



On the use of small training sets for neural network-based characterization of mixed pixels in remotely sensed hyperspectral images

Javier Plaza*, Antonio Plaza, Rosa Perez, Pablo Martinez

Neural Networks and Signal Processing Group (GRNPS), Computer Science Department, University of Extremadura, Avda. de la Universidad s/n, 10.071 Cáceres, Spain

ARTICLE INFO

Article history:

Received 21 November 2006

Received in revised form 12 November 2008

Accepted 15 April 2009

Keywords:

Hyperspectral

Image processing

Mixed pixels

Spectral mixture analysis

Multi-layer perceptron

Automatic training sample generation algorithms

Mixed training samples

Nonlinear spectral unmixing

ABSTRACT

In this work, neural network-based models involved in hyperspectral image spectra separation are considered. Focus is on how to select the most highly informative samples for effectively training the neural architecture. This issue is addressed here by several new algorithms for intelligent selection of training samples: (1) a border-training algorithm (BTA) which selects training samples located in the vicinity of the hyperplanes that can optimally separate the classes; (2) a mixed-signature algorithm (MSA) which selects the most spectrally mixed pixels in the hyperspectral data as training samples; and (3) a morphological-erosion algorithm (MEA) which incorporates spatial information (via mathematical morphology concepts) to select spectrally mixed training samples located in spatially homogeneous regions. These algorithms, along with other standard techniques based on orthogonal projections and a simple Maximin-distance algorithm, are used to train a multi-layer perceptron (MLP), selected in this work as a representative neural architecture for spectral mixture analysis. Experimental results are provided using both a database of nonlinear mixed spectra with absolute ground truth and a set of real hyperspectral images, collected at different altitudes by the digital airborne imaging spectrometer (DAIS 7915) and reflective optics system imaging spectrometer (ROSIS) operating simultaneously at multiple spatial resolutions.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Imaging spectroscopy (i.e., hyperspectral imaging) is a remote sensing technique capable of identifying materials and objects in the air, land and water on the basis of the unique reflectance patterns that result from the interaction of solar energy with the molecular structure of the material [1]. Advances in sensor technology have led to the development of hyperspectral instruments [2] capable of collecting tens or even hundreds of images, corresponding to different wavelength channels, for the same area on the surface of the Earth. As a result, each pixel (vector) in a hyperspectral image has an associated *spectral signature* or “fingerprint” that uniquely characterizes the underlying objects, as shown by Fig. 1.

The wealth of spectral information provided by hyperspectral sensors has opened ground-breaking perspectives in many applications with high computational requirements [3–5], including environmental modeling and assessment, target detection for military and defense/security deployment, urban planning and management studies, risk/hazard prevention and response including wild-land fire

tracking, biological threat detection, monitoring of oil spills. However, the design of processing algorithms for hyperspectral data introduces additional challenges. In particular, conventional supervised classification techniques for hyperspectral imagery were originally developed under the assumption that the classes to be separated are discrete and mutually exclusive, i.e., it is assumed that each pixel vector is “pure” and belongs to a single spectral class. Often, however, this is not a realistic assumption. In particular, most of the pixels collected by hyperspectral imaging instruments contain the resultant mixed spectra from the reflected surface radiation of various constituent materials at a sub-pixel level. The presence of mixed pixels is due to several reasons [6]. First, the spatial resolution of the sensor is generally not high enough to separate different pure signature classes at a macroscopic level, and the resulting spectral measurement can be a composite of individual pure spectra (often called *endmembers* in hyperspectral analysis terminology) which correspond to materials that jointly occupy a single pixel. Second, mixed pixels also result when distinct materials are combined into a microscopic (*intimate*) mixture, independently of the spatial resolution of the sensor.

Spectral mixture modeling (also called *spectral unmixing*) involves the separation of a pixel spectrum into its pure component end-member spectra, and the estimation of the abundance value for each endmember [7]. Several techniques for spectral unmixing have been

* Corresponding author. Tel.: +34 927 257000x82576; fax: +34 927 257202.
E-mail address: jplaza@unex.es (J. Plaza).

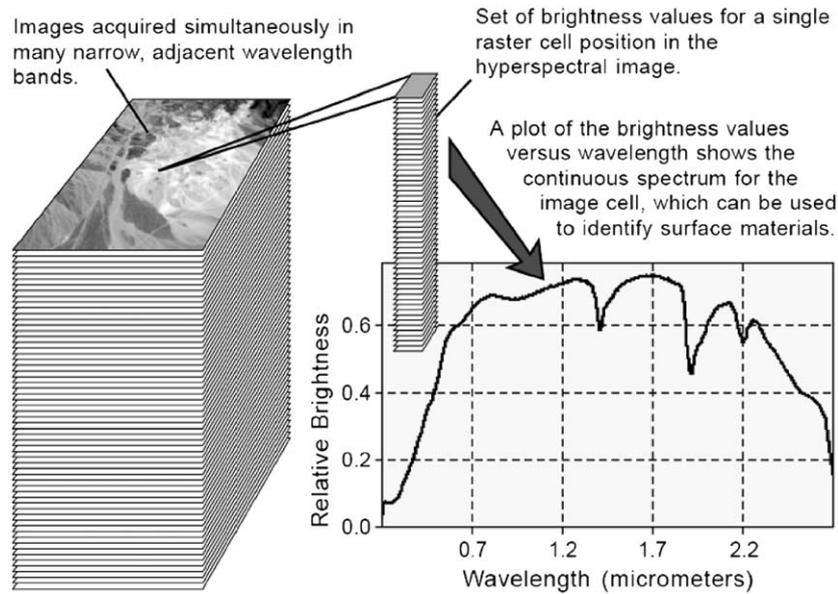


Fig. 1. The concept of hyperspectral imaging.

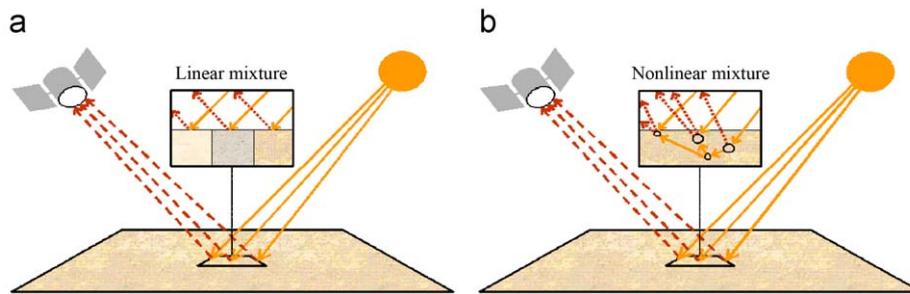


Fig. 2. Schematical description of scattering in linear (single scattering) (a) and nonlinear (multiple scattering) (b) mixtures.

developed in the literature. For instance, the linear mixture model assumes that the collected spectra are linearly mixed [8]. As a result, a linear (macroscopic) mixture is obtained when the endmember substances are sitting side-by-side within the field of view of the imaging instrument (see Fig. 2(a)). The linear model assumes minimal secondary reflections and/or multiple scattering effects in the data collection procedure [6]. Subsequently, the resultant mixed spectrum can be simply expressed as a linear combination of endmember components, weighted by a scalar endmember abundance fraction as follows:

$$\mathbf{r} = \mathbf{E}\boldsymbol{\alpha} + \mathbf{n} = \sum_{i=1}^p \mathbf{e}_i \alpha_i + \mathbf{n}, \quad (1)$$

where \mathbf{r} is a pixel vector given by a collection of values at different wavelengths, \mathbf{E} is a matrix containing p endmember signatures $\{\mathbf{e}_i\}_{i=1}^p$, $\boldsymbol{\alpha}$ is a vector containing the fractional abundance values for each of the p endmembers in \mathbf{r} , and \mathbf{n} is a noise vector.

Although the linear mixture model has practical advantages such as the ease of implementation and flexibility in different applications, there are many naturally occurring situations where nonlinear models may best characterize the resultant mixed spectra for certain endmember distributions [9]. In particular, nonlinear mixtures generally occur in situations where endmember components are randomly distributed throughout the field of view of the

instrument [6], as shown by Fig. 2(b). In those cases, the mixed spectra collected at the imaging instrument are better described by assuming that part of the source radiation is subject to multiple scattering effects before being collected by the sensor. A general expression for the nonlinear mixture model is given by

$$\mathbf{r} = f(\mathbf{E}, \boldsymbol{\alpha}) + \mathbf{n}, \quad (2)$$

where f is an unknown nonlinear function that defines the interaction between \mathbf{E} and $\boldsymbol{\alpha}$. Various learning-from-data techniques have been proposed in the literature to estimate the f in hyperspectral imaging applications. For instance, independent component analysis (ICA) has been proposed in the recent literature as a relevant technique for handling the inversion in Eq. (2) [10,11]. ICA is an unsupervised source separation process [12] that has shown significant success in blind source separation, feature extraction, and unsupervised recognition. Another approach that has demonstrated great potential to decompose mixed pixels is the use of artificial neural networks, which have demonstrated an inherent capacity to approximate complex nonlinear functions [13,14]. Although many neural network architectures exist, for decomposition of mixed pixels in terms of nonlinear relationships mostly feed-forward networks of various layers, such as the multi-layer perceptron (MLP), have been used [15]. The MLP is typically trained using the error back-propagation algorithm, a supervised technique of training with three

phases [16]. In the first one, an initial vector is presented to the network, which leads to the activation of the network as a whole. The second phase computes an error between the output vector and a vector of desired values for each output unit, and propagates it successively back through the network. The last phase computes the changes for the connection weights, which are randomly generated in the beginning.

It has been shown in the literature that MLP-based neural models, when trained accordingly, generally outperform other nonlinear models such as regression trees or fuzzy classifiers [17]. A variety of issues have been investigated to evaluate the impact of the training stage in neural network-based interpretation of mixed pixels, including the size and location of training sites, and the composition of training sets. Most of the attention has been devoted to investigate the number of training samples required for the learning stage [18]. Sometimes the smallness of a training set represents a major problem. This is especially apparent for hyperspectral imaging applications, where the requirement of large volumes of training sites has often been linked to the high-dimensional nature of the data [19]. Even if the endmembers participating in mixtures in a certain area are known, proportions of these endmembers on a per-pixel basis are difficult to be estimated a priori. Therefore, a challenging aspect in the design of neural network-based techniques for spectral unmixing is to reduce the need for large training sets. Studies have investigated a range of issues, including the use of feature selection and feature extraction methods to reduce the dimensionality of the input data [20], the use of unlabeled and semi-labeled training samples [19], the accommodation of spatial dependence in the data to define an efficient sampling design [21], or the use of statistics derived on other locations [22].

In this paper, we hypothesize that the problem of mixed pixel characterization demands intelligent training sample selection algorithms able to seek for the most informative training samples, thus requiring an optimization of the compromise between estimation accuracy (to be maximized) and ground-truth knowledge (to be minimized). In order to address this issue, we develop several unsupervised algorithms for intelligent selection of training samples from the input data, which are then used to train a simple MLP neural network. Our main reason to select the MLP architecture for demonstration purposes is that this type of network has been claimed to be particularly sensitive to free parameters, such as the arrangement and number of neurons in the different layers [9]. However, experimental results in this work show that a simple MLP network configuration can produce very stable results (even in the presence of complex mixtures) when trained accordingly, using small training sets containing highly mixed pixels. This fact leads us to believe that the composition of the training set is a crucial parameter in neural network-based spectral mixture analysis applications, with a more significant impact on the final output than the size of the training set, or the choice of a specific network architecture/configuration.

The remainder of this paper is structured as follows. Section 2 briefly describes the neural network architecture used in this work for spectral unmixing purposes. Section 3 develops several new unsupervised training sample selection algorithms, able to seek for the most informative training samples. Section 4 validates the proposed approach using a database of nonlinearly mixed spectra with absolute ground truth. Section 5 conducts real hyperspectral image experiments using DAIS 7915 and ROSIS spectrometry images collected at multiple resolutions. Finally, Section 6 concludes with some remarks and hints at plausible future research.

2. Multi-layer perceptron for spectral unmixing

The MLP is a type of multi-layer neural network [16] that can be depicted as a directed graph. The architecture used in this work

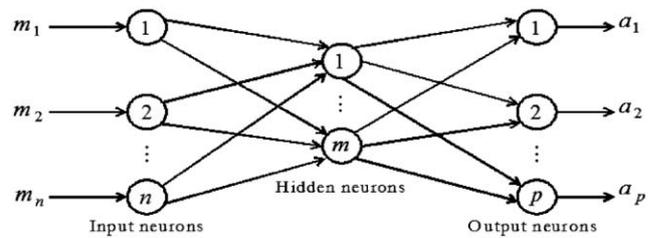


Fig. 3. Architecture of a multi-layer perceptron (MLP) neural network for spectral unmixing.

is shown in Fig. 3, where each column of nodes is a layer and the leftmost layer is the input layer. The second layer is the hidden layer, and the third layer is the output layer. The neuron count at the input layer, n , equals the number of spectral bands. The input patterns are pixel vectors directly obtained from the input data. The number of neurons at the output layer, p , equals the number of spectral endmembers. In this work, we estimate the value of p automatically using the concept of virtual dimensionality (VD), coined by Chang and Du [23]. Specifically, the method used to determine the VD in this paper is the one developed by Harsanyi-Farrand-Chang [24], referred to as HFC method. It should be noted that m , the number of hidden neurons, is generally fine-tuned depending on the problem under consideration. However, in this work we are mainly interested in exploring training mechanisms and their implications, without particular emphasis on careful adjustment of network configuration parameters. Subsequently, finding optimal parameters for the hidden layer(s) is beyond our scope. Based on previous results in the literature [15] and our own experimentation, the considered architecture is based on one hidden layer only, with the number of neurons empirically set to the square root of the product of the number of input features and information classes. The sigmoid function is used as the nonlinear activation function [16]. The configuration for the proposed architecture and its training were introduced in Ref. [25]. One of our main goals in this work is to demonstrate that the architecture shown in Fig. 3 can produce adequate results in different analysis scenarios when trained accordingly, in spite of the fact that further optimizations in the configuration of the hidden layer may lead to improved results.

3. Algorithms for automatic selection of training samples

Conventional approaches for identification of training samples tend to select samples located in exemplar regions of each class only, while atypical cases are often removed or down-weighted in training set refinement operations. Such exemplar training patterns are located near the central “core” of the class in feature space. However, a key concern in the context of mixed pixel interpretation is how to identify and characterize the response of sites that lie away from the class core, and near to the decision boundaries commonly used in conventional (i.e., pure) pixel classification. Therefore, “border” (or, equivalently, mixed) training patterns may be required to incorporate the inherent complexity of mixing systematics into the spectral mixture analysis [26].

In this section, we describe several techniques for automated selection of training samples designed to assist in the selection of descriptive samples for supervised learning. First, an orthogonal subspace projection (OSP) algorithm [2,27] is presented as a method to automatically extract spectrally distinct training samples from the input data. Then, we develop several new algorithms to extract the most representative training samples from the data set according to different criteria, such as the “borderness” (convexity) of those samples or the degree of spectral similarity to other spatially

adjacent training samples. Here, we must emphasize that the main idea behind our proposed algorithms is to adapt the concept of “borderness” to a mixed pixel interpretation scenario, in which a “border” pattern corresponds to a “mixed” pixel and a “non-border” pattern corresponds to a “pure” pixel which is often identified using convex geometry techniques, hence the relationship between “borderness” and “convexity.” Finally, a simple Maximin-distance algorithm, which has been commonly used in pattern recognition applications [28], is also adapted to the problem of extracting useful training samples for supervised learning. Different sets of training samples will be used in the following section to train the MLP neural network architecture described in Section 2 and to further investigate the impact of the training sample generation process on the characterization of mixed pixels in hyperspectral data.

Before describing the proposed training sample selection algorithms in detail, we emphasize that some of these algorithms make use of a point-wise spectral distance that will be denoted by *Dist* in this work. The choice of *Dist* is flexible and several distance measures have been proposed for this purpose in the recent hyperspectral imaging literature [2], including the Euclidean distance, the spectral information divergence (SID), or a hidden Markov model-based information divergence (HMMID). In this work we adopt a widely used measure in hyperspectral imaging applications, the spectral angle distance (SAD), which is invariant in the multiplication of the input vectors by constants and, consequently, is invariant to unknown multiplicative scalings that may arise due to differences in illumination and sensor observation angle [6]. Let us consider two *N*-dimensional spectral signatures $\mathbf{s}_i = (s_{i1}, s_{i2}, \dots, s_{iN})^T$ and $\mathbf{s}_j = (s_{j1}, s_{j2}, \dots, s_{jN})^T$, where the superscript “*T*” denotes the vector transpose operation. The SAD between s_i and s_j is given by the following expression [2]:

$$\begin{aligned} \text{SAD}(\mathbf{s}_i, \mathbf{s}_j) &= \cos^{-1}(\mathbf{s}_i \cdot \mathbf{s}_j / \|\mathbf{s}_i\| \cdot \|\mathbf{s}_j\|) \\ &= \cos^{-1} \left(\frac{\sum_{l=1}^N s_{il}s_{jl}}{\left[\sum_{l=1}^N s_{il}^2 \right]^{1/2} \left[\sum_{l=1}^N s_{jl}^2 \right]^{1/2}} \right). \end{aligned} \quad (3)$$

3.1. Orthogonal subspace projection algorithm (OSP)

The automatic target generation process was developed in Ref. [29] to find potential target pixels that can be used to generate a set of target signatures using an OSP-based approach. It makes use of an orthogonal subspace projector [30] given by the following expression:

$$P_{\mathbf{U}}^{\perp} = \mathbf{I} - \mathbf{U}(\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T, \quad (4)$$

where \mathbf{I} is the identity matrix and \mathbf{U} is a signature matrix that contains a set of spectral signatures. The algorithm can be summarized as follows. If we assume that \mathbf{t}_0 is an initial target pixel vector, the algorithm begins by applying an orthogonal subspace projector $P_{\mathbf{t}_0}^{\perp}$ specified by Eq. (4) with $\mathbf{U} = \mathbf{t}_0$ to all image pixel vectors. It then finds a target pixel vector, denoted by \mathbf{t}_1 with the maximum orthogonal projection in the orthogonal complement space, denoted by $(\mathbf{t}_0)^{\perp}$, which is orthogonal to the space (\mathbf{t}_0) linearly spanned by \mathbf{t}_0 . The reason for this selection is that \mathbf{t}_1 is generally the most distinct feature \mathbf{t}_0 in the sense of orthogonal projection because \mathbf{t}_1 has the largest magnitude of projection in $(\mathbf{t}_0)^{\perp}$ produced by $P_{\mathbf{t}_0}^{\perp}$. A second target pixel vector \mathbf{t}_2 can be found by applying an orthogonal subspace projector $P_{[\mathbf{t}_0, \mathbf{t}_1]}^{\perp}$ with $\mathbf{U} = [\mathbf{t}_0, \mathbf{t}_1]$ to the original image and a target pixel vector that has the maximum orthogonal projection in $(\mathbf{t}_0, \mathbf{t}_1)^{\perp}$ is selected as \mathbf{t}_2 . The above procedure is repeated to find a third target pixel vector \mathbf{t}_3 , a fourth target pixel vector \mathbf{t}_4 , and so on, until a certain stopping rule is satisfied. In this paper, the stopping rule is determined by a maximum number of training samples to

be generated, *t*. The algorithm can be summarized by the following steps:

- (1) *Initial condition*: Select an initial target pixel vector of interest denoted by \mathbf{t}_0 . In order to initialize the algorithm without knowing \mathbf{t}_0 , we select a target pixel vector with the maximum length as the initial target \mathbf{t}_0 , namely $\mathbf{t}_0 = \arg\{\max_{\mathbf{r}} \mathbf{r}^T \mathbf{r}\}$, which has the highest intensity, i.e., the brightest pixel vector in the image scene. It is worth noting that this selection may not be necessarily the best selection. However, according to our experiments it was found that the brightest pixel vector was always extracted later on, if it was not used as an initial target pixel vector in the initialization. Set $k = 1$ and $\mathbf{U}_0 = [\mathbf{t}_0]$.
- (2) At the *k*-th iteration, apply $P_{\mathbf{U}_0}^{\perp}$ via Eq. (4) to all pixel vectors \mathbf{r} in the image and then find the *k*-th target \mathbf{t}_k generated at the *k*-th stage which has the maximum orthogonal projection as follows:

$$\mathbf{t}_k = \arg\{\max_{\mathbf{r}} [(P_{[\mathbf{t}_0, \mathbf{U}_{k-1}]}^{\perp} \mathbf{r})^T (P_{[\mathbf{t}_0, \mathbf{U}_{k-1}]}^{\perp} \mathbf{r})]\}, \quad (5)$$

where $\mathbf{U}_{k-1} = [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_{k-1}]$ is the target matrix generated at the (*k* – 1)st stage.

- (3) *Stopping rule*: If $k < t - 1$, let $\mathbf{U}_k = [\mathbf{U}_{k-1}, \mathbf{t}_k] = [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_k]$ be the *k*-th target matrix, go to Step 2. Otherwise, go to Step 4.
- (4) At this stage, the OSP algorithm is terminated and the final set of target pixel vectors produced after executing the algorithm comprises *t* target pixel vectors, $\{\mathbf{t}_0, \mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_{t-1}\} = \{\mathbf{t}_0\} \cup \{\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_{t-1}\}$, found after repeatedly using Eq. (5).

3.2. Border-training algorithm (BTA)

The algorithm described in the previous subsection searches for spectrally distinct training samples without accounting for their degree of spectral purity or extremity in the data cloud. The separation of a training set into “pure” (non-border) and “mixed” (border) patterns was first explored by Foody [31], who expressed “borderness” as the difference between the two smallest distances measured for each training pattern. Here, membership was indicated by the Mahalanobis distance, which provides a measure of the typicality of a pattern to a certain class. A border-training pattern is expected to be almost as close to its actual class of membership as it is to any other class. Therefore, the difference in the Mahalanobis distances between the two most likely classes of membership would be small for a border pattern. This focus on the vicinity of the hyperplanes that can optimally separate the classes is similar to the aspects of Lee and Landgrebe’s decision boundary feature extraction [20]. Here, we develop an automatic algorithm inspired by the concept proposed by Foody [31], but further adapted to a mixed pixel interpretation scenario. The algorithm consists of a two-stage process, in which a set of “pure” training samples are first automatically extracted from the input data, and then a degree of borderness related to those samples is used to identify highly mixed training samples.

3.2.1. Selection of pure training samples

In order to select pure training samples, the N-FINDR algorithm developed by Winter [32] has been adapted in this work to the generation of training samples from pure pixels in the data. The algorithm assumes that, in *N* spectral dimensions, the *N*-dimensional volume formed by a simplex with vertices specified by purest pixels is always larger than that formed by any other combination of pixels. In this work, we use a modified version of N-FINDR, which assumes that the number of existing endmembers in the data, *p*, is known in advance. In this work, this value is estimated using the VD concept [23].

Estimated, our implementation of N-FINDR is initialized by a

simplex whose vertices are formed by a random set of pixels $\{\mathbf{e}_1^{(0)}, \mathbf{e}_2^{(0)}, \dots, \mathbf{e}_p^{(0)}\}$ used as initial endmembers selected from the data. The algorithm then proceeds by finding $V(\mathbf{e}_1^{(0)}, \mathbf{e}_2^{(0)}, \dots, \mathbf{e}_p^{(0)})$ the volume of the simplex defined by vertices $\mathbf{e}_1^{(0)}, \mathbf{e}_2^{(0)}, \dots, \mathbf{e}_p^{(0)}$, denoted by $S(\mathbf{e}_1^{(0)}, \mathbf{e}_2^{(0)}, \dots, \mathbf{e}_p^{(0)})$. Then, for each image sample pixel vector \mathbf{r} , it recalculates the volumes of p simplices, $S(\mathbf{r}, \mathbf{e}_2^{(0)}, \dots, \mathbf{e}_p^{(0)})$, $S(\mathbf{e}_1^{(0)}, \mathbf{r}, \dots, \mathbf{e}_p^{(0)})$, ..., $S(\mathbf{e}_1^{(0)}, \mathbf{e}_2^{(0)}, \dots, \mathbf{r})$, each of which is formed by replacing one pixel $\mathbf{e}_j^{(0)}$ with the sample vector \mathbf{r} . If none of these p recalculated volumes is greater than $S(\mathbf{e}_1^{(0)}, \mathbf{e}_2^{(0)}, \dots, \mathbf{e}_p^{(0)})$, then no endmember sample in $\mathbf{e}_1^{(0)}, \mathbf{e}_2^{(0)}, \dots, \mathbf{e}_p^{(0)}$ is replaced. Otherwise, the endmember which is absent in the largest volume among the p simplices above is replaced by \mathbf{r} . Let such an endmember be denoted by $\mathbf{e}_j^{(1)}$. A new set of endmembers can be produced by letting $\mathbf{e}_j^{(1)} = \mathbf{r}$ and $\mathbf{e}_i^{(1)} = \mathbf{e}_i^{(0)}$ for $i \neq j$. The same process is repeated again for other sample vectors until all pixel vectors in the original data set are exhausted. In the end, a set of endmember pixels denoted as $\{\mathbf{e}_i\}_{i=1}^p$ is obtained.

3.2.2. Selection of mixed training samples

Based on the set of endmembers $\{\mathbf{e}_i\}_{i=1}^p$ obtained by the N-FINDR algorithm above, a new algorithm has been developed to extract the set of training samples with the highest degree of borderness or convexity in the data cloud:

- (1) Label the training samples $\{\mathbf{e}_i\}_{i=1}^p$ produced by the N-FINDR algorithm 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 by $\{\mathbf{r}_j\}_{j=1}^q$ with $q \geq p$. The selection of the threshold θ depends on the similarity between the pixels in the data, although in our experiments we have found that a value of $\theta=0.1$ provides highly satisfactory results in most cases.
- (3) Associate each signature of the set $\{\mathbf{r}_j\}_{j=1}^q$ to one of the available pure classes, $\{\mathbf{e}_i\}_{i=1}^p$, by computing $\mathbf{r}_j^{(i)} = \text{argmin}_i\{\text{Dist}(\mathbf{r}_j, \mathbf{e}_i)\}$ for all $j = 1, \dots, q$, where Dist is a point-wise spectral distance and the notation of $\mathbf{r}_j^{(i)}$ indicates that the Dist between \mathbf{r}_j and \mathbf{e}_i is the minimum, i.e., \mathbf{e}_i is the most spectrally similar endmember to \mathbf{r}_j . Let $\mathbf{r}_{j,k}^{(i)} \subseteq \{\mathbf{r}_j\}_{j=1}^q$ be the k -th sample associated with class \mathbf{e}_i , and let $|\mathbf{r}_{j,k}^{(i)}|$ be the cardinality of the set $\{\mathbf{r}_{j,k}^{(i)}\}$, which contains all the samples in $\{\mathbf{r}_j\}_{j=1}^q$ associated with \mathbf{e}_i .
- (4) For each potential training sample \mathbf{t}_i (pixel vector in the input data), compute the Mahalanobis distance from pure class \mathbf{e}_i as $MD(\mathbf{t}_i, \mathbf{e}_i) = (\mathbf{t}_i - \boldsymbol{\mu}_i)^T \mathbf{K}_i^{-1} (\mathbf{t}_i - \boldsymbol{\mu}_i)$, where \mathbf{K}_i is the sample covariance matrix of the class given by \mathbf{e}_i and $\boldsymbol{\mu}_i$ is the mean for that class, given by $\boldsymbol{\mu}_i = (1/|\mathbf{r}_{j,k}^{(i)}|) \sum_{k=1}^{|\mathbf{r}_{j,k}^{(i)}|} \mathbf{r}_{j,k}^{(i)}$. Similarly, compute a “borderness” score for each \mathbf{t}_i as the difference between the two smallest values of $MD(\mathbf{t}_i, \mathbf{e}_i)$, $i = 1, \dots, p$.
- (5) Select a final set of $\{\mathbf{t}_i\}_{i=1}^f$ training samples according to their borderness score (in decreasing order).

It should be noted that the Mahalanobis distance computation in the algorithm above may be affected by matrix singularity problems in the presence of limited training samples [33,34]. This issue can be addressed by fine-tuning parameter θ to extract a sufficient number of samples in step (2) of the algorithm.

3.3. Mixed-signature algorithm (MSA)

As an alternative to the algorithm developed in the previous subsection, we describe below an automatic algorithm that iteratively seeks for the most highly mixed training samples. The algorithm

considers all possible combinations of endmembers in order to search for additional highly mixed training samples, where combinations made up of many participating endmembers are prioritized in the process to account for complex mixtures. The algorithm can be summarized by the following steps (for illustrative purposes, Fig. 4 displays a toy example illustrating the behavior of the algorithm in a simple, two-dimensional space which can be described using three spectral endmembers labeled as \mathbf{e}_1 , \mathbf{e}_2 and \mathbf{e}_3):

1. Compute $\mathbf{C}_p = (1/p) \sum_{i=1}^p \mathbf{e}_i$, i.e., the centroid of the simplex defined by spectral endmembers $\mathbf{S} = \{\mathbf{e}_i\}_{i=1}^p$ produced by the N-FINDR endmember extraction algorithm described above.
2. Select a first training sample that corresponds to the sample pixel vector which is spectrally closest to \mathbf{C}_p using $\mathbf{t}_1 = \text{argmin}_j\{\text{Dist}(\mathbf{r}_j, \mathbf{C}_p)\}$ (see Fig. 4(a)). As shown by Fig. 4(a), the centroid may not be a real pixel vector in the scene.
3. Compute the spectral distance from \mathbf{C}_p for each of the endmembers \mathbf{e}_i using $\text{Dist}(\mathbf{e}_i, \mathbf{C}_p)$.
4. For each of the endmembers in the endmember set \mathbf{S} do:
 - 4.1 Remove (temporally) from \mathbf{S} the endmember which is less spectrally similar to \mathbf{C}_p , thus obtaining a new endmember set $\{\mathbf{e}_i\}_{i=1}^{p-1}$.
 - 4.2 Calculate the centroid of the set $\{\mathbf{e}_i\}_{i=1}^{p-1}$, i.e., $\mathbf{C}_{p-1} = (1/(p-1)) \sum_{i=1}^{p-1} \mathbf{e}_i$.
 - 4.3 Select the pixel vector whose signature is the most spectrally similar to \mathbf{C}_{p-1} as the new training sample (see Fig. 4(b)).
 - 4.4 Add the endmember that was removed from the original endmember set \mathbf{S} and repeat from step 4.2 but removing (temporally) from \mathbf{S} the endmember which is less spectrally similar to \mathbf{C}_p and has never been removed from the set, until all endmembers in \mathbf{S} have been removed from the original set once (see Fig. 4(c) and (d)).

The MSA algorithm described above (as it is also the case with the OSP and BTA algorithms) relies on the spectral properties of the data alone. Several studies, however, have investigated the accommodation of spatial dependence in hyperspectral imaging applications [35,36]. In the following subsection, we describe a novel training sample selection approach that incorporates the spatial information in the process.

3.4. Morphological-erosion algorithm (MEA)

Morphological operations [37], when extended to hyperspectral imagery [38], can be very useful for the interpretation of mixed pixels, mainly because they take into account both the spatial and the spectral properties of the image data in simultaneous fashion. Let us denote by $\mathbf{r}(x, y)$ the sample pixel vector at spatial coordinates (x, y) . Similarly, let B be a kernel (i.e., structuring element (SE)) defined in the spatial domain of the image (the x - y plane). The proposed morphological algorithm is based on the following sequential steps:

1. Translate the SE over the spatial domain of the multidimensional image, and extract the most highly mixed pixel vector in the spatial neighborhood of each pixel vector $\mathbf{r}(x, y)$ using a so-called morphological erosion given by $\varepsilon_B[\mathbf{r}(x, y)] = \text{argmin}_{(s,t) \in B} \{\sum_s \sum_t \text{Dist}(\mathbf{r}(x, y), \mathbf{r}(x+s, y+t))\}$ and explained below:
 - 1.1. First, a point-wise spectral distance Dist between each pixel vector $\mathbf{r}(x, y)$ and all other pixels in the neighborhood defined by B is computed by a translation operation in the spatial domain [38].
 - 1.2. Then, a cumulative distance given by the sum of Dist scores [36] is obtained for each pixel vector with regard to all other pixel vectors in the spatial neighborhood defined by B .

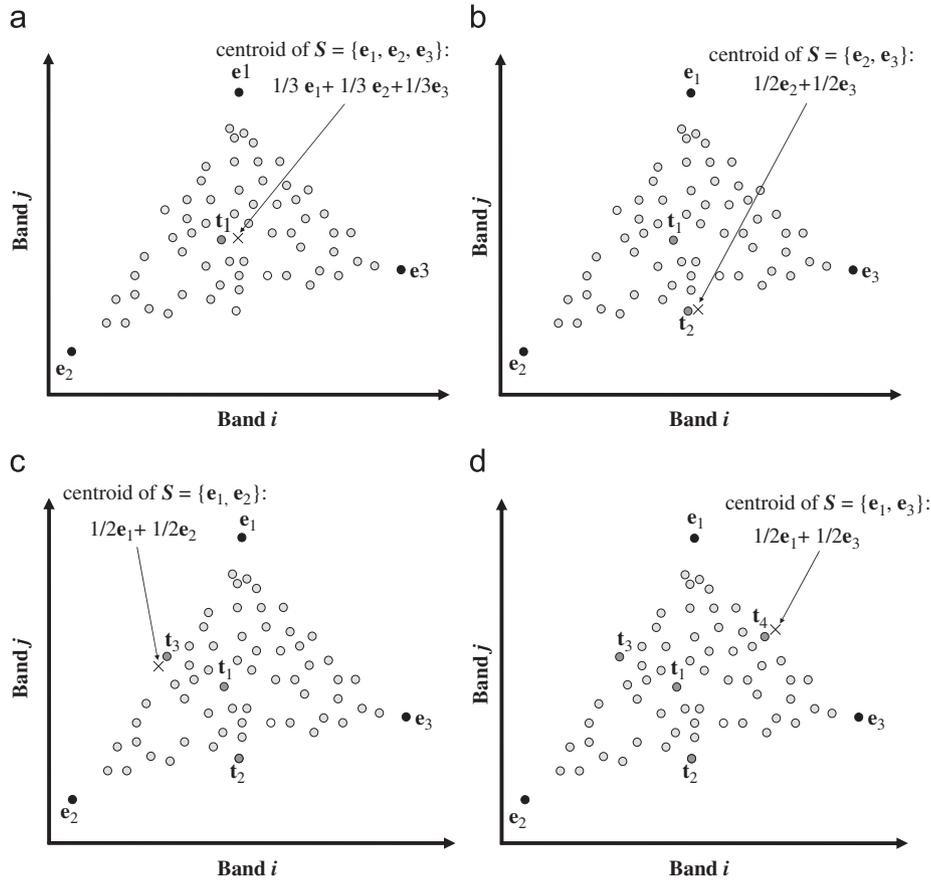


Fig. 4. Toy example illustrating the performance of MSA algorithm in a two-dimensional space. (a) Selection of training sample t_1 by MSA, (b) selection of training sample t_2 by MSA, (c) selection of training sample t_3 by MSA, and (d) selection of training sample t_4 by MSA.

- 1.3. Finally, the pixel with the lowest cumulative distance score is considered as the “most highly mixed” in the spatial neighborhood. This is the pixel which is most spectrally similar to its spatial neighbors, as opposed to the one with highest cumulative score, which is the “most spectrally distinct” pixel.
2. Define a morphological erosion index (MEI) for each pixel $\mathbf{r}(x, y)$ by computing the $Dist$ between the pixel provided by the morphological erosion operation at the local neighborhood and the data centroid $\mathbf{D}_M = (1/M) \sum_x \sum_y \mathbf{r}(x, y)$, where $M = x \cdot y$ is the total number of pixel vectors in the input data set.
3. Select the set of t pixel vectors in decreasing order of MEI score to generate a set of $\{t_i\}_{i=1}^t$ training samples.

It should be noted that the MEA algorithm described above does not require previous knowledge about the spectral endmembers in the data as in the case of BTA and MSA algorithms. Instead, MEA depends on the spatial properties of the input structuring element B (whose size and shape must be set accordingly). This dependence, which is also present in classic mathematical morphology, allows for the tuning of the method for different applications.

3.5. Maximin-distance algorithm (Maximin)

Below we describe a simple Maximin-distance algorithm [9] that can also be used to generate training samples:

1. Let the first training sample t_1 be the pixel vector with the maximum vector length, that is,

$$t_1 = \arg\{\max_r \mathbf{r}^T \mathbf{r}\}. \quad (6)$$

2. For each $2 \leq j \leq t$, find the j -th training sample, t_j , which has the largest distance between t_j and the set $S_{j-1} = \{t_1, t_2, \dots, t_{j-1}\}$, that is,

$$t_j = \arg\{\max_r Dist(\mathbf{r}, S_{j-1})\} \quad (7)$$

with $Dist(\mathbf{r}, S_{j-1})$ defined by

$$\begin{aligned} d(\mathbf{r}, S_{j-1}) &= \min_{1 \leq k \leq j-1} Dist(\mathbf{r}, t_k) \\ &= \min\{Dist(\mathbf{r}, t_1), d(\mathbf{r}, t_2), \dots, d(\mathbf{r}, t_{j-1})\}. \end{aligned} \quad (8)$$

It is worth noting that when $j = 2$, $S_1 = \{t_1\}$. In this case, Eq. (7) is reduced to $t_2 = \arg\{\max_r d(\mathbf{r}, t_1)\}$. To conclude this section, it should be noted that the methods discussed in this section were implemented using various similarity measures [2], such as the SAD, SID and HMMID. In all cases, the results obtained were very similar. As a result, this paper only reports experiments based on using SAD for demonstration purposes.

4. Experimental validation

In this section, we make use of a database consisting of a set of mineral mixtures with absolute ground truth. The data consisted of 20 spectra collected using RELAB, a bi-directional, high-resolution spectrometer able to collect 211 spectral bands in the spectral range from 0.4 to 2.5 μm [14]. Spectra of three individual endmembers: Anorthosite, Enstatite and Olivine are present in the database and labeled, respectively, as p_1 , p_2 , p_3 . In addition, the database contains several binary and ternary mixtures of these endmembers [14]. For illustrative purposes, Fig. 5 shows the spectral signatures of the three endmembers. Within the 20 considered signatures, there were

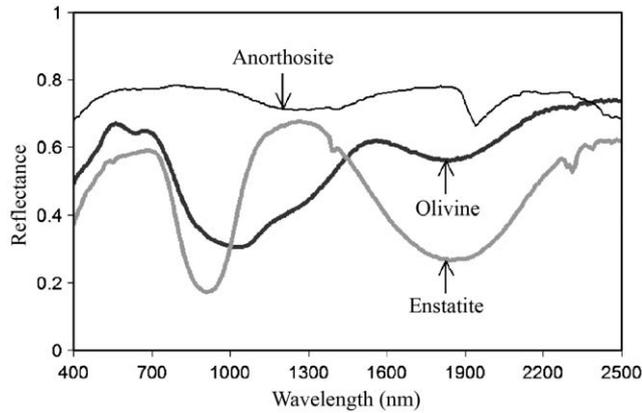


Fig. 5. Spectral signatures of the considered mineral spectra.

Table 1

Known abundance fractions in percentage of spectral signatures in Mustard's data set and selection order of training samples by three different algorithms (MSA, OSP and Maximin).

Signature	True abundance fractions			Sample selection order		
	Anorthosite	Enstatite	Olivine	MSA	OSP	Maximin
p_1	100	–	–	20th	1st	1st
p_2	–	100	–	19th	2nd	5th
p_3	–	–	100	18th	3rd	8th
b_1	90	–	10	17th	19th	4th
b_2	75	–	25	13th	8th	7th
b_3	50	–	50	9th	5th	10th
b_4	25	–	75	12th	12th	6th
b_5	10	–	90	16th	9th	13th
b_6	–	90	10	15th	18th	9th
b_7	–	75	25	11th	13th	2nd
b_8	–	50	50	8th	7th	12th
b_9	–	25	75	10th	16th	14th
b_{10}	–	10	90	14th	14th	3rd
c_1	16.16	16.24	67.60	6th	20th	11th
c_2	16.13	67.85	16.02	7th	11th	16th
c_3	67.81	15.99	16.20	5th	6th	19th
c_4	16.05	41.83	42.12	2nd	4th	18th
c_5	41.92	16.11	41.97	4th	10th	17th
c_6	41.83	41.77	16.40	3rd	15th	15th
c_7	33.61	33.03	33.36	1st	17th	20th

10 binary mixtures, labeled as $\{b_i\}_{i=1}^{10}$, and 7 ternary mixtures, labeled as $\{c_i\}_{i=1}^7$, which are given by the abundance fractions shown in Table 1. The remainder of this section is organized as follows. First, we present the results obtained by the standard linear mixture model in the considered problem. Second, we present the results obtained by the proposed nonlinear methodology, using different training algorithms.

4.1. Linear spectral unmixing

In order to assess the performance of the standard linear mixture model in the considered problem, we first estimated the abundance fractions of pure signatures p_1 , p_2 , p_3 in the 20 available spectra using the fully constrained linear mixture model (FCLSU) in Ref. [39]. The individual results of the estimation are displayed in Fig. 6, which plots scatterplots of true versus FCLSU-estimated fractional abundances for each mineral. The plots in Fig. 6 seem to indicate that binary mixtures in this experiment are mostly linear in nature. As a result, they could be modeled relatively well using a

linear model—it can be seen that binary mixtures of Enstatite/Olivine were very accurately characterized by the standard FCLSU approach (see Fig. 6(b) and (c)). In contrast, binary mixtures containing the Anorthosite endmember were consistently under-predicted or over-predicted by FCLSU (see Fig. 6(a)). This may be explained by the presence of a higher reflectance endmember as the Anorthosite (see Fig. 5). In spite of this, the overall trend in the FCLSU-based abundance estimation revealed a rather linear behavior. However, results in Fig. 6(a) and (b) indicate that nonlinear modeling may be required to characterize ternary mixtures. It should be noted that the overall root mean square error (RMSE) scores in abundance estimation were 0.113 (above 10%) for the Anorthosite, and 0.091 and 0.062 for the Enstatite and Olivine, respectively. For the sake of completeness, we also tested the performance of unconstrained linear mixture models, which consistently showed higher RMSE scores than those reported above.

4.2. Neural network-based nonlinear unmixing

Our next step was to apply our proposed MLP-based model, trained using the algorithms described in Section 3. The order of extraction of available samples by these algorithm is reported in Table 1, which reveals that the MSA algorithm first exhausted all ternary mixtures and then selected all available binary mixtures. It should also be noted that the most highly mixed available samples were always extracted first. For example, binary mixtures made up of 50% of one material and 50% of another were always extracted before the 90–10% or 75–25% mixtures. In fact, the last three signatures extracted by the MSA were the three pure materials. At this point, we emphasize that the Mustard data sets used in experiments were only available to us in the form of a spectral library of signatures with no spatial correlation. As a result, the proposed MEA algorithm was not applied, mainly because this algorithm is specifically designed to exploit spatial information. Similarly, the limited number of spectral signatures contained in the Mustard library prevented us from applying the BTA algorithm for the selection of border training samples, due to observed singularity problems in the covariance estimate of each pure class. In order to assess the performance of the proposed MSA algorithm with other approaches, the OSP and Maximin were also used for selection of spectrally distinct training samples. As shown by Table 1, the OSP first exhausted all three pure signatures, which is consistent with the behavior of the algorithm in other studies [40], and then extracted binary and ternary mixtures with no apparent order. On the other hand, the Maximin extracted both pure and binary mixtures first (the last extracted signatures were all made up of ternary mixtures).

The training samples extracted by the three algorithms above were used (in sequential order, using the first training sample provided by a certain algorithm in first place) to train our MLP-based model. This means that the samples were incorporated one by one to the training set, according to their order of selection indicated in Table 1. Every time a new training sample was incorporated to the training set, it was also automatically removed from the test set. Fig. 7(a)–(c) shows the plots of individual RMSE scores in fractional abundance estimation (in percentage) measured for test signatures, against the number of training samples used in the learning stage of the network. A maximum of 10 samples per training selection algorithm was considered in experiments, in order to guarantee that the test accuracies were measured on a sufficiently representative test set.

From Fig. 7, it is clear that the training samples selected by the mixed-signature selection algorithm were the most informative and useful in terms of decreasing RMSE scores for each endmember material. It can also be seen that the first three training samples generated by OSP (pure training samples) could not accurately model

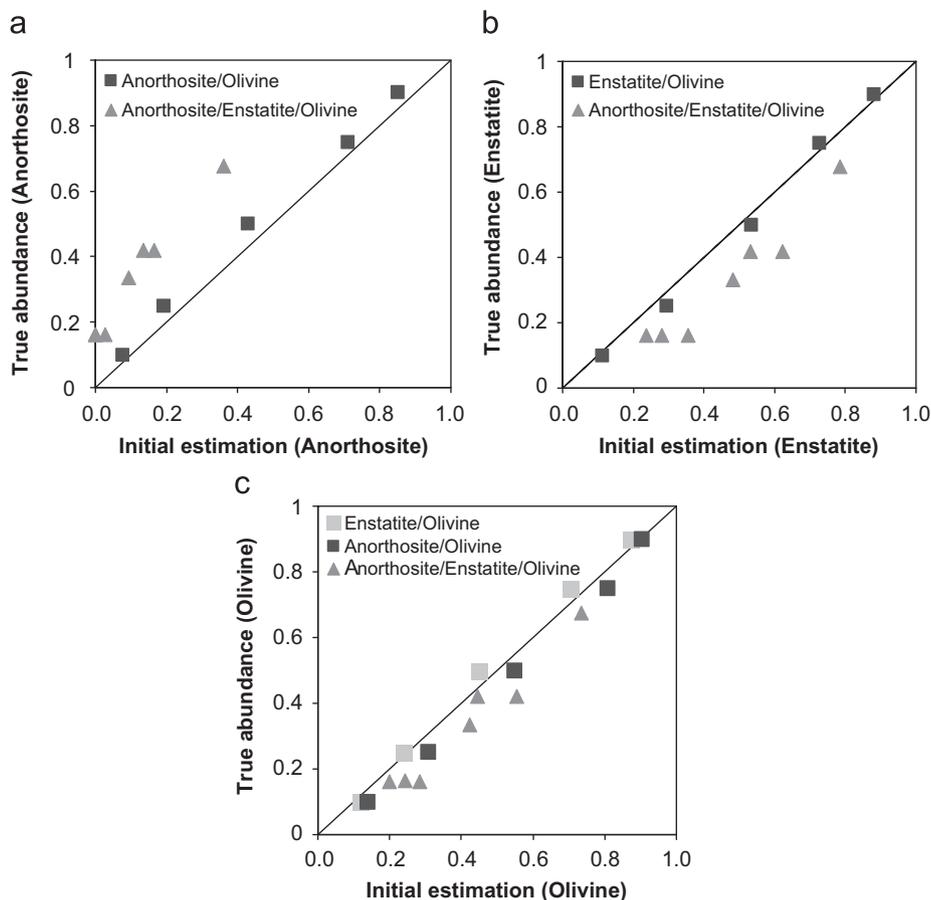


Fig. 6. Scatter-plots of true fractional abundances of mixed spectra in Mustard's data versus the abundance estimations obtained by the linear mixture model (FCLSU). (a) Mixtures containing Anorthosite, (b) mixtures containing Enstatite, and (c) mixtures containing Olivine.

mixing phenomena. Apparently, the explanatory power of the first three or four training samples generated by the MSA (which helped reduced estimation errors to 4% approximately) could not be achieved by a higher number of samples generated using other algorithms. Interestingly, very little improvement was observed when more than three (see Fig. 7(a)) or four (Fig. 7(b)–(c)) samples were used by the proposed MSA, a fact that reveals that the very first signatures extracted by the algorithm were indeed the most useful ones for mixture characterization purposes. Although this experiment cannot be regarded as conclusive due to the limited number of signatures available in the Mustard database, it suggests that the use of intelligent training data selection could assist in the task of exploiting training samples for spectral mixture characterization.

The need for mixed training data does, however, require detailed knowledge on abundance fractions for the considered training sites. In practice, these data are likely to be derived from imagery acquired at a finer spatial resolution than the imagery to be classified, e.g., using data sets acquired by sensors operating simultaneously at multiple spatial resolutions [34]. Such multi-resolution studies may also incorporate prior knowledge, which can be used to help target the location of training sites, and to focus training site selection activities on regions likely to contain the most informative training samples. In the following section, we describe an application of the proposed methodology to a real data analysis scenario, based on the utilization of real hyperspectral images collected at different spatial resolutions.

5. Real hyperspectral image experiments

This section examines the accuracy of the proposed framework using real hyperspectral imagery. Scenes collected over a so-called “Dehesa” semi-arid ecosystem (formed by *quercus ilex* (cork-oak trees), soil and pasture) are used as a case study in experiments [41]. In the Iberian Peninsula, Dehesa systems are used for a combination of livestock, forest and agriculture activity. Determination of fractional land-cover using remote sensing techniques may allow for a better monitoring of natural resources in Dehesa agro-ecosystems. Our choice of this type of landscape was made on several accounts. The first one is the availability of hyperspectral image data sets with accurate geo-registration for a real Dehesa test site in Caceres, SW Spain, collected simultaneously by two instruments operating at multiple spatial resolutions: the digital airborne imaging spectrometer (DAIS 7915) and the reflective optics system imaging spectrometer (ROSIS) of DLR [42]. A second major reason is the simplicity of the Dehesa landscape, which greatly facilitates the collection of reliable field data for model validation purposes.

5.1. Data description

The data used in this study consisted of two main components: image data and field measurements of land-cover fractions, collected at the time of image data acquisition. The image data are formed by a ROSIS scene collected at high spatial resolution, with 1.2-m

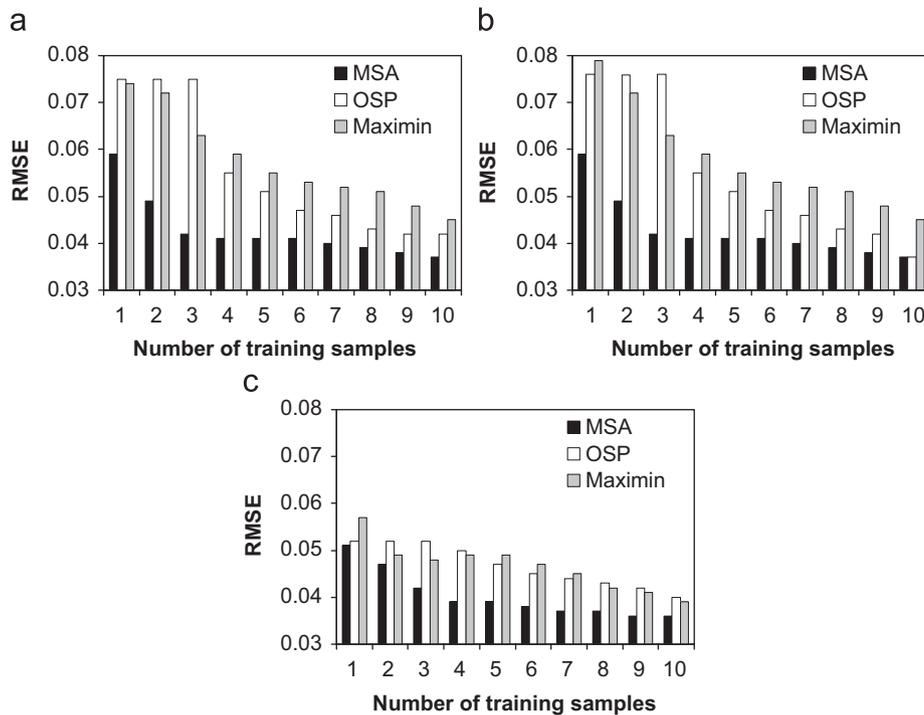


Fig. 7. RMSE scores in fractional abundance estimation for mixed spectra in Mustard's data, obtained after applying the proposed MLP-based model trained with different samples produced by the MSA, OSP and Maximin algorithms. (a) Mixtures containing Anorthosite, (b) mixtures containing Enstatite, and (c) mixtures containing Olivine.

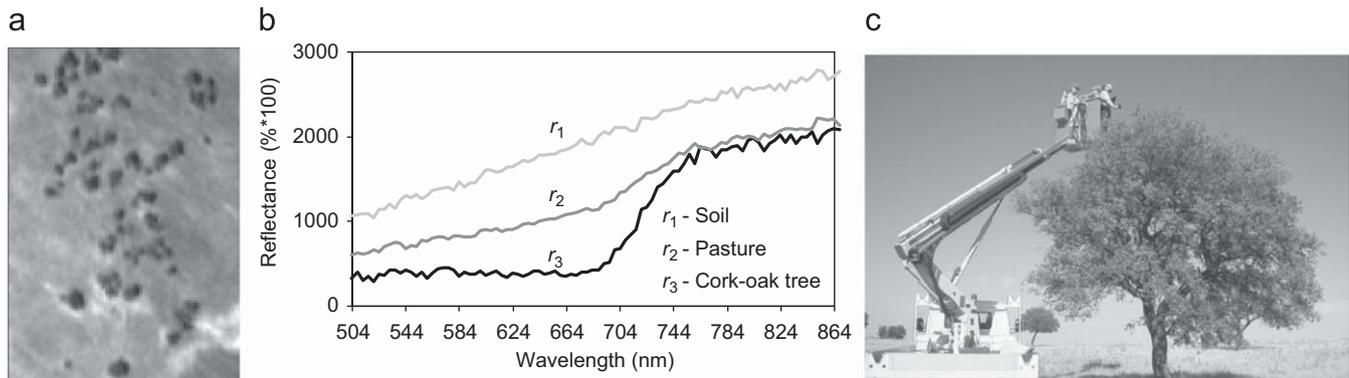


Fig. 8. (a) Spectral band (584nm) of a ROSIS hyperspectral image collected in SW Spain. (b) Endmember signatures of soil (r_1), pasture (r_2) and cork-oak tree (r_3). (c) Ground-truth collection using an ASD FieldSpec Pro spectroradiometer.

pixels, and its corresponding DAIS scene, collected at low spatial resolution with 6-m pixels. The spectral range from 504 to 864 nm was selected for experiments. Fig. 8(a) shows the selected test site, which corresponds to a representative Dehesa test area that contains several cork-oak trees (appearing as dark spots) and pasture (gray) areas on a bare soil (white) background. The following procedure was applied to obtain reliable estimates of the true fractional land-cover for each DAIS pixel in the considered Dehesa test site:

1. First, the ROSIS image was roughly classified into the three land-cover components above using a maximum-likelihood supervised classification approach based on image-derived spectral endmembers, where Fig. 8(b) shows the three endmembers used for mapping that were derived using the automated AMEE algorithm.
2. The classified ROSIS image was registered with the DAIS image using an automated ground control point-based method with sub-pixel accuracy. This method is based on two steps. First, a fully

unsupervised morphological approach for automatic extraction of landmark chips and corresponding windows in the scene is applied. Then, a (hierarchical) robust feature matching procedure, based on a multiresolution overcomplete wavelet decomposition scheme, is used for registration purposes. Additional details about the sub-pixel registration method used in experiments are available in Ref. [43].

3. The classification map was associated with the DAIS image to provide an initial estimation of land-cover classes for each pixel at the DAIS image scale. For that purpose, a 6×6 -m grid was overlaid on the 1.2×1.2 -m classification map derived from the ROSIS scene, where the geographic coordinates of each pixel center point were used to validate the registration procedure at a sub-pixel level.
4. Fractional abundances were initially calculated within each 6×6 -m grid as the proportion or ROSIS pixels labeled as cork-oak tree, pasture and soil within that grid. The abundance maps at the ROSIS level described above were thoroughly refined using field

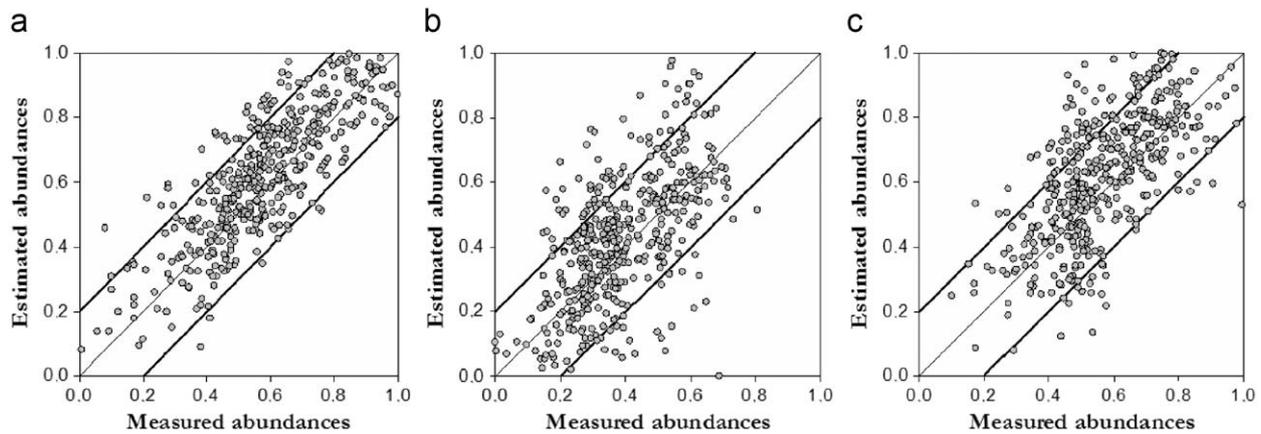


Fig. 9. Abundance estimations of soil, pasture and cork-oak by the linear mixture in the DAIS 7915 image. (a) soil. (b) Pasture. (c) Cork-oak.

data before obtaining the final reference proportions. Specifically, fractional land-cover data were collected on the ground at more than 30 evenly distributed field sites within the test area. These sites were delineated during the field visit as polygons, using high-precision GPS coordinates. At each site, land-cover fractions were estimated using a combination of various techniques. For instance, field spectra were collected for several areas using an ASD FieldSpec Pro spectro-radiometer. Of particular interest were field measurements collected on top of tree crowns (see Fig. 8(c)), which allowed us to model different levels of tree crown transparency [44,45]. Ground estimations of pasture abundance were conducted in selected sites of known dimensions, using pasture harvest procedures supported by visual inspection and laboratory analyses [46]. The final reference estimations were the consensus of a group of experts with experience in similar field data collection and validation activities [47].

5.2. Experimental results

This subsection describes the accuracy results obtained by different spectral unmixing techniques in the considered problem. The subsection is organized as follows. First, we present the results obtained by the standard linear mixture model in the considered problem. Second, we present the results obtained by the proposed nonlinear methodology, using different training algorithms. The subsection concludes with a brief discussion on the non-negativity and sum-to-one constraints in the proposed nonlinear approach.

5.2.1. Linear spectral unmixing

Fig. 9 shows the scatterplots of measured versus FCLSU-estimated fractional abundances for the three endmember materials (successfully estimated by the VD method) in the DAIS (low spatial resolution) image data set, where the diagonal represents perfect match and the two flanking lines represent plus/minus 20% error bound. As shown by Fig. 9, most linear predictions for the soil endmember fall within the 20% error bound (see Fig. 9(a)). On the other hand, the multiple scattering within the pasture and cork-oak tree canopies (and from the underlying surface in the latter case) complicated the spectral mixing in nonlinear fashion [45], which resulted in a generally higher number of estimations lying outside the error bound, as illustrated in Fig. 9(b) and (c). Also, the RMSE scores in abundance estimation for the soil (0.119), pasture (0.153) and cork-oak tree (0.169) were all above 10% estimation error in percentage, which suggested that linear mixture modeling was not flexible enough to

accommodate the full range of spectral variability throughout the landscape.

5.2.2. Neural network-based nonlinear unmixing

Training data are used in this experiment to model nonlinear mixing effects. Fig. 10(a)–(e) shows the training areas automatically extracted from the DAIS scene by the proposed algorithms, while Fig. 10(f) plots the spectral signatures of some selected training samples by different algorithms (specifically, the signature of the first training sample selected by each algorithm is displayed). In all cases, the number of training samples was limited to a very low figure (only six) with the goal of evaluating the possibility of using a very reduced number of training data for practical purposes. For the BTA algorithm, we used $\theta = 0.1$, which is a reasonable spectral similarity value according to our previous experiments in Ref. [35]. The structuring element, B , used in the MEA training sample selection algorithm was an isotropic disk-shaped element of 3-pixel radius, selected in accordance with the size of patterns of interest in the scene [38].

Table 2 compares the performance of the proposed MLP-based model trained with different algorithms and number of samples, where the VD concept [23] was used in all cases to automatically estimate the number of endmembers in the considered DAIS scene. From Table 2, it is clear that using three training samples generated by the BTA, MSA and MEA algorithms always introduced a significant improvement in abundance estimation with regard to the cases in which less training samples were considered. It is also apparent from Table 2 that using additional training samples obtained by BTA, MSA and MEA did not significantly improve the quality of abundance estimations. In this case, the first three samples seemed to be the most informative ones. Quite opposite, the OSP and Maximin algorithms produced rather unstable results, with only moderately acceptable scores (always above 10% estimation error) when all six training samples were used. Interestingly, we experimentally tested that the incorporation of additional samples obtained by those algorithms could never reduce estimation errors below the 10% threshold. This is due to the fact that many of the samples extracted by these algorithms in Fig. 10(d) and (e) were highly pure, spectrally, and could not provide additional information about complex mixtures. Table 2 also reveals that, for both the pasture and the cork-oak tree sites, the MLP-based model trained with the MSA resulted in significantly smaller RMSEs than those found by using the BTA algorithm. Interestingly, when the MEA was used, the proposed model produced the best estimation results for soil and cork-oak tree (around 5% estimation error in percentage using only three training samples),

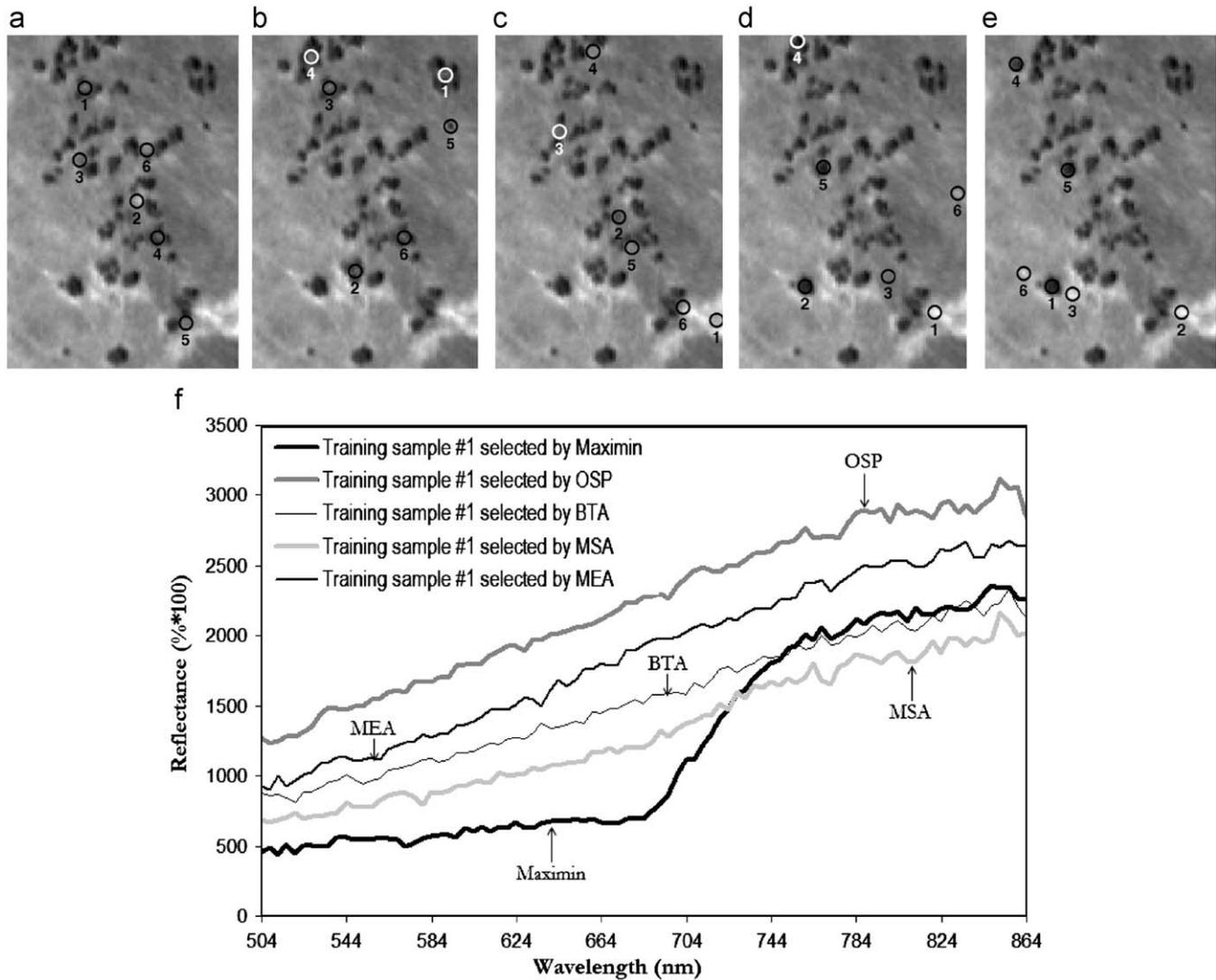


Fig. 10. (a)–(e) Spatial location of training samples extracted by the considered selection algorithms from the DAIS 7915 image. (a) BTA, (b) MSA, (c) MEA, (d) OSP, (e) Maximin and (f) Spectral signature of the first training sample selected by each algorithm.

but the pasture estimations were slightly less accurate than those produced by the MSA perhaps since the pasture component in Dehesa environments generally follow a very irregular spatial pattern, as can be observed comparatively with soil and cork-oak trees in Fig. 10(a). As a result, the introduction of spatial information by the MEA may not be particularly relevant for this land-cover type.

Fig. 11 shows the scatter plots of measured versus predicted fractional abundances for soil, pasture and cork-oak tree by the proposed MLP-based model, trained only with the first three training samples generated by MEA. From Fig. 11, it is clear that the utilization of intelligently selected training samples resulted in fewer points outside the two 20% difference lines, most notably, for both pasture and cork-oak abundance estimates. The pattern of the scatter plots obtained for the soil predictions was similar to the one in Fig. 10(a), in particular, when the soil abundance was high. This suggests that nonlinear effects in Dehesa landscapes mainly result from multiple scattering effects in vegetation canopies.

Fig. 12 plots a curve for each considered training algorithm indicating how many pixel-level predictions fall within a given error bound (in percentage) of the actual field measurements. Linear predictions are also included in the graph. It can be seen that the highest percentage of pixels with prediction errors below 2% was achieved

by MSA. However, with the increase of error bound, the MEA outperformed the MSA, in particular, after the error bound was over 6%. This indicates that the incorporation of spatial information in the process of finding training samples can minimize the global error, although this may come at the expense of sacrificing part of local prediction performance in nonlinear classes.

5.2.3. Non-negativity and sum-to-one constraints in the proposed approach

Although abundance sum-to-one and non-negativity constraints were not imposed in our proposed MLP-based learning stage, negative and/or unrealistic abundance estimations (which usually indicate a bad fit of the model and reveal inappropriate endmember/training data selection) were very rarely found, in particular, when the BTA, MSA or MEA algorithms were used to generate the training set. With the above issue in mind, experimental results in this paper illustrate the advantages of adopting intelligent training mechanisms for the proposed MLP-based model. In particular, the use of intelligent training sample generation algorithms appears to play a very significant role, thus showing the potential to direct training data collection strategies to target the most useful training sites without prior knowledge about the scene.

Table 2

RMSE scores in fractional abundance estimation of soil (r_1), pasture (r_2) and cork-oak tree (r_3) in the DAIS 7915 hyperspectral image using the proposed MLP-based model trained by different algorithms.

Training algorithm	Material	Number of training samples					
		1	2	3	4	5	6
BTA	r_1	0.116	0.109	0.103	0.102	0.102	0.102
	r_2	0.146	0.121	0.093	0.093	0.093	0.090
	r_3	0.159	0.134	0.095	0.094	0.091	0.094
MSA	r_1	0.108	0.092	0.061	0.059	0.059	0.059
	r_2	0.116	0.087	0.040	0.039	0.040	0.041
	r_3	0.144	0.116	0.063	0.060	0.59	0.057
MEA	r_1	0.103	0.089	0.059	0.057	0.055	0.054
	r_2	0.125	0.096	0.046	0.044	0.043	0.044
	r_3	0.136	0.093	0.048	0.047	0.046	0.045
OSP	r_1	0.112	0.114	0.111	0.109	0.107	0.104
	r_2	0.147	0.143	0.140	0.136	0.131	0.124
	r_3	0.165	0.162	0.152	0.143	0.138	0.132
Maximin	r_1	0.115	0.112	0.108	0.109	0.107	0.105
	r_2	0.150	0.143	0.142	0.139	0.137	0.127
	r_3	0.163	0.159	0.156	0.151	0.145	0.139

The numbers marked in bold typeface represent the best compromise values between the number of training samples used and the RMSE score obtained. The scores produced by the linear mixture model (FCLSU) for the above three materials were 0.119, 0.153 and 0.169, respectively.

6. Conclusions and future lines

This paper has focused on the problem of characterizing mixed pixels in remotely sensed hyperspectral images. In order to interpret and model complex nonlinear mixtures often observed in those scenes, we have proposed a simple nonlinear model trained with highly representative training sets which can accurately explain the complex nature of the data using only a few training samples. In order to automate the process of selecting the most useful training samples from the input data set, we have developed several new training sample generation algorithms. Our study reveals that there is a need for nonlinear mixture models coupled with unsupervised algorithms able to seek for the most informative training samples (which are shown in this work to be the most highly mixed signatures in the input data set). Critically, if the regions expected to contain the most highly informative training samples for spectral mixture modeling can be identified in advance, then it is possible to direct the training data acquisition procedures to these regions, and thus reduce the number of required training sites without the loss of prediction accuracy. This issue is of particular importance in real applications, in which the acquisition of large training sets is generally very costly in terms of time and finance.

As with any new approach, there are some unresolved issues that may present challenges over time. For instance, the selection of

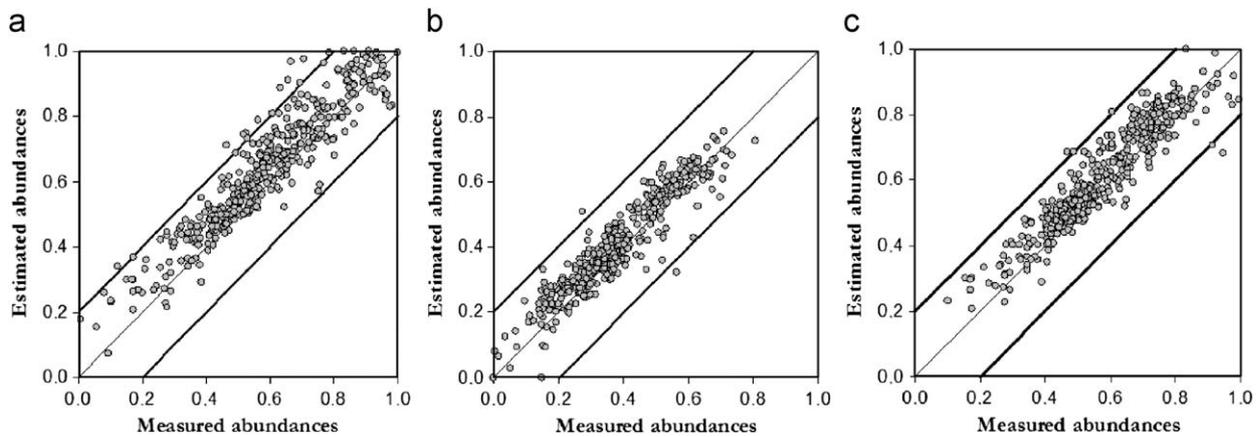


Fig. 11. Abundance estimation results for soil (a), pasture (b) and cork-oak (c) in the DAIS 7915 image using the MLP-based model trained with the first three training samples produced by MEA.

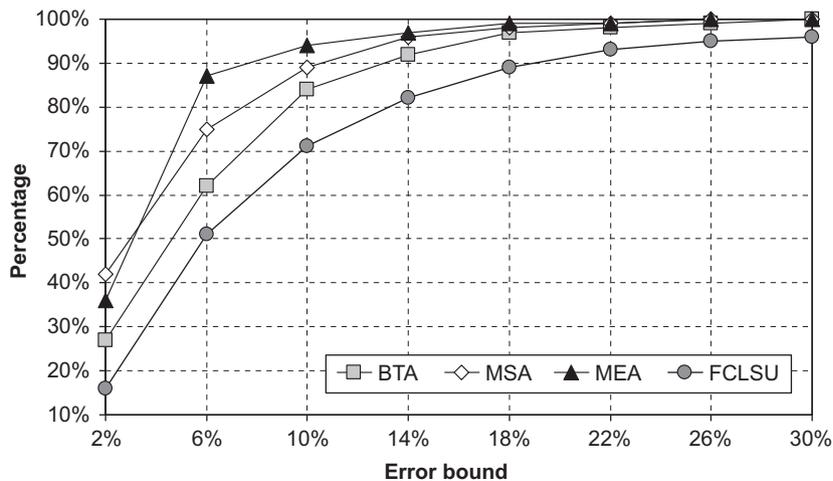


Fig. 12. Percent of pixels in the DAIS 7915 image that lie within a given percent of the actual field measurement after applying the linear mixture model (FCLSU) and the proposed MLP-based model trained with the first three training samples produced by BTA, MSA and MEA algorithms.

endmembers for initialization of two of the proposed algorithms (BTA and MSA) is subjective and, hence, the choice of different end-member sets may lead to different results. However, we generally prefer the model abstraction offered by the concept of spectral end-member to other approaches such as those based on the availability of spectral libraries, which are not very tractable and require very detailed ground information (especially for large areas). Also, the proposed model assumes that the spectral signatures used for the initialization represent pure land-cover classes. In reality, it may be difficult to find such “pure” sites in coarse spatial resolution images, where the spectral response a pixel may be given by several land-cover types. This problem reflects the inherent nature of land-covers and mapping processes, and can only be properly addressed by the utilization of reliable ground data and spectral libraries. In this regard, future analyses should incorporate a denser sampling of sites for training and testing. Another topic deserving future research is the interaction between atmospheric effects and model predictions, which could not be fully addressed in this work due to the limited effect of atmospheric interferers on our test sites.

Acknowledgments

This research was supported by the European Commission through the Marie Curie project entitled “Hyperspectral Imaging Network” (Contract no. MRTN-CT-2006-035927). Funding from the Spanish Ministry of Science and Innovation (HYPERCOMP/EODIX project, reference AYA2008-05965-C04-02) is also gratefully acknowledged. The authors gratefully acknowledge Andreas Müller for his lead of the DLR (*Deutsches Zentrum für Luft- und Raumfahrt*) project that allowed us to obtain the real hyperspectral data sets, and John F. Mustard for providing the laboratory data used in experiments. Last but not least, the authors gratefully acknowledge the anonymous reviewers for their comments and suggestions, which have greatly improved the technical quality and presentation of our manuscript.

References

- [1] R.A. Schowengerdt, Remote sensing: models and methods for image processing, second ed., Academic Press, New York, 1997.
- [2] C.-I. Chang, Hyperspectral imaging: techniques for spectral detection and classification, Kluwer, New York, 2003.
- [3] Q. Du, R. Nekovei, Implementation of real-time constrained linear discriminant analysis to remote sensing image classification, *Pattern Recognition* 38 (2005) 459–471.
- [4] A. Plaza, P. Martínez, R. Pérez, J. Plaza, A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles, *Pattern Recognition* 37 (2004) 1097–1116.
- [5] Q. Du, H. Ren, Real-time constrained linear discriminant analysis to target detection and classification in hyperspectral imagery, *Pattern Recognition* 36 (2003) 1–12.
- [6] N. Keshava, J.F. Mustard, Spectral unmixing, *IEEE Signal Processing Magazine* 19 (2002) 44–57.
- [7] J.B. Adams, M.O. Smith, P.E. Johnson, Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander 1 site, *Journal of Geophysical Research* 91 (1986) 8098–8112.
- [8] J.J. Settle, N.A. Drake, Linear mixing and the estimation of ground cover proportions, *International Journal of Remote Sensing* 15 (1993) 2381–2392.
- [9] W. Liu, E.Y. Wu, Comparison of non-linear mixture models: sub-pixel classification, *Remote Sensing of Environment* 18 (2004) 1976–2003.
- [10] J. Wang, C.-I. Chang, Applications of independent component analysis in endmember extraction and abundance quantification for hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing* 44 (2006) 2601–2616.
- [11] J.M.P. Nascimento, J.M. Bioucas, Does independent component analysis play a role in unmixing hyperspectral data?, *IEEE Transactions on Geoscience and Remote Sensing* 43 (2005) 175–187.
- [12] P. Comon, Independent component analysis: a new concept, *Signal Processing* 36 (1994) 287–314.
- [13] G.G. Wilkinson, Open questions in neurocomputing for earth observation, in: I. Kanellopoulos, G.G. Wilkinson, F. Roli, J. Austin (Eds.), *Neurocomputation in remote sensing data analysis*, Springer, Berlin, 1999, pp. 3–13.
- [14] K. Guilfoyle, M.L. Althouse, C.-I. Chang, A quantitative and comparative analysis of linear and nonlinear spectral mixture models using radial basis function neural networks, *IEEE Transactions on Geoscience and Remote Sensing* 39 (2001) 2314–2318.
- [15] A. Baraldi, E. Binaghi, P. Blonda, P.A. Brivio, A. Rampini, Comparison of the multilayer perceptron with neuro-fuzzy techniques in the estimation of cover class mixture in remotely sensed data, *IEEE Transactions on Geoscience and Remote Sensing* 39 (2001) 994–1005.
- [16] C.M. Bishop, *Neural networks for pattern recognition*, Oxford University Press, Oxford, 1995.
- [17] C. Huang, J.R.G. Townshend, A stepwise regression tree for nonlinear approximation: applications to estimating subpixel land cover, *International Journal of Remote Sensing* 24 (2003) 75–90.
- [18] X. Zhuang, B.A. Engel, D.F. Lozano, R.B. Fernández, C.J. Johannsen, Optimization of training data required for neuro-classification, *International Journal of Remote Sensing* 15 (1999) 3271–3277.
- [19] M. Chi, L. Bruzzone, A semilabeled-sample-driven bagging technique for ill-posed classification problems, *IEEE Geoscience and Remote Sensing Letters* 2 (2005) 69–73.
- [20] C. Lee, D.A. Landgrebe, Decision boundary feature extraction for neural networks, *IEEE Transactions on Neural Networks* 8 (1997) 75–83.
- [21] P.M. Atkinson, Optimal ground-based sampling for remote-sensing investigations—estimating the regional mean, *International Journal of Remote Sensing* 12 (1991) 559–567.
- [22] L. Bruzzone, D.F. Prieto, Unsupervised retraining of a maximum likelihood classifier for the analysis of multitemporal remote sensing images, *IEEE Transactions on Geoscience and Remote Sensing* 39 (2001) 456–460.
- [23] C.-I. Chang, Q. Du, Estimation of number of spectrally distinct signal sources in hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing* 42 (2004) 608–619.
- [24] J.C. Harsanyi, W. Farrand, C.-I. Chang, Detection of subpixel spectral signatures in hyperspectral image sequences, in: *Proceedings of the American Society on Photogrammetry and Remote Sensing*, Reno, NV, 1994, pp. 236–247.
- [25] J. Plaza, P. Martínez, R. Pérez, A. Plaza, Joint linear/nonlinear spectral unmixing of hyperspectral image data, in: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, Barcelona, Spain, 2007, pp. 4037–4040.
- [26] G.M. Foody, M. Arora, Incorporating mixed pixels in the training, allocation and testing stages of supervised classifications, *Pattern Recognition Letters* 17 (1996) 1389–1398.
- [27] C. Brumbley, C.-I. Chang, An unsupervised vector quantization-based target subspace projection approach to mixed pixel detection and classification in unknown background for remotely sensed imagery, *Pattern Recognition* 32 (1999) 1161–1174.
- [28] J.T. Tou, R.C. González, *Pattern Recognition Principles*, Addison-Wesley, Reading, MA, 1974 pp. 92–94.
- [29] H. Ren, C.-I. Chang, Automatic spectral target recognition in hyperspectral imagery, *IEEE Transactions on Aerospace and Electronic Systems* 39 (4) (2003) 1232–1249.
- [30] C. Brumbley, C.-I. Chang, An unsupervised vector quantization-based target subspace projection approach to mixed pixel detection and classification in unknown background for remotely sensed imagery, *Pattern Recognition* 32 (1999) 1161–1174.
- [31] G.M. Foody, The significance of border training patterns in classification by a feedforward neural network using backpropagation learning, *International Journal of Remote Sensing* 20 (1999) 3549–3562.
- [32] M.E. Winter, N-FINDR: an algorithm for fast autonomous spectral endmember determination in hyperspectral data, in: *Proceedings of SPIE Image Spectrometry V*, vol. 3753, 1999, pp. 266–277.
- [33] S. Tadjudin, D.A. Landgrebe, Covariance estimation with limited training samples, *IEEE Transactions on Geoscience and Remote Sensing* 37 (1999) 2113–2118.
- [34] G.M. Foody, A. Mathur, Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification, *Remote Sensing of Environment* 93 (2004) 107–117.
- [35] A. Plaza, P. Martínez, R. Pérez, J. Plaza, A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data, *IEEE Transactions on Geoscience and Remote Sensing* 42 (2004) 650–663.
- [36] A. Plaza, P. Martínez, R. Pérez, J. Plaza, Spatial/spectral endmember extraction by multidimensional morphological operations, *IEEE Transactions on Geoscience and Remote Sensing* 40 (2002) 2025–2041.
- [37] P. Soille, *Morphological Image Analysis: Principles and Applications*, second ed., Springer, Berlin, Germany, 2003.
- [38] A. Plaza, P. Martínez, J. Plaza, R. Pérez, Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations, *IEEE Transactions on Geoscience and Remote Sensing* 43 (2005) 466–479.
- [39] D. Heinz, C.-I. Chang, Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing* 39 (2000) 529–545.
- [40] J. Harsanyi, C.-I. Chang, Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach, *IEEE Transactions on Geoscience and Remote Sensing* 32 (1994) 779–785.
- [41] F.J. Pulido, M. Díaz, S.J. Hidalgo, Size structure and regeneration of Spanish holm oak *quercus ilex* forests and Dehesas: effects of agroforestry use on their long-term sustainability, *Forest Ecology and Management* 146 (2001) 1–13.
- [42] A. Müller, A. Hausold, P. Strobl, HySens-DAIS/ROSIIS imaging spectrometers at DLR, in: *Proceedings of SPIE*, vol. 5508, 2001, pp. 660–670.
- [43] A. Plaza, J. Le Moigne, N.S. Netanyahu, Morphological feature extraction for automatic registration of multispectral scenes, in: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, Barcelona, Spain, 2007, pp. 421–424.

- [44] J. Plaza, A. Plaza, P. Martínez, R. Pérez, Nonlinear mixture models for analyzing laboratory simulated-forest hyperspectral data, in: *Proceedings of SPIE*, vol. 5508, 2003, pp. 660–670.
- [45] J. Plaza, P. Martínez, A. Plaza, R. Perez, Quantifying the impact of spatial resolution on endmember extraction from hyperspectral imagery, in: *Proceedings of third EARSeL Workshop on Imaging Spectroscopy*, Herrsching, Germany, 2003, pp. 117–125.
- [46] G.A. Carpenter, S. Gopal, S. Macomber, S. Martens, C.E. Woodcock, A neural network method for mixture estimation for vegetation mapping, *Remote Sensing of Environment* 70 (1999) 138–152.
- [47] J. Moreno, V. Caselles, J.A. Martínez-Lozano, J. Meliá, J. Sobrino, J.A. Calera, F. Montero, J.M. Cisneros, The measurements programme at Barrax, in: *Proceedings of DAISEX Final Results Workshop*, ESA SP-499, 2001, pp. 43–51.

About the Author—JAVIER PLAZA received the Ph.D. degree in Computer Engineering from the University of Extremadura, Spain, in 2008. Dr. Plaza is an Assistant Professor with the Computer Science Department, University of Extremadura, Spain. He has been a Visiting Researcher at the University of Maryland, Baltimore County, MD. His current research work is focused on the development of efficient implementations of nonlinear mixture model-based algorithms for abundance estimation of materials in hyperspectral images. He is also involved in the design and configuration of commodity cluster computing architectures for high-performance hyperspectral analysis. Other major research interests include the development of quantitative and comparative applications for remote sensing, and design of web-based applications for hyperspectral image processing. He has served as a reviewer for several journals, including the *IEEE Transactions on Image Processing*, the *IEEE Transactions on Geoscience and Remote Sensing*, and the *IEEE Geoscience and Remote Sensing Letters*.

About the Author—ANTONIO PLAZA received the Ph.D. degree in Computer Science from the University of Extremadura, Caceres, Spain, in 2002. Dr. Plaza has been an Associate Professor with the Computer Science Department, University of Extremadura, Spain, since 2000. He has held the Visiting Researcher positions at the Applied Information Sciences Branch, NASA/Goddard Space Flight Center, Greenbelt, MD, the University of Maryland, Baltimore County, MD, and the AVIRIS Data Facility, NASA/Jet Propulsion Laboratory, Pasadena, CA. His main research interests span computer vision, image processing, pattern recognition, and development and efficient implementation of hyperspectral image analysis algorithms on parallel computing facilities. Dr. Plaza has authored or co-authored more than 150 publications including journal papers, book chapters and peer-reviewed conference proceedings. He has edited a book on High Performance Computing in Remote Sensing for Taylor and Francis and holds two patents on hyperspectral imaging. He is the coordinator of Hyper-I-Net, a Marie Curie Research Training Network focused on hyperspectral imaging activities in Europe, and currently serves as an Associate Editor for the *IEEE Transactions on Geoscience and Remote Sensing* journal in the areas of hyperspectral image analysis and signal processing. Additional information about the activities of Dr. Plaza can be found online at <http://www.umbc.edu/rssipl/people/aplaza>.

About the Author—ROSA PEREZ received the Ph.D. degree in Computer Science from the Polytechnic University of Madrid, Spain, in 1995. Dr. Perez has been a Professor of Computer Science at the Computer Science Department, University of Extremadura, Spain, since 1985. Her main research interests include neural networks, systolic array- and FPGA-based design, pattern recognition, and signal processing. She has served as a reviewer for the *IEEE Transactions on Neural Networks*.

About the Author—PABLO MARTINEZ received the Ph.D. degree in Physics from the University of Granada, Spain, in 1992. Dr. Martinez has been a Professor of Computer Science at the University of Extremadura, Spain, since 1985. He is the Head Scientist of the Neural Networks and Signal Processing Group (GRNPS). He has held the Visiting Researcher positions at the Applied Information Sciences Branch, NASA/Goddard Space Flight Center, Greenbelt, MD, and the Department of Electrical Engineering, University of Maryland, College Park, MD. His main research interests include remote sensing, digital image analysis and neural network-based pattern recognition. He has been a reviewer for the *IEEE Transactions on Geoscience and Remote Sensing*.