

A NEURAL NETWORK APPROACH FOR PIXEL UNMIXING IN HYPERSPECTRAL DATA

Giorgio Licciardi, Fabio Del Frate

Earth Observation Laboratory – Tor Vergata University
Via del Politecnico, 1 – 00133 Rome, Italy
licciardi@disp.uniroma2.it, delfrate@disp.uniroma2.it

ABSTRACT

Neural networks algorithms have already shown good capabilities in handling nonlinear inversion problems in hyperspectral remote sensing. In this study we investigate on their potential in solving spectral unmixing. A Multi-Layer Perceptron (MLP) neural network scheme is trained for the implementation of a pixel-based classification algorithm. Subsequently, for the output response, the “winner-takes-all” rule is replaced by a more soft interpretation able to give the percentage with which, each of the considered land cover classes, may be associated to the analysed pixel. In an experimental set-up addressing multi-temporal and multi-angular CHRIS-PROBA imagery, the results obtained with such a technique have been compared with those yielded by Linear Spectral Unmixing (LSU), up to date one of the most frequently used approach for dealing with the unmixing problems.

Index Terms— Spectral Unimixing, Hyperspectral, Neural Networks

1. INTRODUCTION

In hyper-spectral imagery, mixed pixels are a mixture of distinct substances, and they exist for one of two reasons. First, if the spatial resolution of a sensor is low enough that disparate material can jointly occupy a single pixel, the resulting spectral measurement will be some composite of individual spectra. Second, mixed pixels can result when distinct materials are combined into homogeneous mixture. This circumstance can occur independently of the spatial resolution of the sensor.

The basic premise of mixture modeling is that, within a given scene, the surface is dominated by a small number of different materials, all having relatively constant spectral properties, the so-called endmembers. If we assume that most of the spectral variability within a scene results from the varying proportions of the endmembers, it consequently follows that some combinations of their spectral properties can model the spectral variability observed. If the

endmembers in a pixel appear in spatially segregated patterns similar to a square chessboard, the adopted model is basically linear. In this case the spectrum of a mixed pixel is a linear combination of the endmember spectra weighted by the fractional area coverage of each endmember in a pixel. This model can be expressed by:

$$x = \sum_{i=1}^M a_i s_i + w = Sa + w$$

where x is the received pixel spectrum vector, S is the matrix whose columns are the $M = 1, \dots, i$ endmembers, a is the fractional abundance vector and w is the additive observation noise vector. Otherwise, if the components of interest in a pixel are in an intimate association, like sand grains of different composition in a beach deposit, light typically interacts with more than one component as it is multiply scattered, and the mixing between these different components is a nonlinear process. Which character (linear or nonlinear) dominates the spectral signature of mixed pixel is still an unresolved issue. Several applications have demonstrated that the linear approach is a useful technique for interpreting the variability in remote sensing data [1]. Despite the obvious advantages of using a nonlinear approach for intimate mixtures, it has not been widely applied to remotely acquired data, because the particle size, together with composition, and alteration state of the endmembers are essential controlling parameters of the solutions. For this reason, the Linear Mixing Model is considered to be the most frequently used model for representing the synthesis of mixed pixels from distinct endmembers [2].

The complete unmixing problem can be decomposed as a sequence of three consecutive procedures:

- **Dimensionality reduction:** Reduce the dimensionality of the input data vector;

- **Endmember determination:** Estimate the set of distinct spectra in the scene;
- **Inversion:** Estimate the fractional abundances of each mixed pixel from its spectrum and the endmember spectra.

If dimensionality reduction, by itself, is not a necessary step for unmixing, it is natural to retain it as a stage of unmixing because this technique may expedite subsequent processing phases.

2. METHODOLOGY

In the proposed study the implementation of both the dimensionality reduction and the inversion phases rely on neural networks algorithms therefore more in general we can say that the study proposes a neural network approach for pixel unmixing. Neural networks models have been already proven to own good properties in handling complex nonlinear problems. They learn their discriminating relationships directly from the data and do not require particular a-priori knowledge on the quantities to be estimated. It has already been shown how, in the context of hyperspectral imagery, neural networks represent a rather competitive family of algorithms for the classification of the data [3]. They have also been successfully applied for the design of one of the first end-to-end processing scheme dedicated to hyperspectral imagery provided by the Compact High-Resolution Imaging Spectrometer (CHRIS), on board of the Project for On-Board Autonomy (PROBA) satellite [4].

As shown in [5], dimensionality reduction can be performed by NN through an auto-associative architecture. The particular network topology employs 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 at the output layer. Since there are fewer units in the bottleneck layer than the output, the bottleneck nodes must represent or encode the information in the inputs for the subsequent layers to reconstruct the input. Hence a feature extraction from the input vector is performed and the network is said to provide a nonlinear principal component analysis (NLPCA).

As far as the endmember determination is concerned, this has been based on the selection of the classes characterizing the final land cover maps. Finally, the abundances estimation was carried out through the analysis of the output of the classification provided by the neural network algorithm trained using as input the nonlinear principal components provided by the bottleneck layer of the AANN and as output the desired classification response. For this latter, in the learning phase, no mixed training

pixels have been used and the following standard code was considered: the output unit associated to the actual land cover class had value “1” while the remaining ones had value “0”. Although the network is trained considering binary values in the output vector, the activations functions of its processing units are the real valued sigmoidal functions, providing an output value in the range [0,1]. Therefore such a value can be considered as an abundance measurement. More analytically the abundance a_i corresponding to i -esm class is given by the following expression:

$$a_i = \frac{o_i}{\sum_{k=1}^M o_k}$$

where o_k indicates the neural network output associated to the k -esm endmember and M is the total number of endmembers, in this case $M = 10$.

3. MATERIAL AND DATA SET

The chosen test site is the area surrounding Tor Vergata University and the town of Frascati. This is a mainly flat area located in the southeast of Rome, Italy, which represents an interesting heterogeneous landscape. The neural network methodology was applied to a combination of hyper-spectral multi-angular and multi-temporal CHRIS-PROBA mode-3 acquisitions. Each acquisition consisted of 18 measurements and a set of 72 measurements was obtained adding a 36° acquisition to a 3 dates set of measurements taken at nadir.

The multi-temporal dataset was composed by three acquisitions, taken on February 28, 2006, August 19, 2006 and October 9, 2006. Such dates are, in principle, particularly suitable to sample the crops' growth cycle, hence to catch the differences among the multi temporal signatures associated to each land cover type. It should be added that the images underwent atmospheric calibration and other pre-processing stages such as destriping according to the procedures indicated in [4]. The number of nodes in the NLPCA bottleneck layer was chosen through a comparison with a standard PCA. The selected topology consists of 72-25-5-25-72 nodes. In fact, it resulted that the first 5 PCA components contained almost the 99% of the whole statistical information. The choice of 25 units for the two intermediate layers resulted from a comparison of various network topologies. In a successive step the 5 nonlinear components have been used to produce the land cover map of the test area. This time the MLP topology is formed by two hidden layers of 36 units while the following classes have been considered for the output: vineyards

(VY); pasture (PS); permanent crops (PC); industrial (IN); dark asphalt (DA); maize (MA); built-up area (BU); bright asphalt (BA); agricultural area (AA).

4. RESULTS

The neural algorithm classification has been trained using a training set and a test set of 3300 and 1975 patterns, respectively. The number of training epochs necessary to get the network trained is about 130, which is significantly lower respect to the case where the 72 measurements are given straightforward to the net.

The abundances estimation was obtained through the analysis of the output of the classification of the neural network algorithm adopting the procedure explained in the previous section. The value of each output was used as an estimator of the fractional abundances of each class. The technique was compared through a quantitative exercise with the Linear Spectral Unmixing (LSU) method [6]. In particular a ground-truth in terms of percentages of abundances of the considered classes was determined by visual inspection using a very high resolution panchromatic Quickbird image for a certain number of pixels. It should be noted that the considered CHRIS-PROBA imagery pixel size resolution is 18 m, so the panchromatic Quickbird characterized by a less than 1 m resolution can be recognized as suitable for producing the ground-truth reference. Table 1 reports the mean and the standard deviation values of each endmember considering the whole set of measured CHRIS-PROBA pixels.

In Fig. 1 we report some examples of results obtained on selected pixels extracted from the data set. It can be noted that the LSU was unable to correctly estimate the abundances of the elements, providing values always above 0 also in those pixels where the corresponding elements are not present. This behavior can be explained considering that standard unmixing methodologies such as LSU, to produce good result, require the endmembers to be the most uncorrelated as possible. Usually the endmembers extracted from the image, correspond to macroscopic objects in the scene such as water, soil or vegetation. Differently, in our case, the endmembers correspond to classes that may be very correlated each other, leading a standard unmixing technique to a wrong result. On the other hand, the unmixing technique using neural network provides a good accuracy even if the chosen endmembers are closely correlated. In Table 2 we report the quantitative assessment in terms of RMSE (Root Mean Square Error) computed over the entire data set considering all classes. From the reported values we see that NN seems to be definitely more effective than LSU.

	Mean	St. Dev.
VY	0,17	0,26
PS	0,16	0,30
PC	0,17	0,26
IN	0,06	0,21
DA	0,10	0,26
MA	0,00	0,00
BU	0,12	0,24
BA	0,16	0,29
AA	0,07	0,15

Table 1. Statistics of the abundances values for the considered data set. See the text for the class code.

	LSU	NN
RMSE	0.35	0.08

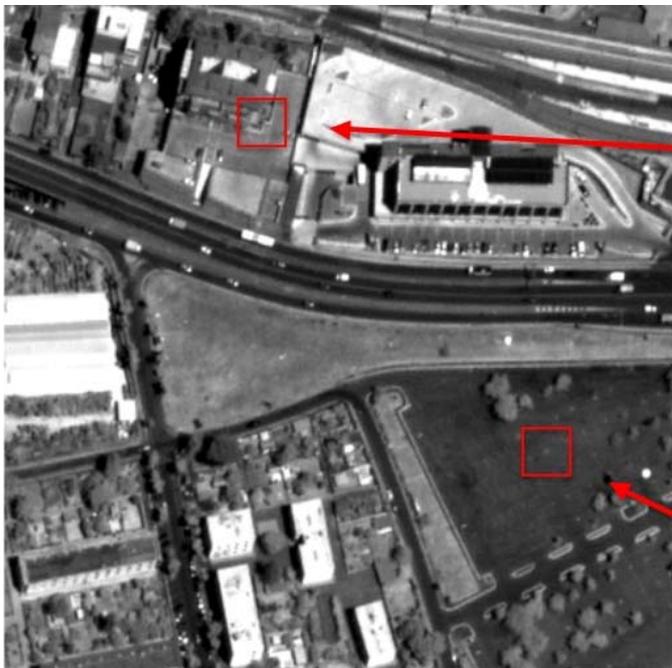
Table 2. RMSE evaluated for the estimation of the abundances of the pixels in the considered data set using the neural network and the LSU approaches.

5. REFERENCES

- [1] P. Mitra, C. A. Murthy, and S. K. Pal, "Unsupervised feature selection using feature similarity," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 3, pp. 301–312, 2002.
- [2] 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, no. 7, pp. 1360–1367, 1994.
- [3] G. Licciardi, F. Pacifici, D. Tuia, S. Prasad, T. West, F. Giacco, 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 contes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3857–3865, 2009.
- [4] R. Duca and F. Del Frate, "Hyperspectral and Multi-Angle CHRIS Proba Images for the generation of land cover maps," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, n. 10, pp. 2857-2866, October 2008
- [5] F. Del Frate, G. Licciardi, R. Duca, "Autoassociative Neural Networks for Features Reduction of Hyperspectral Data," *Proceedings of the First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 26-28 August 2009, Grenoble, France
- [6] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, 2002.



	LSU	NN	GT
VY	1,04%	0,10%	0,00%
PS	9,83%	42,38%	40,00%
PC	21,05%	0,09%	0,00%
IN	0,00%	5,88%	0,00%
DA	7,59%	0,00%	0,00%
MA	2,02%	0,00%	0,00%
BU	18,79%	0,01%	0,00%
BA	12,62%	51,38%	60,00%
AA	27,06%	0,16%	0,00%



	LSU	NN	GT
VY	9,69%	0,04%	0,00%
PS	12,58%	0,02%	0,00%
PC	24,47%	0,04%	0,00%
IN	0,00%	1,61%	0,00%
DA	2,99%	0,00%	0,00%
MA	9,11%	3,19%	0,00%
BU	20,97%	51,15%	50,00%
BA	4,80%	43,57%	50,00%
AA	15,39%	0,38%	0,00%

	LSU	NN	GT
VY	10,36%	5,92%	0,00%
PS	11,34%	0,20%	0,00%
PC	0,00%	1,39%	5,00%
IN	14,83%	0,02%	0,00%
DA	13,17%	0,02%	0,00%
MA	8,19%	0,03%	0,00%
BU	9,18%	0,19%	0,00%
BA	6,52%	92,18%	95,00%
AA	26,42%	0,06%	0,00%

Fig. 1: Examples extracted from the data set of ground-truth and estimated abundances