Automated Generation of Semi-Labeled Training Samples for Nonlinear Neural Network-Based Abundance Estimation in Hyperspectral Data

Javier Plaza, Antonio Plaza, Rosa Pérez and Pablo Martínez
Neural Networks and Signal Processing Group
Department of Computer Science, University of Extremadura
Avda. de la Universidad s/n, E-10071, Cáceres, Spain
jplaza@unex.es

***Abstract*—As the initial stage of a supervised classification, the quality of training has a significant effect on the entire classification process and its accuracy. In hyperspectral data analysis, a judicious selection of training samples can be tremendously difficult due to the presence of subpixel targets and mixed pixels, in particular, when no prior knowledge about the data is available. For instance, the Multi-Layer Perceptron (MLP) neural network can provide very accurate nonlinear estimations of fractional abundances, provided that the training set contains all possible mixture conditions. However, the requirement of large volumes of training data is a serious limitation in remote sensing because, even if classes concurring to a per-pixel cover class mixture are known, proportions of these classes are very difficult to be estimated a priori. This paper investigates, explores and further proposes solutions to resolve the issues above. Specifically, we develop a nonlinear neural network-based mixture model, coupled with unsupervised algorithms for automated generation of semi-labeled samples that can be effectively used for mixed pixel classification. These unsupervised algorithms, intended for situations where ancilliary information is difficult to be collected prior to data analysis, rely on the principle that patterns that lie close to the location of decision boundaries in feature space are more informative than patterns drawn from the class cores. Computer simulations and real experiments are conducted for performance analysis of nonlinear unmixing techniques based on training samples.**

Keywords- Hyperspectral imaging, Nonlinear mixture analysis, Training samples, Semi-labeled samples, Multi-layer perceptron.

I. INTRODUCTION

Most of the pixels collected by hyperspectral imagers contain the resultant mixed spectrum from the reflected surface radiation of subpixel constituent materials within the pixel. Mixed pixels exist for several reasons. First, if the spatial resolution of the sensor is not high enough to separate different pure signature materials at a macroscopic level (endmembers), these can jointly occupy a single pixel, and the resulting spectral measurement is a composite of the individual spectra. Second, mixed pixels can also result when distinct materials are combined into a homogeneous mixture. This circumstance also occurs independent of the spatial resolution of the sensor.

Spectral mixture analysis involves the separation of a pixel spectrum into its component endmember spectra and the estimation of the abundance value for each endmember in the pixel. The use of a linear spectral mixture model assumes that the collected spectra are linearly mixed. The definition of a linear (macroscopic) mixture is that endmember substances are sitting side-by-side within the field of view of the imager. Although the linear model has several advantages including ease of implementation and flexibility in different applications, there are many naturally occurring situations where a nonlinear mixture model may better describe the resultant mixed spectra for certain endmember distributions. In particular, nonlinear mixtures occur in situations where endmember components are randomly distributed throughout the field of view of the instrument.

In recent years, neural networks have demonstrated great potential as a method to decompose mixed pixels due to the inherent capacity of neural architectures to approximate nonlinear functions [1]. Although many types of neural networks exist, for decomposition of mixed pixels in terms of nonlinear relationships mostly feed-forward networks such as the multi-layer perceptron (MLP) have been used. Despite some encouraging results, exploitation of neural networks in mixed pixel classification and, specifically, in pixel unmixing, remains difficult. This is because the requirement of large and representative training sets is a serious limitation in remote sensing due to extremely difficult training data collection, in particular, when the information to be collected has to do with endmember fractional abundances. Therefore, a great challenge in neural network-based analysis of remotely sensed imagery is to find an adequate pool of training samples without prior knowledge for the network so that that these unsupervised training samples can accurately describe the data. This goal is particularly important in spectral mixture analysis applications. However, a judicious selection of training data can be tremendously difficult in hyperspectral image data, due to the presence of mixed pixels and subpixel targets.

While conventional approaches tend to select training samples located in exemplar (i.e., pure) regions of each class, recent studies in the literature demonstrate that patterns that lie away from the class core, and near to the decision boundaries commonly used in conventional, pure pixel classification, are more informative [2]. This is particularly so in the context of mixed pixel classification problems, where such patterns in the boundary correspond to most highly mixed pixels in the data.

0-7803-9050-4/05/$20.00 ©2005 IEEE.

This paper offers a new look at the problem of training sample generation for neural network-based abundance estimation in hyperspectral imagery, and provides an unsupervised algorithm to generate semi-labeled samples [3], i.e., samples whose fractional abundances are estimated during the analysis process, so that they can be effectively used as training samples for neural network-based mixed pixel classification. The algorithm is intended for situations in which ground-truth fractional abundances are difficult to collect prior to data analysis, and is based on soft classifiers that rely on the principle that patterns that lie close to the location of decision boundaries in feature space can lead to significantly more accurate abundance estimation accuracies than patterns drawn from the class cores [4].

The paper is organized as follows. Section II describes the proposed neural network architecture for nonlinear spectral unmixing. Section III develops an unsupervised semi-labeled sample generation algorithm. Section 4 conducts experimental results using computer simulated and real hyperspectral data sets dominated by nonlinear mixing effects. Section 5 concludes with some remarks and hints at plausible future research.

II. Neural Network Architecture

The neural network classifier considered in experiments is based on MLP, a standard multi-layer architecture that can successfully approximate virtually any function when trained correctly. This implies that training is the most important step in MLP-based classification [1]. The architecture of the MLP used in this work comprises a number of identical processing units organized in layers, with those units on one layer connected to those on the next layer by means of weighted connections. The topology of the fully connected three-layer MLP used in this work is shown in Fig. 1, where the neuron count at the input layer, p, equals the number of endmember classes (estimated by a fully constrained linear mixture model), where a vector of linear endmember abundances for each spectral signature $\boldsymbol{r}$, denoted by $\boldsymbol{\gamma}^{(\boldsymbol{r})} = \left(\gamma_1^{(\boldsymbol{r})}, \gamma_2^{(\boldsymbol{r})}, \cdots, \gamma_p^{(\boldsymbol{r})}\right)$, is used as the input pattern. On other hand, the output layer has the same number of neurons as the input layer. Finally, a specific problem domain dictates how to select the number of hidden layers and the number of neurons in those layers. Since this paper is devoted to investigation of training methods, finding optimal network parameters for the MLP is beyond its scope. Based on previous results in the literature and our own experimentation, we set the number of hidden neurons empirically to $2 \times p$, i.e., twice the number of input neurons.

The usage of the MLP neural architecture in Fig. 1 involves two phases: training and classification. MLP models are typically trained using the error back-propagation algorithm [5], a supervised learning-from-data technique of training where a set of labeled input-output training vector pairs are presented to the network, which computes an error between the output vector, that can be seen as an estimate of the endmember abundance fractions of the training vector, and the vector of desired values for each output unit. This error is propagated successively back through the network and the matrix of weights is progressively updated until the network approximates the desired output closely enough. The main problem of this approach is that the accuracy of classification largely depends on the quality of training, which requires high-quality fractional estimates that are very difficult to collect on the ground. In the following section, we develop a semi-labeled sample generation algorithm, which estimates such fractional covers (based on soft classifiers) by searching for the most informative patterns in feature space.

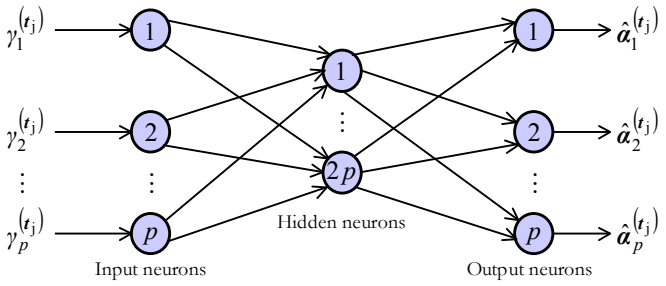

Figure 1. Architecture of multi-layer perceptron (MLP) neural network used for mixed pixel classification.

III. Semi-Labeled Sample Generation Algorithm

Our main goal in this section is to provide an algorithm able to estimate fractional abundances of the most informative spectral signatures. These signatures are likely to be close to the vicinity of the hyperplanes that can separate the classes in conventional classification procedures [4]. The separation of a training set into border and non-border patterns in the context of a pure pixel classification problem was first explored by Foody [1], who expressed 'borderness' as the difference between the two smallest distances measured for each training pattern. Here, membership is indicated by the Mahalanobis distance, which provides a measure of the typicality of a pattern to a certain class. In this section, we develop a novel automatic algorithm that is based on the concepts above, but adapted to a mixed pixel classification scenario. The algorithm is based on the following sequential steps:

1. Extract a set of spectral endmembers $\{\boldsymbol{e}_i\}_{i=1}^{p}$ from the data using an automated extraction algorithm [5], and label those endmembers as class core patterns.
2. Apply a spectral screening algorithm to identify the sample spectral signatures within a small spectral angle θ from any of the p core classes above, denoted from now on as $\{\boldsymbol{r}_j\}_{j=1}^{q}$, with $q \geq p$.
3. Assign each signature of the set $\{\boldsymbol{r}_j\}_{j=1}^{q}$ to one of the available pure classes $\{\boldsymbol{e}_i\}_{i=1}^{p}$, by computing $\boldsymbol{r}_j^{(i)} = \arg\min_i \{\mathrm{SAM}(\boldsymbol{r}_j, \boldsymbol{e}_i)\}$ for all $j = 1, \cdots, q$, where SAM is the spectral angle mapper [5], and the notation of $\boldsymbol{r}_j^{(i)}$ indicates that $\boldsymbol{r}_j$ would be assigned to the class $\boldsymbol{e}_i$ if the SAM between $\boldsymbol{r}_j$ and $\boldsymbol{e}_i$ is the minimum.

4. Let $\boldsymbol{r}_{j,k}^{(i)} \subseteq \{\boldsymbol{r}_j\}_{j=1}^{q}$ be the k-th sample associated with class $\boldsymbol{e}_i$, and let $\left|\boldsymbol{r}_{j,k}^{(i)}\right|$ be the cardinality of $\{\boldsymbol{r}_{j,k}^{(i)}\}$, which is the set of the samples in $\{\boldsymbol{r}_j\}_{j=1}^{q}$ associated with $\boldsymbol{e}_i$.

5. For each sample signature $\boldsymbol{s}_l$, estimate its fractional abundances by computing the Mahalanobis distance to each $\boldsymbol{e}_i$ as $\mathrm{MD}(\boldsymbol{s}_l, \boldsymbol{e}_i) = (\boldsymbol{s}_l - \boldsymbol{\mu}_i)^T \boldsymbol{K}_i^{-1} (\boldsymbol{s}_l - \boldsymbol{\mu}_i)$, where $\boldsymbol{K}_i$ is the sample covariance matrix of the class given by $\boldsymbol{e}_i$, and $\boldsymbol{\mu}_i$ is the mean for that class.

6. Compute a borderness score for each sample $\boldsymbol{s}_l$ as the difference between the two smallest values of $\mathrm{MD}(\boldsymbol{s}_l, \boldsymbol{e}_i)$, with $j = 1, \cdots, p$.

7. Select a set of t border samples, denoted by $\{\boldsymbol{s}_l^{(\mathrm{border})}\}_{l=1}^{t}$, with the lowest borderness score, and use their estimated fractional abundances in step 5 to train the proposed MLP-based neural network classifier.

The following section provides an experimental assessment of the proposed approach in spectral mixture analysis applications using both computer simulations and real data experiments.

IV. Experimental Results

A. Computer simulations

One of the major problems involved in analyzing the quality of fractional abundance estimation methods in remotely sensed imagery is the fact that ground-truth information about the real abundances of materials at sub-pixel levels is very difficult (if not downright impossible) to obtain in real scenarios. In order to avoid this shortcoming, simulation of hyperspectral imagery has been suggested as a simple and intuitive way to perform a preliminary evaluation of analysis techniques. The primary reason for the use of simulated imagery as a complement to real data analysis is that all details of the simulated images are known. These details can be efficiently investigated because they can be manipulated individually and precisely. In this section, we use simulated data based on real spectra collected by the AVIRIS imaging spectrometer. Two AVIRIS imaging spectrometer datasets of the Jasper Ridge Biological Preserve (JRBP) in California have been selected for experiments. The datasets, acquired on April 1998, consist of 512x614 pixels and 224 spectral bands, with a nominal ground resolution of 20 meters and spectral resolution of 10 nm (available online: http://aviris.jpl.nasa.gov). In a previous study of surface materials over JRBP, image endmembers were derived from the scenes above based on extensive ground knowledge. Fig. 2 plots spectral signatures associated with two of the main constituent materials at JRBP. These endmember signatures, denoted as $\boldsymbol{e}_1$ (soil) and $\boldsymbol{e}_2$ (evergreen forest) will be used to simulate nonlinear mixtures using a simple nonlinear function, the logarithmic function. Simulation experiments will be used as a baseline to interpret results with real data, described in the following subsection.

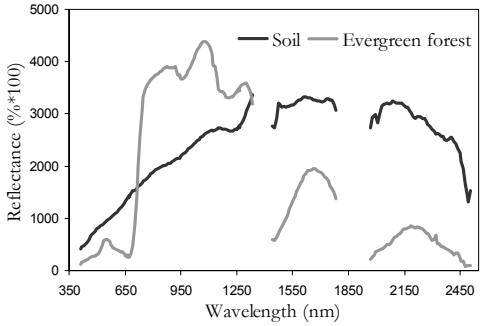

Figure 2. Pure spectral signatures of soil and evergreen forest obtained from an AVIRIS hyperspectral scene collected over Jasper Ridge, California.

Using the spectral signatures in Fig. 2, we have created a simulated image with nonlinear mixtures of $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ using a simple logarithmic function. The logarithmic function is useful to simulate nonlinear effects due to atmospheric absorption phenomena and other sources of non-linearity, although the exploration of further nonlinear functions in computer simulations is a highly desirable topic for future research work. The resulting 90x90-pixel scene consists of nine vertical regions $R_1, \ldots, R_9$ of ten pixels width, containing nonlinear mixtures where the abundances of $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ were first established linearly using the proportions in Table I, and then processed by a logarithmic function by $c_j(x, y) = \log \alpha_j(x, y)$, where $\alpha_j(x, y)$ is the initial (linear) abundance and $c_j(x, y)$ is the nonlinear abundance of $\boldsymbol{e}_j$.

TABLE I. Abundance Assignment for Regions in a Simulated Scene with Simple Logarithmic Nonlinear Mixtures

Region	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9
$\alpha_1(x, y)$	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
$\alpha_2(x, y)$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9

It should be noted that all the pixels in the simulated scene above are mixed in different proportions, with no pure instances of any material. Our main reason to avoid including completely pure spectral signatures in the simulated data is that this is a common situation in hyperspectral analysis, where no completely pure pixels may be present in the scene due to available spatial resolution and mixing phenomena. Random noise was added to the scene above to simulate contributions from ambient (clutter) and instrumental sources. White gaussian noise was created by using numbers with a standard normal distribution obtained from a pseudorandom number generator and added to each pixel. For the simulations, we consider the SNR for each band as the ratio of the 50% signal level to the standard deviation of the noise.

Table II shows an experimental study of the accuracy of linear spectral unmixing (LMM) and the proposed MLP-based approach in fractional abundance estimation using the scene with simulated nonlinear mixtures described above. Two different sets were used for the training of the network. The first one, T1, is composed of 27 manually selected spectral signatures that represent every possible mixture case (3 pixels per region) in the data set. The second one, T2, uses only 4

intelligently selected training samples by the proposed algorithm in section III. In both cases, true fractional abundances (known after the controlled simulation procedure) and Mahalanobis distance-based estimations were used to train the network. The former case is therefore based on a learning procedure based on training samples, while the latter does not consider ground-truth and relies entirely on semi-labeled samples.

TABLE II. RMSE ERROR (PERCENTAGE) IN ABUNDANCE ESTIMATION FOR SIMULATED NONLINEAR MIXTURES USING TRAINING AND SEMI-LABELED SAMPLES

Endmember material	Training samples			Semi-labeled samples	
	LMM	MLP-T1	MLP-T2	MLP-T1	MLP-T2
Soil	11.05%	1.58%	1.69%	1.61%	1.71%
Forest	12.58%	1.79%	1.85%	1.82%	1.87%

From results in Table II, we can conclude that the training samples selected by the proposed unsupervised algorithm (T2) in section III are the most informative and useful in terms of decreasing the overall RMSE. Specifically, only 4 training samples are required to obtain RMSE scores below 2%. It should be noted that, in order to obtain similar scores using manual selection, at least 15 manually selected samples (T1) had to be used. In addition, the scores associated with semi-labeled samples are always very similar to those found using training samples, a fact that seems to indicate that the true fractional abundance proportions were successfully estimated by the proposed Mahalanobis distance-based method. Although the above results are encouraging, further experiments with real data sets are required.

B. Real Data Experiments

In this section, real spectra collected from nonlinear mixtures will be analyzed. These data, provided by Prof. John Mustard at Brown University, consisted of a set of spectral signatures collected using the RELAB spectrometer (a high-resolution, bi-directional spectrometer at Brown University). The measurement precision of the RELAB spectrometer is better than 0.25%, which makes it an ideal candidate to evaluate fractional abundance estimation accuracy. The data included spectra from individual endmembers such as Olivine, Enstatite and Magnetite, where two types of mixtures (Olivine/Enstatite and Olivine/Magnetite were considered in experiments). Table III shows the known abundances for each of the endmembers above in the available mixtures.

TABLE III. NONLINEAR MIXED SIGNATURES IN MUSTARD DATA

Olivine/Enstatite		Olivine/Magnetite	
Olivine	Enstatite	Olivine	Magnetite
0.90	0.10	0.95	0.05
0.75	0.25	0.90	0.10
0.50	0.50	0.75	0.25
0.25	0.75	0.50	0.50
0.10	0.90	0.25	0.75

Using the ten mixed spectra in Table III, we created a synthetic scene made up of 500x500 pixels, where the mixed spectra above were randomly distributed throughout the scene using approximately the same number of pixels for each mixture. Table IV shows the abundance estimation results using the same configuration addressed in computer simulations, i.e., linear mixture analysis (LMM) and MLP-based abundance estimation using training sets T1 (made up of 10 manually selected pixels representing all mixture cases) and T2 (made up of only 3 intelligently selected pixels using the algorithm in section III). As can be seen in Table IV, while LMM could not capture mixture variability in the data, all MLP cases produced better estimation results. In particular, when the MLP was trained with only 3 semi-labeled training samples, it produced almost identical results than those found after training MLP with all mixture cases and using ground-truth fractions. The above results seem to indicate that intelligent incorporation of semi-labeled samples into the classification procedure enables an appropriate representation of mixed classes, as well as a meaningful evaluation of mixed pixel classification accuracy in terms of endmember fractional abundances.

TABLE IV. RMSE ERROR (PERCENTAGE) IN ABUNDANCE ESTIMATION FOR REAL NONLINEAR MIXTURES USING TRAINING AND SEMI-LABELED SAMPLES

Endmember material	Training samples			Semi-labeled samples	
	LMM	MLP-T1	MLP-T2	MLP-T1	MLP-T2
Olivine	11.21%	7.45%	7.89%	8.41%	8.16%
Enstatite	12.70%	8.12%	8.23%	7.99%	8.29%
Magnetite	11.92%	7.98%	8.14%	7.95%	8.92%

V. CONCLUSIONS

Finding training samples without prior knowledge is a very challenging issue in hyperspectral data exploitation. The high spectral resolution provided by modern hyperspectral sensors allows uncovering many *weak*, subpixel targets that cannot be identified a priori. Also, the complexity of intra-pixel mixing phenomena indicates that the most highly mixed pixels may be the most informative ones and, therefore, the best candidates to be selected as training samples for supervised classification. The incorporation of mixed pixels as training samples does, however, require detailed ground data on abundance fractions for the training sites. In this paper, we explored an alternative solution, based on the use of intelligent algorithms aimed at generating semi-labeled samples, i.e., samples whose fractional abundances are estimated during the analysis process.

REFERENCES

[1] A. Baraldi et al., "Comparison of the multi-layer perceptron with neuro-fuzzy techniques in the estimation of cover class mixture in remotely sensed data," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, pp. 994-1005, 2001.

[2] G. M. Foody and A. Mathur, "Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification," *Remote Sens. Environ.*, vol. 93, pp. 107-117, 2004.

[3] M. Chi and L. Bruzzone, "A semilabeled-sample-driven bagging technique for ill-posed classification problems," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, pp. 69-73, 2004.

[4] Q. Jackson and D. A. Landgrebe, "An adaptative method for combined covariance estimation and classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, pp. 1082-1087, 2002.

[5] J. Plaza, P. Martinez, R. Perez and A. Plaza, "Nonlinear neural network mixture models for fractional abundance estimation in AVIRIS hyperspectral images," in *Proc. XIII JPL Airborne Earth Science Workshop*, Pasadena, 2004.