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

$m \times n$ matrix C defined by $A \vee B = C$, where $c_{ij} = a_{ij} \vee b_{ij}$. Similarly, the *pointwise minimum* of the same two matrices is defined as $A \wedge B = C$, where $c_{ij} = a_{ij} \wedge b_{ij}$. If A is $m \times p$ matrix and B is $p \times n$ matrix, then the *max product* of A and B is the matrix $C = A \boxtriangledown B$ whose i,j-th element, c_{ij} is given by Eq. (5). The *min product* of A and B is the matrix $C = A \boxtriangle B$ whose entries are computed following Eq. (6). For $i = 1, \ldots, n$ and $j = 1, \ldots, m$

$$c_{ij} = \bigvee_{k=1}^{p} (a_{ik} + b_{kj}), \tag{5}$$

$$c_{ij} = \bigwedge_{k=1}^{p} (a_{ik} +' b_{kj}). \tag{6}$$

These two matrix products are collectively referred to as *minimax products* [6,7]. A vector $\boldsymbol{x} \in \mathbb{R}^n_{\pm\infty}$ is called a *max fixed point of A* if $A \boxtriangledown \boldsymbol{x} = \boldsymbol{x}$ and a *min fixed point of A* if $A \boxtriangle \boldsymbol{x} = \boldsymbol{x}$. An $n \times n$ matrix A is said to be *diagonally max dominant* if and only if, for each $j \in \{1, \ldots, n\}$, it satisfies condition (7) for $i = 1, \ldots, n$. Dually, A is said to be *diagonally min dominant* if and only if, for each $j \in \{1, \ldots, n\}$, Eq. (8) is verified for $i = 1, \ldots, n$:

$$a_{jj} - a_{ij} = \bigvee_{k=1}^{n} (a_{jk} - a_{ik}), \tag{7}$$

$$a_{jj} - a_{ij} = \bigwedge_{k=1}^{n} (a_{jk} - a_{ik}). \tag{8}$$

3. Lattice auto-associative memories

3.1. Fundamental properties

Suppose $X = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k\} \subset \mathbb{R}^n$ and $Y = \{\boldsymbol{y}^1, \ldots, \boldsymbol{y}^k\} \subset \mathbb{R}^m$ are two finite sets of pattern vectors with desired association given by the diagonal $\{(\boldsymbol{x}^\xi, \boldsymbol{y}^\xi) : \xi \in K\}$ of $X \times Y$ where $K = \{1, \ldots, k\}$. The goal is to store these pattern pairs in some memory $\mathfrak{M}$ such that for $\xi \in K$, $\mathfrak{M}$ recalls $\boldsymbol{y}^\xi$ when presented with the pattern $\boldsymbol{x}^\xi$. With each pair of pattern associations (X, Y) we define two canonical lattice based associative memories [23,24], the *min memory* W_{XY} and the *max memory* M_{XY}, both of size $m \times n$, whose elements w_{ij} and m_{ij}, for $i = 1, \ldots, m$ and $j = 1, \ldots, n$, are given by

$$w_{ij} = \bigwedge_{\xi=1}^{k} (y_i^\xi - x_j^\xi), \quad m_{ij} = \bigvee_{\xi=1}^{k} (y_i^\xi - x_j^\xi). \tag{9}$$

The memories are called *lattice auto-associative memories* (LAMs) whenever $Y = X$. Observe that the diagonals of the matrices W_{XX} and M_{XX} consist entirely of zeros, i.e. $w_{ii} = m_{ii} = 0$ for all $i \in \{1, \ldots, n\}$. Also, when speaking of fixed points of the matrices W_{XX} and M_{XX}, we always mean a fixed point of W_{XX} with respect to the operation $\boxtriangledown$ and of M_{XX} with respect to the operation $\boxtriangle$. The next lemma is a fundamental tool to prove several properties of LAMs [22].

Lemma 3.1. *Given $X = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k\} \subset \mathbb{R}^n$, let w_{ij} (resp. m_{ij}) denote the i,j-entry of the min auto-associative memory W_{XX} (resp. max auto-associative memory M_{XX}). If $i,j,\ell \in \{1, \ldots, n\}$, then $w_{ij} + w_{j\ell} \leqslant w_{i\ell}$ (resp. $m_{ij} + m_{j\ell} \geqslant m_{i\ell}$).*

Proof. We give an argument only for the min memory W since a similar proof, for the max memory M, is immediate by using the right expression in Eq. (9) and changing the sense of all inequalities. From the first expression of Eq. (9), entries w_{ij} and $w_{j\ell}$ satisfy, respectively, the following inequalities for all $\gamma = 1, \ldots, k$:

$$w_{ij} = \bigwedge_{\xi=1}^{k} (x_i^\xi - x_j^\xi) \leqslant x_i^\gamma - x_j^\gamma, \tag{10}$$

$$w_{j\ell} = \bigvee_{\xi=1}^{k} (x_j^\xi - x_\ell^\xi) \leqslant x_j^\gamma - x_\ell^\gamma. \tag{11}$$

Therefore, for all $\gamma \in \{1, \ldots, k\}$,

$$w_{ij} + w_{j\ell} \leqslant (x_i^\gamma - x_j^\gamma) + (x_j^\gamma - x_\ell^\gamma) = x_i^\gamma - x_\ell^\gamma. \tag{12}$$

Hence, $w_{ij} + w_{j\ell} \leqslant \bigwedge_{\xi=1}^{k} (x_i^\gamma - x_\ell^\gamma) = w_{i\ell}$. □

Lemma 3.2. *Given $X = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k\} \subset \mathbb{R}^n$, let w_{ij} (resp. m_{ij}) denote the i,j-entry of the min auto-associative memory W_{XX} (resp. max auto-associative memory M_{XX}). Then,*

$$w_{ij} = \bigwedge_{\ell=1}^{n} (w_{i\ell} - w_{j\ell}), \quad m_{ij} = \bigvee_{\ell=1}^{n} (m_{i\ell} - m_{j\ell}). \tag{13}$$

Proof. From Lemma 3.1, it follows for fixed values of $i,j \in \{1, \ldots, n\}$ that $w_{ij} \leqslant w_{i\ell} - w_{j\ell}$ for all $\ell = 1, \ldots, n$. Therefore,

$$w_{ij} \leqslant \bigwedge_{\ell=1}^{n} (w_{i\ell} - w_{j\ell}). \tag{14}$$

If $w_{ij} < \bigwedge_{\ell=1}^{n} (w_{i\ell} - w_{j\ell})$, then $w_{ij} < w_{i\ell} - w_{j\ell}$ for any $\ell \in \{1, \ldots, n\}$; let $\ell = j$, so $w_{ij} < w_{ij} - w_{jj} = w_{ij}$ which is a contradiction. Whence our lemma follows. The right equality in Eq. (13) is proven in a similar manner. □

Let $X = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k\} \subset \mathbb{R}^n$ be a finite pattern set and let $\boldsymbol{x} \in \mathbb{R}^n$, then the following statements, proven in [22], characterize the algebraic behavior of LAMs:

P1. $W_{XX} \boxtriangledown \boldsymbol{x}^\xi = \boldsymbol{x}^\xi = M_{XX} \boxtriangle \boldsymbol{x}^\xi$, $\forall \xi \in K$.
P2. $W_{XX} \boxtriangledown \boldsymbol{x} = \boldsymbol{x}$ if and only if $M_{XX} \boxtriangle \boldsymbol{x} = \boldsymbol{x}$.
P3. $\boldsymbol{x}$ is a fixed point of W_{XX} if and only if $\boldsymbol{x}$ is lattice dependent on X.
P4. W_{XX} is diagonally max dominant; M_{XX} is diagonally min dominant.

According to P1, both W_{XX} and M_{XX} are *perfect recall* memories for uncorrupted input. Also, property P2 says that W_{XX} and M_{XX} share the same set of fixed points [30,31], defined as

$$F(X) = \{\boldsymbol{x} : W_{XX} \boxtriangledown \boldsymbol{x} = \boldsymbol{x} = M_{XX} \boxtriangle \boldsymbol{x}\}. \tag{15}$$

Statement P3 provides an algebraic classification of the set $F(X)$ in terms of linear minimax combinations. Closely related to the notion of diagonally max or min dominance mentioned in proposition P4 is the notion of max or min dominance of a set of vectors.

The set $X = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k\} \subset \mathbb{R}^n$ is said to be *max dominant* if and only if for every $\lambda \in K$ there exists an index $j_\lambda \in \{1, \ldots, n\}$ such that Eq. (16), or equivalently, Eq. (17) is satisfied for $i = 1, \ldots, n$

$$x_{j_\lambda}^\lambda - x_i^\lambda = \bigvee_{\xi=1}^{k} (x_{j_\lambda}^\xi - x_i^\xi), \tag{16}$$

$$x_{j_\lambda}^\lambda - x_i^\lambda \geqslant x_{j_\lambda}^\xi - x_i^\xi \quad \forall \xi \in K. \tag{17}$$

Similarly, X is said to be *min dominant* if and only if for every $\lambda \in K$ there exists an index $j_\lambda \in \{1, \ldots, n\}$ such that Eq. (18), or equivalently Eq. (19) is satisfied for $i = 1, \ldots, n$

$$x_{j_\lambda}^\lambda - x_i^\lambda = \bigwedge_{\xi=1}^{k} (x_{j_\lambda}^\xi - x_i^\xi), \tag{18}$$

$$x_{j_\lambda}^\lambda - x_i^\lambda \leqslant x_{j_\lambda}^\xi - x_i^\xi \quad \forall \xi \in K. \tag{19}$$

Assume that X is a max or min dominant set and let $X' = \{\boldsymbol{x}^\lambda | \lambda \in J \subset K\}$. Since Eqs. (16)–(19) remain valid for all $j_\lambda, i \in \{1,\ldots,n\}$ for $\lambda, \xi \in J$, then any proper subset X' of X is also max or min dominant. The notion of max or min dominance is the key to the concept of strong lattice independence. A set of LI vectors $X = \{\boldsymbol{x}^1,\ldots,\boldsymbol{x}^k\} \subset \mathbb{R}^n$ is said to be *strong LI* (SLI) if and only if X is max dominant or min dominant (or both). It is important to emphasize that for a set X to be strongly LI it needs to satisfy two properties, namely lattice independence *and* either Eq. (17) or Eq. (19) or both [22,26].

Henceforth, the set of vectors consisting of the columns of the min memory matrix W_{XX} will be denoted by W and the set of vectors consisting of the columns of the max memory matrix M_{XX} will be denoted by M. In particular, $\mathbf{w}^j \in W$ corresponds to the j-th column vector of W_{XX} and its i-th component is represented by $w_i^j = w_{ij}$. Similarly, $\mathbf{m}^j \in M$ corresponds to the j-th column vector of M_{XX} with entries denoted by $m_i^j = m_{ij}$. It follows from property P4 that the sets W and M are max and min dominant, respectively. However, as shown in the following numerical example, W and M need not be LI and, hence, *not* strongly LI.

Example 3.1. Suppose $X = \{\boldsymbol{x}^1, \boldsymbol{x}^2, \boldsymbol{x}^3\}$, where $\boldsymbol{x}^1 = (-1,0,1)^{\mathrm{T}}$, $\boldsymbol{x}^2 = (1,2,3)^{\mathrm{T}}$, and $\boldsymbol{x}^3 = (3,4,5)^{\mathrm{T}}$. Then

$$W = W_{XX} = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{pmatrix} = (\mathbf{w}^1, \mathbf{w}^2, \mathbf{w}^3).$$

However, $\mathbf{w}^1 = 1 + \mathbf{w}^2$, $\mathbf{w}^1 = 2 + \mathbf{w}^3$, and $\mathbf{w}^2 = 1 + \mathbf{w}^3$. Therefore W is not LI and neither is M since $M_{XX} = W_{XX}$. If $W^1 = (\mathbf{w}^2, \mathbf{w}^3)$, $W^2 = (\mathbf{w}^1, \mathbf{w}^3)$, and $W^3 = (\mathbf{w}^1, \mathbf{w}^2)$, then $W_{W^1W^1} = W_{W^2W^2} = W_{W^3W^3} = W$. Besides, it is easy to verify that, $W_{W^\lambda W^\lambda} \boxtimes \mathbf{w}^\lambda = \mathbf{w}^\lambda$ for $\lambda = 1,2,3$. By construction, $\boldsymbol{x}^2 = 2 + \boldsymbol{x}^1$ and $\boldsymbol{x}^3 = 4 + \boldsymbol{x}^1$, hence X itself is lattice dependent.

3.2. Computational procedures

In fact, properties P2 and P3 provide the basis for a simple computational procedure used to test lattice independence in a vector set. A working algorithm described in mathematical pseudocode is given below with key steps (S) consecutively numbered for reference. The pattern set $X \subset \mathbb{R}^n$ is given as a matrix of size $n \times k$, where n represents the number of rows (dimensionality) and k represents the number of patterns; the output is a binary variable $b \in \{0,1\}$, if $b = 1$ then set X is LI otherwise it is not.

Algorithm 1 (*Lattice independence set test*).

[Initialize counter]
S1 $\sigma \leftarrow 0$
[Scan all vectors in X]
S2 **for** $\lambda = 1$ **to** k
  [assume current vector is lattice dependent]
  $b_\lambda \leftarrow 0$
  [build associative memory from X^λ]
  $A \leftarrow W_{X^\lambda X^\lambda}$
  [compute output vector]
  $\boldsymbol{x} \leftarrow A \boxtimes \boldsymbol{x}^\lambda$
  [output vector is a fixed point?]
  $b_\lambda \leftarrow 1$ **if** $\boldsymbol{x} \neq \boldsymbol{x}^\lambda$
  [add flag value of tested vector]
  $\sigma \leftarrow \sigma + b_\lambda$
[Return test result]
S3 $b \leftarrow (\sigma = k)$

Notice that the value assigned to A, within step S2, can be replaced with the max-auto-associative memory $M_{X^\lambda X^\lambda}$ and correspondingly, instead of using the max product, the output vector $\boldsymbol{x}$ must be computed with the min product of A and $\boldsymbol{x}^\lambda$, i.e., $\boldsymbol{x} = A \boxtimes \boldsymbol{x}^\lambda$. Alternative algorithms have also been developed based on criteria derived from the definition of a LI set [32].

To obtain from a given vector set X, a reduced set X' that is LI we present next a procedure, based on Algorithm 1, that uses a simple selection and elimination mechanism.

Algorithm 2 (*Lattice independence set generation*).

[Select vectors in X sequentially]
S1 **for** $\lambda = 1$ **to** k
  [compute reduced set]
  $X^\lambda \leftarrow X \backslash \{\boldsymbol{x}^\lambda\}$
  [compute output vector]
  $\boldsymbol{x} \leftarrow W_{X^\lambda X^\lambda} \boxtimes \boldsymbol{x}^\lambda$
  [delete vector if output is a fixed point]
  $X \leftarrow X^\lambda$ **if** $\boldsymbol{x} = \boldsymbol{x}^\lambda$
[Return lattice independent subset X']
S2 X

In the second instruction of the main cycle S1, the max-memory can be used instead of the min-memory W. In that case, the output vector $\boldsymbol{x}$ should be computed with the min-product of $M_{X^\lambda X^\lambda}$ and $\boldsymbol{x}^\lambda$, i.e., $\boldsymbol{x} = M_{X^\lambda X^\lambda} \boxtimes \boldsymbol{x}^\lambda$. A more general scheme would use a random selection mechanism to probe vectors in X for lattice independence but additional steps are required to keep a register of the column indices that have been tested. Also, due to randomization, the resulting subset X' could be different each time the procedure is executed.

4. Endmember determination using LAMs

4.1. Theoretical foundation

The linear unmixing model described by Eqs. (1)–(2), assumes that the endmembers are affinely independent and the importance of strong lattice independence in the determination of endmembers is due to the following conjecture.

Conjecture 4.1. *If $X = \{\boldsymbol{x}^1,\ldots,\boldsymbol{x}^k\} \subset \mathbb{R}^n$ is strongly LI, then X is affinely independent.*

This conjecture was stated as a theorem in [22], nevertheless a gap has been recently found in its proof and the result awaits further scrutiny to establish its validity. Also, since we have not been able to disprove it, we can use Conjecture 4.1 as a working platform for the techniques developed here. As mentioned in Example 3.1, the sets of column vectors W and M derived from W_{XX} and M_{XX} need not be strongly LI and, therefore, not affinely independent. There is, however, an easy method to reduce these sets to affinely independent sets. This method is derived from the constructive proof of the next theorem. In the proof we assume the following fact.

Lemma 4.1. *Let $X = \{\boldsymbol{x}^1,\ldots,\boldsymbol{x}^k\} \subset \mathbb{R}^n$, then $W_{WW} = W_{XX}$ and $M_{MM} = M_{XX}$.*

Proof. From Eq. (9), the ij-element of the memory matrix W_{XX} is $w_{ij} = \bigwedge_{\xi=1}^k (x_i^\xi - x_j^\xi)$ for all $i,j = 1,\ldots,n$. Similarly, the ij-element of the memory matrix W_{WW} is defined as $\psi_{ij} = \bigwedge_{\ell=1}^n (w_i^\ell - w_j^\ell)$, where the upper index has changed to n because the $W = W_{XX}$ matrix has size $n \times n$. Since $\bigwedge_{\ell=1}^n (w_i^\ell - w_j^\ell)$ can be written as $\bigwedge_{\ell=1}^n (w_{i\ell} - w_{j\ell})$, it follows from Lemma 3.2, that $\psi_{ij} = w_{ij}$ for all $i,j \in \{1,\ldots,n\}$. Hence, $W_{WW} = W_{XX}$. Using the definition for the

max memory, the same argument can be given to show that $M_{MM} = M_{XX}$. □

Theorem 4.1. *There exist sets of vectors $V \subseteq W$ and $N \subseteq M$ such that V and N are strongly LI and $F(V) = F(N) = F(X)$ or, equivalently, $W_{VV} = W_{XX}$ and $M_{NN} = M_{XX}$.*

Proof. We only prove the existence of the set V as the existence of the set N can be proven in an analogous fashion. The set of vectors W is max dominant and any subset V of W is, therefore, also max dominant. Thus, all we need to show is the existence of a LI set V such that $W_{VV} = W_{XX}$. Let $W = \{\mathbf{w}^1, \ldots, \mathbf{w}^n\}$ and set $W_1 = W \backslash \{\mathbf{w}^1\}$. If $W_{W_1 W_1} = W_{XX}$, set $V_1 = W_1$, otherwise set $V_1 = W$. Hence, if $V_1 = W_1$, then $\mathbf{w}^1 \notin V_1$, and if $V_1 = W$, then $\mathbf{w}^1 \in V_1$. In other words, $\mathbf{w}^1 \notin V_1$ whenever $\mathbf{w}^1$ is lattice dependent, and $\mathbf{w}^1 \in V_1$ whenever $\mathbf{w}^1$ is LI. In either case, we have $W_{V_1 V_1} = W_{XX}$. Next, set $V_2 = V_1 \backslash \{\mathbf{w}^2\}$ if $W_{V_2 V_2} = W_{XX}$, otherwise set $V_2 = V_1$. Again, in either case we will have $W_{V_2 V_2} = W_{XX}$. Continue in this manner until $V_n = V_{n-1} \backslash \{\mathbf{w}^n\}$ if $W_{V_n V_n} = W_{XX}$, otherwise set $V_n = V_{n-1}$. Once W has been exhausted, set $V = V_n$. By construction, $V \subseteq W$ is a LI subset for which $W_{VV} = W_{XX}$. Besides, V is also max dominant, therefore V is strongly LI. □

The computational algorithm derived from the above proof is shown next. The input set $X \subset \mathbb{R}^n$ is given as a matrix of size $n \times k$ where k is the number of patterns, each of dimension n, and the output is a strongly LI subset V of W. Notice that, after step S3 is executed, V is a matrix of size $n \times n$, hence the main cycle in S4 goes from column 1 to n. To abbreviate instruction comments, we use the word *column* instead of "column index" or "current column index";similarly, we use the word *displacement* instead of "column displacement counter".

Algorithm 3 (*Strong lattice ind. set generation*).

```
[Compute associative memory from input set X]
S1 A ← W_XX
[Initialize column and displacement]
S2 λ ← 1;c ← 1
[Assume all vectors in V = W_XX are SLI]
S3 V ← A
[Scan all columns in V]
S4 for ξ = 1 to n
   [remove column]
   X' ← V\{v^λ}
   [build new memory from X']
   B ← W_X'X'
   [matches original memory W_XX?]
   if B = A
      [compute reduced set]
      then V ← X'
      [column to delete is last displacement]
      λ ← c
      [next column to bedeleted]
      else λ ← λ + 1
      [update displacement]
      c ← λ
[Return strong lattice independent subset]
S5 V
```

In steps S1 and S4 of Algorithm 3, the assignment to matrix variables A and B can be replaced, respectively, with the max-auto-associative memories M_{XX} and $M_{X'X'}$. A technique to test or generate SLI sets based on definitions has been described elsewhere [32]. Recently, a complementary discussion of lattice independence with proofs presented in array oriented programming languages, such as APL and J, appears in [28].

Using Lemma 3.1 with $j = \lambda$ and $\ell = \xi$, it follows that $w_{i\lambda} + w_{\lambda\xi} \leqslant w_{i\xi}$, or equivalently, $-w_{i\lambda} \geqslant w_{\lambda\xi} - w_{i\xi}$. Since $w_{\lambda\lambda} = 0$, then $w_{\lambda\lambda} - w_{i\lambda} \geqslant w_{\lambda\xi} - w_{i\xi}$. Therefore, taking the *row* index λ as any $i_\lambda \in \{1, \ldots, n\}$, we have that $w^\lambda_{i_\lambda} - w^\lambda_i = \bigvee_{\xi=1}^n (w^\xi_{i_\lambda} - w^\xi_i)$ for $i = 1, \ldots, n$. It turns out that the column vectors of W_{XX} form a set that is diagonally max-dominant. A dual argument can be given to establish that the columns of M_{XX} form a vector set that is diagonally min-dominant. The previous discussion makes clear why Algorithm 3 is concerned only with lattice independence since minimax dominance is inherent to LAMs.

Using the method outlined in the proof of 4.1 and embodied in Algorithm 3, we are able to find a set of affinely independent points. However, the relationship of this set of affinely independent points to the set X is not directly obvious. To obtain affinely independent points that are related to the data set points, we have to do the following *additive scaling*. Let $\mathbf{v} = \bigwedge_{\xi=1}^k \boldsymbol{x}^\xi, \mathbf{u} = \bigvee_{\xi=1}^k \boldsymbol{x}^\xi$ denote, respectively, the *minimum* and *maximum vector bounds* of X whose entries are defined for all $i = 1, \ldots, n$ by

$$v_i = \bigwedge_{\xi=1}^k x_i^\xi, \quad u_i = \bigvee_{\xi=1}^k x_i^\xi. \tag{20}$$

Define two new matrices $\overline{M}$ and $\overline{W}$ by setting, for all $i = 1, \ldots, n$

$$\overline{\mathbf{m}}^i = v_i + \mathbf{m}^i, \quad \overline{\mathbf{w}}^i = u_i + \mathbf{w}^i. \tag{21}$$

Lemma 4.2. *Given $X = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k\} \subset \mathbb{R}^n$, let $\overline{w}_{ij}$ (resp. $\overline{m}_{ij}$) denote the i,j-entry of matrix $\overline{W}$ (resp. $\overline{M}$). If $i,j,\ell \in \{1, \ldots, n\}$, then $\overline{w}_{ij} + \overline{w}_{j\ell} \leqslant u_j + \overline{w}_{i\ell}$ (resp. $\overline{m}_{ij} + \overline{m}_{j\ell} \geqslant v_j + \overline{m}_{i\ell}$).*

Proof. From the second expression in Eq. (21), we have that, $(u_j + w_{ij}) + (u_\ell + w_{j\ell}) \leqslant u_j + (u_\ell + w_{i\ell})$, simplifies to $w_{ij} + w_{j\ell} \leqslant w_{i\ell}$, which follows immediately from Lemma 3.1. In a similar way, $(v_j + m_{ij}) + (v_\ell + m_{j\ell}) \geqslant v_j + (v_\ell + m_{i\ell})$ is obtained after substitution using the first part of Eq. (21), and reduces to $m_{ij} + m_{j\ell} \geqslant m_{i\ell}$ (dual statement of Lemma 3.1). □

Using Lemma 4.2 with $j = \lambda$ and $\ell = \xi$, it follows that $\overline{w}_{i\lambda} + \overline{w}_{\lambda\xi} \leqslant u_\lambda + \overline{w}_{i\xi}$, or equivalently, $-\overline{w}_{i\lambda} \geqslant \overline{w}_{\lambda\xi} - \overline{w}_{i\xi} - u_\lambda$. Since $\overline{w}_{\lambda\lambda} = u_\lambda$, then $\overline{w}_{\lambda\lambda} - \overline{w}_{i\lambda} - u_\lambda \geqslant \overline{w}_{\lambda\xi} - \overline{w}_{i\xi} - u_\lambda$. Therefore, taking the *row* index λ as any $i_\lambda \in \{1, \ldots, n\}$, we have that $\overline{w}^\lambda_{i_\lambda} - \overline{w}^\lambda_i = \bigvee_{\xi=1}^n (\overline{w}^\xi_{i_\lambda} - \overline{w}^\xi_i)$ for $i = 1, \ldots, n$. It turns out that the column vectors of $\overline{W}$ form a set that is diagonally max-dominant. A dual argument can be given to establish that the columns of $\overline{M}$ form a vector set that is diagonally min-dominant. In other words, max and min dominance is an invariant property of the LAMs W_{XX} and M_{XX} with respect to the additive scaling defined by Eq. (21). An analogous result for LI sets was proven in [22]. Therefore, it is not difficult to see that $\overline{\mathbf{w}}^i \leqslant \mathbf{u}$ and $\overline{\mathbf{m}}^i \geqslant \mathbf{v}$ for all $i = 1, \ldots, n$; observe that, the maxmin vector bounds, $\mathbf{u}$ and $\mathbf{v}$, are equal to the main diagonals of $\overline{W}$ and $\overline{M}$, respectively. Geometrically, it means that $\mathbf{u}$ is the *max-envelope* of all columns of $\overline{W}$ and $\mathbf{v}$ is the *min-envelope* of all columns of $\overline{M}$.

It is important to remark that the sets $\overline{M}$ and $\overline{W}$ contain, respectively, n minimum and n maximum points, where the couple of points $\{\overline{\mathbf{m}}^i, \overline{\mathbf{w}}^i\}$ occur *along* the ith coordinate. Furthermore, we have

Lemma 4.3. *The set of points, $\mathscr{P} = \overline{M} \cup \overline{W} \cup \{\mathbf{v}, \mathbf{u}\}$, forms a convex polytope $\mathfrak{P}$ with $2(n+1)$ vertices that contains X.*

Proof. Since $\{\overline{\mathbf{m}}^1, \overline{\mathbf{w}}^1, \ldots, \overline{\mathbf{m}}^n, \overline{\mathbf{w}}^n, \mathbf{v}, \mathbf{u}\}$ is a finite set of points, the *convex hull* generated with these $2n+2$ points is the desired polytope, i.e., $\mathfrak{P} = C(\mathscr{P})$. From Eq. (20), it follows that $v_i \leqslant x_i^\xi$ and $u_i \geqslant x_i^\xi$ for any ξ and all i, hence, $\mathbf{v} \leqslant \mathbf{x}^\xi \leqslant \mathbf{u}$ for all $\xi \in K$. Take an arbitrary entry of pattern $\mathbf{x}^\xi \in X$, say x^ξ_λ, then, for some $i \in \{1, \ldots, n\}$ we have that, $x^\xi_\lambda - x^\xi_i \leqslant \bigvee_{\xi=1}^k (x^\xi_\lambda - x^\xi_i)$. Equivalently, $u_i + (x^\xi_\lambda - x^\xi_i) \leqslant u_i + w^i_\lambda = \overline{w}^i_\lambda$, or, $x^\xi_\lambda + (u_i - x^\xi_i) = x^\xi_\lambda + \varepsilon = \overline{w}^i_\lambda$ where $\varepsilon \geqslant 0$ (see Eq. (21)). Therefore, $x^\xi_\lambda \leqslant \overline{w}^i_\lambda$. Similarly, $x^\xi_\lambda - x^\xi_i \geqslant \bigwedge_{\xi=1}^k (x^\xi_\lambda - x^\xi_i)$ and $v_i + (x^\xi_\lambda - x^\xi_i) \geqslant v_i + m^i_\lambda = \overline{m}^i_\lambda$, or, $x^\xi_\lambda + (v_i - x^\xi_i) = x^\xi_\lambda + \delta = \overline{m}^i_\lambda$ where $\delta \leqslant 0$ (see Eq. (21)). Therefore, $x^\xi_\lambda \geqslant \overline{m}^i_\lambda$ for some

$i \in \{1,\ldots,n\}$. Thus, along each coordinate, the λ-entry of any $\mathbf{x}^\xi \in X$ is bounded from above (resp. below) by a λ entry of some $\overline{\mathbf{w}}^i \in \overline{W}$ (resp. $\overline{\mathbf{m}}^i \in \overline{M}$). Hence, $X \subseteq \mathfrak{P}$. □

Example 4.1. Consider a finite set $X \subset \mathbb{R}^2$ whose *convex hull* $C(X)$ is spanned by the points of $X' \subset X$, where

$$X' = (\boldsymbol{x}^1,\ldots,\boldsymbol{x}^6) = \begin{pmatrix} 2.5 & 2 & 2.5 & 4 & 5 & 4.5 \\ 3.5 & 2 & 1 & 2 & 4 & 5 \end{pmatrix}.$$

It follows from statements P1 and P2, that W_{XX} and M_{XX} are completely determined by X'. In this case,

$$W_{XX} = \begin{pmatrix} 0 & -1 \\ -2 & 0 \end{pmatrix} \quad \text{and} \quad M_{XX} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}.$$

Clearly, each set $W = \{(0,-2)^{\mathrm{T}}, (-1,0)^{\mathrm{T}}\}$ and $M = \{(0,1)^{\mathrm{T}}, (2,0)^{\mathrm{T}}\}$ is strongly LI and hence affinely independent but neither set contains points of $C(X)$. However, from Eq. (20), $\mathbf{v} = (2,1)^{\mathrm{T}}$ and $\mathbf{u} = (5,5)^{\mathrm{T}}$. Then, using Eq. (21) for additive scaling, we obtain

$$\overline{M} = \begin{pmatrix} 2 & 3 \\ 3 & 1 \end{pmatrix} \quad \text{and} \quad \overline{W} = \begin{pmatrix} 5 & 4 \\ 3 & 5 \end{pmatrix},$$

where $\overline{M} = (\overline{\mathbf{m}}^1, \overline{\mathbf{m}}^2)$ and $\overline{W} = (\overline{\mathbf{w}}^1, \overline{\mathbf{w}}^2)$. The inequalities, $\overline{\mathbf{m}}^i \geqslant \mathbf{v}$ and $\overline{\mathbf{w}}^i \leqslant \mathbf{u}$ for $i = 1,2$, are readily confirmed.

Fig. 1 indicates the physical relationship between the vertices of the convex polygon determined by $\mathfrak{P} = \{\overline{\mathbf{m}}^1, \overline{\mathbf{m}}^2, \overline{\mathbf{w}}^1, \overline{\mathbf{w}}^2, \mathbf{v}, \mathbf{u}\}$ and the set X representing the data. Observe that $\{\overline{\mathbf{m}}^i, \overline{\mathbf{w}}^i\}$ are extreme points, minimum and maximum, along both coordinates. The intercept points below and above the origin on the x_2 axis ($x_1 = 0$), drawn with small circles, correspond to the first column of W_{XX} and M_{XX}, respectively; similarly, the intercept points to the left and right of the origin on the x_1 axis ($x_2 = 0$), represent the second column of W and M_{XX}. To obtain a large simplex containing most of the data in X, one can use the points of $\overline{W}$ and use $\mathbf{v}$ as the dark point. However, other simplexes can also be formed that may prove useful in image segmentation and object identification. In the pictorial sample, triangles such as $(\overline{\mathbf{w}}^1, \overline{\mathbf{m}}^2, \overline{\mathbf{w}}^2)$ or $(\overline{\mathbf{m}}^1, \overline{\mathbf{w}}^2, \overline{\mathbf{m}}^2)$ could also be used. Although visualization is possible only for points in $\mathbb{R}^2$ and $\mathbb{R}^3$, we remark that the geometrical description and interpretation of the picture shown in Fig. 1 generalizes to any dimension.

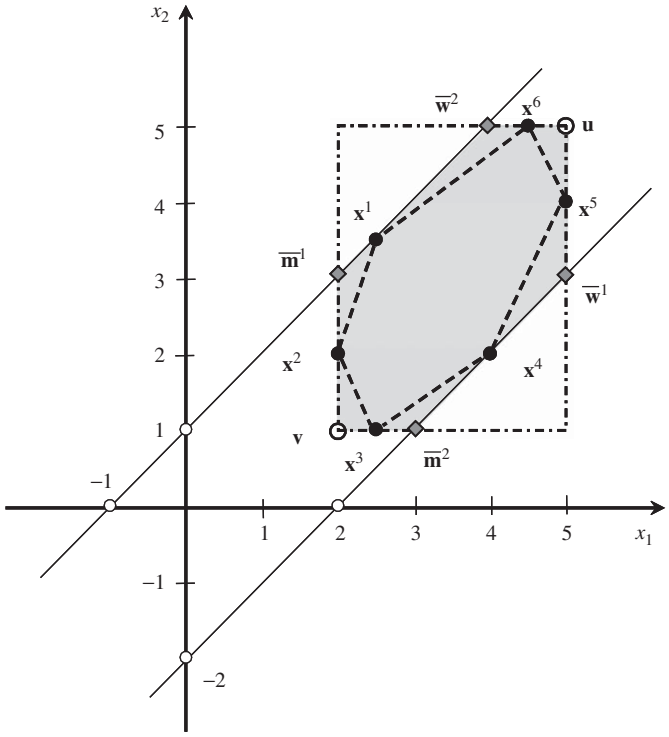


Fig. 1. The fixed point set $F(X)$ is the infinite strip bounded by the two lines of slope 1. The intersection of $F(X)$ with the box determined by $\mathbf{v}$ and $\mathbf{u}$ equals $\mathfrak{P}$ (shaded area).

4.2. Application example

In our approach to endmember determination we consider the set X to be a subset of pixels obtained from the hyperspectral image cube. This cube may have been significantly reduced in spectral dimensionality by applications of a chosen technique such as principal component analysis, minimum noise fraction transform, or adjacent band removal of highly correlated bands [10,17]. Such reductions are often necessary to eliminate undesirable effects produced during data acquisition as well as to diminish computational requirements.

For example, the *maximal* storage space required for a single hyperspectral image scene, of size $614 \times 512 \times 224$ (pixels, lines and bands), captured by NASA-JPL's Airborne Visible and Infrared Imaging Spectrometer (AVIRIS) is 134.3125 MB [37]. More specifically, we consider for our application example, a subcube corresponding to scene no. 4 of the 1997 AVIRIS flight over the Cuprite mining site in Nevada, displayed in Fig. 2. With respect to the image format used by AVIRIS, pixels 81–614 and bands 169–220, covering the wavelength range from 1.95 to 2.47 μm, where selected to compare our results with the USGS (United States Geographical Survey) Cuprite 1995 mineral distribution made by Clark and Swayze. Thus, the hyperspectral image subcube is of size $534 \times 512 \times 52$ and the data set X has 273,408 vectors of dimension 52.

The first step is to form the memories W_{XX} and $M_{XX} = W^*_{XX} = -W^{\mathrm{T}}_{XX}$. We then form the sets of vectors W and M from the columns of W_{XX} and M_{XX}, respectively. Using the vectors $\mathbf{v}$ and $\mathbf{u}$ computed with Eq. (20), the elements of W and M are

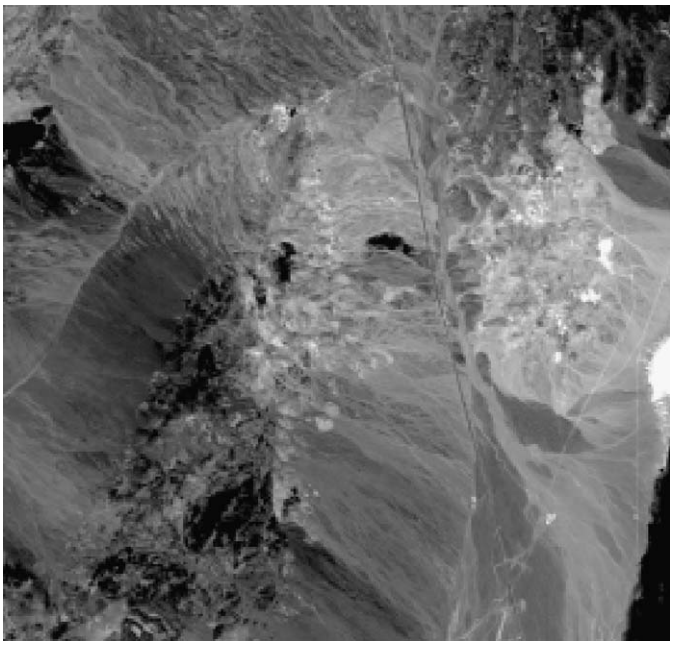


Fig. 2. Scene #4 of the Cuprite Nevada site. Enhanced grayscale image obtained from an RGB pseudocolor composition of bands 207 (red), 123 (green), and 36 (blue).

then scaled in order to obtain $\overline{W}$ and $\overline{M}$. It is these sets that can then be reduced to affinely independent sets using the procedure outlined in Algorithm 3. It turns out that $\overline{W}$ and $\overline{M}$ (of size 52×52) are strongly LI sets and by Conjecture 4.1 they are affinely independent. Therefore, each memory matrix provides us with 52 "candidate" endmembers. Since contiguous columns are highly correlated, most of these potential endmembers can be discarded using appropriate techniques. For example, minimum mutual information has been used to obtain a final set of endmembers [13]; here, pattern elimination is based on minimal Chebyshev distance or angle between pairs of vectors (see Appendix for details). In addition, the scaling based on Eq. (21), generates an "upward spike" in endmembers selected from $\overline{W}$ since $\overline{w}_{ii} = u_i$, or a "downward spike" if endmembers come from $\overline{M}$ because $\overline{m}_{ii} = v_i$. The anomalous spikes can be smoothed so that, the global shape of each final endmember agrees with available library reference spectra. A simple smoothing procedure considers the nearest one or two spectral samples next to $\overline{w}_{ii}$ or $\overline{m}_{ii}$ [20]. It is given, for any $i \in \{1, \ldots, n\}$, by

$$e_{ii} = \begin{cases} e_{2,1} & \Leftrightarrow \ i = 1, \\ \frac{1}{2}(e_{i-1,i} + e_{i+1,i}) & \Leftrightarrow \ 1 < i < n, \\ e_{n-1,n} & \Leftrightarrow \ i = n, \end{cases} \tag{22}$$

where the endmember vector $\boldsymbol{e}$ can be a selected column vector $\overline{\mathbf{w}}$ or $\overline{\mathbf{m}}$.

Fig. 3 displays five final endmembers extracted from scene no. 4 of Cuprite. Note that a correspondence between endmembers and laboratory spectra cannot be exact, since some wavelength values over the entire spectral range are different between airborne imaging spectrometers and ground instrumentation. Thus, small "shifts" of the absorption bands characteristic of each mineral may be noticed in endmember curves. The approximate mineral tags have the following endmember column correspondence: $\overline{\mathbf{w}}^{33}$ is alunite, $\overline{\mathbf{m}}^{16}$ is buddingtonite, $\overline{\mathbf{w}}^{25}$ is calcite, $\overline{\mathbf{w}}^{2}$ is kaolinite, and $\overline{\mathbf{m}}^{27}$ is muscovite. In comparison, Fig. 4 shows the subsampled spectra in the range 2.0–2.5 μm of similar minerals obtained from the USGS spectral library (splib06a, 2007) corresponding to: Alunite-al706, Buddingtonite-gds85, Calcite-hs483b, Kaolinite-kga2, and Muscovite-gds108. We remark that spectral characteristics are common to a family of minerals that

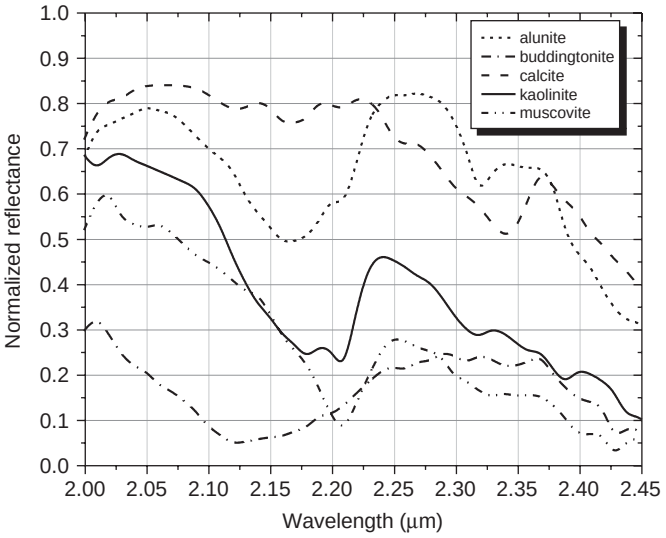


Fig. 3. Final endmembers determined from sets $\overline{W}$ and $\overline{M}$, computed with data of the Cuprite site (scene 4).

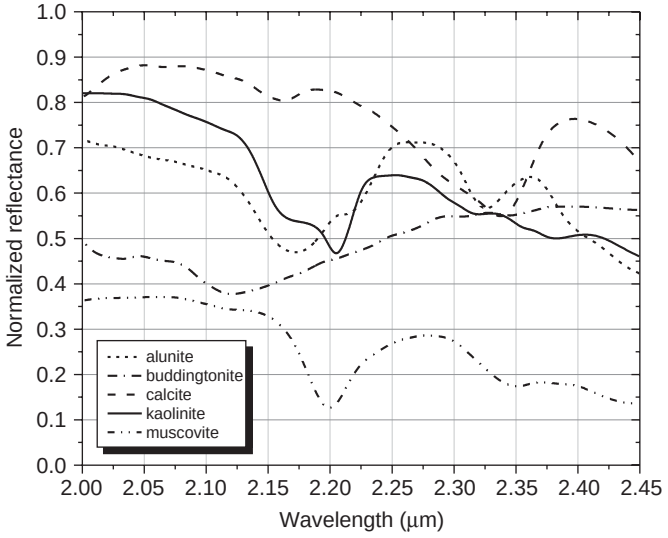


Fig. 4. Mineral reference spectra similar to the final endmembers shown in Fig. 3.

are variants or mixtures, and hence, are not restricted to a single material. Therefore, an abundance map calculated from a set of final endmembers usually represents a class of minerals with analogous spectral behavior.

For the unmixing stage, recall that Eq. (1) is an *overdetermined* system of linear equations ($n > m$), subject to the restrictions of full additivity and non-negativity of abundance coefficients as stated in Eq. (2). In the present case, both matrices $\overline{W}$ and $\overline{M}$ have full rank, thus their columns are linear independent vectors. Also, it happens that, the set of final endmembers, $E = \{\overline{\mathbf{w}}^{33}, \overline{\mathbf{m}}^{16}, \overline{\mathbf{w}}^{25}, \overline{\mathbf{w}}^{2}, \overline{\mathbf{m}}^{27}\} \subset \overline{W} \cup \overline{M}$, is a linear independent set whose pseudoinverse matrix is unique. Although, the *unconstrained* solution corresponding to Eq. (1), where $n = 52 > 5 = m$, has a single solution, some coefficients turn out to be negative for many pixel spectra and do not sum up to unity. If full additivity is enforced, again negative coefficients appear. Therefore, the best approach consists of imposing non-negativity for the abundance proportions and simultaneously, relaxing full additivity by considering the inequality $\sum_{k=1}^{m} a_k < 1$. Specifically, we use the *non-negative least squares* (NNLS) algorithm that solves the problem of minimizing $\|E\boldsymbol{a} - \boldsymbol{x}\|_2$ (Euclidean norm) subject to the condition $\boldsymbol{a} \geqslant \mathbf{0}$. The details related to the NNLS algorithm can be found in [16,19]. Fig. 5 illustrates the abundance maps corresponding to three of the five approximate endmembers. The maps shown where obtained with the NNLS method as implemented in Matlab. They have been enhanced for visual clarity by incrementing their brightness and contrast in 15%. The maps correlate well with the standard USGS reference map (after color thresholding) and with the results presented in [35].

5. Conclusions

In this paper we have described a novel method that approximates all potential endmembers of a hyperspectral image after a single scan, from which a subset of representative endmembers can be selected. Our approach is based on a scaled version of the LAMs, W_{XX} and M_{XX}, namely the sets $\overline{W}$ and $\overline{M}$, whose algebraic properties as well as geometric characterization are fundamental to the success of the computational technique derived from it. The computations involved are very fast since

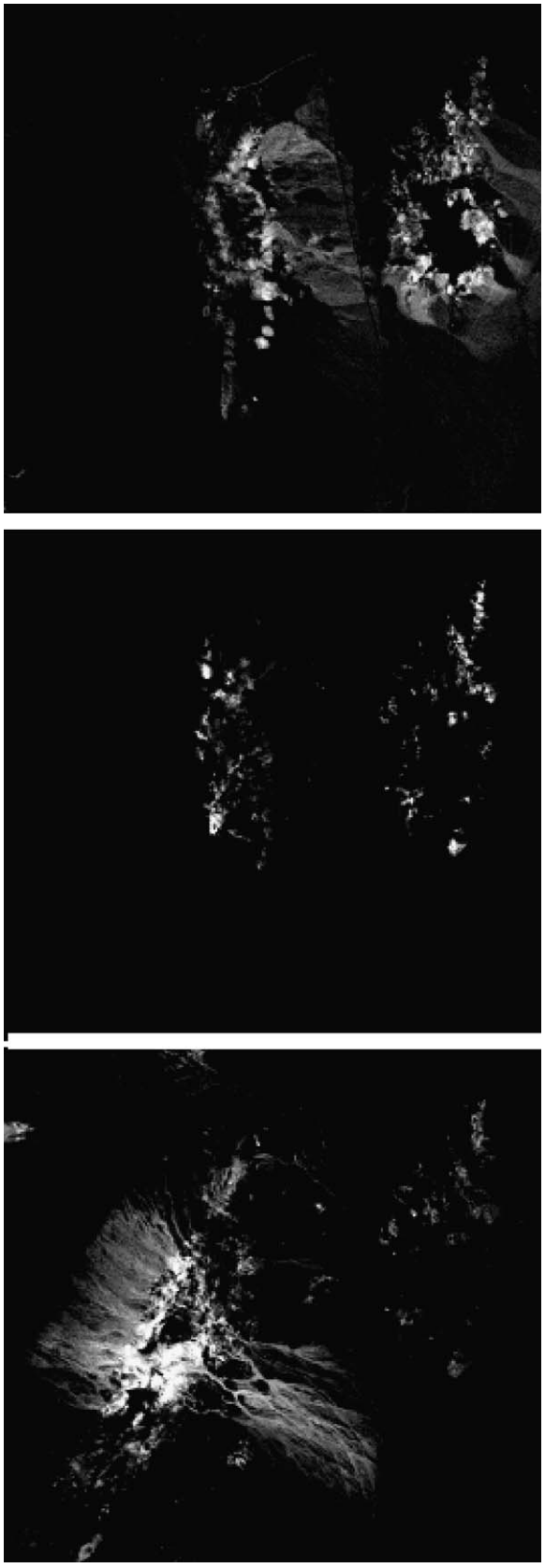


Fig. 5. Top to bottom: abundance maps of alunite, kaolinite, and muscovite. Grayscale values indicate relative abundance proportion; brighter zones signal high mineral content, darker zones signal low content or mineral mixtures.

Table 1
Potential endmember groups and representatives.

$\overline{W}$ columns	Rep. $\overline{\mathbf{w}}^j$	$\overline{M}$ columns	Rep. $\overline{\mathbf{m}}^j$
{1,...,15}	{2*, 5}	{1,...,10}	{1}
{16,...,27}	{25*}	{11,...,20}	{16*}
{28,...,35}	{33*}	{21,...,32}	{26, 27*}
{36,...,41}	{39}	{34,...,41}	{38}
{42,...,48}	{45}	{42,...,48}	{47}
{49,...,52}	{51}	{49,...,52}	{51}

Final endmembers are marked with an asterisk.

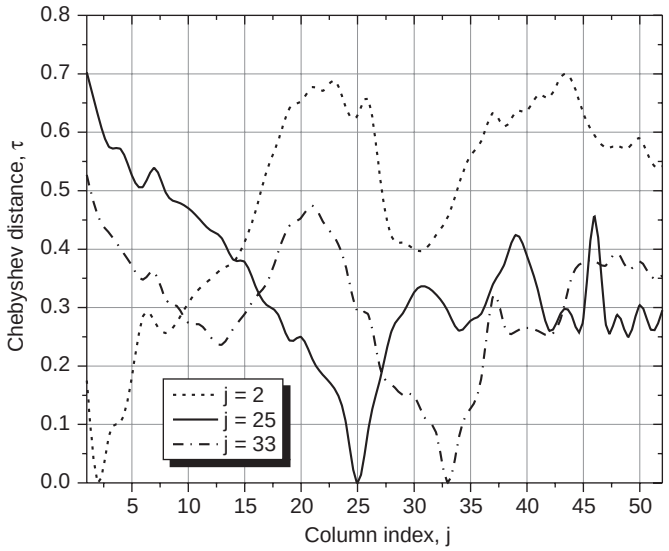


Fig. 6. Chebyshev distance curves for $\overline{\mathbf{w}}^2$, $\overline{\mathbf{w}}^{25}$, and $\overline{\mathbf{w}}^{33}$.

simple arithmetical operations are required. Since the cardinality of either set $\overline{W}$ or $\overline{M}$ equals n, the dimensionality of pixel spectra, the LAMs based algorithm gives $2n$ possible endmembers. The search for final endmembers, although based on quantitative criteria such as minimal Chebyshev distance or minimal spectral angle, is still performed interactively. However, the mathematical theory provided and the AVIRIS application results, demonstrate that our approach is effective and competitive with other available methods. Unsupervised methods based on the same theoretical ground as presented in this paper are described in [15], and an alternative development based on LAMs and endmember linear independence has been proposed recently in [25]. As a future line of research, we will consider dimension reduction techniques on the pixel data set before computation of the LAMs as well as cardinality reduction techniques on the set of potential endmembers, derived from $\overline{W}$ and $\overline{M}$, to find automatic procedures that match final endmembers against library spectra.

Acknowledgments

Gonzalo Urcid thanks the National System of Research (SNI) and the National Council of Science and Technology (CONACyT) in Mexico for partial support through Grant no. 22036. We thank the anonymous reviewers for their critical and constructive suggestions.

Appendix

Given any two vectors $\mathbf{x}^{\xi}, \mathbf{x}^{\gamma} \in \mathbb{R}^n$, the expressions given by

$$\tau(\mathbf{x}^{\xi}, \mathbf{x}^{\gamma}) = \bigvee_{i=1}^{n} |x_i^{\xi} - x_i^{\gamma}|, \tag{23}$$

$$\theta(\mathbf{x}^{\xi}, \mathbf{x}^{\gamma}) = \cos^{-1} \frac{\mathbf{x}^{\xi} \cdot \mathbf{x}^{\gamma}}{\|\mathbf{x}^{\xi}\| \|\mathbf{x}^{\gamma}\|}, \tag{24}$$

where $\|\cdot\|$ denotes Euclidean norm, define, respectively, the *Chebyshev distance*, Eq. (23), and the *angle* (in radians) between the two vectors, Eq. (24). Each expression is symmetric in its arguments for $\gamma \neq \xi$ and, if $\gamma = \xi$, then $\tau = 0$ and $\theta = 0$. Two vectors are said to be "similar" if τ or θ are less than a given tolerance error. In our application example, to discard correlated potential endmembers, Eqs. (23) and (24) were computed between every pair of column vectors of $\overline{W}$ as well as between columns of $\overline{M}$, for $\xi = 1, \ldots, 51$ and $\gamma = \xi + 1, \ldots, 52$.

Vector grouping based on distance values and calculated angles is shown in Table 1, where each set of indices indicates strong correlation between corresponding patterns. In a second step, representative endmembers were selected from each group based on additional criteria such as, overall pattern shape, number of local minima and maxima, and amplitude range. The third and last step involves a matching process between available library spectra and representative endmembers from which the set of final endmembers is obtained. As an example, Fig. 6 displays the Chebyshev distance curves of each final endmember selected from $\overline{W}$. Notice that, $\tau(\overline{\mathbf{w}}^2, \overline{\mathbf{w}}^j)$ for $j \in [1, 15]$, $\tau(\overline{\mathbf{w}}^{25}, \overline{\mathbf{w}}^j)$ for $j \in [16, 27]$, and $\tau(\overline{\mathbf{w}}^{33}, \overline{\mathbf{w}}^j)$ for $j \in [28, 52]$, are all less than 0.4. Similarly, $\tau(\overline{\mathbf{m}}^{16}, \overline{\mathbf{m}}^j)$ for $j \in [2, 20]$ and $\tau(\overline{\mathbf{m}}^{27}, \overline{\mathbf{m}}^j)$ for $j \in [21, 46]$, are also less than 0.4.

References

[1] J.B. Adam, M.O. Smith, P.E. Johnson, Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander 1 site, J. Geophys. Res. 91 (B8) (1986) 8098–8112.

[2] G. Birkhoff, Lattice Theory, third edn, American Mathematical Society, Providence, RI, 1967.

[3] J.W. Boardman, Automated spectral unmixing of AVIRIS data using convex geometry concepts, in: AVIRIS Workshop Proceedings, JPL Publication, 1993, vol. 1, no. 4, pp. 11–14.

[4] J.W. Boardman, Analysis, understanding and visualization of hyperspectral data as convex sets in n-space, in: Imaging Spectrometry, Proceedings of SPIE, vol. 2480, 1995, pp. 14–22.

[5] N. Bourbaki, Affine Spaces and Projective Spaces, in: Algebra I, Elements of Mathematics, Springer, Berlin, 1989.

[6] R. Cuninghame Green, Minimax Algebra, in: Lectures Notes in Economics and Mathematical Systems, vol. 166, Springer, New York, NY, 1979.

[7] R. Cuninghame Green, Minimax algebra and applications, in: P. Hawkes (Ed.), Advances in Imaging and Electron Physics, vol. 90, Academic Press, New York, NY, 1995, pp. 1–121.

[8] W.H. Farrand, Hyperspectral remote sensing of land and the atmosphere, in: B. Gunther, D. Steel (Eds.), Encyclopedia of Modern Optics, vol. 1, Academic Press, San Diego, CA, 2005, pp. 395–403.

[9] J. Gallier, Geometric Methods and Applications for Computer Science and Engineering, in: Texts in Applied Mathematics, vol. 38, Springer, New York, 2001.

[10] M. Graña, M.J. Gallego, Associative morphological memories for spectral unmixing, in: Proceedings of the 11th European Symposium on Artificial Neural Networks, 2003, pp. 481–486.

[11] M. Graña, J. Gallego, Associative morphological memories for endmember induction, in: Proceedings of the IEEE, International Geoscience and Remote Sensing Symposium, vol. 6, 2003, pp. 3757–3759.

[12] M. Graña, J. Gallego, Hyperspectral image analysis with associative morphological memories, in: Proceedings of the IEEE, International Conference on Image Processing, vol. 2, 2003, pp. III-549–552.

[13] M. Graña, J.L. Jiménez, C. Hernández, Lattice independence, autoassociative morphological memories and unsupervised segmentation of hyperspectral images, in: Proceedings of the 10th Joint Conference on Information Sciences, NC-III, 2007, pp. 1624–1631.

[14] M. Graña, P. Sussner, G.X. Ritter, Associative morphological memories for endmember determination in spectral unmixing, in: Proceedings of the IEEE, 12th International Conference on Fuzzy Systems, vol. 2, 2003, pp. 1285–1290.

[15] M. Graña, I. Villaverde, J.O. Maldonado, C. Hernández, Two lattice computing approaches for the unsupervised segmentation of hyperspectral images, Neurocomputing (2009), this issue, doi:10.1016/j.neucom.2008.06.026.

[16] Kh.D. Ikramov, M. Matin far, Computer-algebra implementation of the least squares method on the nonnegative orthant, J. Math. Sci. 132 (2) (2006) 156–159.

[17] N. Keshava, A survey of spectral unmixing algorithms, Lincoln Lab. J. 14 (1) (2003) 55–78.

[18] N. Keshava, J.F. Mustard, Spectral unmixing, IEEE Signal Process. Mag. 19 (1) (2002) 44–57.

[19] C.L. Lawson, R.J. Hanson, Solving Least Squares Problems, Prentice-Hall, Englewood Cliffs, NJ, 1974.

[20] D.S. Myers, Hyperspectral endmember detection using morphological autoassociative memories, Master Thesis, University of Florida, Gainesville, FL, 2005.

[21] G.X. Ritter, Lattice algebra & Minimax algebra, in: Image Algebra, Unpublished Manuscript available via anonymous ftp, CISE Department, University of Florida, Gainesville, FL, 1999.

[22] G.X. Ritter, P. Gader, Fixed Points of Lattice Transforms and Lattice Associative Memories, in: P. Hawkes (Ed.), Advances in Imaging and Electron Physics, vol. 144, Academic Press, San Diego, CA, 2006, pp. 165–242.

[23] G.X. Ritter, P. Sussner, Associative memories based on lattice algebra, in: Proceedings of IEEE, International Conference on Systems, Man, and Cybernetics, 1997, pp. 3570–3575.

[24] G.X. Ritter, P. Sussner, J.L. Díaz de León, Morphological associative memories, IEEE Trans. Neural Networks 9 (2) (1998) 281–293.

[25] G.X. Ritter, G. Urcid, Fast autonomous endmember determination using lattice algebra, IEEE Trans. Pattern Anal. Mach. Intell., submitted for publication.

[26] G.X. Ritter, G. Urcid, L. Iancu, Reconstruction of patterns from noisy inputs using morphological associative memories, J. Math. Imaging Vision 19 (2) (2003) 95–111.

[27] D.A. Roberts, M. Gardner, R. Church, S. Ustin, G. Sheer, R.O. Green, Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models, Remote Sensing Environ. 65 (3) (1998) 267–279.

[28] R. Selfridge, A small part of lattice theory, ACM SIGAPL APL Quote Quad 35 (3) (2007) 24–29.

[29] L.A. Skornjakov, Elements of Lattice Theory, Adam Hilger, Bristol, UK, 1977.

[30] P. Sussner, Fixed points of autoassociative morphological memories, in: Proceedings of the International Joint Conference on Neural Networks, 2000, pp. 611–616.

[31] P. Sussner, M.E. Valle, Gray-scale morphological associative memories, IEEE Trans. Neural Networks 17 (3) (2006) 559–570.

[32] G. Urcid, J.C. Valdiviezo, Generation of lattice independent vector sets for pattern recognition applications, in: Mathematics of Data/Image Pattern Recognition, Compression, Coding, and Encryption X with Applications, Proceedings of SPIE, vol. 6700, 2007, pp. 67000C:1–12.

[33] M.E. Winter, Fast autonomous spectral endmember determination in hyperspectral data, in: Proceedings of the 13th International Conference on Applied Remote Sensing, vol. 2, 1999, pp. 367–344.

[34] M.E. Winter, An algorithm for fast autonomous spectral endmember determination in hyperspectral analysis, in: Imaging Spectrometry, Proceedings of SPIE, vol. 3753, 1999, pp. 266–275.

[35] M.E. Winter, Comparison of approaches for determining end-members in hyperspectral data, in: Proceedings of IEEE, Aerospace Conference, vol. 3, 2000, pp. 305–313.

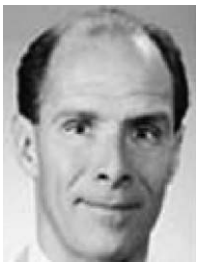
[36] J.C. Valdiviezo, G. Urcid, Hyperspectral endmember detection based on strong lattice independence, in Applications of Digital Image Processing XXX, Proceedings of SPIE, vol. 6696, 2007, 669625:1–12.

[37] G. Vane, R.O. Green, T.G. Chrien, H.T. Enmark, E.G. Hansen, W.M. Porter, The airborne visible/infrared imaging spectrometer (AVIRIS), Remote Sensing Environ. 44 (1993) 127–143.

Gerhard X. Ritter received the B.A. (1966) and Ph.D. (1971) degrees from the University of Wisconsin-Madison. He is currently Professor Emeritus of Computer Science of the Computer and Information Science and Engineering Department (CISE), the Director of the Center for Computer Vision and Visualization, and Professor Emeritus of Mathematics at the University of Florida. Dr. Ritter is the Chair of the Society of Industrial and Applied Mathematics (SIAM) Activity Group in Imaging Science and of the American Association of Engineering Societies (AAES) R & D task Force. He is the Editor-in-Chief of the *Journal of Mathematical Imaging and Vision*, and a member of the Editorial Boards for both the *Journal of Electronic Imaging* and the *Journal of Pattern Analysis and Applications*. Since 1995 he is a Fellow of SPIE and he was the recipient of the 1998 General Ronald W. Yates Award for Excellence in Technology Transfer by the Air Force Research Laboratory and the 1989 International Federation for Information Processing (IFIP) Silver Core Award. He is the author of two books and more than 90 refereed publications in computer vision, mathematics, and neural networks. His current research interests include artificial

neural networks, pattern recognition, and the mathematical foundations of image processing and computer vision.

Gonzalo Urcid received his B.E. (1982) and M.Sc. (1985) both from the University of the Americas and his Ph.D. (1999) in optics from the National Institute of Astrophysics, Optics, and Electronics (INAOE), Tonantzintla, Mexico. He is a Research Scientist in the Optics Department at INAOE since 2003. He holds the appointment of National Researcher from the Mexican National Council of Science and Technology (SNI-CONACYT) since 2001. His present research interests include applied mathematics, artificial neural networks, and optical-digital image processing.



Mark S. Schmalz received his Ph.D. (1996) degree from the University of Florida. He is currently an Associate Scientist of the Computer and Information Science and Engineering Department (CISE). His current research interests are in massively parallel computing, automated processing and understanding of imagery and natural language, data compression, and cryptology. He is also among the CISE faculty affiliated with UF's Digital Arts and Sciences program, with research interests in psychoacoustics and computer-generated music. He has authored or co-authored over 140 research papers in open conference proceedings and journals, and currently is authoring two books.