



Classification of regional land cover in ALOS PALSAR's FBD data based on support vector machines

Hongfu Wang and Xiaorong Xue

School of Computer and Information Engineering, Anyang Normal University, Anyang, Henan, China

ABSTRACT

This paper presents the assessment of ALOS PALSAR orthorectified FBD data for regional land cover classification. The SVM approach is mainly used throughout this paper. It is shown that the SVM-RFE algorithm is effective for providing an optimized set of textural parameters to be computed at large scale. In addition to this, an original methodology has been implemented with the intention to give a real insight about the usefulness of textural parameters within the SVM based classification. An optimization of the independent SVM-based classifiers and a clustering procedure complete the methodology.

Keywords: Synthetic aperture radar(SAR), FBD Data, classification, SVM

INTRODUCTION

Tropical forests are also habitat to about half of the world's species and provide a livelihood for millions of people. With the importance of avoiding deforestation and associated emissions, the issue of Reducing Emissions from Deforestation and Degradation (REDD) in developing countries has been initiated [4]–[6]. A monitoring activity in support of REDD necessitates a capacity to estimate the changes throughout all forest within a country's boundary. Satellite remote sensing plays a key role in estimating loss of the forest cover and land cover change. It is able to provide this information accurately and up-to-date, over wide areas, at a uniform sampling scheme, and repeatedly over time.

This paper investigates the abilities and the limitations of ALOS PALSAR 50-m mosaic data for land cover classification in tropical rainforest as part of the ALOS K&C Initiative Project. Incorporated with World Wildlife Fund (WWF), Riau province in central Sumatra was selected as a test site. Riau hosts some of the most biodiversity ecosystem with unique species in the world. It is covered by vast peat lands estimated to hold Indonesia's largest stock of carbon.

2. Data and legend development

Riau province has a significant amount of tropical peat lands (about 50% of its surface). The study area includes both organic and mineral soils. Before human influence, the study area was covered by different types of swamp and peat land forests and coastal mangroves. In addition, the mineral soils were occupied by several lowland and upland evergreen forest types.

The ALOS PALSAR was developed by JAXA and launched in January 2006. As the reference data set by the ALOS K&C Initiative Project, PALSAR 50-m orthorectified mosaic product is created from the ascending path using the fine-beam dual polarization (FBD) mode.

The conversion between the amplitude (DN) and σ^0 is given as follows [7]:

$$\sigma^0 = 10 \times \log_{10}(\text{DN}^2) + \text{CF} \quad (1)$$

With the calibration factor (CF) dependent on the processing date, figure 1 shows the PALSAR 50-m orthorectified mosaic product acquired in 2007 over the Riau province [6].

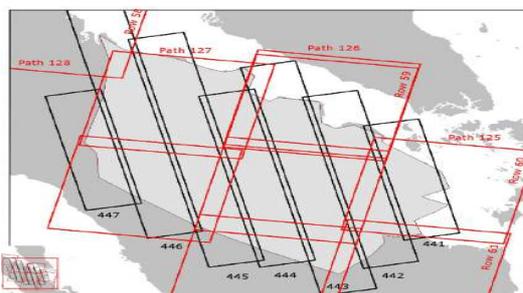


Figure1. Illustration of the different acquisitions used in this study with their corresponding footprint in Riau: PALSAR strips in black and Landsat scenes used for the WWF land cover classification in red.

3. Tools for an operational classifier at regional scale

It is very difficult to assess the complex nature of textures and use it in a comprehensive manner into a classifier. The characterization of an optimal processing window is not straightforward for a given land cover in addition to the fact that a multiscale approach has been proven to be very useful. Features Based on Spatial Information

SVM theory has been first described for binary classification computing an optimal hyper plane in a high-dimensional space. Its implementation is beyond the scope of this paper and is extensively documented in some dedicated books such as [8],

1) Binary Classifier: In the framework of this paper, 50-m mosaic PALSAR products are only used, restricting the features to be the backscattering coefficients plus some textural information. In the Euclidean space of dimension p , SVM computes a decision function (hyperplane) which separates the positive from the negative samples. The decision function can be written as $f(x) = w \cdot x + b$ where $w \in R^p$ is a weight vector and b a scalar. $b/\|w\|$ is the distance between the hyperplane and the origin where $\|w\|$ the Euclidean norm of w . An equivalent problem consists in solving the following constraint:

$$\langle w, x_i \rangle + b \geq \pm 1 \quad \forall i \quad (1)$$

where $x_i \in R^p$ is the i th sample vector. After some manipulations, the distance between the closest positive and negative samples is equal to $2/\|w\|$. Consequently, the optimal hyperplane is given by maximizing this margin, subject to the constraints defined by (1). In general, input data are nonseparable and positive slack variables ξ should be introduced in the aforementioned constraints. ξ_i is a measure of the misclassification error for the sample vector and exceeds unity in that case. The optimization problem can be written as

$$\underset{w, \xi}{\text{minimize}} \quad J = \frac{1}{2} \|w\|^2 + c \sum_i \xi_i^2 \quad (2)$$

2) Multiclass Classifier: In the first case, a set of $N_c(N_c - 1)/2$ classifiers is trained, each classifier separating a pair of classes. A simple majority vote is usually applied to compute the final class membership. In the latter case, N_c binary classifiers $f_c (c=1 \dots N_c)$ are computed, each splitting one class against the remaining $N_c - 1$ classes. The final decision y_i for a sample x_i corresponds to the maximal value (i.e., the uncelebrated distance measurement from the projected sample to the hyperplane), as follows:

$$y_i = \arg \max_c f_c(x_i) \quad (3)$$

3) Feature Selection for the Binary Case: In the case of a binary and linear problem, the squared weight $\|w\|^2$ with

$w \in R^p$ can be used as ranking criterion since it is inversely proportional to the margin. Given the p initial features, the SVM hyperplane is trained through the computation of $w_j^2 (j = 1 \dots p)$. Thus, the feature with the smallest value is discarded and the operation is iteratively operated until a smaller subset $S (\dim S < p)$ is reached. To do so, after computing the Lagrange multiplier with (3) for the whole set of p parameters, the weight is successively calculated for the data set without feature r (noted x^{-r} hereafter)

$$w_{-r}^2(\alpha) = \sum_{ij}^{N_{sv}} \alpha_i \alpha_j y_i y_j K(x_i^{-r}, x_j^{-r}) \quad (4)$$

4. Experimental results over the ROIs

This section presents some assessments for the estimation of the best textural parameters to be computed at regional scale. It also shows the intrinsic capabilities and limitations of the SVM theory for land cover classification using the single L-band dual-polarized data set.

4.1 Experimental Protocol and Relevance of the SVM-RFE Process

In order to be properly usable by the SVM approach, all the parameters are first standardized with a mean of zero and a standard deviation of one over the whole data set. The 896 spatial parameters are computed over all the ROIs. It is obvious that our reached compromise ($C = 10$ and $\sigma = 1$) is not optimized for all configurations and all binary classifiers. Nevertheless, this pair of retrieved values is consistent in our case both for the time consumption during the training/testing phases and for the classification accuracy.

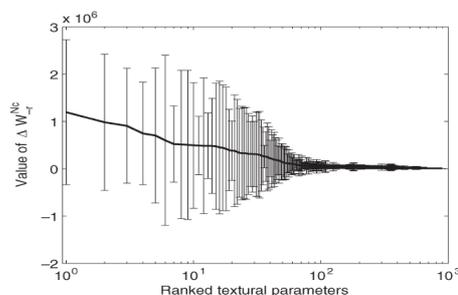


Figure 2. Evolution of $W_{-r}^{N_c}$ over the $N_c = 13$ classes during the SVM-RFE process. The bar plot indicates the mean value with its standard deviation among the ten tests.

4.2 Stability Issue and Classification Experiments

In this part, we would like to retrieve the best set of parameters from the previous experiment. In addition, all of them are not always informative depending on the selected samples: The relative high values for the standard deviation for the first ranked parameters shows that some dozens or so are really complementary. The best set of parameters must be independent of the sampled training data set. However, if selected from figure 2, the first textural features ranked on their averaged values of $\Delta W_{-r}^{N_c}$ over the ten iterations would neglect their mutual information.

This combinatorial optimization problem is solved in this paper by the Munkres assignment algorithm [3]. Knowing the weights $\Delta W_{-r}^{N_c}$ resulting from the different N_{biter} tests of the SVM-RFE process, this system can be formulated in the form of a matrix. In figure 3, the overall accuracy for the classification over all the ROIs is assessed depending on the number of parameters used as input of the SVM classifier. Two simulations are realized: The first one uses the averaged values of $\Delta W_{-r}^{N_c}$ over all the iterations as a criteria for selecting the best textural parameters, the second one using the Munkres assignment algorithm.

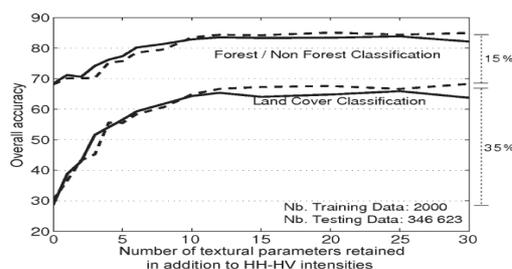


Figure3. Accuracy of the SVM-based classification at local scale depending on the number of textural parameters used as input in addition to the HH/HV channels.

When comparing the two sets of curves (forest and land cover classification), the two-class assessment leads obviously to better results (85%). The sills around 85% and 67% are confirmed when hundreds of parameters are used (not shown in figure 3.) [9].

4.3 Analysis of Textural Properties

In the framework of a binary approach, we can determine a list of features with increasingly impact on the detection of one given class. Knowing the attributes of each feature, namely, its quantization ($N_q = 8, \dots, 64$), its polarization (HH or HV) and so on, we can study this given class by analyzing the values of $\Delta W_{-r}^{N_c}$ depending on these attributes. As an example, given the best N_q textural features computed from a binary SVM-RFE, the ratio of weights coming from features with HV polarization is calculated as follows:

$$\Psi_{HV} = \frac{\sum_{r \in \{HV\}}^{N_p} r = 1 \Delta W_{-r}}{\sum_{r=1}^{N_p} \Delta W_{-r}} \quad (5)$$

5. Land covers Classification at regional scale

In this paper, we use an original twofold approach. First of all, the k-means method partitions the HH/HV channels into 15 clusters, which are used to identify homogeneous regions within the image. The different homogeneous areas are used as seed clusters. In order to produce the smooth final decision, each output of the SVM classifier is used as criteria for merging these seed clusters. For a given seed cluster with its averaged $f_c(x_i)$ value, the L^1 norm metric computes the distance with the cancroids of neighboring seed clusters and merge the closest one (this method is usually called the “average group linkage”).

The overall agreement is about 69.9% if the six classes are considered and reaches 86.5% in the case of the natural forest discrimination[10]. In a parallel study using a multiscale maximum likelihood approach and five classes without any textural information, the land cover classification has been tested over some part of strips 443/444 and its accuracy was about 60% and 53% for strip 443 and 444, respectively. As compared to our results (cf. TABLE II),

TABLE II ACCURACY OF THE LAND COVER CLASSIFICATION WITH RESPECT TO PALSAR ACQUISITION DATE AND ELEVATION

Strip number	Acquisition Date (yyymmdd)	DEM Mean / Std (in m)	OA (in %)	MPA (in%)
441	20071125	9 / 5	69.84	44.11
442	20070911	18 / 19	76.48	65.78
443	20070628	54 / 74	72.28	64.86
444	20071130	50 / 47	63.08	62.11
445	20071101	62 / 76	63.45	57.74
446	20070703	150 / 209	72.47	64.64
447	20070720	450 / 466	75.76	61.74

The gain coming from this SVM-based methodology appear to be non-negligible. Even if the Pulsar for the dry and wet forests, the oil palm plantation and the “others” are good (over 70%), the discrimination of the clear-cut areas and the acacia plantations are much more difficult. Two reasons may explain this result. As it has been explained in Section II, acacia plants grow very fast. In addition, the time difference between the Landsat and PALSAR acquisitions may lead to some misclassifications. In a similar vein, clear-cut areas may change between acquisitions.

CONCLUSION

This paper investigates the relevancy of PALSAR orthorectified FBD product at 50-m resolution for regional land cover classification by the support vector machines (SVM). The SVM-based classifier is carried out across the whole Riau province and its results are compared with a Landsat-based estimation. The agreement is over 65% with six classes and 80% for the natural forest map. These results are remarkable since only one PALSAR FBD product is used and this assessment is performed on more than 40 million pixels. The results confirm the high potential of the PALSAR sensor for forest monitoring at regional, if not global scale.

Acknowledgment

This work was supported by the Project of National Natural Science Foundation of China (U1204402) and the Natural Science Research Project of Henan Province education Department (14B520019)

REFERENCES

- [1] F. Achard, R. DeFries, H. Eva, M. Hansen, and P. Mayaux, *Environ. Res. Lett.*, vol. 2, no. 4, pp. 045022-1–045022-11, Oct.–Dec. **2007**.
- [2] Amores, J., Sebe, N., Radeva, P.: *Pattern Recogn. Lett.*, **2006**, 27, (3), pp. 201–209
- [3] Bordes, J.-B., Ma?tre, H.: ‘Semantic annotation of satellite images’. Fifth Int. Conf. on Machine Learning and Data Mining in Pattern Recognition (MLDM), **2007**, pp. 120–133
- [4] S. Saatchi, B. Nelson, E. Podest, and J. Holt, *Int. J. Remote Sens.*, vol. 21, no. 6/7, pp. 1201–1234, **2000**.
- [5] D. H. Hoekman and M. A. M. Vissers, *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 12, pp. 2881–2889, Dec. **2003**.
- [6] C.-W. Hsu and C.-J. Lin, *IEEE Trans. Neural Networks*, vol. 13, no. 2, pp. 415–425, Mar. **2002**.
- [7] A. Kalousis, J. Prados, and M. Hilario, *Knowl. Inf. Syst.*, vol. 12, no. 1, pp. 95–116, May **2007**.
- [7] T. N. Tran, R. Wehrens, and L. M. C. Buydens, *Chemom. Intell. Lab. Syst.*, vol. 77, no. 1/2, pp. 3–17, May **2005**.
- [8] M. Rosa-Zurera, A. M. Cóbrecas-Álvarez, J. C. Nieto-Borge, M. P. Jarabo-Amores, and D. Mata-Moya, ‘Wavelet denoising with edge detection for speckle reduction in SAR images,’ in Proc. EUSIPCO, Poznan, Poland, **2007**.
- [9] D. Comaniciu and P. Meer, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–618, Aug. **2002**.
- [10] P. Jarabo-Amores, M. Rosa-Zurera, D. Mata-Moya, and R. Vicen-Bueno, ‘Mean-shift’ filtering to reduce speckle noise in SAR images,’ in Proc. Int. Instrum. Meas. Technol. Conf., Singapore, **2009**, pp. 1188–1193.