Pattern Recognition 42 (2009) 3032 – 3045

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

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 Caceres, 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.

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 endmember 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).

0031-3203/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2009.04.008

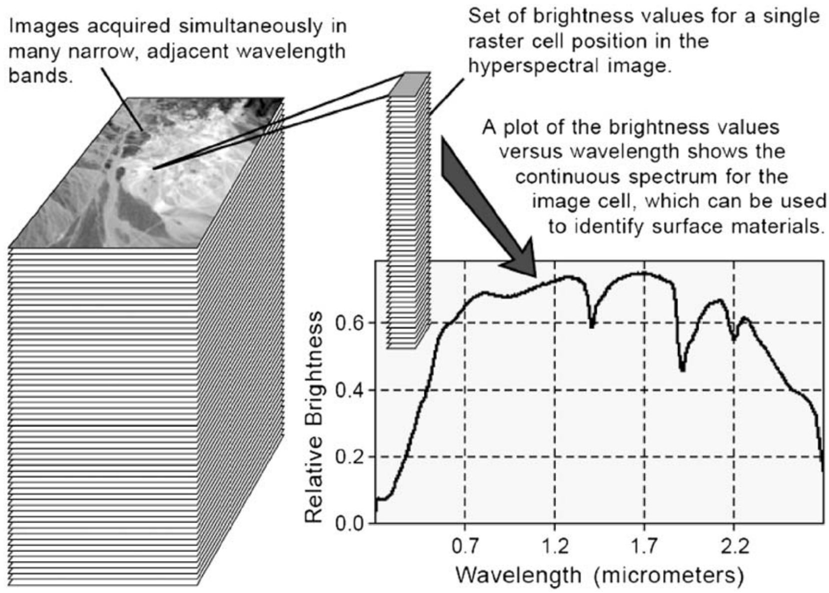

Fig. 1. The concept of hyperspectral imaging.

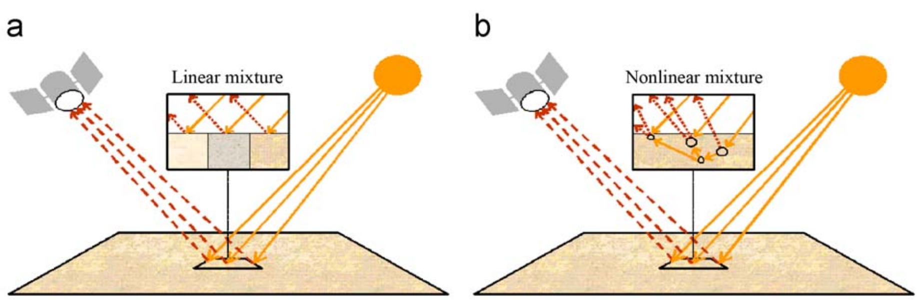

Fig. 2. Schematical description of scattering in linear (single scattering) (a) and nonlinear (multiple scattering) (b) mixtures.

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:

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

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

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

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