RELEVANCE VECTOR MACHINE FOR EFFICIENT CLASSIFICATION OF SCATTERED PATTERNS IN HYPERSPECTRAL IMAGERY

Fereidoun A. Mianji, Student Member, IEEE, Yuhang Zhang, Ye Zhang

School of Electronics and Information Technique, Harbin Institute of Technology
fmianji@ieee.org

1. INTRODUCTION

Analysis of hyperspectral data for defining the land-cover classes through classification techniques, in particular for small patches and scattered land-covers, is not a trivial task. Factors such as high spatial variability of land-cover signatures, the “boundary effect” between neighboring land-covers, and the curse of dimensionality make this task more challenging [1]. As the integrity of a land-cover class decreases in an image, i.e. it becomes more scattered and distributed in smaller segments, its heterogeneity increases due to the presence of more mixed pixels (stronger “boarder effect”) and as a consequence the classification accuracy decreases [2].
Variety of techniques has been applied to supervised classification of remotely sensed hyperspectral imagery to deal with the classification accuracy and curse of dimensionality issue in classification of hyperspectral imagery in last decade. Among them feature reduction techniques [3], adaptive statistics estimation by exploitation of classified (semilabeled) samples [4], regularization of the sample covariance matrix [5], analysis of the spectral signatures to model the classes [6], and support vector machines [7] can be referred as the main categories of approaches. Although these approaches have obtained many achievements, they seldom take the real situation of presence of small land-cover patches or insufficiency of available training samples into account. This paper proposes a new efficient classification approach to tackle the problems of complexity and accuracy, in particular for small and scattered land-covers, through relevant vector machine (RVM).

2. MATERIALS AND METHODS

The proposed approach adopts a combined system using an appropriate data transformation technique with a multiclass RVM design to realize a noncomplex and accurate classifier.

2.1. Key information preserving feature reduction

The adopted class-oriented discriminant analysis method for this paper is Fisher linear discriminant analysis (FLDA), which is a standard supervised technique for dimension reduction in pattern recognition [8]. FLDA transformation maximizes the ratio of between class variance to within class (intraclass) variance. The model which presents the FLDA is the generalized eigen-problem specified by

$$S_W^{-1} S_B w = \lambda w \quad (1)$$

where S_B and S_W are the between and within class scatter matrices, respectively, and λ is a generalized eigenvalue.

2.2. Multiclass relevance vector machine classification

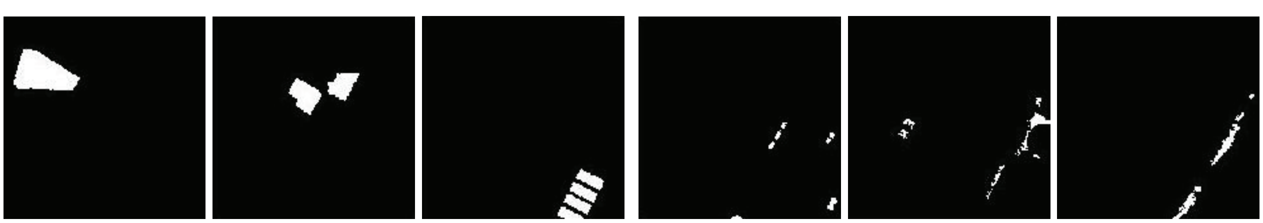
RVM is a probabilistic sparse kernel model identical to support vector machine (SVM) in functional form which doesn't suffer the drawbacks of SVM such as necessity to be a continuous symmetric kernel of a positive integral operator (Mercer's condition) and liberal use of kernel functions (insufficient sparsity) [9].

To adopt a binary classifier like RVM to the classification task of hyperspectral remote sensing data which usually involves simultaneous discrimination of numerous information classes, the general strategies are based on combining an ensemble of binary classifiers, a set of two-class problems, according to some decision rules. The adopted design for the proposed method is the one-against-one strategy with parallel architecture. In this design all possible pair-wise classifications are modeled by using $M(M-1)/2$ RVMs. Each pixel is analyzed by a discriminant function to define its belonging to one of information classes C_i , C_j where $C_i, C_j \in C, i \neq j$ and a positive score is given to the winner class of each one-against-one competition. The final decision about the class of each pixel is made on the total score each class obtains.

3. RESULTS AND DISCUSSION

3.1. The real data for experiments

The test image is a remotely sensed hyperspectral image with a ground resolution of 3.5 meter and 126 available



(a) (b) (c) (d) (e) (f)

Fig. 1. (a), (b), and (c) Large and integrated land-cover classes (C1, C2, and C3, respectively). (d), (e), and (f) Small and scattered land-cover classes (C4, C5, and C6, respectively).

bands which is collected by the AVIRIS. A ground truth map of the image is available too. For experiments, 3 land-cover classes of large and integrated shape (C1, C2, and C3) and 3 land-cover classes of scattered form including many small patches (C4, C5, and C6) are selected. Fig. 1 depicts the image and selected classes.

3.2. Experiments and results

The selected land-cover classes are used to generate sets of training samples in different sizes for each class. The proposed method (RVM+FLDA) is compared with RVM in join with the most common feature transformation technique in remote sensing, namely principal component analysis (PCA) and also SVM which is one of the most accurate supervised classification methods. Using PCA the original hyperspectral data in transformed to 3 different reduced spaces with 10, 20, and 30 bands. SVM is applied with two most efficient kernel types, namely polynomial kernel (SVM-Poly) and Gaussian radial basis function kernel (SVM-RBF). Tables 1 and 2 tabulate the obtained results for training to test sample ratios of 1/15 and 1/120, respectively.

Table 1. Single class classification accuracies and the computational times achieved through the different approaches. Train to test sample ratio: $1/15$.

Method	Classification Accuracy (%)						Times (s)	
	C1	C2	C3	C4	C5	C6	Train	Test
RVM+PCA10	99.57	96.74	98.49	90.38	88.90	94.38	403	1.8
RVM+PCA20	99.55	96.63	98.42	89.81	92.39	94.85	412	2.3
RVM+PCA30	99.57	96.76	98.52	90.15	93.14	95.18	461	2.9
SVM-Poly	99.87	97.90	99.16	91.75	92.33	95.69	4.2	3.1
SVM-RBF	99.86	97.67	98.95	90.61	90.68	95.22	1.6	6.8
RVM+FLD	99.57	98.64	99.16	97.37	94.79	96.91	662	1.3

Table 2. Single-class classification accuracies, and the computational time achieved through the different approaches. Train to test sample ratio: $1/120$.

Method	Classification Accuracy (%)						Times (s)	
	C1	C2	C3	C4	C5	C6	Train	Test
RVM+PCA10	99.84	95.00	89.45	81.56	77.75	85.95	21	1.7
RVM+PCA20	99.84	94.81	89.45	74.91	78.15	85.48	21	1.9
RVM+PCA30	99.84	95.02	89.45	81.90	77.80	86.00	18.5	2.2
SVM-Poly	99.86	94.26	95.89	82.13	85.87	87.59	1	1.1
SVM-RBF	99.84	93.43	93.41	76.63	81.52	78.13	1	2.2
RVM+FLD	99.63	97.88	98.57	94.16	89.65	96.67	44	1

In low training to test sample ratio (1/15) the proposed method is comparable with the other techniques for integrated land-cover classes (C1-C3) while is superior to the other approaches for scattered land-cover classes (C4-C6). For lower training to test sample ratio (1/120) which is a much more realistic case, it dramatically outperforms the other techniques (except for C1) in particular for scattered land-cover classes. In terms of the computational time, it is slower in training phase but is faster in testing phase which makes it an appropriate option for real-time applications with *a priori* training.

4. CONCLUSION

A new method to enhance the classification accuracy of small and scattered patches of land-covers which often contain the key information in hyperspectral imaging is proposed in this paper. It is efficient in terms of complexity too.

ACKNOWLEDEMENT

This work was supported by the Natural Science Foundation of China under the grants 60972143 and 60972144, and research fund for the doctorial program of higher education of China under the grant 20092302110033.

5. REFERENCES

[1] P. Hsieh, L. C. Lee, and N. Chen, " Effect of spatial resolution on classification errors of pure and mixed pixels in remote sensing," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, no. 12, pp 2657-2663, Dec. 2001.

[2] J. H. Smith, J. D. Wickham, S. V. Stehman, and L. Yang, "Impacts of patch size and land-cover heterogeneity on thematic image classification accuracy," *Photogrammetric Engineering & Remote Sensing*, vol. 68, no. 1, pp. 65-70, Jan. 2002.

[3] S. B. Serpico and L. Bruzzone, "A new search algorithm for feature selection in hyperspectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, pp. 1360–1367, July 2001.

[4] Q. Jackson and D. A. Landgrebe, "An adaptive classifier design for highdimensional data analysis with a limited training data set," *IEEE Trans. Geosci. Remote. Sens.*, vol. 39, pp. 2664–2679, Dec. 2001.

[5] C.-I Chang and B.-H. Ji, "Weighted abundance constrained linear spectral mixture analysis," *IEEE Trans. Geosci.Remote Sens.*, vol. 44, no. 2, pp. 378-388, 2006.

[6] F. Tsai and W. D. Philpot, "A derivative-aided hyperspectral image analysis system for land-cover classification," *IEEE Trans. Geosc. Remote. Sensing*, vol. 40, pp. 416–425, Feb. 2002.

[7] C. Huang, L. S. Davis, and J. R. G. Townshend, "An assessment of support vector machines for land cover classification," *Int. J. Remote Sens.*, vol. 23, pp. 725–749, 2002.

[8] Q. Du, "Modified Fisher's linear discriminant analysis for hyperspectral image dimension reduction and classification," *in Proc. SPIE,* 2006, vol. 6378 63781D, pp. 1-8.

[9] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, 2001.