Neurocomputing 72 (2009) 2101–2110

Contents lists available at ScienceDirect

Neurocomputing

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

Autonomous single-pass endmember approximation using lattice auto-associative memories ☆

Gerhard X. Ritter [a], Gonzalo Urcid [b,*], Mark S. Schmalz [a]

[a] CISE Department, University of Florida, Gainesville, FL 32611, USA
[b] Optics Department, INAOE, Tonantzintla, Pue 72000, Mexico

ARTICLE INFO

Available online 16 December 2008

PACS:
87.19.lv
87.85.dq
93.85.Bc
93.85.Pq

Keywords:
Neural networks: associative memories
Lattice associative memories
Lattice algebra: lattice independence
Hyperspectral image analysis: endmember search
Spectral unmixing
Abundance maps

ABSTRACT

We propose a novel method for the autonomous determination of endmembers that employs recent results from the theory of lattice based auto-associative memories. In contrast to several other existing methods, the endmembers determined by the proposed method are physically linked to the data set spectra. Numerical examples are provided to illustrate lattice theoretical concepts and a hyperspectral image subcube, from the Cuprite site in Nevada, is used to find all endmember candidates in a single pass.

1. Introduction

Advances in passive remote sensing has produced imaging devices with ever growing spectral resolution. The high spectral resolution produced by current hyperspectral imaging devices facilitates identification of fundamental materials that make up a remotely sensed scene and thus supports discrimination between them. A typical pixel of a multispectral or hyperspectral image generally represents a region on the ground consisting of several square meters. For example, each Landsat Thematic Mapper pixel represents a $30 \times 30\,m^2$. Thus, a hyperspectral image pixel can have all or parts of many different objects in it. The collection of measured reflectance values associated with the pixel is called the *spectrum* of the pixel. It is, therefore, useful to know the percentage of different, fundamental object parts that are most represented in the spectrum of a given pixel. The most widely used spectral mixing model is the *linear mixing model*, which assumes that the observed reflectance spectrum of a given pixel is a linear combination of a small number of unique constituent deterministic signatures known as *endmembers*. This model has been used by a multitude of researchers ever since Adam et al. [1] analyzed an image of Mars using four endmembers. In the cited reference and various other applications, hyperspectral image segmentation and analysis takes the form of a pattern recognition problem as the segmentation problems reduces to matching the spectra of the hyperspectral image to predetermined spectra stored in a library. In many cases, however, endmembers cannot be determined in advance and must be selected from the image directly by identifying the pixel spectra that are most likely to represent the fundamental materials. This compromises the *autonomous endmember detection* problem. Unfortunately, the spatial resolution of a sensor makes it often unlikely that any pixel is composed of a single endmember. Thus, the determination of endmembers becomes a search for image pixels with the least contamination from other endmembers. These are also referred to as *pure* pixels. The pure pixels exhibit *maximal* reflectance in certain spectral bands and correspond to vertices of a high dimensional simplex. This simplex, hopefully, encloses most if not all the pixel spectra.

In this paper we assume the *linear mixing model*, which is based on the fact that points on a simplex can be represented as a linear sum of the vertices that determine the simplex [8,17,18]. The mathematical equations of the model and its constraints are

☆ Expanded version of a paper presented at the 10th Joint Conference on Information Sciences (Salt Lake City, July 2007).

* Corresponding author. Tel.: +52 222 266 3100; fax: +52 222 247 2940.
E-mail addresses: ritter@cise.ufl.edu (G.X. Ritter), gurcid@inaoep.mx (G. Urcid), mssz@cise.ufl.edu (M.S. Schmalz).

0925-2312/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2008.06.025

given, respectively, by Eqs. (1) and (2):

$$\boldsymbol{x} = \sum_{k=1}^{m} a_k \boldsymbol{e}^k + \boldsymbol{n} = E\boldsymbol{a} + \boldsymbol{n}, \tag{1}$$

$$\sum_{k=1}^{m} a_k = 1 \quad \text{and} \quad a_k \geqslant 0 \quad \forall k, \tag{2}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ is the measured spectrum over n bands of an image pixel, $E = (\boldsymbol{e}^1, \boldsymbol{e}^2, \ldots, \boldsymbol{e}^m)$ is an $n \times m$ matrix whose columns are the m endmember spectra assumed to be affinely independent, $\boldsymbol{a} = (a_1, a_2, \ldots, a_m)^{\mathrm{T}}$ is an m-dimensional column vector whose entries are the corresponding fractional abundances or, equivalently, the percentages of endmember spectra present in $\boldsymbol{x}$, and $\boldsymbol{n} \in \mathbb{R}^n$ is an additive noise vector.

Endmembers may be obtained from spectral libraries for certain specific materials, or autonomously from the image by a variety of techniques [3,4,27,33,34]. Autonomous endmember detection has received wide attention since signatures of various objects that may be present in an image are unknown before hand. Boardman [3,4] uses the framework of the geometry of convex sets to identify the $m+1$ endmembers as the vertices of the smallest simplex that bounds the measured data. A major problem is that the vertices need not be image pixels (which in most cases they are not) and, hence, need not have any physical connection to actual image data.

Winter's N-FINDR method [33,34] is based on inflating a simplex within the data set to determine the largest simplex *inscribed* within the data. It is not clear how pixels outside the inscribed data are handled and the exact algorithm is not available in print or on the web. Additionally, the algorithm is computationally intensive despite claims to the contrary. Individual pixels need to be examined and simplex volume recalculated for each image pixel. In contrast, the autonomous endmember determination proposed in this paper is extremely fast and carries little computational overhead. The method is derived from examining a lattice based auto-associative memory that stores the hyperspectral image cube in its memory. Graña et al. [10–12,14] was the first to propose the use of lattice based auto-associative memories for autonomous endmember determination. Specifically, he employs the notion of morphological independence which does not necessarily lead to finding an affinely independent set of vectors that in some sense provides a maximal simplex within the data set. Graña's algorithm forces the user to choose a starting pixel and different starting pixels can produce different results. The method described in this paper is different and will always provide the *same* sets of endmembers for a given hyperspectral image. Recent works based on strong lattice independence and alternative criteria to get a set of final endmembers appear in [13,36].

2. Mathematical background

2.1. Linear and affine independence

If $X = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k\} \subset \mathbb{R}^n$, denotes a finite set of real vectors, recall that a *linear combination* of X is an expression of the form $\sum_{\xi=1}^{k} a_\xi \boldsymbol{x}^\xi$ where the a_ξ's are scalars, i.e., $a_\xi \in \mathbb{R}$ for all $\xi \in K = \{1, \ldots, k\}$. Then, X is said to be a *linearly independent* set if the unique solution to the equation $\sum_{\xi=1}^{k} a_\xi \boldsymbol{x}^\xi = \boldsymbol{0}$ is given by $a_\xi = 0$ for $\xi \in K$. Otherwise, the vectors in X are said to be linearly dependent. The next lemma states a basic result in linear algebra [9].

Lemma 2.1. *Let $K^\gamma = K\backslash\{\gamma\}$ denote the index set from which index γ has been deleted. If the set of vector differences, $X' = \{\boldsymbol{x}^\xi - \boldsymbol{x}^\gamma : \xi \in K^\gamma\}$ is linearly independent for some $\gamma \in K$, then X' is a linearly independent set $\forall \gamma \in K$.*

Thus, to form set X', any vector $\boldsymbol{x}^\gamma$ in X, considered as a "point", can be selected as an origin for the remaining vectors in $X\backslash\{\boldsymbol{x}^\gamma\}$. From a geometrical point of view, an *affine combination* or *barycentre* is a linear combination of X subject to the condition $\sum_{\xi=1}^{k} a_\xi = 1$. Furthermore, a *convex combination* is an affine combination such that, $a_\xi \geqslant 0\ \forall \xi \in K$, and the set of all convex combinations formed with elements of X is known as the *convex hull* of X, denoted here as $C(X)$. In effect, an affine combination is a weighted average of the points in question. For example, the unique point $\boldsymbol{x} \in C(X)$ computed as $(1/k)\sum_{\xi=1}^{k} \boldsymbol{x}^\xi$, is the convex combination known as the *center of mean distances* of X.

With the help of Lemma 2.1, it is possible to characterize the notion of affine independence as follows: $X = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k\} \subset \mathbb{R}^n$ is said to be an *affinely independent* set if $X' = \{\boldsymbol{x}^\xi - \boldsymbol{x}^\gamma : \xi \in K^\gamma\} \subset \mathbb{R}^n$ is a linearly independent set for some $\gamma \in K$ [9]. Notice that, although set X has k elements, there are only $k-1$ points in X'. Also, it is not difficult to justify that, the vectors $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k \in \mathbb{R}^n$ are *affinely independent* if the unique solution to the simultaneous equations $\sum_{\xi=1}^{k} a_\xi \boldsymbol{x}^\xi = \boldsymbol{0}$ and $\sum_{\xi=1}^{k} a_\xi = 0$ is given by $a_\xi = 0$ for all $\xi = 1, \ldots, k$ [5]. Hence, linear independence implies affine independence but not vice versa.

2.2. Basic concepts from lattice theory

Computational concepts for neural networks based on lattice theory [2,21,29] are governed by the *bounded lattice ordered group* $(\mathbb{R}_{\pm\infty}, \vee, \wedge, +, +')$ or $\mathbb{R}_{\pm\infty}$*-blog*, where $\mathbb{R}$ denotes the set of real numbers, $\mathbb{R}_{\pm\infty} = \mathbb{R} \cup \{-\infty, \infty\}$ is the set of *extended real numbers*, $\vee$ and $\wedge$ denotes, respectively, the binary operations of maximum and minimum, and $+, +'$ denotes addition and its dual operation defined by

$$\begin{aligned} &x +' y = y + x \quad \forall x \in \mathbb{R},\ y \in \mathbb{R}_{\pm\infty}, \\ &\infty +' (-\infty) = \infty = (-\infty) +' \infty, \\ &\infty + (-\infty) = -\infty = (-\infty) + \infty. \end{aligned} \tag{3}$$

If $x \in \mathbb{R}_{\pm\infty}$, then its *additive conjugate* is given by $x^* = -x$. In a similar fashion, for a given vector $\boldsymbol{x} \in \mathbb{R}^n_{\pm\infty}$, its conjugate is defined by $\boldsymbol{x}^* = -\boldsymbol{x}^{\mathrm{T}}$, where T denotes transposition. Scalar addition in the $\mathbb{R}^n_{\pm\infty}$-blog, where $\mathbb{R}^n_{\pm\infty}$ denotes the n-fold Cartesian product of $\mathbb{R}_{\pm\infty}$, is defined component wise. That is, if $a \in \mathbb{R}_{\pm\infty}$ and $\boldsymbol{x} \in \mathbb{R}^n_{\pm\infty}$, then $a + \boldsymbol{x} = (a + x_1, \ldots, a + x_n)^{\mathrm{T}}$; the dual operation, $a +' \boldsymbol{x}$, is defined similarly. As our application domain concerns only with finite sets of real valued vectors, $X = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k\} \subset \mathbb{R}^n_{\pm\infty}$ for which $\boldsymbol{x}^\xi \in \mathbb{R}^n$ for each $\xi \in K$ where $K = \{1, \ldots, k\}$. With this restriction the operation of scalar addition is *self-dual* since $a +' \boldsymbol{x}^\xi = a + \boldsymbol{x}^\xi$ for any $a \in \mathbb{R}_{\pm\infty}$ and for all $\xi \in K$. Henceforth, we suppose that $X = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k\} \subset \mathbb{R}^n$.

A *linear minimax combination* of vectors from the set X is any vector $\boldsymbol{x} \in \mathbb{R}^n_{\pm\infty}$ of the form

$$\boldsymbol{x} = \mathfrak{S}(\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k) = \bigvee_{j \in J} \bigwedge_{\xi \in K} (a_{\xi j} + \boldsymbol{x}^\xi), \tag{4}$$

where J is a finite set of indices and $a_{\xi j} \in \mathbb{R}_{\pm\infty}$, $\forall j \in J$ and $\forall \xi \in K$. The expression $\mathfrak{S}(\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k)$ given by (4) is also called a *linear minimax sum*. A vector $\boldsymbol{x} \in \mathbb{R}^n$ is *lattice dependent* on X if and only if $\boldsymbol{x} = \mathfrak{S}(\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k)$ for some linear minimax sum of vectors from X. The vector $\boldsymbol{x}$ is said to be *lattice independent* (LI) of X if and only if it is not lattice dependent on X. The set X is said to be LI if and only if $\forall \lambda \in \{1, \ldots, k\}$, $\boldsymbol{x}^\lambda$ is LI of the reduced set X^λ defined as $X\backslash\{\boldsymbol{x}^\lambda\} = \{\boldsymbol{x}^\xi \in X : \xi \neq \lambda\}$ [22].

Given two $m \times n$ matrices $A = (a_{ij})$ and $B = (b_{ij})$ with entries from $\mathbb{R}_{\pm\infty}$, then the *pointwise maximum*, $A \vee B$, of A and B, is the