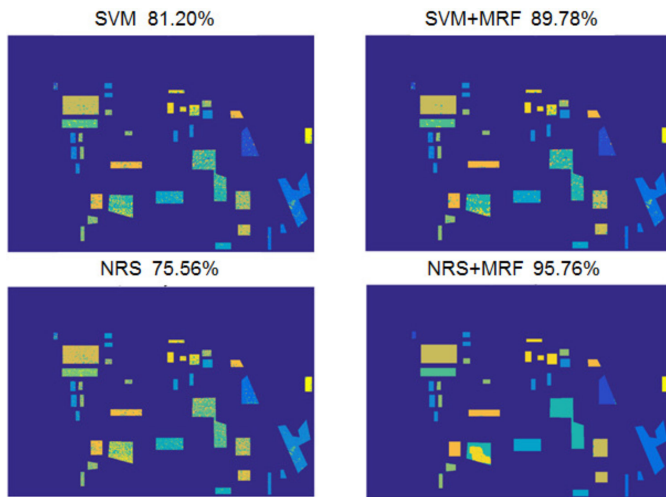


Fig. 6. Classification results using ten features selected by the proposed method with the CR: (top left) SVM, (top right) SVM+MRF, (bottom left) NRS, and (bottom right) NRS+MRF.

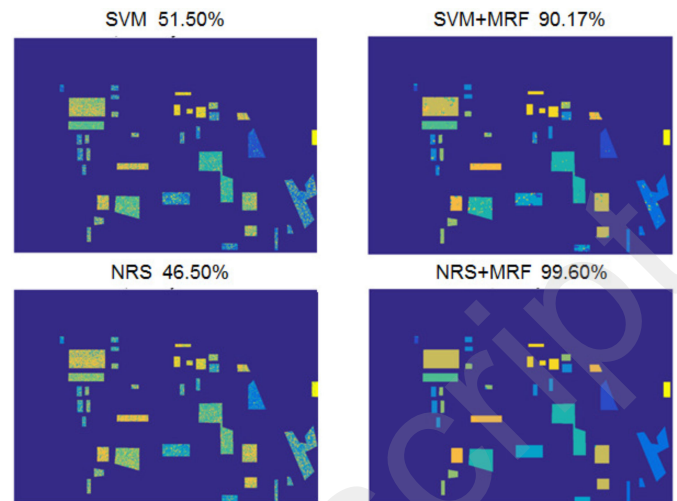


Fig. 8. Classification results using ten features selected by OSP: (top left) SVM, (top right) SVM+MRF, (bottom left) NRS, and (bottom right) NRS+MRF.

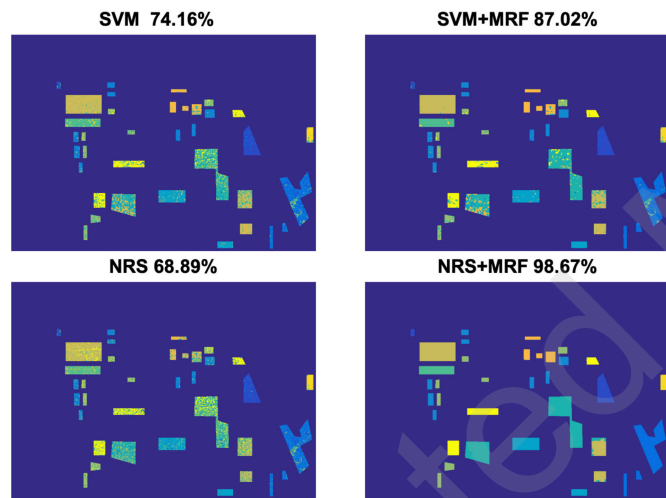


Fig. 7. Classification results using ten features selected by the proposed method with the Euclidean distance: (top left) SVM, (top right) SVM+MRF, (bottom left) NRS, and (bottom right) NRS+MRF.

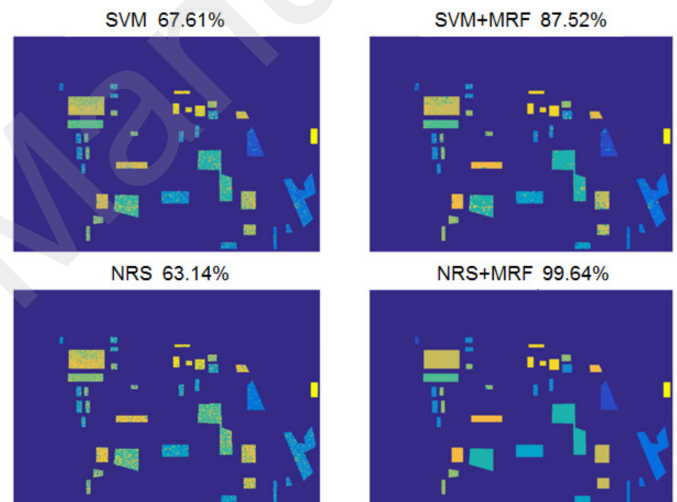


Fig. 9. Classification results using ten features constructed by PCA: (top left) SVM, (top right) SVM+MRF, (bottom left) NRS, and (bottom right) NRS+MRF.

For the NRS, the classification accuracy is even higher than the result of 107 features. This means that the feature subset preserves the polarimetric scattering information at the pixel level very well. Therefore, these ten polarimetric features can be considered as the effective feature combinations for this certain dataset.

The classification experiments with the Euclidean distance in our proposed framework are also implemented. The result can be seen in Fig. 7. We can see that by using the Euclidean distance, the selected feature subset has a bit lower classification performance compared to the CR. When the classifiers are SVM, NRS, SVM+MRF, and NRS+MRF, their overall classification accuracy is 74.16%, 68.89%, 87.02%, and 98.67%, respectively.

3) *Classification of Feature Combination by OSP*: The classification results of the feature subset selected by OSP and selection criteria are shown in Fig. 8. One obvious property

of this feature combination is that the classification results by SVM+MRF and NRS+MRF are very high even up to those with input of the whole feature set. However, by use of the simple classifiers SVM and NRS, the results are around 50% or lower than that. It means that at the pixel level, the features selected by the OSP method contain less discriminative information for classifications than the CR approach.

4) *Classification of Constructed Feature Combination by PCA*: For the aim of comparison, we also constructed feature combinations using the PCA method, although in this way there is no corresponding direct feature related to certain scattering components. The classification results of the feature subset selected by PCA are shown in Fig. 9. The performance of this feature subset is close to the one chosen by OSP when using SVM+MRF and NRS+MRF, but a bit better when employing SVM or NRS.

TABLE VI
CLASSIFICATION OVERALL ACCURACY OF FIVE FEATURE COMBINATIONS
WITH DIFFERENT CLASSIFIERS

Combinations	SVM	NRS	SVM+MRF	NRS+MRF
107	81.31	77.68	96.10	99.58
10CR	81.20	75.56	89.78	95.76
10ED	74.16	68.89	87.02	98.67
10OSP	51.50	46.50	90.17	99.60
10PCA	67.61	63.14	87.52	99.64

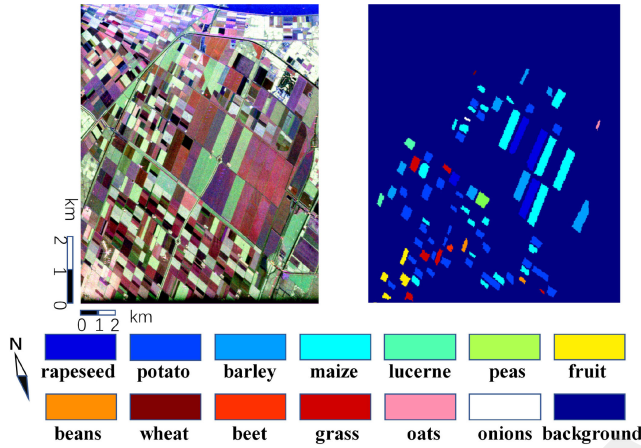


Fig. 10. Flevoland dataset II: (left) Pauli image and (right) ground truth image.

The classification results of above four feature combinations with various classifiers are displayed in Table VI. To conclude, with complex classifiers like NRS+MRF/SVM+MRF, which utilize spatial information, the feature subset selection by OSP and constructed features by PCA get very good performance, while, on the other hand, with simple classifiers using pixel-level polarimetric information, the feature combination obtained by CR criteria performs better. It means the CR approach keeps more polarimetric scattering properties useful for classifications.

Another PolSAR dataset is acquired in the same area by the NASA/JPL AIRSAR system in June 1991. There are 14 different terrain types measured with ground truth, which are rapeseed, lucerne, peas, wheat, grass, oats, onions, potato, beet, maize, fruit, barley, beans, and background. The size of this dataset is 1024×1279 pixels, and the number of pixels with the ground truth data is around 122 928. The Pauli image and the ground truth image are shown in Fig. 10.

Similar data processing steps are implemented in this dataset, including speckle filtering and polarimetric feature extraction. First, the polarimetric SAR feature space is constructed by 107 features for each pixel. Then, the classification experiments with the whole feature space are done. The parameters inside the classifiers are basically the same with the first dataset. The difference is that the number of training samples is set to 1% for each class, instead of 300 for each class. This is because there are some small classes containing less pixels than 300. The number of test samples is shown in Table VII. The classification results of

TABLE VII
NUMBER OF TEST SAMPLES FOR EACH CLASS IN FLEVOLAND II DATA

Class	Test
rapeseed	17621
potato	32854
barley	16475
maize	40019
lucerne	2336
peas	2409
fruit	4156
beans	1258
wheat	120
beet	559
grass	5800
oats	360
onions	247

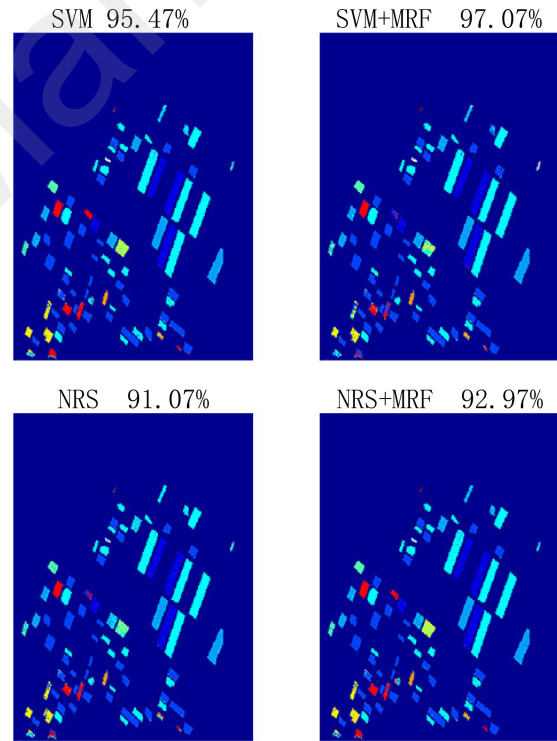


Fig. 11. Classification results of dataset II using 107 features: (top left) SVM, (top right) SVM+MRF, (bottom left) NRS, and (bottom right) NRS+MRF.

all 107 features using four classifiers are given in Fig. 11. Using the SVM, the classification result is 95.47%, while the NRS gets lower accuracy of 91.07%, SVM+MRF and NRS+MRF classifiers have better overall classification accuracies of 97.07% and 92.97%, respectively.

Following that, feature combination by the CR method is implemented. Nine features are chosen out of the total 107

TABLE VIII
NINE FEATURES SELECTED BY THE CR WITH DATASET II

Number	Feature
1	T22
2	lamda
3	VanZyl3 Odd
4	Yamaguchi4 Odd
5	SPAN
6	MCSM wire
7	Yamaguchi3 Odd
8	MCSM Odd
9	Huynen T33

T22 corresponds to double bounce scattering, Huynen T33 is volume scattering, and MCSM wire refers to wire scattering. SPAN is the whole scattering power, while lamda is an important parameter of eigenvalue decomposition.

The classification results of the feature subset selected by CR criteria are shown in Fig. 12. It can be seen that for all these four classifiers, this feature combination can achieve good classification performances. Even a bit lower than the results of 107 features, however, it should be noted that here only nine features are used as input of the classifiers. Using the SVM, the classification result is 83.40%, while NRSs gets the accuracy of 84.03%. SVM+MRF and NRS+MRF classifiers have a bit better classification accuracy of 87.12% and 88.32%, respectively.

Even without MRF improvement, only SVM or NRS can still get good results from this feature subset. This means that the feature subset preserves the polarimetric scattering information at the pixel level very well. Therefore, these nine polarimetric features can be considered as the effective feature combinations for this certain dataset.

VI. CONCLUSION AND FUTURE WORK

Studying the selection strategy of the polarimetric feature combination is of significance, since the computation cost can be reduced a lot. In fact, it depends on both the features of specific datasets and the classifier. In this article, we focused on the polarimetric features of PolSAR data. A set of 107 polarimetric features is extracted, and the polarimetric feature vector is constructed, which almost covers all the representative and recent ones. Two polarimetric feature selection methods are used to obtain a subset of ten features, which preserve the information of different scattering mechanisms. The proposed framework uses the CR and the Euclidean distance as criteria for similarity measurement and selects the representative ones within each feature group above the threshold; meanwhile, it keeps the unrelated features under the threshold. Besides, the influence of the threshold is also analyzed. OSP is introduced to PolSAR data from the hyperspectral area, but the performance is not as good as the proposed one. Besides, the computation time is longer. As to PCA, it is not the kind of method to keep the features with physical meaning, because the constructed feature corresponds to no scattering mechanisms. Experiments are implemented to assess these two methods, as well as to compare them with the principal components constructed by the PCA approach. With the subset of ten features, good classification results are achieved by using SVM+MRF and NRS+MRF classifiers. However, when the SVM and NRS are used as classifiers without combining spatial information by using the MRF, the ten features selected by CR criteria have a much better classification performance than that of the other OSP criteria. It means that better polarimetric scattering information at the pixel level is preserved through the proposed selection method for the application of vegetation classification. Another dataset in Flevoland area is tested using the proposed method to select nine polarimetric features, and the experiments show that less than 9% of features could achieve the comparable classification accuracy with the whole feature set.

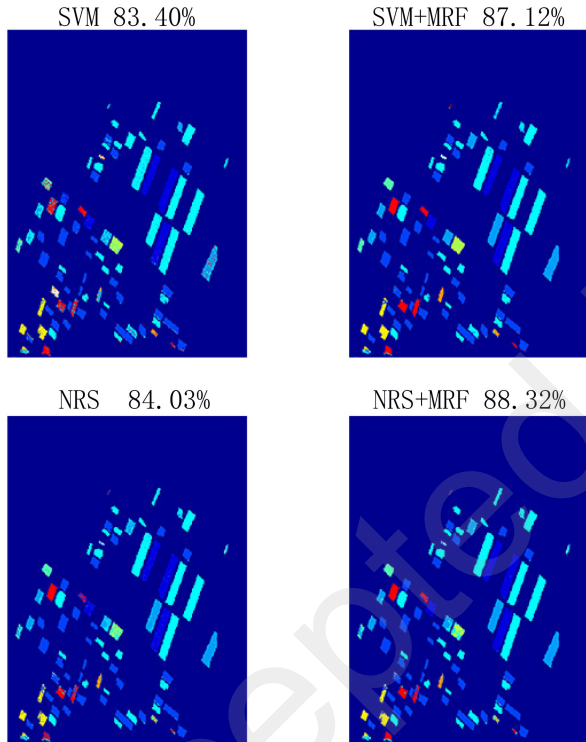


Fig. 12. Classification results of dataset II using nine features selected by the CR: (top left) SVM, (top right) SVM+MRF, (bottom left) NRS, and (bottom right) NRS+MRF.

features under the correlation threshold of 50%, as shown in Table VIII. Here, in this experiment, nine instead of ten features are kept due to the fact that there are nine independent elements of the coherency or covariance matrix, hence nine independent components of the whole extracted feature groups. It can be seen that the subset contains seven decomposition components (lamda, Huynen T33, MCSM wire, MCSM Odd, VanZyl3 Odd, Yamaguchi3 Odd, and Yamaguchi4 Odd), one matrix element T22, and one polarimetric descriptor SPAN. Within these nine features, different kinds of scattering mechanisms are preserved. Odd components refer to single bounce or surface scattering,

In the future, other similarity criteria could be tested in our proposed framework, in order to find out the best ones especially for describing the relations among lots of polarimetric features. Besides, new polarimetric features inside PolSARpro v.5.1 will be included in experiments. The experiments on other datasets will also be shown in future work, in order to see the difference in feature combinations for various data, and to reveal the underlying reason due to scattering mechanisms.

REFERENCES

- [1] Y. Bai, D. Peng, X. Yang, L. Chen, and W. Yang, "Supervised feature selection for polarimetric SAR classification," in *Proc. Int. Conf. Signal Process.*, 2014, pp. 1006–1010.
- [2] H. Wang, Y. Pi, and Z. Cao, "Unsupervised classification of polarimetric SAR images based on ICA," in *Proc. Int. Conf. Natural Comput.*, 2007, vol. 3, pp. 576–582.
- [3] V. V. Chamundeswari, D. Singh, and K. Singh, "An analysis of texture measures in PCA-Based unsupervised classification of SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 214–218, Apr. 2009.
- [4] L. Chen, W. Yang, Y. Liu, and H. Sun, "Feature evaluation and selection for polarimetric SAR image classification," in *Proc. Int. Conf. Signal Process.*, 2010, pp. 2202–2205.
- [5] G. A. Haddadi, M. R. Sahebi, and A. Mansourian, "Polarimetric SAR feature selection using a genetic algorithm," *Can. J. Remote Sens.*, vol. 37, no. 1, pp. 27–36, 2011.
- [6] M. Tao, F. Zhou, and Z. Zhang, "Tensor based dimension reduction for polarimetric SAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2014, pp. 2802–2805.
- [7] B. Banerjee, A. Bhattacharya, and K. M. Buddhiraju, "A generic land-cover classification framework for polarimetric SAR images using the optimum Touzi decomposition parameter subset—An insight on mutual information-based feature selection techniques," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1167–1176, Apr. 2014.
- [8] S. Gupta, S. Kumar, A. Garg, D. Singh, and N. S. Rajput, "Class wise optimal feature selection for land cover classification using SAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2016, pp. 68–71.
- [9] S. Fukuda and H. Hirose, "Support vector machine classification of land cover: Application to polarimetric SAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2001, vol. 1, pp. 187–189.
- [10] B. Koetz, F. Morsdorf, S. V. Der Linden, T. Curt, and B. Allgower, "Multi-source land cover classification for forest fire management based on imaging spectrometry and LiDAR data," *Forest Ecol. Manage.*, vol. 256, no. 3, pp. 263–271, 2008.
- [11] G. Pajares, C. Lopez-Martinez, F. Sanchez-Liado, and I. Molina, "Improving Wishart classification of polarimetric SAR data using the Hopfield neural network optimization approach," *Remote Sens.*, vol. 4, no. 11, pp. 3571–3595, 2012.
- [12] A. Masjedi, M. J. V. Zoj, 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.
- [13] F. Zhang *et al.*, "Nearest-regularized subspace classification for PolSAR imagery using polarimetric feature vector and spatial information," *Remote Sens.*, vol. 9, no. 11, 2017, Art. no. 1114. [Online]. Available: <http://www.mdpi.com/2072-4292/9/11/1114>
- [14] A. Schick, M. Bauml, and R. Stiefel, "Improving foreground segmentations with probabilistic superpixel Markov random fields," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 27–31.
- [15] P. Ghamisi, J. A. Benediktsson, and M. O. Ulfarsson, "Spectral-spatial classification of hyperspectral images based on hidden Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2565–2574, May 2014.
- [16] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton, FL, USA: CRC Press, 2009.
- [17] J. Li, K. Cheng, S. Wang, F. Morstatter, and H. Liu, "Feature selection: A data perspective," *ACM Comput. Surv.*, vol. 50, no. 6, 2016, Art. no. 94.
- [18] J. R. Huynen, "Phenomenological theory of radar targets," Ph.D. dissertation, Faculty Elect. Eng., Math. Comput. Sci., Tech. Univ. Delft, Delft, The Netherlands, 1970.
- [19] R. M. Barnes, "Roll-invariant decompositions for the polarization covariance matrix," in *Proc. Polarimetry Technol. Workshop*, Redstone Arsenal, AL, USA, 1988, pp. 1–15.
- [20] J. Yang, Y.-N. Peng, Y. Yamaguchi, and H. Yamada, "On Huynen's decomposition of a Kennaugh matrix," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 3, pp. 369–372, Jul. 2006.
- [21] S. R. Cloude, "Target decomposition theorems in radar scattering," *Electron. Lett.*, vol. 21, no. 1, pp. 22–24, 1985.
- [22] A. Freeman, J. J. van Zyl, J. D. Klein, H. A. Zebker, and Y. Shen, "Calibration of Stokes and scattering matrix format polarimetric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 3, pp. 531–539, May 1992.
- [23] 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.
- [24] W. A. Holm and R. M. Barnes, "On radar polarization mixed target state decomposition techniques," in *Proc. IEEE Nat. Radar Conf.*, 1988, pp. 249–254.
- [25] B. Zou, D. Lu, L. Zhang, and W. M. Moon, "Eigen-decomposition-based four-component decomposition for PolSAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 3, pp. 1286–1296, Mar. 2016.
- [26] A. Freeman, "Fitting a two-component scattering model to polarimetric SAR data from forests," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 8, pp. 2583–2592, Aug. 2007.
- [27] M. Neumann, L. Ferro-Famil, and E. Pottier, "A general model-based polarimetric decomposition scheme for vegetated areas," in *Proc. Int. Workshop Sci. Appl. SAR Polarimetry & Polarimetric Interferometry*, 2009, p. 668.
- [28] E. Krogager, "New decomposition of the radar target scattering matrix," *Electron. Lett.*, vol. 26, no. 18, pp. 1525–1527, Aug. 1990.
- [29] Y. Yamaguchi, Y. Yajima, and H. Yamada, "A four-component decomposition of PolSAR images based on the coherency matrix," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 3, pp. 292–296, Jul. 2006.
- [30] 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.
- [31] R. Touzi, "Target scattering decomposition of one-look and multi-look SAR data using a new coherent scattering model: The TSVM," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2004, vol. 4, pp. 2491–2494.
- [32] S. Allain, L. Ferro-Famil, and E. Pottier, "A polarimetric classification from PolSAR data using SERD/DERD parameters," in *Proc. 6th Eur. Conf. Synth. Aperture Radar*, Dresden, Germany, 2006. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-00101294>
- [33] T. L. Ainsworth, S. R. Cloude, and J. S. Lee, "Eigenvector analysis of polarimetric SAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2002, vol. 1, pp. 626–628.
- [34] S. L. Durden, J. J. V. Zyl, and H. A. Zebker, "The unpolarized component in polarimetric radar observations of forested areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 2, pp. 268–271, Mar. 1990.
- [35] J. Praks and M. Hallikainen, "A novel approach in polarimetric covariance matrix eigendecomposition," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. Taking Pulse Planet: Role Remote Sens. Manag. Environ. Proc.*, 2000, vol. 3, pp. 1119–1121.
- [36] J. J. V. Zyl, H. A. Zebker, and C. Elachi, "Imaging radar polarization signatures: Theory and observation," *Radio Sci.*, vol. 22, no. 4, pp. 529–543, 1987.
- [37] D. C. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2002.
- [38] H. Ren and C. I. Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1232–1249, Oct. 2003.
- [39] Q. Du and H. Yang, "Similarity-based unsupervised band selection for hyperspectral image analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 564–568, Oct. 2008.
- [40] W. Li, E. W. Tramel, and S. Prasad, "Nearest regularized subspace for hyperspectral classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 477–489, Jan. 2014.
- [41] Y. Y. Boykov and M. P. Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in ND images," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2001, pp. 105–112.
- [42] European Space Agency, PolSARpro version 5.1, 2017. [Online]. Available: <http://earth.esa.int/web/polsarpro/home>
- [43] J. S. Lee, M. R. Grunes, and G. de Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2363–2373, Sep. 1999.



Qiang Yin (S'06–M'08) received the B.S. degree in electronic and information engineering from the Beijing University of Chemical Technology, Beijing, China, in 2004, and the M.S. and Ph.D. degrees in signal and information processing from the Institute of Electronics, Chinese Academy of Science, Beijing, in 2008 and 2016, respectively.

From 2008 to 2013, she was a Research Assistant with the Institute of Electronics, Chinese Academy of Sciences. From 2014 to 2015, she was a Research Fellow with the European Space Agency, Roma, Italy.

She is currently an Associate Professor with the College of Information Science and Technology, Beijing University of Chemical Technology. Her research interests include polarimetric/polarimetric interferometric synthetic aperture radar processing and applications.



Fan Zhang (S'07–M'10–SM'17) received the B.E. degree in communication engineering from the Civil Aviation University of China, Tianjin, China, in 2002, the M.S. degree in signal and information processing from Beihang University, Beijing, China, in 2005, and the Ph.D. degree in signal and information processing from the Institute of Electronics, Chinese Academy of Science, Beijing, in 2008.

He is currently a Full Professor of Electronic and Information Engineering with the College of Information Science and Technology, Beijing University of Chemical Technology, Beijing. His research interests include synthetic aperture radar signal processing, image processing, and high-performance computing.

Dr. Zhang is an Associate Editor for IEEE ACCESS and a Reviewer for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING,

IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, and *International Journal of Antennas and Propagation*.



Wen Hong (M'03–SM'16) received the M.S. degree in electronic engineering from Northwestern Polytechnical University, Xi'an, China, in 1993, and the Ph.D. degree from Beihang University, Beijing, China, in 1997.

From 1997 to 2002, she was a Faculty Member in signal and information processing with the Department of Electrical Engineering, Beihang University. In between, she was a Guest Scientist with DLR-HF, Wessling, Germany, from 1998 to 1999, for one year. Since 2002, she has been a Scientist

with the Science and Technology on Microwave Imaging Laboratory and a Supervisor of the Graduate Student Program with the Chinese Academy of Sciences, Beijing. Her main research interests include polarimetric/polarimetric interferometric synthetic aperture radar (SAR) data processing and application, three-dimensional SAR signal processing, circular SAR signal processing, SAR polarimetry application, and sparse microwave imaging with compressed sensing.



Eric Pottier (M'95–SM'06–F'11) received the M.Sc. and Ph.D. degrees in signal processing and telecommunication from the University of Rennes 1, Rennes, France, in 1987 and 1990, respectively, and the Habilitation from the University of Nantes, Nantes, France, in 1998.

Since 1999, he has been a Full Professor with the University of Rennes 1, where he is currently the Director of the Institute of Electronics and Telecommunications of Rennes (CNRS UMR 6164). He has authored ten chapters in books, more than 80 papers in refereed journals, and has presented more than 380 papers during international conferences, symposiums, and workshops. He has presented advances courses and seminars on radar polarimetry to a wide range of organizations and events. He has authored/coauthored a book with Dr. J.-S. Lee: *Polarimetric Radar Imaging: From Basics to Applications* (CRC Press, 2009). His research interests include analog electronics, microwave theory, radar imaging with an emphasis on radar polarimetry, wide spectrum of areas from radar image processing (synthetic aperture radar (SAR) and inverse SAR), polarimetric scattering modeling, supervised/unsupervised polarimetric segmentation, and classification to fundamentals and basic theory of polarimetry.

Dr. Pottier was recognized as an IEEE Geoscience and Remote Sensing Society (GRSS) Distinguished Lecturer in 2014. He was the recipient of the award for a Very Significant Contribution in the Field of Synthetic Aperture Radar (2000 European Conference on Synthetic Aperture Radar), the 2007 IEEE GRSS Letters Prize Paper Award, the 2007 IEEE GRSS Education Award in recognition of his significant educational contributions to Geoscience and Remote Sensing, the 2012 Einstein Professorship from the Chinese Academy of Science, and the CNFRS-URSI Medal under the auspices of the French Academy of Sciences in 2015.