proceedings

MDPI

Proceedings

Classifying UAVSAR Polarimetric Synthetic Aperture Radar (PolSAR) Imagery Using Target Decomposition Features †

Ghazaleh Alijani [1,*], Mahdi Hasanlou [2] and Zahra Azizi [1]

[1] Department of GIS/RS, Science and Research Branch, Islamic Azad University, Tehran 1477893855, Iran; zazizi@srbiau.ac.ir

[2] School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran 1439957131, Iran; hasanlou@ut.ac.ir

* Correspondence: gh.alijani_8016331@yahoo.com; Tel.: +98-912-846-4519

† Presented at the 2nd International Electronic Conference on Remote Sensing, 22 March–5 April 2018; Available online: https://sciforum.net/conference/ecrs-2.

Received: 22 March 2018

Abstract: Changes in the earth's surface significantly increase natural disasters, resulting in severe damage to man-made objects, such as roads, buildings, bridges, and so on. Radar techniques have advantages, such as lack of sensitivity to weather conditions, to night and day, and to cloud cover conditions, which can be used to identify, alert, and mitigate these damages. Because of the importance of these areas and the need to care for them, land-use classification, one of the important applications of remote sensing, is performed. Polarimetric synthetic aperture radar (PolSAR) images have many capabilities, having the scattering information on four polarized levels (HH, HV, VH and VV) and consequently depending on the shape and structure of the environment. In this study, unmaned aerial vehicle (UAVSAR) image is used. The support vector machine (SVM) model is a well-known classification method, able to run on different types of features and to distinguish classes that are not linearly separable. On the other hand, it is possible to use data mining methods to facilitate data analysis, like classifications. In this regards, it is recommended to use the random forest (RF) technique. The RF is one of the useful methods for data classification which uses a tree structure for decision-making. This method uses strategies to enhance the probability of reaching goals with conditional probability. In this study, by incorporating a variety of target decomposition methods in PolSAR images, images producing the land cover types were generated. Then, 70 features were obtained by applying the support vector machine (SVM), random forest (RF) , and K-nearest neighbor (KNN) classification methods. In order to estimate accuracy, the output of these methods was evaluated by reference data.

Keywords: polarimetric synthetic aperture radar; classification; support vector machine; random forest; K-nearest neighbors

---

1. Manuscript

General Instructions

Today, the use of remote sensing data as an ideal source of precision and speed of operation has become one of the most important means of data collection. In the meantime, radar remote sensing with the capability of capturing images in different weather conditions and throughout the day is becoming widespread. On the other hand, the use of radar polarimetric imaging systems has been widely considered [1] because the derived images can improve applications since they provide more distributive information about the effects of the image. Classification is one of the most important techniques for identifying and distinguishing ground-type classes that are widely used in the field of geosciences, including in determining vegetation, determining thermal heat islands, detecting alterations. Radar images classification is still interesting to researchers. Yakkekhani et al. (2014) [2] examined the use of the support vector machine (SVM) method with a variety of kernels for classification purposes of UAVSAR polarimetric data. In 2006, Lardeux et al. [3] proposed an SVM algorithm to categorize all polarimetric data and tested this method on P-band data. By using the Covariance matrix elements, they showed that the classification of the SVM is better in L band. Khosravi et al. (2014) [4] used a multiple classification systems (MCS) based on the SVM algorithm to classify hyperspectral images. In this research, a comparison was also made between the proposed system and the AdaBoost, Begging, and Randomized Forest (RF) methods.

In the previous paragraph, an overview of the research has been briefly described in a variety of classification methods. Although these studies have succeeded in classifying radar images, they focus on only one classification algorithm. Each of the classification methods has its own special features and applications. Many algorithms are presented for classifying polarimetric images. The most important of these are the K-nearest neighboring (KNN) methods and the SVM and RF algorithms. These three classification methods are recognized as the most suitable models for optimizing the process of classification of remote sensing images [5]. Because of the diversity of methods and the importance of classification when using radar data, it is necessary to examine different classification methods so that the users can choose the preferred methodology for classification. This research intends to examine the types of classification algorithms using UAVSAR radar data. Different types of distribution matrix elements are used for the production of the features. These algorithms used a special manner for producing aclassification map-based training data.

The polarization target decomposition is divided into four main categories [6]: the first category is based on the dichotomy of Kennaugh matrix that is included Yang, Huynen, Holm, and Brans. The second category is developed on the basis of the Covariance matrix C3 or coherent matrix of the Freeman Dong decomposition, Durden, and Yamaguchi methods. The third category is obtained on the basis of the eigenvector and eigenvalues of the covariance matrix (or coherence matrix). Some of these methods are used according to their application, such as Holm, Van, Cloud, Zyl and Could, Pottier, and Could. The fourth category of polarization target decomposition is related to the coherent decomposition of *S* scattering matrix. Some of these methods are Krogager, Cameron, and Touzi [6]. Due to, the presented types of algorithms such as Cloud Pottier, Freeman, Krogager, Van Zyl; presented to produce the features.

2. Proposed Method

2.1. Support Vector Machine

The support vector machine algorithm is one of the supervised training of pattern detection algorithms which was presented by a Russian mathematician called Vapnik in 1995, and its principles are based on the statistical training theory [7]. The main basis of this method is a linear classification of data, by taking safety margins into account, and is basically considered as a binary separator with the main goal of reaching the optimized hyper-plane to increase the boundary of two classes. If the data are not linearly separated, they are transmitted to a higher dimension space using nonlinear kernels, and a hyperplane is formed. Assume that p is a training datum defined as $(x_i.y_i)$, in which x_i is an n-dimensional attribution vector, and $y_i \in \{-1.1\}$ is its tag. This hyper-plane is defined by Equation (1):

$$w^T \Phi(x) + b = 0 \tag{1}$$

where w is the weight vector which is perpendicular to the intended hyper-plane, b is the bias vector which is a constant value, showing the distance between hyper-plane and origin, and $\Phi(x)$ is a kernel to transfer the data to a higher dimension space. As discussed, the aim of this classification is to find a hyper-plane by maximizing the margin and minimizing the overall error of Equation (2).

$$min(\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{k}\xi_i)$$
$$subjectto: y_i(w\Phi(x_i) + b) > 1 - \xi_i \ i = 1, \dots, k \quad (2)$$

C is the adjusting parameter/factor that adjusts generalization. To consider the noise in the data and interruptions between training data, ξ_i is used.

2.2. K-Nearest Neighbour (KNN)

KNN has been used in statistical estimation and pattern recognition. The training sample of c pairs of a random sample $(x_i.y_i)$, in which i = 1, 2, …, n and y takes {1, 2, …, n} values, can be defined as Equation (3):

$$train = \{(x_1.y_1), (x_2.y_2), \dots, (x_c.y_c)\}$$
$$y = \{1, 2, \dots, n\} \quad (3)$$

where y_i determines the class of x_i among the c probable classes. For this reason, for classification, firstly the nearest neighbor x' of X is determined in the training samples (Equation (4)) [8].

$$d(x_i.y_i) = \sqrt{\sum_{r=1}^{n}(a_r(x_i) - a_r(x_j))^2} \quad (4)$$

The popularity of this method ensues from two properties, i.e., simple application and determined error boundaries. However, high analytical load and high sensitivity to k values are some of its drawbacks. Therefore, the k values have a key role in this method. If k is too small, the algorithm is sensitive to noise, and if k is too large, it is possible that, among the nearest neighbors, a tag of other classes is entailed.

2.3. Random Forest Algorithm (RF)

The random forest algorithm is one of the recent methods of image classification invented by Breiman in 2001 [9], by developing the Bagging method. This method differs from Bagging in random feature selection. When creating a decision tree, RF firstly selects a random subset of features in each selection branch. The number of trees should be sufficient to fix the error rate [10]. The RF provides a more flexible classification, because of randomly selecting subsets for producing each decision tree [11].

3. Study Area

The data which was used in this research was taken from full polarized data of Panama, the capital city of Panama, in South America. These data were collected on 6 February of 2010, by the UAVSAR Airborne system of the jet propulsion laboratory (JPL) institute of NASA. These images have with a spatial resolution of 1.6 m in each pixel. The dimension of this data is $12{,}756 \times 12{,}773$ pixels. To reduce the data volume, and consequently reduce the calculation load, a subset with dimension of 276×266 pixels including the urban area, water, and vegetation was selected. Figures 1 and 2 show the related Pauli false color image and corresponding true color optical image of the considered region on Google Earth. Panama is located at 85°48′20″ W (−85.80556) and 30°10′36″ N (30.17667) geographical longitude and latitude, respectively. The considered region calculation load, a subset with dimension of 276×266 pixels including the urban area, water, and vegetation, was selected. The considered region consists of three prominent classes: (1) water, (2) vegetation, and (3) urban area. A ground truth map was produced based on visual comparison and application of high-resolution Google Earth images (Figure 3).

Figure 1 Overview of the study area, Pauli false color image.

Figure 2. True color image of the study area (Panama).

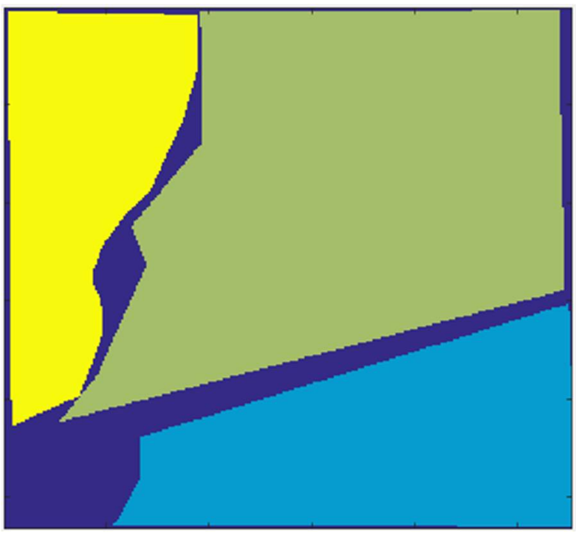

Figure 3. Overview of the proposed method.

4. Implementation

4.1. Extracted Decomposition Descriptors

To classify the polarimetric image, the decomposition descriptor was firstly extracted from the image in the PolSARpro_v4.2.0. To reduce the effect of speckle noise, a filter of 3×3 size on the image for coherent target decomposition descriptors was used (Table 1).

Table 1. Features extracted from decomposition.

Alpha	Holm1_T11	TSVM_alpha_s2
anisotropy	Holm1_T22	TSVM_alpha_s3
beta	Holm1_T33	TSVM_phi_s
combination_1mH1mA	Holm2_T11	TSVM_phi_s1
combination_1mHA	Holm2_T22	TSVM_phi_s2
combination_H1mA	Holm2_T33	TSVM_phi_s3
combination_HA	Freeman_Dbl	TSVM_psi
delta	Freeman_Odd	TSVM_psi1
entropy	Freeman_Vol	TSVM_psi2
gamma	Freeman2_Ground	TSVM_psi3
lambda	Freeman2_Vol	TSVM_tau_m
Huynen_T11	HAAlpha_T11	TSVM_tau_m1
Huynen_T22	HAAlpha_T22	TSVM_tau_m2
Huynen_T33	HAAlpha_T33	TSVM_tau_m3
Barnes1_T11	Krogager_Kd	VanZyl3_Dbl
Barnes1_T22	Krogager_Kh	VanZyl3_Odd
Barnes1_T33	Krogager_Ks	VanZyl3_Vol
Barnes2_T11	Neumann_delta_mod	Yamaguchi3_Dbl
Barnes2_T22	Neumann_delta_pha	Yamaguchi3_Odd
Barnes2_T33	Neumann_psi	Yamaguchi3_Vol
Cloude_T11	Neumann_tau	Yamaguchi4_Dbl
Cloude_T22	TSVM_alpha_s	Yamaguchi4_Hlx
Cloude_T33	TSVM_alpha_s1	Yamaguchi4_Odd
		Yamaguchi4_Vol

4.2. Steps for Implementation

When the features used in decomposition algorithm were extracted, the resuting features were overlaid on each other to be applied to the classification algorithm. Using the ground optical image matched by radar images, the training and test data were randomly extracted. In this research, 30% and 70% of the values were devoted to the training and testing data, respectively. Figure 3 illustrates the ground truth map in three classes. Figure 4 presents the proposed method flowchart.

In this stage, the extracted features of the radar image were introduced as three different classifiers, including SVM, KNN, and RF. Consequently, the results of each classifier were examined and evaluated.

One of the supervised image classification methods in this paper is the KNN method. This classification was performed on a synthetic aperture radar (SAR) image with 70 extracted features. The results of the classification using visual and numerical algorithms were evaluated using an overall accuracy (OA) criteria.

The SVM method has a C parameter, and its kernel function, which is of the radial basis function (RBF) type, has a parameter γ, which needs to be optimized. For this purpose, a Grid search (GS) was used to determine its optimal values [12]. Also, the RF method has two parameters: the number of trees (N_{tree}) and the number of features in each tree (M_{try}); where necessary, their optimal values were determined. Table 2 presents the optimal values of the parameters of the two methods mentioned.

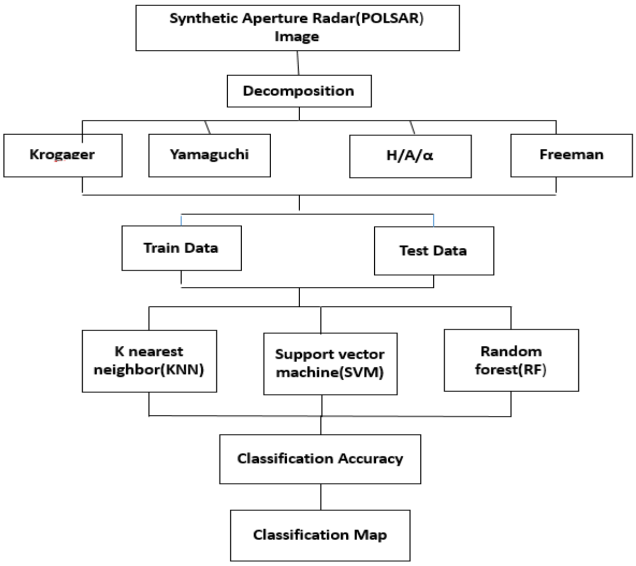

Figure 4. The ground truth dataset.

Table 2. The optimal classification parameters.

Method	Parameter	
SVM	C = 2	$\Gamma = 2/4414 \times 10^{-4}$
RF	$N_{tree} = 100$	$M_{try} = 8$
KNN	K = 1	-

The quantitative results from the use of the classification algorithms are shown in Table 3. On the basis of the numerical results, it is clear that the RF algorithm has a higher degree of accuracy than the other two methods.

Table 3. Classification performance accuracy.

Method	Overall Accuracy (%)
RF	88.65
SVM	77.38
KNN	73.29

Figures 5–7 show the visual representation of the algorithms used on the data. Similar to the numerical analysis, the visual result presenting the performance of the three of algorithms, show that the best is the RF algorithm.

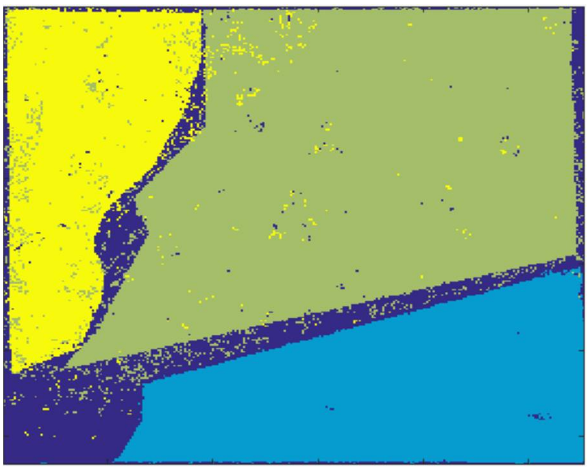

Figure 5. The classified map using the RF method.

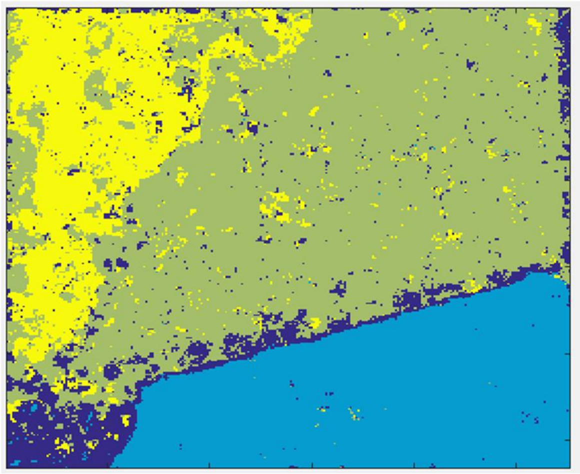

Figure 6. The classified map using the SVM method.

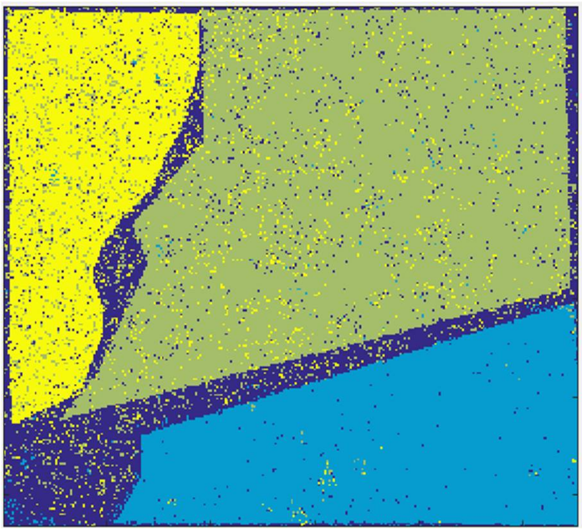

Figure 7. The classified map using the KNN method.

5. Conclusions

SAR radar data, and especially all-polar data, are used as a relatively new and very useful source of information from the ground because of the features and benefits of the imaging process. In this paper, we examined the classification methods, including SVM, KNN, and RF, in order to compare them with respect to classification speed with appropriate accuracy. For this purpose, the entire polarized image was used in three major classes: urban area, vegetation, and water. After extraction of the characteristic features, each of the mentioned methods for the purpose of classification was applied to the data. The visual and numerical results from the classifications are presented in this study. The RF classification showed better classification accuracy than the other two methods. In addition, the RF classification method required less time to process the data than the other methods. The SVM method in the KNN methodology gave a better accuracy, but it required a high processing time due to the optimization of the parameters.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Krogager, E.; Czyz, Z.H. Properties of the Sphere, Diplane Helix Decomposition. In Proceedings of the Third International Workshop on Radar Polarimetry, (JIPR'95), RESTE, Nantes, France, 1 March 1995; pp. 106–114.
2. Yekkehkhany, B.; Homayouni, S.; Mcnairn, H.; Safari, A. Multi-temporal Full Polarimetry L-Band SAR data classification for Agriculture Land cover mapping. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Quebec City, QC, Canada, 13–18 July 2014.
3. Lardeux, C.; Frison, P.-L.; Rudant, J.-P.; Souyris, J.-C.; Tison, C.; Stoll, B. Use of the SVM classification with polarimetric SAR data for land use cartography. In Proceedings of the Geoscience and Remote Sensing Symposium, IGARSS 2006, Denver, CO, USA, 31 July–4 August 2006; pp. 493–496.
4. Khosravi, I.; Mohammad-Beigi, M. Multiple Classifier Systems for Hyperspectral Remote Sensing Data Classification. *J. Indian Soc. Remote Sens.* **2014**, *42*, 423–428.
5. Noi, P.T.; Kappas, M. Comparison of Random Forest, K-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors* **2018**, *18*, 18, doi:10.3390/s18010018.
6. Lee, J.S.; Pottier, E. *Polarimetric Radar Image: From Basics to Applications*; CRC Press: Boca Raton, FL, USA, 2009.
7. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 2000.
8. Barandela, R.; Gasca, E. *Decontamination of Training Samples for Supervised Pattern Recognition Methods*; Springer: Berlin/Heidelberg, Germany, 2000; pp. 621–630.
9. Breiman, L. Manual for setting up, using, and understanding random forest. *Remote Sens. J.* **2003**, *4*, 23–41.
10. Jin, J. A Random forest based method for urban land cover classification using LiDAR data and aerial imagery. *Remote Sens. Environ.* **2012**, *81*, 190–211.
11. Zhou, Z.H. *Ensemble Methods: Foundations and Algorithms*; Taylor & Francis: Abingdon, UK, 2012; 236p.
12. Hasanlou, M.; Samadzadegan, F.; Homayouni, S. SVM-based hyperspectral image classification using intrinsic dimension. *Arabian J. Geosci.* **2015**, *8*, 477–487.