Hybrid Sparse Linear and Lattice Method for Hyperspectral Image Unmixing

Ion Marques, Manuel Graña

Computational Intelligence Group, UPV/EHU[1]

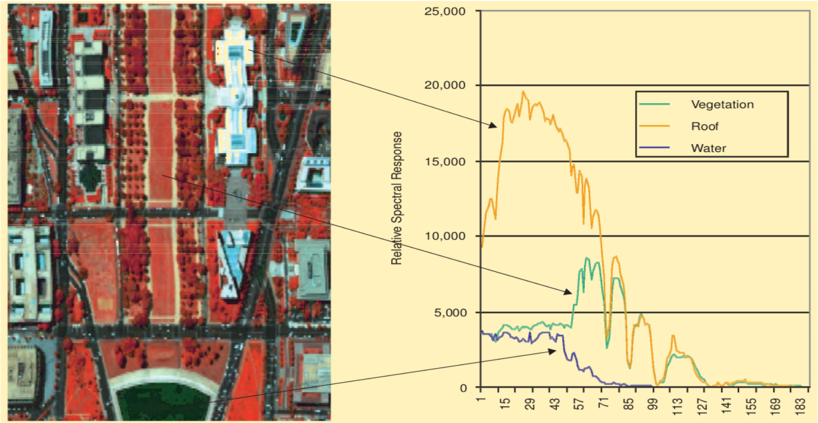
June 10, 2014

[1]

Summary

- The Linear Mixing Model (LMM): pixel spectra are affine combinations of endmembers.
- The WM algorithm: endmembers extracted from the rows and columns of dual Lattice Autoassociative Memories (LAAM) built on the image spectra.
  - The number of endmembers is huge.
- Additional endmember selection steps
  - Clustering
  - Unmixing with linear sparse regression techniques: Conjugate Gradient Pursuit (CGP)

Hyperspectral data[a]



[a] D. Landgrebe, «Hyperspectral image data analysis», Signal Processing Magazine, IEEE, vol. 19, nº. 1, pp. 17–28, 2002.

Contents

Contents

Introduction

- The Linear Mixing Model (LMM)

$$\mathbf{x} = \mathbf{E}\alpha + \mathbf{n} \quad (1)$$

- set of endmembers, $\mathbf{E} = \{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_q\}$.
- α is an $q \times 1$ vector of fractional abundances resulting from the unmixing process.

Introduction

The WM algorithm is a Lattice Computing based EIA finding a collection of affine independent vectors that define a convex polytope covering the data of the image in high dimensional spectral space. The algorithm is very fast, using only lattice operators and the resulting endmember set has a direct relation with the image data. However, it has the inconvenient of producing too many endmembers, which are strongly correlated. Therefore, some endmember selection method is needed to find the relevant endmembers which produce the most parsimonious explanation of the data. In this paper we propose a clustering step followed by the application of greedy sparse methods, based on gradient pursuit .

Introduction

The aim of the sparse methods is to find the minimal set of contributions from a dictionary that make up the data with minimal loss. Formally, if we denote a sparse fractional abundance vector α, the unmixing problem is then

$$\min_{\alpha} \|\alpha\|_0 \text{ subject to } \|\mathbf{x} - \mathbf{E}\alpha\|_2 \leq \delta,\ \alpha \geq 0,\ \mathbf{1}^T\alpha = 1, \quad (2)$$

where $\mathbf{1}^T$ is a line vector of 1's, $\|\alpha\|_0$ denotes the number of nonzero components of α, and $\delta \geq 0$ is the error tolerance.

Contents

WM algorithm

1. Compute $\mathbf{v} = [v_1, \ldots, v_L]$ and $\mathbf{u} = [u_1, \ldots, u_L]$,
$$v_k = \min_{\xi} x_k^{\xi}; u_k = \max_{\xi} x_k^{\xi}$$
for all $k = 1, \ldots, L$ and $\xi = 1, \ldots, N$,
2. Compute the LAAMs
$$\mathbf{W}_{XX} = \bigwedge_{\xi=1}^{N} \left[\mathbf{x}^{\xi} \times \left(-\mathbf{x}^{\xi} \right)' \right]; \mathbf{M}_{XX} = \bigvee_{\xi=1}^{N} \left[\mathbf{x}^{\xi} \times \left(-\mathbf{x}^{\xi} \right)' \right]$$
where $\times$ is any of the $\boxtimes$ or $\boxtimes$ operators.
3. Build $W = \left\{ \mathbf{w}^1, \ldots, \mathbf{w}^L \right\}$ and $M = \left\{ \mathbf{m}^1, \ldots, \mathbf{m}^L \right\}$ such that
$$\mathbf{w}^k = u_k + \mathbf{W}^k; \mathbf{m}^k = v_k + \mathbf{M}^k; k = 1, \ldots, L.$$
4. Return the set $V = W \cup M \cup \{\mathbf{v}, \mathbf{u}\}$.

Contents

Endmember selection via k-means

- To induce a reduced set of endmembers $\mathbf{E}^{\star} \subset \mathbf{E}$, we perform k-means
- Two "closeness" measurements in the clustering process:
    1. Pearson correlation
    $\mathbf{dist}(\boldsymbol{x}, \boldsymbol{y}) = 1 - \mathbf{corr}(\boldsymbol{x}, \boldsymbol{y}) = 1 - \frac{\sum_{i=1}^{n}(x_i - \mu_x)(y_i - \mu_y)}{(n-1)\sigma_x \sigma_y}$.
    2. Cosine dissimilarity
    $\mathrm{b}\mathbf{dist}(\boldsymbol{x}, \boldsymbol{y}) = 1 - \cos\theta = 1 - \frac{\sum_{i=1}^{n}(x_i \cdot y_i)}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}}$.
        1. Each centroid is the component-wise mean of the points in that cluster, after centering and normalizing those points to zero mean and unit standard deviation.
- We perform the clustering I times, selecting random initial cluster points at each iteration.
- The set $\mathbf{E}^{\star}$ will consist of the endmembers that are closer to the centroids of said clusters.

Sparse unmixing

- data matrix $\mathbf{X}$.
- dictionary matrix $\Phi \in \mathbb{R}^{q \times L}$.
  - The q columns of Φ are referred as atoms: induced endmembers $\Phi = \mathbf{E}^{\star}$
- find a mixing matrix $\mathbf{M}$

$$\mathbf{X} = \Phi\mathbf{M} + \varepsilon, \quad (3)$$

optimizing a sparsity measure. $\mathbf{M}$ collection of abundance images obtained by the unmixing process.

Conjugate Gradient Pursuit

- $\mathbf{r}^0 = X$ is the initial residual error. $\Gamma^0 = \emptyset$ is an index set. $\mathbf{y}^0_{\Gamma^0} = 0$ is the initial set of output sparse vectors. $b_0 = 1$ is a term needed to calculate new conjugate gradients.
- We denote $\mathbf{D}_{\Gamma^i}$ the matrix containing all conjugate update directions from iteration $i - 1$, with an additional row all zeros.

Conjugate Gradient Pursuit

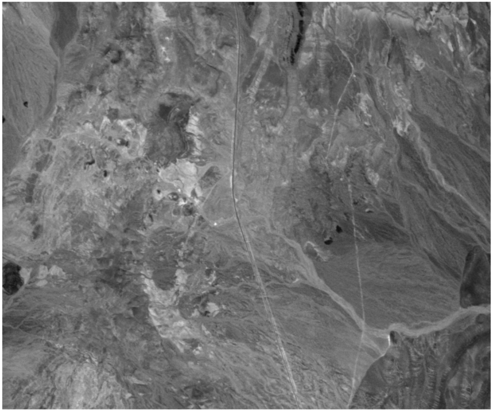
1. For $i = 1, 2, 3, ...$ until stopping criterion is met:
   1. Calculate gradient $\mathbf{g}$ for $\mathbf{y}$ restricted to Γ^i: $\mathbf{g}_{\Gamma^i} = \Phi_{\Gamma^i}^T \left(X - \Phi_{\Gamma^i} \mathbf{y}_{\Gamma^i}^{i-1} \right)$.
   2. Select the best element index: $\gamma^i = \arg_j \max \left| \mathbf{g}_{\Gamma^i} \right|$.
   3. Update the index set: $\Gamma^i = \Gamma^{i-1} \cup \gamma^i$.
   4. Calculate the gram matrix $\mathbf{G}_{\Gamma^i} = \Phi_{\Gamma^i}^T \Phi_{\Gamma^i}$.
   5. We calculate the update direction $\mathbf{d}_{\Gamma^i} = b_0 \mathbf{g}_{\Gamma^i} + \mathbf{D}_{\Gamma^i} \mathbf{b}$ where $\mathbf{b} = \left(\mathbf{D}_{\Gamma^i}^T \mathbf{G}_{\Gamma^i} \mathbf{D}_{\Gamma^i} \right)^{-1} \left(\mathbf{G}_{\Gamma^i}^T \mathbf{D}_{\Gamma^i} \mathbf{g}_{\Gamma^i} \right)$,
   6. Calculate new set of vectors $\mathbf{y}_{\Gamma^i}^i = \mathbf{y}_{\Gamma^i}^{i-1} + a^i \mathbf{d}_{\Gamma^i}$. where $\mathbf{c}^i = \Phi_{\Gamma^i} \mathbf{d}_{\Gamma^i}$, and $a^i = \frac{\left\langle \mathbf{r}^i, \mathbf{c}^i \right\rangle}{\|\mathbf{c}^i\|_2^2}$,
   7. Calculate new residual error$\mathbf{r}^i = \mathbf{r}^{i-1} - a^i \mathbf{c}^i$.
2. Output $\mathbf{r}$ and $\mathbf{y}$.

Contents

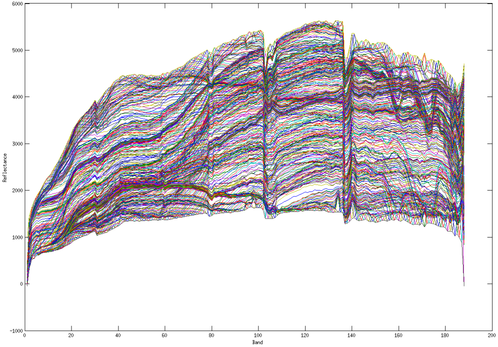
Experimental design

- A sub-image of the AVIRIS Cuprite dataset.
- 512×614 pixels.
- 224 spectral bands between 0.4 μm and 2.5 μm, with spectral resolution of 10 nm.
- preprocessing leaves 187 spectral bands.
- The Cuprite site is well understood mineralogically and is broadly used as a trusted benchmark for hyperspectral research.

Cuprite

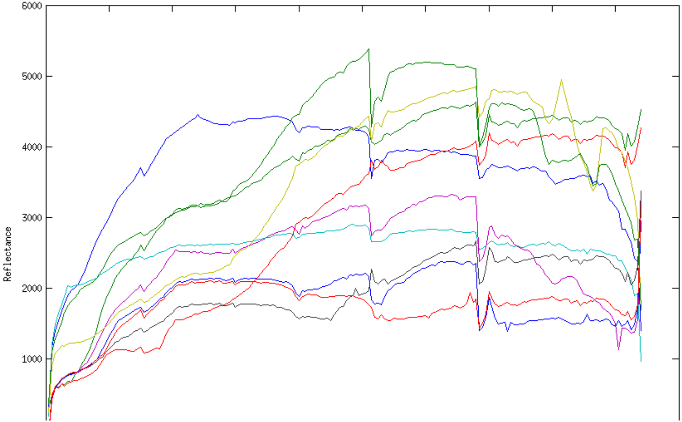


WM endmembers



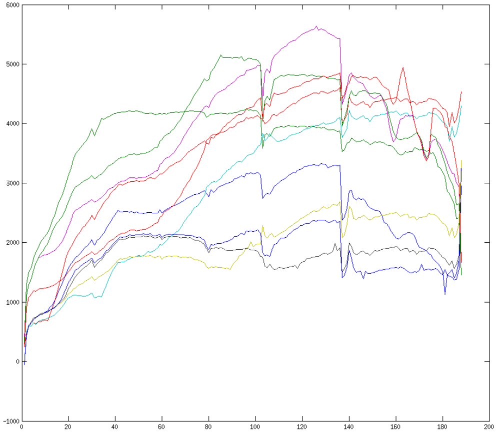
Clustering of WM endmembers

Endmembers obtained using $k = 10$ and Pearson correlation distance

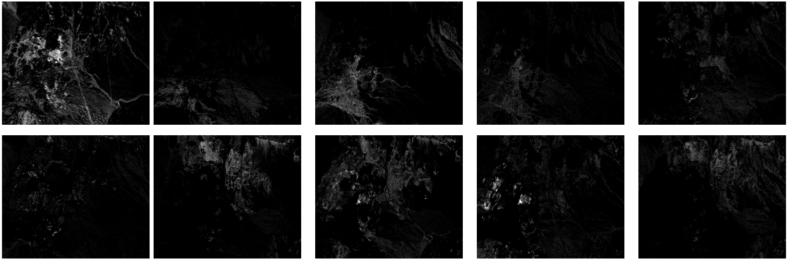


Clustering of WM endmembers

Endmembers obtained using $k = 10$ and cosine similarity.



Sparse linear unmixing abundances



Conclusions

- Ritter's WM Algorithm basic endmember induction
  - clustering step to reduce the number of endmember prior to the unmixing process.
  - CGP algorithm to calculate sparse abundances.
- Experiments on a complex and well documented hyperspectral image s
- The visual assessment of the results disjoint endmembers that are present in disparate abundances on the pixels of the scene.
- Future work
  - selecting k with some unsupervised criterion,
  - comparing our unmixing results with those obtained using USGS Spectral Library as the sparse algorithm's dictionary.