FEATURE SELECTION FOR HYPERSPECTRAL DATA BASED ON MODIFIED RECURSIVE SUPPORT VECTOR MACHINES

Rui Zhang[1,2] *, Jianwen Ma*[3]*, Xue Chen*[1]*, Qingxi Tong*[1]

1. Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China

2. Graduate University, Chinese Academy of Sciences, Beijing, China

3. Center for Earth Observation and Digital Earth, Chinese Academy of Sciences, Beijing, China

david.zhangrui@gmail.com

1. INTRODUCTION

Feature selection, which is often employed as a pre-processing step prior to hyperspectral data classification, is an important issue in remote sensing applications, and it is an active research field in the pattern recognition and machine learning community. Many feature selection algorithms have been proposed in literatures. Among these algorithms, support vector machine (SVM) recursive feature elimination (SVM-RFE)[1] has been shown to provide superior performance for many feature selection applications, including hyperspectral data feature selection[2]. SVM-RFE utilizes the weight value calculated in the SVM training stage as the ranking criterion. A variety of methods for computing the weights has been proposed, resulting in a number of SVM-RFE variants, including recursive support vector machines (R-SVM)[3].

2. METHOD

In this paper, a new algorithm call Modified R-SVM (MR-SVM) is proposed. It follows the scheme of standard SVM-RFE, but uses a new ranking criterion derived from the R-SVM. Before the feature ranking iteration, an automatic model selection (AMS)[4] algorithm using radius margin bound is employed to eliminate noisy bands. After initializing the ranked feature set R and selected feature subset S, the feature ranking process will run iteratively until all the features are ranked, which are summarized as follows: (a) Train an SVM with features in set S as input variables; (b) compute the weight vector and the measure of discriminatory power; (c) compute the ranking scores for each feature in set S; (d) find the feature with the smallest ranking score; (e) update the set R and S.

3. DATA

To evaluate the proposed algorithm, a benchmark hyperspectral data set is used. This data set, acquired by AVIRIS in June 1992, covers an agricultural area in the state of Indiana of United States of America. The data set includes 220 spectral bands covering 0.4-2.5μm and comprises 145×145 pixels represent 16 cover classes. The nine largest of the 16 classes are used. The training samples and test samples are selected to 4671 pixels and 4674 pixels respectively. For this data set, 20 bands can be identified as dominated by noise (bands 104-108, 150-163, and 220) due to atmospheric water absorption but we do not remove these bands deliberately to evaluate the robustness of the algorithm.

4. RESULTS

The classification accuracy is used to evaluate the effectiveness of the algorithm, and the elapsed time for the feature selection measures the computational efficiency. First, we use the AMS algorithm to obtain a subset of bands that may be dominated by noise, and we remove these bands before the next step. In this paper, 50 bands are removed by the AMS. We compare the selected bands to the reference noisy bands to assess the effectiveness of the AMS algorithm. The result is presented in Table 1. Second, we use SVM to classify the extracted feature space according to the selected feature list, and the classification results for several subsets of features are recorded (the dimensions are equal to 150, 100, 50, 20, 10, and 5, respectively). The parameters used in classification are acquired by the grid search and 10-fold cross validation. The comparisons between the MR-SVM and the SVM-RFE in classification accuracies and execution time are included

This study was supported financially by the Knowledge Innovation Program of the Chinese Academy of Sciences (No.kzcx2-yw-313-3), the National High Technology Research and Development Program of China (No. 2007AA12Z157).

in the experiment. The results are presented in Table 2.

Table 1 50 Noisy bands obtained by the AMS

Selected 50 noisy bands by AMS:
154,**105**,**156**,**162**,**153**,1,**157**,**151**,**104**,**220**,**159**,**150**,109,**155**,**163**,**108**,**106**,**160**,219,**161**,**107**,165,**158**,218,**152**,96,149,86,2,83,36,84,88,80,87,98,95,97,164,3,85,82,110,93,99,90,81,103,91,78
Reference noisy bands:
104,**105**,**106**,**107**,**108**,**150**,**151**,**152**,**153**,**154**,**155**,**156**,**157**,**158**,**159**,**160**,**161**,**162**,**163**,**220**

Table 2 Classification results obtained by the MR-SVM and SVM-RFE for the AVIRIS data set. OA is overall accuracy, FS Time is elapsed time for feature selection, and Accu. % is accuracy in percent

Method		Number of Bands							OA %	FS Time
		220	150	100	50	20	10	5		
MR-SVM	Accu. %	91.7	92.8	93.3	91.3	89.0	87.8	78.4	89.2	743s
SVM-RFE	Accu. %	91.7	90.8	89.5	91.0	89.3	88.8	76.4	88.2	21242s

Table 1 shows the selected 50 noisy bands obtained by the AMS process, which runs before the feature ranking iteration. Obviously, all the 20 noisy bands due to atmospheric water absorption are found by the AMS algorithm effectively in terms of reference noisy bands list, and this step will increase the robustness of the entire feature selection process.

From Table 2, the overall accuracy of the MR-SVM outperforms the result obtained by the SVM-RFE. When the dimensions of the features are equal to 150, 100, 50 and 5, the classification accuracy obtained by the MR-SVM is higher than the SVM-RFE. Only when the dimensions are equal to 20 and 10, the accuracy of the SVM-RFE outperforms the MR-SVM a little. The execution time is a very important indicator for feature selection algorithms. The time needed by the MR-SVM dramatically reduced compared to the SVM-RFE; hence, the MR-SVM is a more practical algorithm than the SVM-RFE in the feature selection for hyperspectral data.

5. CONCLUSIONS

An SVM based feature selection algorithm for hyperspectral data MR-SVM is developed in this paper. For the test case presented, it is competitive with the state-of-the-art feature selection algorithm SVM-RFE in classification accuracy, and it shows robustness when presented with noisy data. Moreover, the processing time for feature selection using the MR-SVM algorithm is significantly reduced compared to that of SVM-RFE.

REFERENCES

[1] I. GUYON, J. WESTON, S. BARNHILL and V. VAPNIK, "Gene selection for cancer classification using support vector machines," *Machine Learning*, vol. 46, no. 1-3, pp. 389-422, 2002.

[2] Y. BAZI and F. MELGANI, "Toward an optimal SVM classification system for hyperspectral remote sensing images, " *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3374-3385. 2006.

[3] X. G. ZHANG, X. LU, Q. SHI, X. Q. XU, H. C. E. LEUNG, L. N. HARRIS, D J. IGLEHART, A. MIRON, J. S. LIU and W. H. WONG, "Recursive SVM feature selection and sample classification for mass-spectrometry and microarray data," *Bmc Bioinformatics*, vol. 7, no. 197, Apr 2006.

[4] O. CHAPELLE, V. VAPNIK, O. BOUSQUET, AND S. MUKHERJEE, "Choosing multiple parameters for support vector machines," *Machine Learning,* vol. 46, pp. 131-159, 2002.