DIMENSIONALITY REDUCTION OF HYPERSPECTRAL DATA: ASSESSING THE PERFORMANCE OF AUTOASSOCIATIVE NEURAL NETWORKS

G. Licciardi, F. Del Frate, G. Schiavon, D. Solimini

Earth Observation Laboratory – DISP
Tor Vergata University, Rome, Italy
Tel. +390672597734, delfrate@disp.uniroma2.it

1. INTRODUCTION

Hyperspectral data vectors are composed by a relatively high number of components, given the narrowness of the spectral channels with respect to the instrument observational bandwidth. A majority of image processing techniques present an overall complexity that increases with increasing number of spectral bands. Hence, the design of effective procedures aiming at lowering the size of the data, but able to preserve the data information content, is crucial. Indeed, the high correlation among channels introduces redundancy into the measurement vector, so that including a large number of spectral samples often contributes more noise than independent pieces of information, in addition to adding computational cost. As a result, the information relevant to the user may be expected to be obscured and the accuracy of the final product degraded. Both issues call for a dimensionality reduction as a key issue towards enhanced performance of hyperspectral observations.

2. FEATURE EXTRACTION

Feature extraction reduces data dimensionality by mapping the feature space onto a lower-dimensional space. Both unsupervised and supervised linear transformation methods have been developed [1]. Two among the most common unsupervised techniques are the Principal Component Analysis (PCA) [2] and the Maximum Noise Fraction (MNF) [3]. A recently proposed technique [4] is based on Neural Networks (NN), already recognized as a rather competitive family of algorithms for hyperspectral data analysis [5]. The NN features extraction is carried out by Auto-Associative Neural Networks AANN, which generate nonlinear features from the data [6], [7]. AANN's have feed-forward connections and sigmoidal nodal transfer functions, trained by back-propagation or similar algorithms [8]. The implemented network architecture has three hidden layers, including an internal "bottleneck" layer of smaller dimension than either input or output. The network is trained to perform the identity mapping, where the input is approximated by the output. Since the number of units in the bottleneck layer is lower than those in both input and output, the bottleneck nodes are able to encode the input information for subsequent reconstruction in the output. This process results in a feature extraction from the input vector.

3. METHODOLOGY AND EXPERIMENT

The performance of the aforementioned AANN technique has been quantitatively evaluated and critically analyzed either from the point of view of its capability of reproducing the set of hyperspectral data by a reduced number of components or in terms of the accuracy of the derived final product. This latter consists of a land cover map generated by a dedicated NN composed by Multi-Layer Perceptrons (MLP), accepting as input the reduced vector provided by the AANN. We observe that features extraction plays an even more crucial role in a NN classification scheme. In fact, it has been observed that minimizing the number of inputs of a NN still avoiding significant loss of information, generally affects positively its mapping ability and computational efficiency.

Two test sites have been selected for the performance assessment. The first one is the area of DEMMIN (Durable Environmental Multidisciplinary Monitoring Information Network). This includes a group of farms with an extension of approximately 25,000 ha. The field sizes in this area are quite large, 200–250 ha, on average and the main crops are wheat, barley, rape, maize and sugar-beet. The range of elevation does not exceed 50 m over the whole test site. A set of data was acquired by the INTA-AHS instrument, in the framework of the ESA AGRISAR measurement campaign [9]. The second test site is in the Frascati-Tor Vergata area located South East of Rome, Italy. This is a mainly flat zone with an interesting heterogeneous landscape. Parcels with permanent plants such as vineyards, orchards and olive groves, are mixed with fields of seasonal crops, mainly maize, and uncultivated areas or pasture. Man-made land cover consists of residential urban areas, industrial and commercial units, and different kinds of road networks. The experiment on this area utilizes the data acquired by the Compact High-Resolution Imaging Spectrometer (CHRIS), included in the payload of the Project for On-Board Autonomy (PROBA)-1 satellite [10].

4. RESULTS

A preliminary analysis based on PCA was carried out to determine the number of PCA components containing most of the statistical information. The computations show that the first 5 PCA components contain almost the 99.9% of the whole statistical information. For this reason, the five principal components form the benchmark for the comparison. An example of the obtained results is given in Fig. 1A, which shows the first principal component produced by the different features extraction methodologies from the AHS data taken over the DEMMIN test site. The MNF component is apparently disturbed. In fact, due to the intrinsic light dispersion properties of grating spectrometers and to minor misalignment of optical components, the wavelength mask for pixels near the center of an array can slightly differ from that relative to pixels on the sides [11]. This effect is often referred as the ‘‘smile’’ or ‘‘frown’’ and appears to severely affect the MNF approach, at least in the case at hand. Conversely, the NLPCA technique appears to be rather robust to this type of noise, while a slight disturbing pattern due to the *smile* effect might be the cause of the brighter area on the right side of the first PCA component map. This study carries out a comparative analysis of the performance of the considered techniques in reducing dimensionality of the hyperspectral data. Results on the accuracy of land cover pixel-based maps yielded by the reduced vector and the MLP-NN algorithm are also quantitatively discussed by confusion matrices computed over extended ground-truths for both test-sites. A final issue concerns the computer time for the training phase of the MLP classifier: we report on the computational efforts associated with the considered features extraction approach, always considering the same number of training pixels and the same NN topology.

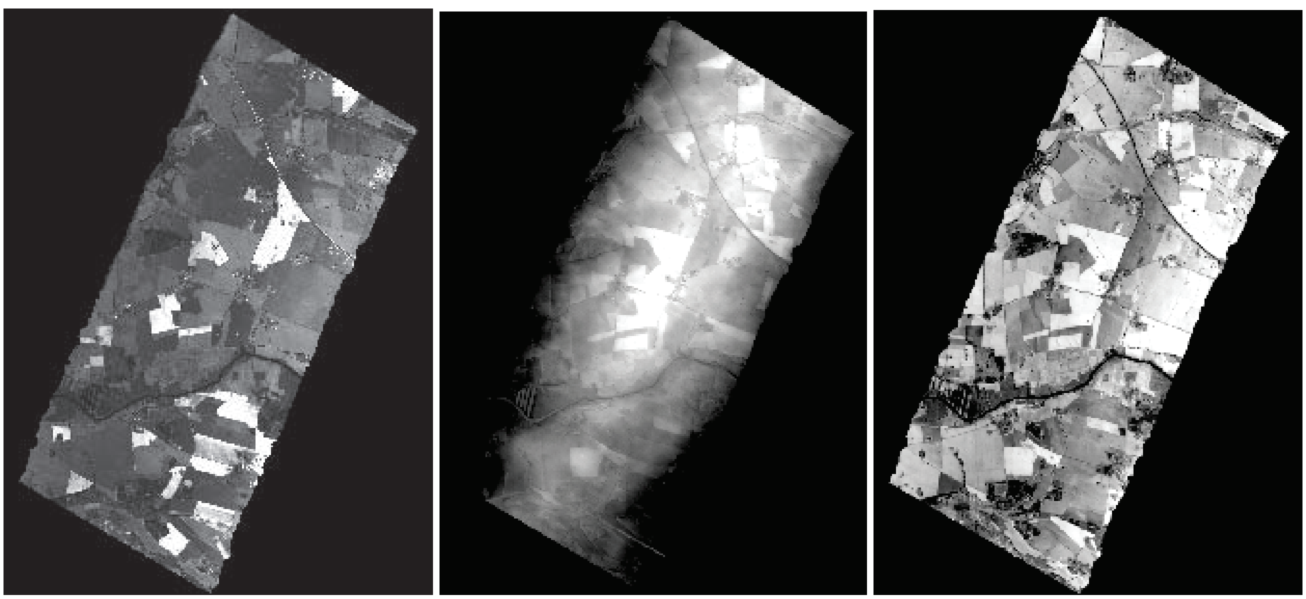

Fig. 1: Images illustrating the 1st component for NLPCA (left), MNF (center), PCA (right)

5. REFERENCES

[1] Bor-Chen Kuo, Cheng Hsuan Li, and Jinn-Min Yang, “Kernel Nonparametric Weighted Feature Extraction for Hyperspectral Image Classification,” *IEEE Trans. Geosci. Remote Sens.,* vol. 47, no. 4, pp. 1139–1155, Apr. 2009.

[2] Singh A, Harrison A., “Standardized principal components,” *Int J Remote Sensing*; 6:883–96, 1985.

[3] Green, A., Berman, M, Switzer, P., Craig, D., “A Transformation For Ordering Multispectral Data In Terms Of Image Quality With Implications For Noise Removal,” *IEEE Trans. Geosci. Rem. Sens.,* 26(1): 65-74, 1988.

[4] G. Licciardi, F. Del Frate, R. Duca, "Feature Reduction of Hyperspectral Data Using Autoassociative Neural Networks Algorithms," *Proceedings of International Geoscience and Remote Sensing Symposium*, Cape Town, South Africa, 13-17 July, 2009

[5] G. Licciardi, F. Pacifici, D. Tuia, S. Prasad, T. West, F. Giacco, C. Thiel, J. Inglada, E. Christophe, J. Chanussot, and P. Gamba, "Decision Fusion for the Classification of Hyperspectral Data: Outcome of the 2008 GRS-S Data Fusion Contest," *IEEE Trans.Geoscience and Remote Sensing*, vol 47, n. 11, pp. 3857-3865, November 2009

[6] F. Del Frate, G. Schiavon, "Nonlinear principal component analysis for the radiometric inversion of atmospheric profiles by using neural networks," *IEEE Transactions on Geoscience and Remote Sensing* , vol. 37, n. 5, pp. 2335--2342, 1999.

[7] M. A. Kramer, "Nonlinear principal component analysis using autoassociative neural networks," *J. Amer. Inst. Chem. Eng. (AIChE),* vol. 37, pp. 233–243, 1991.

[8] C. Bishop, *"Neural Networks for Pattern Recognition,"* Clarendon Press Oxford, 1995

[9] AGRISAR 2006 Final Report, Document ID 19974/06/I-LG

[10] R. Duca and F. Del Frate, "Hyperspectral and Multi-Angle CHRIS Proba Images for the generation of land cover maps," *IEEE Trans. Geosci. Remote Sens.,* vol. 46, n. 10, pp. 2857-2866, Oct. 2008

[11] Bo-Cai Gaoa, Marcos J. Montesa, Curtiss O. Davis, "Refinement of wavelength calibrations of hyperspectral imaging data using a spectrum-matching technique," *Remote Sensing of Environment,* 90 (2004) 424–433