

# SUPERVISED MANIFOLD-LEARNING ALGORITHM FOR POLSAR FEATURE EXTRACTION AND LULC CLASSIFICATION

Wei Wang\*, Zhuangzhuang Tian, Biao Tian, Jun Zhang

College of Electronic Science and Technology, National University of Defense Technology, 410073, Changsha, China -  
wangwei\_nudt@hotmail.com, gkt\_cn@outlook.com, tbncsz@126.com, zhangjun@nudt.edu.cn

**KEY WORDS:** Polarimetric synthetic aperture radar (PolSAR), Classification, Feature extraction, Manifold learning

## ABSTRACT:

In this paper, a supervised manifold-learning method is proposed for PolSAR feature extraction and classification. Based on the tensor algebra, the proposed method characterizes each pixel with a local neighbourhood centered at it, thereby combining the spatial and polarimetric information within the image. The inherent spatial information is beneficial to alleviate the influence of speckle noise and improve the stability of the extracted features. In addition, the label information of training samples is utilized in feature extraction, therefore the discriminability of different classes can be well preserved. The tensor discriminative locality alignment (TDLA) method is applied to find the multilinear transformation from the original feature space to the low-dimensional feature space. Based on the extracted features in the low-dimensional space, the SVM classifier is applied to achieve the final classification result. A real PolSAR data set acquired over San Francisco is adopted for performance evaluation. The experimental results show that the proposed method can not only improve the classification accuracy, but also alleviate the influence of speckle noise. In addition, the spatial details can be well preserved, demonstrating the superior performance of the proposed method.

## 1. INTRODUCTION

Polarimetric synthetic aperture radar (PolSAR) can obtain much more information about the land covers and targets than single polarized SAR systems (Lee, Pottier, 2009), thus it is increasingly used in various remote sensing applications, such as change detection, urban planning, biomass estimation, etc. Among them, land use/land cover (LULC) classification is of the key importance, and many supervised and unsupervised classification methods have been developed (Wang et al., 2016, Yu et al., 2012). Feature extraction and classification techniques are the two essential elements which determine the result. Thereinto, feature extraction plays a significant role since that some suitable features can achieve accurate classification results, even if only a simple classifier is applied; in contrast, it could be quite difficult to get a satisfactory result without elaborately extracted features, even though using a complex and advanced classifier.

In the past few decades, many polarimetric features which reveal specific scattering mechanisms have been obtained, including H/A/alpha (Cloude, Pottier, 1997), Freeman-Durden (Freeman, Durden, 1998), Yamaguchi (Yamaguchi et al., 2011), Neumann, Huynen, etc. However, a single feature or feature set may cause misclassifications in cases where different land covers possess the similar scattering mechanism and where the same class of land cover incur different scattering mechanisms. For example, the buildings not aligned orthogonal to radar line of sight (namely oriented buildings) also cause cross-polarized scattering, which is mainly generated by vegetation (Xiang et al., 2015). Therefore, these two land covers cannot be well separated if the features obtained by a single polarimetric decomposition method are employed. Recently, some new decomposition methods can overcome this problem to some extent, however, the orthometric and oriented buildings are

usually assigned to only one class in the classification map. In fact, it is also necessary and meaningful to separate these two kinds of buildings for urban planning or other applications. In order to achieve more detailed and accurate LULC classification, various polarimetric features are needed, which can reflect the characteristics of the land covers from different perspectives.

It is shown that PolSAR classification can be improved by combining multiple features together. However, various features have information redundancies, which may impact the performance of the classifier. In addition, simple combination of the features can greatly increase the computation complexity. Therefore, it is necessary to reduce the redundancy among various features and preserve the discriminative information at the same time. Generally, there are two ways for dimension reduction, i.e., feature selection and feature extraction. This paper investigates feature extraction by transforming the original feature space to a lower dimensional subspace, with most of the salient information preserved. In this field, there are many linear and nonlinear techniques, including principal component analysis (PCA) (Licciardi et al., 2014), independent component analysis (ICA) (Yamada et al., 2012), linear discriminated analysis (LDA), locally linear embedding (LLE) (Roweis, Saul, 2000), isometric feature mapping (ISOMAP) (Balasubramanian, Schwartz, 2002), Laplacian eigenmaps (LEs) (Belkin, Niyogi, 2003), t-distributed stochastic neighbor embedding (tSNE) etc., which have been successfully applied in PolSAR feature extraction. However, these methods deal with each pixel independently, and the spatial relationship among neighboring pixels is neglected.

In recent years, tensor which can represent high-order data is drawing increasing attention, and has been applied in hyperspectral image analysis, computer vision, and so on (Ai et

\* Corresponding author

al., 2013, Renard, Bourennane, 2009, Lu et al., 2011). Based on multilinear algebra, many tensor subspace learning algorithms have been developed and applied in dimension reduction (Yan et al., 2007, Xu et al., 2008). For example, using the original feature tensor of PolSAR data, the tensor decomposition techniques have been incorporated with ICA for feature extraction (Tao et al., 2015). However, the proposed tensor-based feature extraction is an unsupervised procedure. In fact, the discriminative information from the given training samples should also be considered, so that the distance between the extracted feature set of different land covers can be maximized, leading to more accurate classification results.

To make the best of the information of training samples while extracting the intrinsic features, this paper proposes a supervised manifold-learning algorithm for PolSAR feature extraction and classification. The tensor discriminative locality alignment (TDLA) method is introduced to extract the low-dimensional features that are embedded in the high-dimensional feature space (Zhang et al., 2013). Based on a wide range of informative polarimetric features, the tensor representation of each pixel is constructed by combining itself and its surrounding neighbors. The tensor structure can help to better characterize the features of an individual pixel, and it is also beneficial to alleviate the influence of speckle noise. In the TDLA method, the discriminative information from the given training samples is considered, therefore the separability of different classes can be improved in the low-dimensional feature space.

The remainder of this paper is organized as follows. Section 2 introduces the fundamentals of the TDLA algorithm, and then presents the proposed PolSAR feature extraction and classification method. In Section 3, the experimental results are provided to validate the effectiveness of the proposed method. Section 4 concludes this paper.

## 2. THE PROPOSED METHOD

### 2.1 Tensor discriminative locality alignment algorithm

The TDLA algorithm can find a multilinear transformation from the original high-order feature space  $X \in R^{L_1 \times L_2 \times \dots \times L_M}$  to the reduced feature space  $X^* \in R^{d_1 \times d_2 \times \dots \times d_M}$  (Zhang et al., 2013). The formula can be expressed as

$$X^* = X \prod_{i=1}^M U_i^T \quad (1)$$

where  $U_i \in R^{L_i \times d_i}$  ( $i=1,2,\dots,M$ ,  $d_i \leq L_i$ ) are a set of projection matrices.  $X \times_i U_i^T$  means the mode- $i$  product of tensor  $X$  and matrix  $U_i$ . In the following, we consider the second-order tensor  $X \in R^{L_1 \times L_2}$  for simplicity.

The input data of the TDLA algorithm is a set of training samples  $X_i$  ( $i=1,2,\dots,N$ ) and their labels  $y_i$  ( $y_i \in [1,2,\dots,C]$ ), where  $N$  is the number of training samples and  $C$  is the class number. For each sample  $X_i$ , we divide the other  $N-1$  samples into two groups by their class labels: the samples of the same class and the samples belonging to others. Then, we sort the samples in each group based on their Euclidean distances to  $X_i$ . The  $n_1$  nearest samples of the group "same class" and the

$n_2$  nearest samples of the group "different classes" are selected to build a patch of  $X_i$

$$Patch(X_i) = \{X_i, S_{i(1)}, \dots, S_{i(n_1)}, D_{i(1)}, \dots, D_{i(n_2)}\} \in R^{L_1 \times L_2 \times (1+n)} \quad (2)$$

where  $n = n_1 + n_2$ ,  $S_i$  means the samples in the same class, and  $D_i$  the samples in the different classes. For each  $Patch(X_i)$ , the corresponding representation in the reduced feature space is denoted by

$$Patch(X_i^*) = \{X_i^*, S_{i(1)}^*, \dots, S_{i(n_1)}^*, D_{i(1)}^*, \dots, D_{i(n_2)}^*\} \in R^{d_1 \times d_2 \times (1+n)} \quad (3)$$

In the reduced feature space, in order to preserve the discriminability of different classes, TDLA suggests that the distances between  $X_i^*$  and  $S_i^*$  should be as small as possible, while the distances between  $X_i^*$  and  $D_i^*$  should be as large as possible. On this basis, the following optimizations upon  $X_i^*$  can be constructed.

$$\arg \min_{X_i^*} \sum_{j=1}^{n_1} D^2(X_i^* - S_{i(j)}^*) \quad (4)$$

$$\arg \max_{X_i^*} \sum_{j=1}^{n_2} D^2(X_i^* - D_{i(j)}^*) \quad (5)$$

Formula (4) and (5) can be combined together as

$$\arg \min_{X_i^*} \sum_{j=1}^{n_1} D^2(X_i^* - S_{i(j)}^*) - \alpha \sum_{j=1}^{n_2} D^2(X_i^* - D_{i(j)}^*) \quad (6)$$

where  $\alpha$  is a weighting factor. By setting

$$\beta = [\underbrace{1, \dots, 1}_{n_1}, \underbrace{-\alpha, \dots, -\alpha}_{n_2}] \quad (7)$$

Formula (6) can be simplified as

$$\begin{aligned} & \arg \min_{X_i^*} \sum_{j=1}^{n_1} \beta_j D^2(X_i^* - S_{i(j)}^*) + \sum_{j=1}^{n_2} \beta_{n_1+j} D^2(X_i^* - D_{i(j)}^*) \\ & = \arg \min_{X_i^*} \sum_{j=1}^n \beta_j \|X_{P_i(1)}^* - X_{P_i(j+1)}^*\|^2 \end{aligned} \quad (8)$$

If we define

$$Q = \begin{bmatrix} \sum_{j=1}^n \beta_j & -\beta^T \\ -\beta & \text{diag}(\beta) \end{bmatrix} \in R^{(1+n) \times (1+n)} \quad (9)$$

Then the final representation of the patch optimization can be denoted by

$$\arg \min_{X_i^*} \sum_{g=1}^{n+1} \sum_{h=1}^{n+1} (Q_{g,h} \cdot [X_{P_i(g)}^* \otimes X_{P_i(h)}^*; (1:2)(1:2)]) \quad (10)$$

where  $X_{P_i(g)}^* \otimes X_{P_i(h)}^*$  means the operation of tensor contraction. The whole alignment of TDLA is obtained by summing over the patch optimizations of all the training samples. For each

sample, the local patch provided by (2) is unique in order, here we unite the samples in each patch to a unified system by assuming that the  $n + 1$  samples in (2) are selected from the training samples

$$Whole(X) = \{X_1, \dots, X_N\} \in R^{L_1 \times L_2 \times N} \quad (11)$$

The unification can be achieved by using a selection matrix  $E_i \in R^{N \times (1+n)}$  defined by

$$E_{i(a,b)} = \begin{cases} 1, & \text{if } a = F_i\{b\} \\ 0, & \text{else} \end{cases} \quad (12)$$

Where  $F_i \in \{i, i_1, \dots, i_n\}$  denotes the set of global indices of samples in  $Patch(X_i)$ . Then, the sum of all the patch optimizations described in (10) can be written as

$$\arg \min_{X_1^*, \dots, X_N^*} \sum_{g=1}^N \sum_{h=1}^N (\Omega_{g,h} \cdot [X_g^* \otimes X_h^*; (1:2)(1:2)]) \quad (13)$$

$$\Omega = \sum_{i=1}^N E_i Q_i E_i^T \in R^{N \times N} \quad (14)$$

The full optimization (13) aims to obtain an optimal subspace from the original high-order feature space. By putting (1) into (13) and imposing the constraints that  $U_i^T U_i = I$ , the solution of  $U_i$  can be derived as

$$\arg \min_{U_k^T U_k = I} \text{tr}(U_k^T F(k) U_k) \quad (15)$$

$$F(1) = \sum_{g=1}^N \sum_{h=1}^N \Omega_{g,h} \cdot \text{Mat}_1(X_g \times_2 U_2^T) \cdot \text{Mat}_1^T(X_h \times_2 U_2^T) \quad (16)$$

$$F(2) = \sum_{g=1}^N \sum_{h=1}^N \Omega_{g,h} \cdot \text{Mat}_2(X_g \times_1 U_1^T) \cdot \text{Mat}_2^T(X_h \times_1 U_1^T) \quad (17)$$

The solution of (15) is acquired by combining the eigenvectors associated with the smallest  $d_k$  eigenvalues of matrix  $F(k)$ . Then an alternating optimization procedure is applied to obtain  $U_1$  and  $U_2$ .

## 2.2 Tensor modeling of PolSAR data

The PolSAR data is usually represented by the covariance matrix after multilook processing, on the basis of the complex scattering matrix. The covariance matrix  $C$  is Hermitian and only contains six independent elements: three real diagonal and three complex off-diagonal elements. The covariance matrix itself can only provide limited information if it is directly utilized for land cover classification. Therefore, many polarimetric decomposition methods have been investigated to provide an interpretation of the covariance matrix as the combination of canonical scattering mechanisms. Table 1 lists the polarimetric features extracted using different decomposition methods. In addition, 9 elements obtained from the covariance matrix are also included. Then, the land covers can be well characterized by altogether 48 informative polarimetric features.

Based on the polarimetric features, it is straightforward to represent each pixel by a 48-dimensional vector, which is then

Method	Polarimetric feature		
	C11	C22	C33
Elements of Covariance matrix (9)	C12_modulus	C12_phase	C13_phase
	C13_modulus	C23_modulus	C23_phase
	Cloude_C11	Cloude_C22	Cloude_C33
Freeman (3)	Freeman_Odd	Freeman_Dbl	Freeman_Vol
VanZyl(3)	VanZy_Odd	VanZy_Dbl	VanZy_Vol
Krogager(3)	Krogager_Ks	Krogager_Kd	Krogager_Kh
Yamaguchi (4)	Yamaguchi_Odd	Yamaguchi_Dbl	Yamaguchi_Vol
	Yamaguchi_Hlx		
H/A/Alpha (6)	Entropy	Alpha	Anisotropy
	Lambda1	Lambda2	Lambda3
Huynen(3)	Huynen_C11	Huynen_C22	Huynen_C33
An&Yang (4)	An_Yang_Odd	An_Yang_Dbl	An_Yang_Vol
	An_Yang_Hlx		
Zhang (5)	Zhang_Odd	Zhang_Dbl	Zhang_Vol
	Zhang_Hlx	Zhang_Wire	
Xiang (5)	Xiang_Odd	Xiang_Dbl	Xiang_Vol
	Xiang_Hlx	Xiang_Cros	

Table 1. Polarimetric features obtained using different methods

processed and classified independently, without considering the spatial constraints. In this paper, by introducing the tensor algebra for feature representation, each pixel is characterized by combining the spatial and polarimetric information. Specifically, each pixel is described using the neighborhood centered at the pixel. Figure 1 shows the structures of different neighborhood.

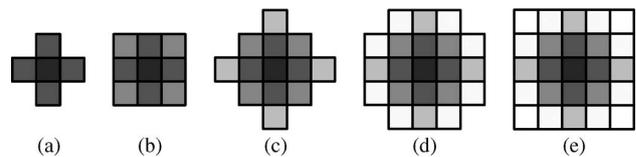


Figure 1. The structures of different neighborhoods

For each pixel, suppose its feature vector is  $x \in R^L$  ( $L=48$ ). By combining  $x$  and the feature vectors of its local neighbors, the feature representation can be constructed as

$$X = [x, x_1, x_2, \dots, x_k] \in R^{L \times (k+1)} \quad (18)$$

where  $x_i (i=1, 2, \dots, k)$  is the feature vector of a neighbor pixel. The feature matrix is actually a second-order tensor, in which one mode corresponds to the polarimetric dimension and the other mode denotes the spatial dimension. Since that the pixels within a local neighborhood are likely from the same land cover, the polarimetric-spatial representation helps to alleviate the effect of speckle noise, and to better preserve the spatial information.

## 2.3 TDLA-based PolSAR feature extraction

Based on the second-order feature tensor of all the training samples, the TDLA algorithm is applied to find two projections  $U_1$  and  $U_2$  for feature extraction.  $U_1 \in R^{L_1 \times d_1}$  reduces the feature dimension in the polarimetric mode, and  $U_2 \in R^{L_2 \times d_2}$  reduces the feature dimension in the spatial mode. Then, we can obtain the feature in the reduced space for each pixel in the PolSAR image.

$$X^* = X \times_1 U_1^T \times_2 U_2^T \in R^{d_1 \times d_2} \quad (19)$$

Using the extracted features, the SVM classifier is applied to get the final classification result.

### 3. EXPERIMENTAL RESULTS

In this section, a real PolSAR data set is used to evaluate the performance of the proposed method. In addition, the classical PCA method and an unsupervised manifold learning algorithm, i.e., tSNE are also employed for comparison. In the proposed method, the neighborhood in Figure 1(b) is utilized, then the spatial dimension  $L_2$  equals to 9. In the reduced feature space,  $d_1$  and  $d_2$  are set as 3 and 1, respectively. The numbers  $n_1$  and  $n_2$ , which determine the patch size, are set as 5. We fix the weighting factor  $\alpha$  at 2.

#### 3.1 Data set description

This paper adopts a fully polarimetric data set acquired with Radarsat-2 system on 11 November 2009. The data set has a spatial resolution of 8 m, and a size of 1500×1400 pixels. The study area is located in San Francisco, USA, which has coverage of different kinds of land covers, including sea, vegetation, and buildings. Figure 2(a) shows the Pauli RGB image of this area. In addition, it can be seen that the building areas can be further divided into low density urban, high density urban, and oriented building, according to the backscatter responses, block size and orientation angles. In Figure 2(a), the vegetation can be confused with the oriented buildings due to the scattering similarity. Figure 2(b) presents the ground truth of the image. For each class, 100 pixels are randomly selected for training, and the others in the labelled areas are used as the testing set.

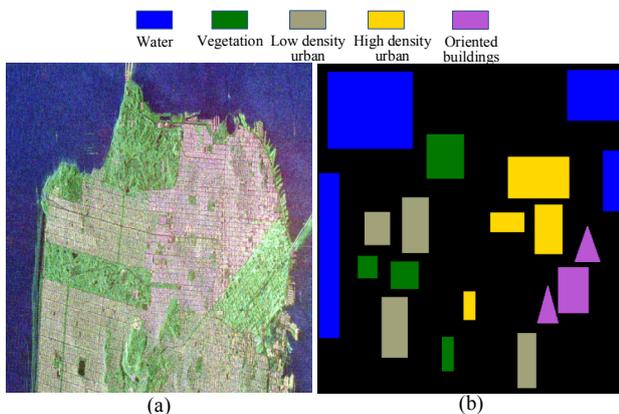


Figure 2. San Francisco data set. (a) Pauli RGB image. (b) The ground truth.

#### 3.2 Feature distribution in the reduced space

Figure 3 shows the scatterplots of the extracted features obtained using different methods. Figure 3(a) corresponds to 3 original features, i.e., C11, C22, and C33 in the covariance matrix. It can be seen that the scatters of different classes are totally in a mess. Figure 3(b) shows the scatterplots of the principal components extracted by PCA. We can find that there is an obvious gap between the red plots (water) and the others. However, the components of the other four classes are very close in the feature space, and much correlation exists among

them, making the following classification a difficult task. tSNE method is a nonlinear feature extraction technique. The scatterplots of the extracted features are shown in Figure 3(c). It can be seen that the separability of different classes is improved to some extent. However, some samples are mixed together. Using the proposed method, the scatterplots of the extracted features are shown in Figure 3(d). By comparison, the features of different classes are more discriminable, and the samples of each class are better clustered. This is mainly because that, the discriminability information of different classes is utilized in feature extraction. The optimization procedure insures that the distances within the same class are minimized, and the distances between different classes are maximized.

#### 3.3 Classification result

Based on the extracted features, the final classification result can be easily obtained, and Figure 4 gives the results using different methods. From the whole perspective, we can see that the main areas are correctly classified by all the three methods. However, there are many misclassifications in Figure 4(a) and (b). For example, in Figure 4(a), a part of the oriented buildings are classified as high density urban. This phenomenon is eliminated in Figure 4(b), however, many pixels are misclassified as oriented buildings, especially in vegetation and low density urban areas. By contrast, the proposed method generates the best classification result, with quite a small part of misclassified areas, as shown in Figure 4(c).

In addition, from Figure 4(a) and (b), we can observe a large amount of noises in the classification result, even in the ocean areas. In PCA and tSNE methods, the feature extraction procedure are implemented based on the polarimetric features of a single pixel, therefore it can be easily affected by the speckle noise. Whereas the proposed method represents each pixel with a polarimetric-spatial feature tensor, and the spatial information can be utilized during feature extraction. The polarimetric-spatial representation is beneficial to alleviate the effect of speckle noise, and to improve the stability of the extracted features. The classification result in Figure 4(c) is visually smooth, and the shape details can be well preserved. Taking the area in the red ellipse as an example, we can hardly observe the spatial details of the coastline in Figure 4(a) and (b). However, the coastline and the port shapes are very obvious in Figure 4(c), demonstrating the effectiveness of the proposed method.

Table 2 presents the accuracy statistics of different methods. The accuracy is calculated by performing 100 independent classification experiments. For each experiment, a fixed number of samples are randomly selected for training, and the other labelled samples are used for testing. From table 2 we can find that the proposed method achieves an accuracy improvement of 3.4% than tSNE.

Method	PCA	tSNE	Proposed Method
Sea	0.993	0.991	<b>0.998</b>
Vegetation	0.825	0.834	<b>0.876</b>
Low density urban	0.859	0.876	<b>0.921</b>
High density urban	0.893	<b>0.905</b>	0.887
Oriented buildings	0.726	0.893	<b>0.912</b>
Overall accuracy	0.856	0.886	<b>0.920</b>

Table 2. Classification accuracy of different methods

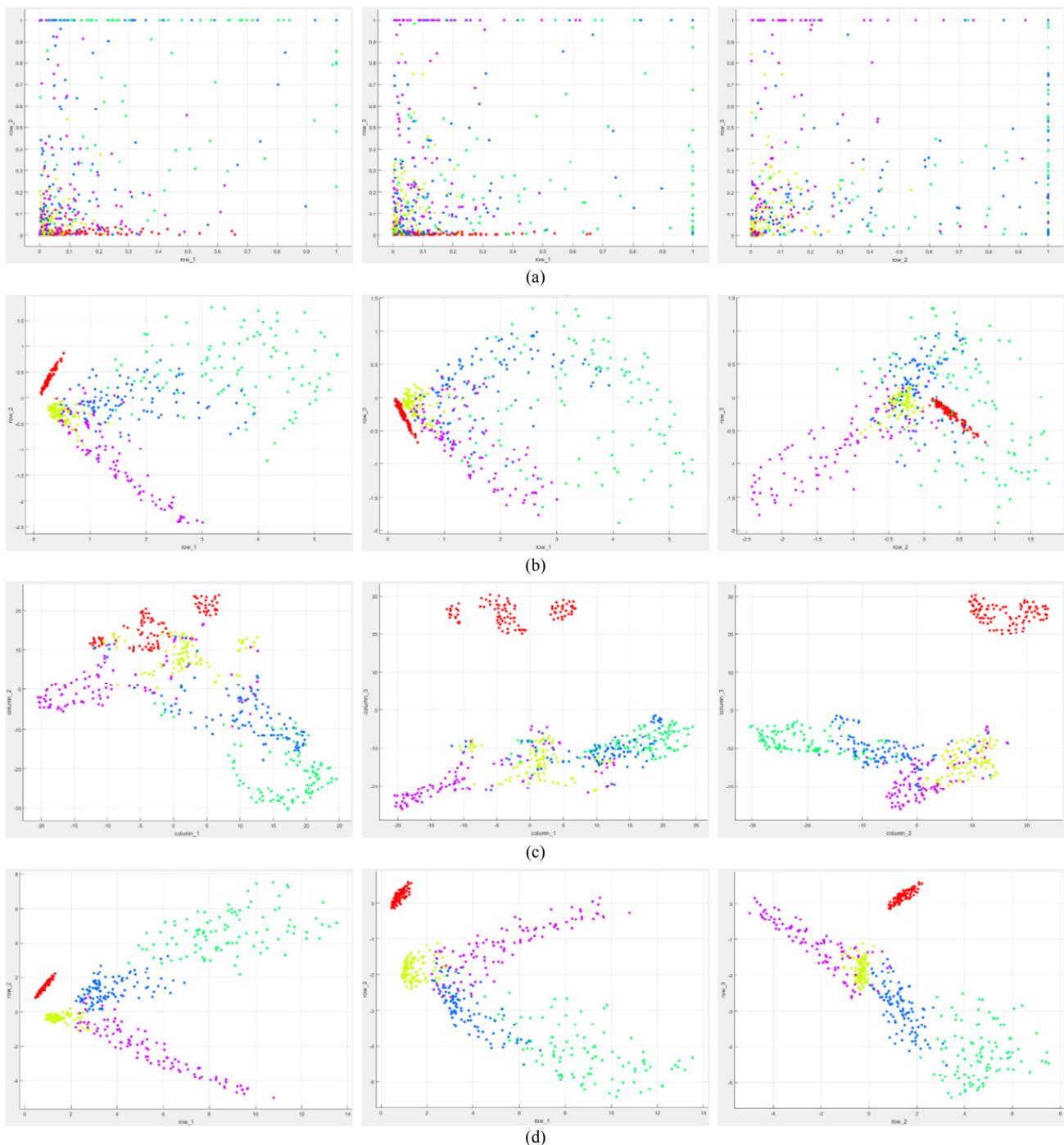


Figure 3. Scatterplots of the extracted features using different methods. (a) The original features C11, C22, and C33. (b) PCA. (c) tSNE. (d) The proposed method.

#### 4. CONCLUSION

This paper proposes a supervised manifold-learning method for PolSAR feature extraction and classification. In this method, each pixel is represented by a second-order tensor, by combining the polarimetric features within a local neighborhood. The polarimetric-spatial representation is helpful to preserve the spatial information and to alleviate the effect of speckle noise. In addition, the discriminability information of different classes is utilized during feature extraction. The TDLA method is applied to extract the low-dimensional

features that are embedded in the high-dimensional feature space. The experimental results on the Radarsat-2 data set show that the proposed method can classify the different land covers with high accuracy, with well-preserved spatial details of the image. In the future, more efforts will be made to improve the extracted features by integrating spatial information.

#### ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China under Grant 61901500.

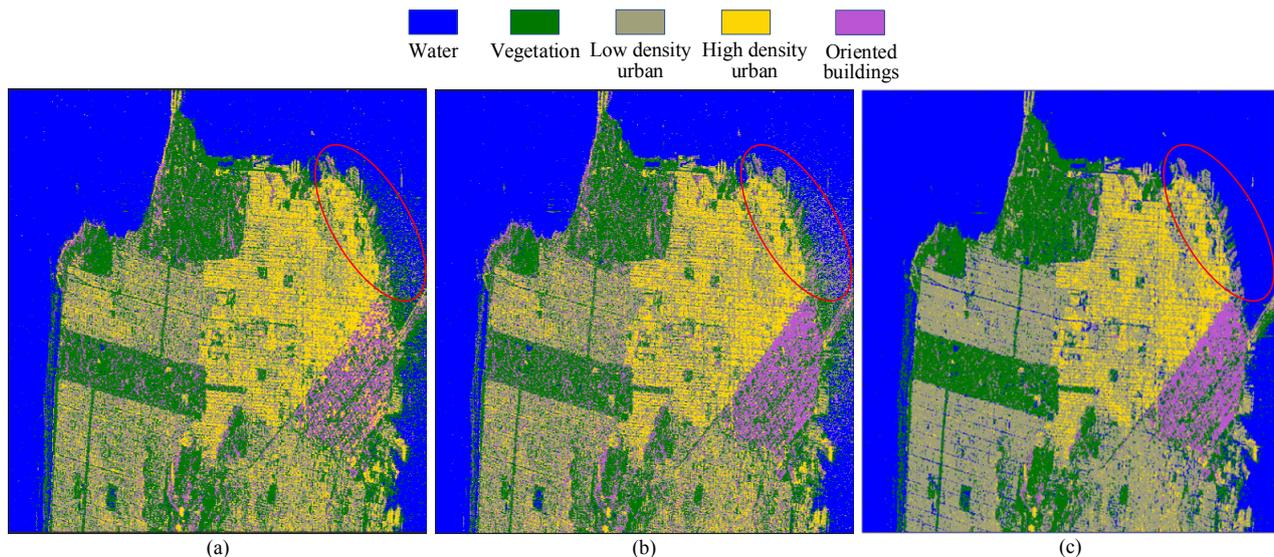


Figure 4. The classification results obtained using different methods. (a) PCA. (b) tSNE. (c) The proposed method.

## REFERENCES

- Ai, D., Duan, G., Han, X., Chen, Y., 2013. Generalized N-dimensional independent component analysis and its application to multiple feature selection and fusion for image classification. *J. Neurocomput.*, 103, 186–197.
- Balasubramanian, M., Schwartz, E. L., 2002. The Isomap algorithm and topological stability. *Science*, 295(5552), 7.
- Belkin, M., Niyogi, P., 2003. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.*, 15(6), 1373–1396.
- Cloude, S. R., Pottier, E., 1997. An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Trans. Geosci. Remote Sens.*, 35(1), 68–78.
- Freeman, A., Durden, S., 1998. A three-component scattering model for polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 36(3): 963-73.
- Lee, J. S., Pottier, E., 2009: *Polarimetric Radar Imaging: From Basics to Applications*. CRC Press, Boca Raton, FL, USA.
- Licciardi, G., Avezano, R. G., Frate, F. D., Schiavon, G., Chanussot, J., 2014. A novel approach to polarimetric SAR data processing based on nonlinear PCA. *Pattern Recognit.*, 47(5), 1953–1967.
- Lu, H., Plataniotis, K. N., Venetsopoulos, A. N., 2011. A survey of multilinear subspace learning for tensor data. *Pattern Recognit.*, 44(7), 1540–1551.
- Renard, N., Bourenane, S., 2009. Dimensionality reduction based on tensor modeling for classification methods. *IEEE Trans. Geosci. Remote Sens.*, 47(4), 1123–1131.
- Roweis, S. T., Saul, L. K., 2000. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500), 2323–2326.
- Tao, M., Zhou, F., Liu, Y., Zhang Z., 2015. Tensorial independent component analysis-based feature extraction for polarimetric SAR data classification. *IEEE Trans. Geosci. Remote Sens.*, 53(5), 2481–2495.
- Wang, W., Xiang, D., Zhang, J., Wan, J., 2016. Integrating Contextual Information With  $H\alpha$  Decomposition for PolSAR Data Classification. *IEEE Geoscience and Remote Sensing Letters*, 13(12), 2034-2038.
- Xiang, D., Ban, Y., Su, Y., 2015. Model-based decomposition with cross scattering for polarimetric SAR urban areas. *IEEE Geosci. Remote Sens. Lett.*, 12(12), 2496–2500.
- Xu, D., Yan, S., Zhang, L., Lin, S., Zhang, H. J., Huang, T. S., 2008. Reconstruction and recognition of tensor-based objects with concurrent subspaces analysis. *IEEE Trans. Circuits Syst. Video Technol.*, 18(1), 36–47.
- Yamada, H., Yamaguchi, Y., Sato, R., 2012. Fundamental study on semiblind scattering component decomposition for PolSAR data using ICA. *9th EUSAR*, Nuremberg, Germany, 753–756.
- Yamaguchi, Y., Sato, A., Boerner, W. M., Sato, R., Yamada, H., 2011. Four-component scattering power decomposition with rotation of coherency matrix. *IEEE Trans. Geosci. Remote Sens.*, 49(6), 2251–2258.
- Yan, S., Xu, D., Yang, Q., Zhang, L., Tang, X., Zhang, H. J., 2007. Multilinear discriminant analysis for face recognition. *IEEE Trans. Image Process.*, 16(1), 212–220.
- Yu, P., Qin, A. K., Claudi, D. A., 2012. Unsupervised polarimetric SAR image segmentation and classification using region growing with edge penalty. *IEEE Trans. Geosci. Remote Sens.*, 50(4), 1302–1317.
- Zhang, L., Zhang, L., Tao, D., Huang X., 2013. Tensor discriminative locality alignment for hyperspectral image spectral-spatial feature extraction. *IEEE Trans. Geosci. Remote Sens.*, 51(1), 242–256.