AN EMPIRICAL MODE DECOMPOSITION AND COMPOSITE KERNEL APPROACH TO INCREASE HYPERSPECTRAL IMAGE CLASSIFICATION ACCURACY

Begüm Demir, Sarp Ertürk

Kocaeli University Laboratory of Image and Signal Processing (KULIS)
Electronics and Telecomm. Eng. Dept., Veziroglu Campus, Kocaeli, Turkey
{begum.demir, sertur}@kocaeli.edu.tr

1. INTRODUCTION

This paper proposes to apply empirical mode decomposition (EMD) to each hyperspectral image band and use composite kernels to combine the information of separate IMFs. EMD is a signal decomposition algorithm that decomposes signals into several Intrinsic Mode Functions (IMFs) and a final residue [1, 2]. Readers are referred to [1, 2] for detail on EMD, which will not be provided in this abstract due to space limitations. In this paper, initially, EMD is applied to each hyperspectral image band individually in the spatial domain and IMFs of hyperspectral image bands are obtained. Then, the information contained in the first IMFs and second IMFs of each band are combined using composite kernels and classification using kernel based Support Vector Machine (SVM) algorithm is performed. Composite kernel SVM classification that combines spectral and spatial information has formerly been proposed in [3] using direct summation and weighted summation kernels to combine spectral and spatial information. In this paper, for each hyperspectral image pixel, two different feature vectors are constructed, one corresponding to the spectral information obtained from the first IMF of each band, and the other corresponding to the spectral information obtained from the second IMF of each band. The feature vectors corresponding to the first and second IMFs can be expressed as

$$\begin{aligned} \mathbf{x}_u^{imf1} &= IMF_{l,1}(ui,uj) \;\; l=1,2,...,L \\ \mathbf{x}_u^{imf2} &= IMF_{l,2}(ui,uj) \;\; l=1,2,...,L \end{aligned} \tag{1}$$

where $IMF_{l,1}$ and $IMF_{l,2}$ show the values of the first IMF and second IMF of the l-th hyperspectral image band, respectively and (ui,uj) shows the spatial location, and L is the total number of bands. After $\mathbf{x}_u^{imf1}$ and $\mathbf{x}_u^{imf2}$ are constructed, the individual kernel matrices are computed. In the proposed composite kernel based approach, the direct summation kernel as shown in (2) and the weighted summation kernel given in (3) are used to combine these kernels.

$$K(\mathbf{x}_u,\mathbf{x}_v) = K(\mathbf{x}_u^{imf1},\mathbf{x}_v^{imf1}) + K(\mathbf{x}_u^{imf2},\mathbf{x}_v^{imf2}) \tag{2}$$

$$K(\mathbf{x}_u,\mathbf{x}_v) = \mu K(\mathbf{x}_u^{imf1},\mathbf{x}_v^{imf1}) + (1-\mu) K(\mathbf{x}_u^{imf2},\mathbf{x}_v^{imf2}) \tag{3}$$

Here, $K(.,.)$ is a kernel function and μ provides a tradeoff between the information of first IMF and second IMF to classify a given pixel and can takes values between zero and one.

2. EXPERIMANTAL RESULTS

The Indian Pine Hyperspectral image [4] is used in experiments. The total number of samples corresponding to each selected class are as follows: Corn-no till 1434 samples, Corn-min till 834 samples, Grass/Pasture 497 samples, Grass/Trees 747 samples, Hay-windrowed 489 samples, Soybean-no till 968 samples, Soybean-min till 2468 samples, Soybean-clean till 614 samples, and Woods 1294 samples (Total of 9345 samples). The classification performances of the proposed approach are demonstrated using SVM classification with Radial Basis Function (RBF) kernel. In the experiments, the penalty parameter of SVM is set to 40 and the gamma parameter of the RBF kernel is tested between [0.1-2] by using a five fold cross validation. The classification accuracies are evaluated in terms of overall accuracy (OA) as well as kappa coefficient (k). Classification results are presented with respect to training data rates (TDR) of 10% which illustrates the case where 10% of the total data samples are used as training data. Proposed algorithm is compared to direct SVM and composite kernel SVM classification which combines spatial and spectral information (denoted as CK-SS-SVM) [3], as well Morphological Profile based classification (denoted as EMP) [5]. Spatial feature vectors are obtained using the mean of the neighborhood pixels of a corresponding feature vector in a 5×5 sized window for CK-SS-SVM. In Table I, it is seen that proposed approach (denoted as CK-EMD-SVM) gives superior performance compared to direct SVM as well as CK-SS-SVM. Because the μ parameter of the weighted summation kernel is varied between [0-1], the mean and standard deviation of OA and k values are given in case of weighted summation kernels. Fig. 1 shows a sample original band of the Indian Pine hyperspectral image with its first IMF and second IMF. Experimental results show that the proposed approach significantly improve classification accuracy. Note that results for different TDR rates as well different data sets (such as the DC Mall image) can not be provided in the abstract because of space limitations but will be included in the full paper.

3. REFERENCES

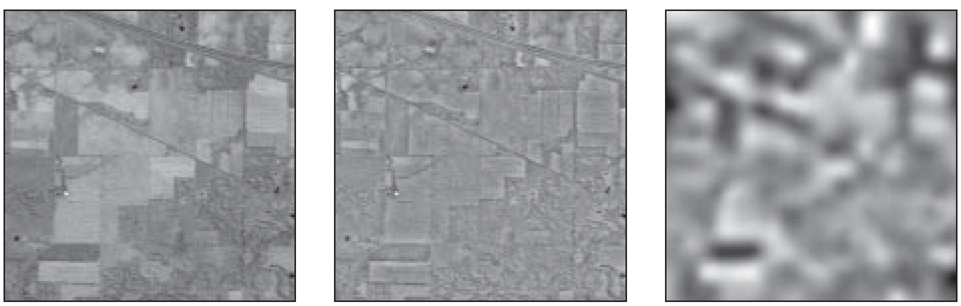
[1] N. E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N-C. Yen, C.C. Tung, and H.H. Liu, "The empirical mode decomposition and the hilbert Spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. London. A*., vol. 454, 1998, pp. 903-995.

[2] A. Linderhed, "Adaptive Image Compression with Wavelet Packets and Empirical Mode Decomposition", Linköping Studies in Science and Technology, Dissertation No. 909, ISBN 91-85295-81-7, 2004.

[3] G. Camps-Valls, L. G´omez-Chova, J. Mu˜noz-Mar´ı, J. Vila-Franc´es, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 1, pp. 93–97, Jan 2006.

[4] D. Landgrebe, 'AVIRIS NW Indiana's Indian Pines 1992 data set,'http://dynamo.ecn.purdue.edu/~biehl/MultiSpec/documentation.html

[5] J. A. Palmason, J. A. Benediktsson, J. R. Sveinsson, and J. Chanussot, "Classification of hyperspectral data from urban areas using morphological preprocessing and independent component analysis," *in Proc. IGARSS*, Jul. 2005, vol. 1, pp. 176–179.



(a) (b) (c)

Fig 1. Indian Pine Image Band #102: (a) original band (b) First IMF (c) Second IMF.

TABLE I

OA AND K VALUES OF DIRECT SVM , EMP, CK-SS-SVM AND CK-EMD-SVM USING 10 % TDR FOR INDIAN PINE DATA

Composite Kernel Approach	Method		10% TDR	
			OA	k
-	SVM		82.24	0.70
-	EMP [5]		94.56	0.73
Direct	CK-SS-SVM [3]		88.69	0.73
	CK-EMD-SVM		**94.98**	0.77
Weighted	CK-SS-SVM [3]	mean	86.15	0.68
		std	1.06	0.02
	CK-EMD-SVM	mean	**94.64**	0.76
		std	0.88	0.007