SUPERVISED POLSAR IMAGE CLASSIFICATION BY COMBINING MULTIPLE FEATURES

Xiayuan Huang[1], *Xiangli Nie*[1], *Hong Qiao*[1], *Bo Zhang*[2]

[1]State Key Lab of Management and Control for Complex Systems, Institute of Automation,
Chinese Academy of Sciences, Beijing 100190, China,
[2]State Key Lab of Scientific and Engineering Computing and Institute of Applied Mathematics,
AMSS, Chinese Academy of Sciences, Beijing 100190, China

ABSTRACT

For polarimetric synthetic aperture radar (PolSAR) image classification, each pixel can be represented by multiple features from different perspectives, such as polarimetric feature (PF), texture feature (TF) and color feature (CF). Both multi-view canonical correlation analysis (MCCA) and multi-view spectral embedding (MSE) are two unsupervised multi-view subspace learning methods which search for different projection matrices for different features to combine multiple features in a common low-dimensional feature space. However, MCCA emphasizes the correlation of multiple features and MSE learns the complementarity of multiple features. To deeply learn the relation of multiple features, we incorporate MCCA with MSE based on the label information and a symmetric version of revised Wishart (SRW) distance for supervised PolSAR image feature extraction. Experimental results confirm that the proposed method can improve the classification performance.

***Index Terms*—** multiple features, MCCA, MSE, feature extraction, PolSAR image classification

1. INTRODUCTION

Polarimetric synthetic aperture radar (PolSAR) uses multiple channels to obtain abundant information of objects. Land cover classification is an important application for PolSAR [1, 2]. Feature extraction is a crucial factor of PolSAR image classification. The classification performance largely depends on whether the extracted features are appropriate. This paper deals with supervised PolSAR image feature extraction based on the label information.

Some previous works directly use the PolSAR data as the feature for image classification [1]. Afterwards, many parameters produced by target decomposition methods such as Freeman decomposition, Krogager decomposition, Pauli decomposition and Huynen decomposition are applied as the feature [3, 4]. We call the above feature as polarimetric feature (PF) which mainly describes the scattering and physical characteristics of targets.

This work was partly supported by the National Natural Science Foundation of China under Grant 61802408, Grant 61379093, Grant 91648205 and Grant 61602483.

Texture feature (TF) is an important visual feature and has also served as the feature for PolSAR image classification [5, 6]. Moreover, TF was usually incorporated with PF for PolSAR image classification [5, 6]. Recently, color feature (CF) has also been used for PolSAR image classification [7]. In [7], Uhlmann *et al.* summarized the aforementioned three features and evaluated the role of CF for PolSAR image classification. CF was extracted from the Pauli RGB image [1] and experimental results showed that CF helped improve the classification performance [7].

To deal with multiple features, some conventional dimensionality reduction methods concatenated multiple feature vectors together as a new vector, such as principle components analysis (PCA) [8], independent component analysis (ICA) [9], Laplacian eigenmaps (LE) [3] and supervised graph embedding (SGE) [10], which ignored the correlation and difference of multiple features. To this end, some multi-view subspace learning methods have been proposed to learn the relation of multiple features [11, 12, 13]. Canonical correlation analysis (CCA) was a typical unsupervised multi-view subspace learning method for two features, which searched for two projection matrices to separately project two features to a common low-dimensional space by maximizing the correlation of the two features [11]. In our previous work [14], we proposed local discriminant canonical correlation analysis (LDCCA) to extract features based on PF and TF of PolSAR images. LDCCA exploited the label information and preserved the local structure of the data, which resulted in a good classification performance. Multi-view CCA (MCCA) [12] extended CCA to multiple features by maximizing the total correlations between any two features. Multi-view discriminative analysis (MvDA) [13] used the label information and also dealt with multiple features. Xia *et al.* proposed multi-view spectral embedding (MSE), which utilized adaptive weights to explore the complementarity of different features [15]. And MSE has been used to combine multiple features for hyperspectral remote sensing image classification [16].

In this paper, we propose a novel method to combine PF, TF and CF for supervised PolSAR image classification. We firstly modify MSE to make it appropriate for the PolSAR data by using the label information and a symmetric version of revised Wishart (SRW) distance. Then the modified MSE is combined with MCCA to learn deeply the relation of multiple features. Experimental results show the superiority of the proposed method.

2. THE PROPOSED METHOD

We exploit three features (i.e., PF, TF and CF) for supervised PolSAR image classification. The details of PF and TF can be found in our previous work [14]. Moreover, CF which is extracted from the Pauli RGB image [1] consists of six descriptors (mean, variance, skewness, kurtosis, energy and entropy) of the RGB color histogram and the HSV color histogram ($6 \times 3 \times 2 = 36$); the four dominant color ratios of the HSV color histogram [7]. Therefore, the number of CF is 40. Similarly to TF, CF is also computed in a sliding 11×11 window [14]. Therefore, PF, TF and CF are denoted as $X_1 \in R^{58 \times n}$, $X_2 \in R^{56 \times n}$ and $X_3 \in R^{40 \times n}$ for n samples, i.e. n pixels. We aim to search for three projection matrices W_1, W_2, W_3 to project X_1, X_2, X_3 to a common low-dimensional feature space for the final classification. In the following description of the proposed method, we use v features instead of 3 features for general cases.

We use v matrices $\{X_i \in R^{D_i \times n}\}_{i=1}^{v}$ to represent v features for n samples, where D_i is the dimensionality for the ith feature. The proposed method aims to search for v projection matrices $\{W_i = [w_{i1}, w_{i2}, \cdots, w_{id}] \in R^{D_i \times d}\}_{i=1}^{v} (d < D_i)$, where d is the reduced dimensionality.

Similar to MSE, we firstly construct graphs to represent the structure information of the data. MSE constructs v graphs corresponding to v features based on the Euclidean distance. That is, the neighboring samples of each pixel are sought based on the Euclidean distance of different feature vectors. However, for the PolSAR data, the Wishart distance which is derived from the distribution of the PolSAR data is commonly used for PolSAR image classification. Then it is more appropriate to use the Wishart distance to construct the graph [1]. Specifically, we exploit a symmetric version of revised Wishart (SRW) distance to seek for neighbouring samples and then construct the common graph for different features [14, 17]. The detailed definition of the SRW distance can be found in [17]. For pixel i and j, which can be characterized by $p \times p$ covariance matrix C_i and C_j, the SRW distance is computed as follows:

$$d_{SRW}(C_i, C_j) = \frac{1}{2}(\mathrm{tr}(C_i^{-1} C_j) + \mathrm{tr}(C_j^{-1} C_i)) - p. \quad (1)$$

In addition, to exploit the label information, the neighboring samples are searched in one class. We search for k neighbouring samples of each pixel based on the SRW distance in one class. If sample j is one of sample j's k neighbouring samples, $G_{ij} = 1$, else $G_{ij} = 0$. Then, the modified MSE based on the SRW distance and the label information is as follows:

$$\begin{aligned} \min_{\{W_i\}_{i=1}^{v}, \{\alpha_i\}_{i=1}^{v}} & \sum_i \alpha_i^r \mathrm{tr}\left(W_i^T X_i L X_i^T W_i\right) \\ \text{s.t.} & \sum_i \alpha_i = 1, \alpha_i \geq 0 \end{aligned} \quad (2)$$

where $L = D - G$, D is a diagonal matrix and $D_{ii} = \sum_{j=1}^{n} G_{ij}$. $\alpha_1, \alpha_2, \cdots, \alpha_v$ are weights for v features which are employed to learn the complementarity of different features. A larger α_i means that the ith feature plays a more important role in PolSAR image classification. r is a parameter and usually $r > 1$. If $r = 1$, the solution will be $\alpha_i = 1$ corresponding to the minimum $\mathrm{tr}\left(W_i^T X_i L X_i^T W_i\right)$ and other view weights equal to 0.

Afterwards, problem (2) is combined with MCCA by the parameter β. Therefore, the final objective function of the the proposed method is:

$$\begin{aligned} \min_{\{W_i\}_{i=1}^{v}, \{\alpha_i\}_{i=1}^{v}} & \mathrm{tr}\left(\sum_i \alpha_i^r W_i^T X_i L X_i^T W_i - \beta \sum_{i<j}^{v} W_i^T X_i X_j^T W_j\right) \\ \text{s.t.} \quad & W_i^T X_i X_i^T W_i = I, i = 1, 2, \cdots, v \\ & \sum_i \alpha_i = 1, \alpha_i \geq 0 \end{aligned} \quad (3)$$

In the above problem, both the v projection matrices $W_1, W_2, \cdots, W_v$ and v weights $\alpha_1, \alpha_2, \cdots, \alpha_v$ need to be solved. It is difficult to solve problem (3) directly. Therefore, we adopt an alternating optimization method to obtain a solution by iteratively updating $\{W_i\}_{i=1}^{v}$ and $\{\alpha_i\}_{i=1}^{v}$. Initially, $\alpha_i = 1/v, i, = 1, 2, \cdots, v$. Then the alternating process is as follows:

1) Fix $\{\alpha_i\}_{i=1}^{v}$ and optimize $\{W_i\}_{i=1}^{v}$. The problem (3) becomes

$$\begin{aligned} \min_{\{W_i\}_{i=1}^{v}} & \mathrm{tr}\left(\sum_i \alpha_i^r W_i^T X_i L X_i^T W_i - \beta \sum_{i<j}^{v} W_i^T X_i X_j^T W_j\right) \\ \text{s.t.} \quad & W_i^T X_i X_i^T W_i = I, i = 1, 2, \cdots, v \end{aligned} \quad (4)$$

For arbitrary column vectors $w_1, w_2, \cdots, w_v$ corresponding to v projection matrices $W_1, W_2, \cdots, W_v$, the Lagrangian function is

$$\begin{aligned} L(\{w_i\}_{i=1}^{v}, \mu) = & \sum_i \alpha_i^r w_i^T X_i L X_i^T w_i - \beta \sum_{i<j}^{v} w_i^T X_i X_j^T w_j - \\ & \mu(w_1^T X_1 X_1^T w_1 - 1) - \cdots - \mu(w_v^T X_v X_v^T w_v - 1) \end{aligned} \quad (5)$$

where μ is the is the Lagrangian multiplier. Set $\partial L / \partial w_i = 0$, i.e.,

$$\frac{\partial L}{\partial w_i} = 2\alpha_1^r X_i L X_i^T w_i - \beta \sum_{j>i} X_i X_j^T w_j - \beta \sum_{j<i} X_i X_j^T w_j - 2\mu X_i X_i^T w_i = 0 \quad (6)$$

The above equation can be written in a matrix form as follows:

$$Aw \quad = \quad 2\mu B \quad (7)$$

$$A = \begin{bmatrix} 2\alpha_1^r X_1 L X_1^T & -\beta X_1 X_2^T & \cdots & -\beta X_1 X_v^T \\ -\beta X_2 X_1^T & 2\alpha_2^r X_2 L X_2^T & \cdots & -\beta X_2 X_v^T \\ \vdots & \vdots & \ddots & \vdots \\ -\beta X_v X_1^T & -\beta X_v X_2^T & \cdots & 2\alpha_v^r X_v L X_v^T \end{bmatrix}$$

$$B = \begin{bmatrix} X_1 X_1^T & & & \\ & X_2 X_2^T & & \\ & & \ddots & \\ & & & X_v X_v^T \end{bmatrix}$$

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_v \end{bmatrix}$$

Through the eigenvalue decomposition, we acquire d eigenvectors for smallest d eigenvalues. $W_1, W_2, \cdots, W_v$ are generated as follows:

$$\begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_v \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1d} \\ w_{21} & w_{22} & \cdots & w_{21} \\ \vdots & \vdots & \vdots & \vdots \\ w_{v1} & w_{v2} & \cdots & w_{vd} \end{bmatrix}$$

2) Fix $\{W_i\}_{i=1}^v$ and optimize $\{\alpha_i\}_{i=1}^v$. The problem (3) is equivalent to

$$\min_{\{\alpha_i\}_{i=1}^v} \sum_i \alpha_i^r \text{tr}\left(W_i^T X_i L X_i^T W_i\right) \quad \text{s.t.} \sum_i \alpha_i = 1, \alpha_i \geq 0 \quad (8)$$

Without regard to $\alpha_i \geq 0$,

$$L(\{\alpha_i\}_{i=1}^v, \lambda) = \sum_i \alpha_i^r \text{tr}\left(W_i^T X_i L X_i^T W_i\right) - \lambda(\sum_i \alpha_i - 1) \quad (9)$$

where λ is the Lagrangian multiplier. Sequently, set $\partial L / \partial \alpha_i = 0$, i.e.,

$$\frac{\partial L}{\partial \alpha_i} = \sum_i r \alpha_i^{r-1} \text{tr}\left(W_i^T X_i L X_i^T W_i\right) - \lambda = 0 \quad (10)$$

Table 2. Classification accuracies on two data sets based on PF, TF and CF

	Flevoland	San Francisco Bay
PF	0.8420	0.8390
TF	0.8777	0.7445
CF	0.8641	0.8100

Additionally, $\sum_i \alpha_i = 1$, then we obtain

$$\alpha_i = \frac{\left(\text{tr}\left(W_i^T X_i L X_i^T W_i\right)\right)^{\frac{1}{1-r}}}{\sum_i \left(\text{tr}\left(W_i^T X_i L X_i^T W_i\right)\right)^{\frac{1}{1-r}}} \quad (11)$$

Moreover, tr $\left(W_i^T X_i L X_i^T W_i\right) \geq 0$, thus the result meets the the non-negative condition. The convergence and the computational complexity analysis of the proposed can refer to previous works [18, 15, 13].

Finally, we get the projection matrices $\{W_i\}_{i=1}^v$ via the above alternative optimization algorithm. Based on the obtained v projection matrices, we concatenate the v low-dimensional features like $[W_1^T X_1, W_2^T X_2, \cdots, W_v^T X_v]^T$ for the following classification.

3. EXPERIMENTAL RESULTS

Experiments are conducted on two PolSAR data sets. The first data set is the Flevoland data set with size 200×320 pixels. This data set is a part of the cropland from Flevoland, which consists of nine classes of land covers: rapeseed, grass, stem beans, lucerne, winter wheat I, potatoes, winter wheat II, bare soil, sugar beat. The Pauli RGB image and the ground truth are displayed in Figs. 1(a-b). The second one is the San Francisco Bay data set consisting of 900×1024 pixels. We consider four classes of land covers (grass, buildings, sea and mountains). Some pathes from four classes are used for experiments as shown in Fig. 2(a). The ground truth is shown in Fig. 2(b). The refined Lee filter is utilized to denoise two data sets in a window of size 7×7 [1]. 100 pixels from each class are used as the training set and the rest of pixels are used as the testing set.

Firstly, PF, TF and CF are separately used for PolSAR classification with the nearest neighbor (NN) classifier and the classification accuracies are shown in Tab. 2. Furthurmore, we employ MCCA, MSE, MvDA and multi-view LDCCA (MLDCCA), supervised Wishart classifier (SWC) to compare with the proposed method. MLDCCA is the multi-view extension of our LDCCA [14] like MCCA. SWC is always used as a baseline when comparing the performance of PolSAR image classification methods. About some parameters, we empirically set $d = 10$, $r = 10$, $k = 10$ for a good performance. In addition, $\beta = 1$ for the the first data set. $\beta = 5$ for the second one. The numerical and visual results of the

Table 1. Classification accuracies comparison on the Flevoland data set

Method	Potatoes	Grass	Beet	Lucerne	Wheat I	Wheat II	Stem beans	Bare soil	Rapeseed	Accuracy
MCCA	0.9923	0.8294	0.9895	0.5405	0.9190	0.9856	0.9967	1.0000	0.8750	0.9067
MSE	0.9849	0.8178	0.9942	0.7549	0.9777	0.9940	0.9767	1.0000	0.9649	0.9298
MvDA	0.9919	0.8319	0.9935	0.7654	0.9519	0.9952	0.9967	1.0000	0.9825	0.9318
MLDCCA	0.9911	0.9608	0.9931	0.8875	0.9410	0.8194	1.0000	1.0000	0.8180	0.9565
SWC	0.9814	0.8636	0.9971	0.7524	0.9013	0.9880	1.0000	1.0000	0.9803	0.9268
The proposed method	0.9929	0.9779	0.9967	0.9304	0.9690	0.8876	0.9867	1.0000	0.9276	**0.9750**

Table 3. Accuracies comparison on the San Francisco Bay data set

Method	Sea	Mountains	Grass	Buildings	Accuracy
MCCA	0.9996	0.6341	0.7481	0.9123	0.8639
MSE	0.9907	0.7042	0.8122	0.9145	0.8855
MvDA	1.0000	0.6693	0.7289	0.9407	0.8747
MLDCCA	1.0000	0.9213	0.7075	0.9390	0.9117
SWC	0.9999	0.5231	0.9228	0.8824	0.8691
The proposed method	0.9989	0.8355	0.8499	0.9673	**0.9335**

first data set are shown in Tab. 1 and Fig. 1. And those of the second data set are shown in Tab. 3 and Fig. 2.

From these experimental results, it can be seen that the combination of three features indeed improves the classification performance. Moreover, MSE and MvDA perform nearly and better than MCCA. MCCA does not utilize the label information and mainly considers the correlation of different features. Although MSE is also unsupervised, it considers the preservation of the structure information of the data and explores the complementarity of different features. MLDCCA performs better than MSE and MvDA because it can preserve the correlation of different features, the label and local structure information of the data. But it neglects the complementarity of different features. The proposed method outperforms other comparing methods because 1) we adopt the multi-view subspace learning method to take full advantage of three features and their relations between each other; 2) the proposed method combines MCCA and MSE to learn deeply the relation of different features.

4. CONCLUSION

This paper proposes a multi-view subspace learning method to combine PF, TF and CF for supervised PolSAR image classification. By introducing the label information and the SRW distance, MSE is modified and then combined with MCCA to preserve the correlation and complementarity of different features. Finally the classification accuracies and the visible classification results show the superiority of the proposed method. In addition, we will extend the proposed method for other types of multi-channel images in the future work.

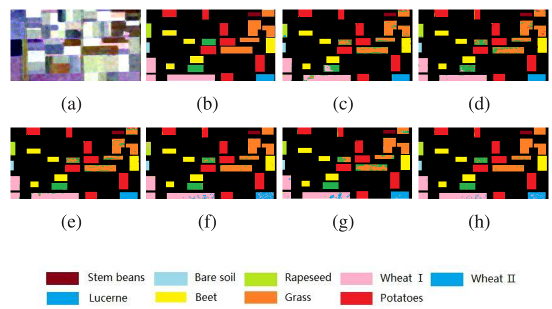

Fig. 1. Classification results comparison on the Flevoland data set: (a) the denoised image, (b) the ground truth, (c) MCCA, (d)MSE, (e) MvDA, (f) MLDCCA (e) SWC and (f) the proposed method.

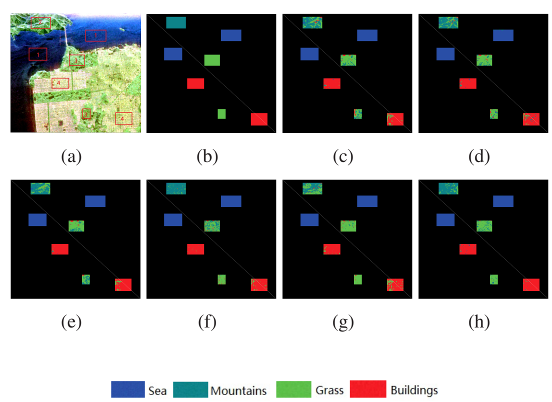

Fig. 2. Classification results comparison on the San Francisco Bay data set: (a) the denoised image and some experimental areas, (b) the ground truth, (c) MCCA, (d)MSE, (e) MvDA, (f) MLDCCA (e) SWC and (f) the proposed method.

5. REFERENCES

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

[2] J.S. Lee, M.R. Grunes, and E. Pottier, "Quantitative comparison of classification capability: Fully-polarimetric versus dual and single-polarization SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 11, pp. 2343–2351, Nov. 2001.

[3] S. T. Tu, J. Y. Chen, W. Yang, and H. Sun, "Laplacian eigenmaps-based polarimetric dimensionality reduction for sar image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 1, pp. 170–179, Jan 2012.

[4] L. Zhang, L. Sun, B. Zou, and W. M. Moon, "Fully polarimetric sar image classification via sparse representation and polarimetric features," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 8, pp. 3923–3932, Aug 2015.

[5] A. Masjedi, M. J. Valadan Zoej, and Y. Maghsoudi, "Classification of polarimetric sar images based on modeling contextual information and using texture features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 2, pp. 932–943, Feb 2016.

[6] B. Ren, B. Hou, J. Zhao, and L. Jiao, "Unsupervised classification of polarimetirc sar image via improved manifold regularized low-rank representation with multiple features," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 2, pp. 580–595, Feb 2017.

[7] S. Uhlmann and S. Kiranyaz, "Integrating color features in polarimetric SAR image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 4, pp. 2197–2216, Apr. 2015.

[8] Y.D. Zhang, L. Wu, and G. Wei, "A new classifier for polarimetric SAR images," *Progress In Electromagnetics Research*, vol. 94, pp. 83–104, 2009.

[9] O. Tannous and D. Kasilingam, "Independent component analysis of polarimetric sar data for separating ground and vegetation components," in *2009 IEEE International Geoscience and Remote Sensing Symposium*, July 2009, vol. 4, pp. IV–93–IV–96.

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

[11] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3-4, pp. 321–377, 1936.

[12] J. Rupnik and J. Shawe-Taylor, "Multi-view canonical correlation analysis," in *Proc. Slovenian KDD Conf. Data Mining Data Warehouses*, 2010, pp. 1–4.

[13] M. Kan, S. Shan, H. Zhang, S. Lao, and X. Chen, "Multi-view discriminant analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 1, pp. 188–194, Jan 2016.

[14] X. Huang, B. Zhang, H. Qiao, and X. Nie, "Local discriminant canonical correlation analysis for supervised polsar image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 11, pp. 2102–2106, Nov 2017.

[15] T. Xia, D. Tao, T. Mei, and Y. Zhang, "Multiview spectral embedding," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 6, pp. 1438–1446, Dec 2010.

[16] L. Zhang, L. Zhang, D. Tao, and X. Huang, "On combining multiple features for hyperspectral remote sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 3, pp. 879–893, March 2012.

[17] S.N. Anfinisen, R. Jenssen, and T. Eltoft, "Spectral clustering of polarimetric SAR data with Wishart-derived distance measures," in *Proc. POLInSAR 2007*, Esrin, Italy, 2007.

[18] C. Hou, F. Nie, H. Tao, and D. Yi, "Multi-view unsupervised feature selection with adaptive similarity and view weight," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 9, pp. 1998–2011, Sep. 2017.